

N° d'ordre : 27/2017-C/MT

**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<sup>ème</sup> Cycle**

**En : Mathématiques**

**Spécialité : Probabilités et Statistique**

**Option : Mathématiques Financières et Actuariat**

**Par : MENNI Nassira**

**Sujet**

**Estimateur à noyau de la régression  
dans les modèles censurés et associés**

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

Mr. Kamal BOUKHETALA	Professeur	à l'U.S.T.H.B, Algérie	Président
Mr. Abdelkader TATACHAK	Professeur	à l'U.S.T.H.B, Algérie	Directeur de thèse
Mr. Mohamed Kadi ATTOUCH	Professeur	à l'Univ. Sidi Bel-Abbès, Algérie	Examinateur
Mme. Zohra GUESSOUM	MC.A	à l'U.S.T.H.B, Algérie	Examinatrice
Mr. Khaled KHALDI	Professeur	à l'Univ. Boumerdès, Algérie	Examinateur
Mr. Elias OULD SAÏD	Professeur	à l'Univ. du Littoral, France	Examinateur

---

## Remerciements

*À mon père, à ma mère  
et à mes chers frères.*

Tout d'abord, je remercie Dieu le tout puissant de m'avoir donné le courage, la volonté et la force d'être là où je suis aujourd'hui et d'avoir mener à bien ce travail.

Je tiens à remercier très vivement mon directeur de thèse Monsieur **Abdelkader Tatachak** pour m'avoir encadré durant ces années de thèse. Ses qualités humaines, son aide et son soutien constant m'ont permis de mener à bien cette thèse. Je ne pourrai jamais le remercier assez. Qu'il trouve ici l'expression de ma profonde gratitude.

J'exprime ma gratitude à Monsieur **Kamal Boukhetala**, Professeur à l'USTHB, pour l'honneur qu'il me fait en présidant le jury de cette thèse.

Je tiens à exprimer ma sincère reconnaissance à Monsieur **Elias Ould Saïd**, MCF-HDR à l'université du Littoral, pour l'honneur qu'il m'a fait en acceptant d'examiner mon travail.

Je remercie très chaleureusement Messieurs **Mohammed Kadi Attouch**, Professeur à l'université de Sidi Bel-Abbès, et **Khaled Khaldi**, Professeur à l'université M'Hamed Bougara de Boumerdes, pour avoir accepté d'examiner ce travail et faire partie du jury. Je leur suis très reconnaissante du temps qu'ils ont consacré à l'évaluation de ce travail.

Mes vifs remerciements sont également adressés à Madame **Zohra Guessoum**, Docteur à l'USTHB, pour l'intérêt qu'elle a porté à mon travail depuis mes premiers pas, et je suis très sensible à sa présence au sein du jury.

A toutes mes copines **Nassima, Farida, Wafaa, Asma, Sara, Dalila, Mariam, Mouna, Saliha, Amira, Naima** qui ont toujours été là pour moi, notamment pour surmonter les épreuves de la vie et me remotiver dans les moments difficiles, et en particulier à ma copine et ma sœur **Latifa**, son aide, ses conseils et son soutien m'ont été précieux.

Enfin, je souhaite de tout cœur remercier ma famille qui a su me faire une confiance inconditionnelle et sans qui de longues études n'auraient jamais été envisageables, plus particulièrement mon père, à qui je serais reconnaissante jusqu'à la fin de mes jours, car c'est grâce à lui que j'ai pu accomplir ce travail et arriver où j'en suis aujourd'hui. Sans oublier ma mère, qui est ma raison d'être et qui veille toujours sur mon bien-être. Je remercie aussi mes frères et sœurs **Saada, Fatiha, Nora**,

---

Sara, Samir, Madjed ainsi que mes beaux frères, ma belle sœur Hafida, mes neveux et nièces pour leur amour et leur soutien sans faille.

Je les remercie tous de m'avoir supportée et encouragée pendant les moments de doute. Je n'aurai jamais pu faire cette thèse sans eux.

---

# Résumé

Dans cette thèse, nous nous intéressons à l'estimateur non paramétrique de la fonction de régression dans le cas de données censurées à droite et présentant une forme de dépendance appelée association . Dans une première partie nous rappelons brièvement les résultats établis sur le comportement asymptotique de quelques estimateurs non paramétriques de la fonction de régression, dans le cas de données complètement observées i.i.d. et présentant des structures de dépendance ( $\alpha$ -mélange et associées), aussi dans le cas où les données ne sont pas complètement observées (censurées à droite) dans le cas i.i.d. et  $\alpha$ -mélange .

Dans une second partie, nous présentons nos résultats en considérant une suite de variable aléatoire (v.a.)  $\{Y_i, i \geq 1\}$  stationnaire associée de fonction de répartition  $F$ , censurée à droite par une suite  $\{C_i, i \geq 1\}$  supposée indépendante entre elles et indépendantes des variables aléatoires  $\{Y_i, i \geq 1\}$ , et la covariable  $\{X_i, i \geq 1\}$  associée, de manière à observer le triplet  $(X_i, Z_i, \delta_i)$  où  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ , et on estime la fonction de régression  $r(x) = \mathbb{E}[Y|X = x]$ .

Le premier résultat obtenu concerne la convergence uniforme presque sûre de l'estimateur en question, ces derniers ont fait l'objet d'un article publié dans Statistical Papers.

Comme deuxième résultat, nous établissons la normalité asymptotique de l'estimateur étudié et nous obtenons des intervalles de confiance.

Afin d'illustrer la partie théorique de la thèse, nous présentons des simulations pour les deux résultats obtenus.

**Mots-clés** Association, Censure droite, Convergence uniforme presque sûre, Estimation de Kaplan-Meier, Estimateur à noyau, Régression non paramétrique, Normalité asymptotique.

---

# Abstract

In this thesis, we are interested to the nonparametric estimator of the regression function in the case of the right censoring data and presenting a form of dependence called association. In a first part we call back briefly some results established on the asymptotic behavior of some nonparametric estimators of the regression function, in the case of completely observed data i.i.d and presenting structure of dependence ( $\alpha$ -mixing et association), then we go on with the case where the data are censored on the right in the case i.i.d and  $\alpha$ -mixing.

In the second part, we present our results considering a strictly stationary and associated sequence of random variables r.v.  $\{Y_i, i \geq 1\}$  with distribution function  $F$ , censored on the right by a sequence  $\{C_i, i \geq 1\}$  supposed to be independent between them and independent from the r.v.  $\{Y_i, i \geq 1\}$ . The covariable  $\{X_i, i \geq 1\}$  is associated, in a way that we observe the triplet  $(X_i, Z_i, \delta_i)$  where  $Z_i = Y_i \wedge C_i$  and  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ , then we estimate the regression function  $r(x) = \mathbb{E}[Y|X = x]$ .

The first result concerns the uniform convergence almost sure of the considered estimator. These results are the object of a paper published in Statistical Papers.

As a second result, we establish the asymptotic normality of the studied estimator and we obtain the confidence intervals.

To illustrate the theoretical part of the thesis we present simulation for both obtained results.

**Keywords** Association, right censoring, almost sure uniform convergence, Kaplan-Meier estimator, kernel estimator, nonparametric regression, asymptotic normality.

---

## Publications

1. Menni, N., Tatachak, A. (2016) A note on estimating the conditional expectation under censoring and association : strong uniform consistency. *Journal of Statistical Papers*, DOI 10.1007/s00362-016-0801-8.

2. Menni, N., Tatachak, A. (under consideration) Asymptotic normality for a regression function estimate under right censored data and association.

## Communications internationales

1. Menni, N., Tatachak, A. Simulation of the kernel estimator of the regression function for right censored data. Colloque International (MSS 2014), Alger (USTHB), Algérie.

2. Menni, N., Tatachak, A. Estimateur à noyau de la fonction de régression dans un modèle de censure à droite associés. Colloque Internationale de statistiques et ses applications (ICSA 2015) du 06 au 07 mai 2015, Saida, Algérie.

3. Menni, N., Tatachak, A. Normalité asymptotique de l'estimateur à noyau de la fonction de régression pour des données censurées associées. Congrès de la Société Mathématique algérienne (CMA 2016) du 08 au 09 novembre 2016, Batna, Algérie.

# Table des matières

<b>1</b>	<b>Introduction générale</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.1.1	Estimation de la fonction de densité . . . . .	3
1.2	Données incomplètes . . . . .	5
1.2.1	Durée de vie . . . . .	5
1.2.1.1	Censure . . . . .	5
1.2.1.2	Troncature . . . . .	7
1.2.2	Caractérisation de la loi d'une durée de vie $Y$ . . . . .	7
1.2.3	Estimation de la fonction de survie . . . . .	9
1.2.3.1	Estimateur de Nelson-Aalen . . . . .	9
1.2.3.2	Estimateur de Kaplan-Meier . . . . .	10
1.3	Estimation de la fonction de régression pour des données complètes . . . . .	11
1.4	Estimation de la fonction de régression pour des données incomplètes . . . . .	13
1.4.1	censure à droite . . . . .	13
1.5	Données Dépendantes . . . . .	14
1.5.1	Notions de mélange . . . . .	14
1.5.1.1	Les conditions de mélange fort . . . . .	15
1.5.2	Association . . . . .	16
1.5.2.1	Un bref historique sur l'association . . . . .	17
1.5.2.2	Quelques formes d'associations . . . . .	18
1.6	Présentation des résultats existants . . . . .	21
1.6.1	Cas des données complètes . . . . .	21
1.6.1.1	Le cas i.i.d. . . . .	21
1.6.1.2	Le cas $\alpha$ -mélange . . . . .	21
1.6.1.3	Le cas associé . . . . .	21
1.6.2	Cas des données incomplètes . . . . .	21
1.6.2.1	Données tronquées à gauche . . . . .	21
1.6.2.2	Données censurées à droite . . . . .	22
1.7	Apport de la thèse . . . . .	26

1.7.1	Présentation des résultats obtenus . . . . .	26
<b>2</b>	<b>Convergence uniforme presque sûre</b>	<b>28</b>
2.1	Introduction . . . . .	29
2.2	Notations and main results . . . . .	31
2.2.1	Assumptions . . . . .	32
2.2.2	Comments on the assumptions . . . . .	32
2.2.3	Main results . . . . .	33
2.3	Simulation results . . . . .	34
2.3.1	Models and procedure . . . . .	34
2.3.2	Comments on the simulation results . . . . .	35
2.4	Proofs . . . . .	36
<b>3</b>	<b>Normalité asymptotique</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Définitions et Notations . . . . .	47
3.3	Hypothèses et résultats . . . . .	48
3.3.1	Discussion sur les hypothèses . . . . .	49
3.3.2	Principaux résultats . . . . .	49
3.4	Application aux intervalles de confiance . . . . .	50
3.5	Simulations . . . . .	50
3.5.1	Normalité asymptotique . . . . .	51
3.5.2	Intervalles de confiance . . . . .	54
3.6	Preuves . . . . .	56
<b>4</b>	<b>Conclusion et perspectives</b>	<b>80</b>
4.1	Conclusion . . . . .	80
4.2	Perspectives . . . . .	81

# Chapitre 1

## Introduction générale

### 1.1 Introduction

L'analyse de régression, dont le nom est dû à Sir Francis Galton (1885), est un domaine très vaste et très important dans l'analyse statistique non paramétrique. Le problème de l'estimation en utilisant la régression a une longue histoire comme l'atteste la littérature. Déjà en 1632, Galileo Galilei utilisa une procédure qui peut être interprétée comme étant un ajustement linéaire de données. Un tel ajustement d'un nuage de points par une droite est le principe du problème classique de la régression linéaire. Une solution à ce problème est fournie par le célèbre principe des moindres carrés, mis en oeuvre indépendamment par A.M. Legendre et C.F. Gauss, et publié en (1805) et (1809), respectivement. Les premiers qui ont construit un estimateur à noyau de la fonction de régression sont Nadaraya (1964) et Watson (1964), ensuite Stone (1977) a établi un résultat de consistance faible de l'estimateur de la fonction de régression. Devroye (1978) a établi la convergence uniforme presque sûre de l'estimateur de Nadaraya-Watson (1964). Récemment l'estimateur de la fonction de régression a connu un réel essor en statistique non paramétrique, on cite l'ouvrage de Schimek(2000) qui fournit beaucoup de connaissances sur le sujet.

Le modèle de la régression linéaire simple est l'un des modèles les plus fréquemment rencontrés en statistique non paramétrique, il décrit la relation entre deux variables aléatoires, il est de la forme suivante :

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, \dots, n,$$

où  $X_i$  représente la variable aléatoire explicative,  $Y_i$  la variable aléatoire expliquée, et  $\varepsilon_i$  est une suite d'erreurs indépendantes des  $X_i$ . Cette méthode possède l'avantage d'être facile à interpréter et, lorsque les postulats sur les résidus  $\varepsilon_i$  sont vérifiés, permet de faire des tests d'hypothèses statistiques formels sur les paramètres. Par contre, il arrive que les linéarités de la relation ne soient pas toujours respectées. Dans ce cas, il est préférable de choisir un modèle plus flexible qui reflète mieux la relation entre  $X$  et  $Y$ . Le modèle de régression non paramétrique suivant peut alors être employé :  $Y_i = r(X_i) + \varepsilon_i, i = 1, \dots, n$ , où  $r(X_i)$  représente la moyenne conditionnelle de la courbe de régression,

c'est-à-dire  $r(x) = \mathbb{E}[Y|X = x]$ , et où les résidus  $\varepsilon_i$  représentent la variation de  $Y$  autour de  $r(x)$ . Les postulats sur les termes d'erreur  $\varepsilon_i$  sont les mêmes que ceux du modèle linéaire et, à part certaines hypothèses de continuité et de lissage, il n'y a habituellement aucune contrainte associée à  $r(x)$ . Ainsi soit la définition :

$$\begin{aligned} r(x) &= \arg \min_a \mathbb{E}[(Y - a)^2|X = x] \\ &= \mathbb{E}[Y|X = x]. \end{aligned}$$

La preuve de cette égalité est trouvée en différenciant l'espérance  $\mathbb{E}[(Y - a)^2|X = x]$  par rapport à  $a$ , ce qui permet d'avoir

$$\begin{aligned} \frac{\partial}{\partial a} \mathbb{E}[(Y - a)^2|X = x] &= -2\mathbb{E}[(Y - a)|X = x] \\ &= -2\mathbb{E}[Y|X = x] + 2a. \end{aligned}$$

Donc  $a = \mathbb{E}[Y|X = x]$ .

Le fait que la dérivée seconde, soit positive mène à la conclusion que cette valeur  $a$  est bien un minimum.

L'objectif principal de notre étude est l'estimation de la fonction de régression pour des données incomplètes.

Dans le cas de troncature on peut citer le papier de Ould saïd et Lemdani (2006) sur l'estimateur à noyau de la fonction de régression pour un modèle tronqué à gauche avec des données indépendantes et identiquement distribuées (i.i.d.)

Plus précisément dans notre travail on s'intéresse aux censures à droite, dans ce domaine on cite les travaux de Guessoum et Ould saïd (2008,2010,2012) sur le comportement asymptotique de la fonction de régression dans le cas où les données sont i.i.d. et  $\alpha$ -mélangeantes.

Dans toute la suite de notre travail, nous aurons besoin de certaines notions utilisées dans l'inférence non paramétrique que nous rappelons brièvement ci-dessous.

### 1.1.1 Estimation de la fonction de densité

Soit  $X_1, \dots, X_n$  un échantillon de variables aléatoires identiquement distribuées, à valeur dans  $\mathbb{R}$ , de densité commune  $\ell$ , que nous cherchons à estimer. Ce problème tire son intérêt du fait qu'on peut déduire l'estimateur, de plus cela constitue une aide précieuse au choix d'un modèle de probabilité. Une solution intuitive a été proposée par Roseblatt(1956). Pour  $h_n > 0$  assez petit on a :

$$\ell(x) \simeq \frac{F(x + h_n) - F(x - h_n)}{2h_n},$$

en remplaçant  $F$  par son estimateur  $F_n$  d'où

$$\ell_n(x) \simeq \frac{F_n(x + h_n) - F_n(x - h_n)}{2h_n},$$

où  $F_n$  est la fonction de répartition empirique, cet estimateur peut être encore écrit :

$$\begin{aligned}\ell_n(x) &= \sum_{i=1}^n \frac{\mathbb{1}_{\{x-h_n < X_i \leq x+h_n\}}}{2nh_n} \\ &= \frac{1}{2nh_n} \sum_{i=1}^n \mathbb{1}_{\{-1 < \frac{x-X_i}{h_n} \leq 1\}} \\ &= \frac{1}{nh_n} \sum_{i=1}^n K_0\left(\frac{x-X_i}{h_n}\right),\end{aligned}$$

avec  $K_0(u) = \frac{1}{2} \mathbb{1}_{\{-1 < u \leq 1\}}$ .

Parzen (1962) a suggéré une généralisation de cet estimateur, en posant :

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

où  $h := h_n$  appelé le paramètre de lissage tendant vers zéro à l'infini.

$K$  est une fonction mesurable définie de  $\mathbb{R} \rightarrow \mathbb{R}$ , appelée noyau avec  $\int_{\mathbb{R}} K(t)dt = 1$ .

### Exemples de noyaux :

- Noyau rectangulaire

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

- Noyau triangulaire

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

- Noyau d'épanechnikov

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

- Noyau gaussien

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

**Remarque 1.1.** Nous pouvons définir l'estimateur de Parzen-Rosenblatt dans le cadre multivarié.

Lorsque la variable  $X$  est à valeur dans  $\mathbb{R}^d$  l'estimateur de la densité s'écrit,

$$\ell_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K_d\left(\frac{x-X_i}{h}\right), x \in \mathbb{R}^d.$$

où  $K_d : \mathbb{R}^d \rightarrow \mathbb{R}$  est un noyau multivarié, tels que

$$\text{pour tout } x = (x_1, \dots, x_d) \in \mathbb{R}^d, K_d(x) = \prod_{j=1}^d K(x_j).$$

## 1.2 Données incomplètes

### 1.2.1 Durée de vie

Nous appelons durée de vie une variable aléatoire généralement positive  $Y$ . Cette variable peut en fait représenter différentes durées dans des domaines divers. En médecine, cette variable peut, par exemple, constituer effectivement la durée de vie concrète de patients atteints par une certaine maladie ou encore le temps de rémission de cette maladie. Nous parlons aussi de durée de vie dans le cas de tests de fiabilité imposés à des systèmes plus ou moins complexes. Un échantillon de ceux-ci sera testé et la durée de vie d'un élément sera le temps entre la mise en marche et la première panne de celui-ci. Plus récemment, les durées de vie sont aussi utilisées dans le domaine de l'économie pour calculer différentes statistiques. Ainsi, une durée de vie pourra représenter le temps entre la perte d'un emploi et l'embauche suivant pour un individu faisant partie d'un échantillon, ce qui permettra de dresser des statistiques sur le temps moyen d'inscription à l'ANPE. Une durée de vie sera donc, dans le cas général, le temps qu'il faut pour passer d'un état A à un état B.

Lorsqu'on dispose d'un échantillon  $\{Y_i; 1 \leq i \leq n\}$  de la v.a.  $Y$  tel que chaque donnée  $Y_i$  de l'échantillon, durée de vie du  $i^{\text{ème}}$  individu, est observée dans sa totalité, nous dirons que les données sont complètement observées ou qu'il s'agit de données complètes. Dans ce cas, un estimateur naturel de la fonction de répartition (f.r.)  $F$  de  $Y$  est alors l'estimateur empirique  $F_n$ , défini par

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

Les propriétés de  $F_n(\cdot)$  sont largement étudiées dans la littérature statistique. Notre travail porte sur des données incomplètes, pour lesquelles la variable d'intérêt n'est pas complètement observée pour toutes les données de l'échantillon.

Nous présentons dans ce qui suit deux cas de données incomplètes : censurées et tronquées.

#### 1.2.1.1 Censure

**Définition 1.1.** (*censure à droite*)

*La durée de vie est dite censurée à droite si l'individu n'a pas subi l'événement à sa dernière observation. En présence de censure à droite, les durées de vie ( $Y$ ) ne sont pas toutes observées, pour certaines d'entre elles, on sait seulement qu'elles sont supérieures à une certaine valeur connue. Soit  $C$  une variable aléatoire de censure, au lieu d'observer la variables  $Y$  qui nous intéresse on observe le couple de variables  $(Z, \delta)$  avec  $Z = Y \wedge C$  et  $\delta = \mathbb{1}_{\{Y \leq C\}}$ ,  $\delta$  est appelé indicateur de censure puisque ses valeurs nous informent sur le fait que l'observation est complète (si  $\delta = 1$ ) ou censurée à droite (si  $\delta = 0$ ).*

Un exemple illustratif est lorsqu'on s'intéresse à la durée de vie d'un genre de machines précis mais que ces dernières tombent en panne s'il se produit une surtension d'électricité. Ici, la durée de vie de la machine est censurée à droite par l'instant auquel se produit la surtension. C'est le type de censure la plus fréquent dans la pratique.

**Définition 1.2.** (*censure à gauche*) Il y a censure à gauche lorsque nous observons la censure  $C$  (et non pas la durée de vie  $Y$ ) et que nous savons que  $Y < C$ .

Un phénomène symétrique au précédent se produit, le patient a déjà subi l'événement avant l'instant où on commence l'étude. Ce modèle est par exemple adapté au cas où l'on s'intéresse à l'âge auquel un individu commence à accomplir une tâche. Tout ce qu'on sait chez l'individu censuré est que le véritable âge est inférieur à la valeur observée (l'âge au moment de l'étude par exemple).

**Définition 1.3.** (*censure par intervalles*) Il y a censure par intervalles lorsque la censure à droite et la censure à gauche sont conjuguées. Dans ce cas l'information apportée par l'expérience se traduit par l'appartenance de la durée de vie à un intervalle de temps ( $C_1 < Y < C_2$ ).

Ce modèle est adapté au cas de suivis périodiques de patients et généralise aussi bien le modèle de censure à droite que celui de censure à gauche.

**Définition 1.4.** (*censure mixte*) Il y a censure mixte lorsque deux phénomènes de censure (l'un à gauche et l'autre à droite) peuvent empêcher l'observation du phénomène d'intérêt sans qu'on puisse nécessairement déterminer un intervalle auquel il appartient.

Les modèles de censure peuvent prendre d'autres formes dont les suivantes :

### La censure non-aléatoire (ou fixe) de type I

Etant donnée un nombre positif fixé  $C$  et un n-échantillon  $Y_1, \dots, Y_n$ , les observations consistent en  $\{(Z_i, \delta_i), i = 1, 2, \dots, n\}$ , où  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ . Ce modèle, souvent utilisé dans les études épidémiologiques, peut correspondre à l'observation de la durée de survie de  $n$  patients au cours d'une expérience de durée prédéterminée  $C$ .

### La censure aléatoire de type I

Etant donné un n-échantillon  $Y_1, \dots, Y_n$ , il existe une variable aléatoire n-dimensionnelle  $C_1, \dots, C_n$ , de  $\mathbb{R}$  telle que les observations consistent en  $\{(Z_i, \delta_i),$  où  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ . Ce modèle est typiquement utilisé pour les essais thérapeutiques. Nous nous intéressons à une cause de décès ayant lieu au bout d'un temps  $Y$  et nous désirons connaître la loi de  $Y$ , cependant, une autre cause aléatoire (décès dû à une autre cause aléatoire, abandon, ...) peut survenir auparavant et empêcher par conséquent l'observation de  $Y$ .

## La censure aléatoire de type II

Etant donné un nombre positif fixé  $r$  et un  $n$ -échantillon  $Y_1, \dots, Y_n$ , les observations consistent en  $\{(Z_i, \delta_i)\}$ , où  $Z_i = Y_i \wedge Y_r$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq Y_r\}}$  avec  $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$  est la statistiques d'ordre associée à  $Y_1, \dots, Y_n$ . Ce modèle, souvent utilisé dans les études de fiabilité, correspond à l'observation de la durée de fonctionnement de  $n$  machines tant que  $r$  d'entre elles ne sont pas tombées en panne.

### 1.2.1.2 Troncature

Nous parlons de troncature à droite (respectivement à gauche) lorsque la variable d'intérêt n'est pas observable quand elle est supérieure (respectivement inférieure) à un seuil  $S$  fixé. Dans le cadre de la censure, la variable  $C$  est observée alors que dans le cas de la troncature à droite (respectivement à gauche) l'analyse porte uniquement sur la loi de  $Y$  conditionnellement à l'événement  $\{Y < S\}$  (respectivement  $\{Y > S\}$ ) et une donnée tronquée ne peut faire partie de l'échantillon. Si une maison de retraite n'accepte que des personnes âgées d'au moins soixante ans, aucun individu décédé avant cet âge n'a la possibilité d'y avoir été admis et est de ce fait tronqué à gauche.

#### Remarque 1.2. (Identifiabilité d'un modèle)

*Dans la suite de cette thèse, nous ferons l'hypothèse que la censure est indépendante de l'événement c'est-à-dire, que  $Y_i$  est indépendant de  $C_i$  : on parle de censure non-informative. Cette hypothèse est très utile d'un point de vue mathématique et indispensable aux modèles classiques d'analyse de survie et aussi règle le problème d'identifiabilité. Autrement dit, si on connaît la loi des observations on peut déterminer de façon unique la loi de  $Y$ . En pratique, cela veut dire que les individus ne doivent pas être censurés parce qu'ils ont un risque de décès particulièrement élevé (ou faible). En d'autres termes, les individus exclus-vivants ou perdus de vue à une date  $t$  doivent être représentatifs des individus encore à risque à cet instant  $t$ .*

## 1.2.2 Caractérisation de la loi d'une durée de vie $Y$

Supposons que la durée de survie  $Y$  soit une variable positive ou nulle, et absolument continue, alors sa loi de probabilité peut être définie par l'une des cinq fonction équivalentes suivantes (chacune des fonction ci-dessous peut être obtenue à partir de l'une des autres fonctions) :

### 1. Fonction de répartition $F$

La fonction de répartition représente, pour  $t$  fixé, la probabilité de mourir avant l'instant  $t$ , c'est-à-dire

$$F(t) = \mathbb{P}(Y \leq t), t > 0,$$

## 2. Fonction de survie $S$

La fonction de survie est, pour  $t$  fixé, la probabilité de survivre jusqu'à l'instant  $t$  c'est-à-dire :

$$\begin{aligned} S(t) &= \mathbb{P}(Y > t) \\ &= 1 - \mathbb{P}(Y \leq t) \\ &= 1 - F(t) \\ &= \bar{F}(t) \end{aligned}$$

## 3. Fonction de densité $f$

C'est la fonction  $f(t) \geq 0$  telle que pour tout  $t \geq 0$

$$f(t) = \frac{dF(t)}{dt} = -S'(t)$$

avec  $F(t) = \int_0^t f(u)du$ ,

pour  $t$  fixé, la densité de probabilité représente la probabilité de mourir dans un petit intervalle de temps après l'instant  $t$ .

## 4. Taux de hasard

Le taux de hasard ou la fonction de risque  $\lambda$  est définie comme la probabilité qu'un individu fasse l'événement considéré durant un intervalle de temps très court sachant qu'il a survécu jusqu'au début de l'intervalle

$$\begin{aligned} \lambda(t) &= \lim_{h \rightarrow 0} \frac{\mathbb{P}(Y < t + h / Y \geq t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\mathbb{P}(Y < t + h, X \geq t) / \mathbb{P}(Y \geq t)}{h} \\ &= \frac{1}{\mathbb{P}(Y \geq t)} \lim_{h \rightarrow 0} \frac{\mathbb{P}(t \leq Y < t + h)}{h} \\ &= \frac{f(t)}{S(t)} \\ &= \frac{f(t)}{1 - F(t)} \end{aligned}$$

La fonction de risque mesure le risque instantané de survenue de l'événement.

## 5. Taux de hasard cumulé

Le taux de hasard cumulé ou la fonction de risque cumulative évaluée au temps  $t$  est l'intégrale de la fonction de risque entre 0 et  $t$

$$\Lambda(t) = \int_0^t \lambda(u) \quad du = -\log S(t)$$

On peut déduire la fonction de survie à partir du taux de hasard cumulé par la relation

$$S(t) = \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(u) \quad du\right).$$

### 1.2.3 Estimation de la fonction de survie

Soit  $Y_1, \dots, Y_n$  une suite de variables aléatoires réelles d'intérêt i.i.d, de fonction de répartition (f.d.r) commune  $F$ ,

soit  $C_1, \dots, C_n$  une suite de variables aléatoires de censure i.i.d de f.d.r. continue  $G$ . Les  $C_i$  sont aussi supposées être indépendantes des  $Y_i$ .

Dans le modèle de censure à droite, on observe les couples  $(Z_1, \delta_1), (Z_2, \delta_2), \dots, (Z_n, \delta_n)$ , où  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ .

Notons

$$\bar{F}(t) = \mathbb{P}(Y > t) = 1 - \mathbb{P}(Y < t),$$

$$\bar{G}(t) = \mathbb{P}(C > t) = 1 - \mathbb{P}(C < t),$$

et les  $Z_i$  sont f.d.r  $H$  donnée par

$$\begin{aligned} \bar{H}(t) &= \mathbb{P}(Z > t) = \mathbb{P}(Y \wedge C > t) \\ &= \mathbb{P}(Y > t)\mathbb{P}(C > t) \\ &= \bar{F}(t).\bar{G}(t) \end{aligned}$$

Posons  $\tau_L = \sup\{y/L(y) > 0\}$  où  $L$  est une f.d.r

#### 1.2.3.1 Estimateur de Nelson-Aalen

Avant de définir l'estimateur de Nelson-Aalen de la fonction de hasard cumulative, nous avons besoin de donner quelques définitions :

$Z_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Z_i \leq t\}}$ , le nombre de sujets présents jusqu'à  $t$

$N_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Z_i \leq t, \delta_i=1\}}$ , le nombre d'observations non censurées présents jusqu'à  $t$

et  $M_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Z_i \leq t, \delta_i=0\}}$ , le nombre d'observations censurées présents jusqu'à  $t$

Donc l'estimateur de Nelson-Aalen s'écrit sous la forme :

$$\Lambda_{NA}(t) = \int_0^t \frac{dN_n(s)}{Z_n(s)} = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i \mathbb{1}_{\{Z_{(i)} \leq t\}}}{1 - N_n(Z_i^-)} \quad (1.1)$$

où  $dN_n(s) = N_n(s) - N_n(s^-)$

cet estimateur de  $\Lambda$  est fondé sur la remarque suivante

$$\begin{aligned} \Lambda(s + ds) - \Lambda(s) &\simeq \lambda(s)ds \\ &= \mathbb{P}(\text{événement ait lieu dans } (s, s + ds) / \text{à risque en } s) \end{aligned}$$

Il est alors naturel d'estimer cette quantité par  $(N_n(s+ds) - N_n(s))/Z_n(s)$ . En sommant ces quantités sur les intervalles de  $[0, t]$  et en faisant tendre ces intervalles vers 0, de telle sorte que chacun ne contient qu'un seul événement, on obtient (1.1).

### 1.2.3.2 Estimateur de Kaplan-Meier

L'estimateur de la fonction de survie le plus utilisé en l'absence d'hypothèses paramétriques sur la distribution de temps de survie est l'estimateur de Kaplan-Meier (EKM). Cet estimateur est aussi appelé estimateur PL (Produit-Limite) car il apparaît comme la limite d'un produit. En effet pour  $t' < t$

$$\begin{aligned} S(t) &= \bar{F}(t) = \mathbb{P}(Y > t) \\ &= \frac{\mathbb{P}(Y > t, Y > t')}{\mathbb{P}(Y > t')} \cdot \mathbb{P}(Y > t') \\ &= \mathbb{P}(Y > t / Y > t') \cdot \mathbb{P}(Y > t') \\ &= \mathbb{P}(Y > t / Y > t') \cdot S(t') \end{aligned}$$

Pour  $t'' < t'$  on obtient

$$S(t') = \mathbb{P}(Y > t' / Y > t'') \cdot S(t'').$$

D'où

$$\begin{aligned} S(t) &= \mathbb{P}(Y > t / Y > t') \mathbb{P}(Y > t' / Y > t'') \cdot \mathbb{P}(Y > t'') \\ &= \mathbb{P}(Y > t / Y > t') \mathbb{P}(Y > t' / Y > t'') \cdot S(t'') \end{aligned}$$

et ainsi de suite.

En considérant les temps d'événements (décès ou censure) distincts  $Z_{(i)} (i = 1, \dots, n)$  rangés par ordre croissant, on obtient

$$\begin{aligned} \mathbb{P}(Y > Z_{(i+1)}) &= \prod_{j=1}^i \mathbb{P}(Y > Z_{(j+1)} / Y > Z_{(j)}) \\ &= \prod_{j=1}^i p_j \\ &= \prod_{j=1}^i (1 - q_j), \end{aligned}$$

où  $p_j$  est la probabilité qu'un individu survive pendant l'intervalle  $I_j = ]Z_{(j)}, Z_{(j+1)}]$ , sachant qu'il était vivant au début de cet intervalle et  $q_j$  est la probabilité qu'un individu meurt pendant l'intervalle  $I_j$  sachant qu'il était vivant au début de cet intervalle. Une estimation naturelle de  $q_j$  est donnée par

$$\hat{q}_j = \frac{M(Z_{(j)})}{R(Z_{(j)})},$$

où  $M(Z_{(j)}) = \sum_{i=1}^n \mathbb{1}_{\{Z_i=Z_{(j)}\}}$  est le nombre de morts réelles observés à l'instant  $Z_{(j)}$  et  $R(Z_{(j)}) = \sum_{i=1}^n \mathbb{1}_{\{Z_i \geq Z_{(j)}\}}$  est le nombre d'individus à risque (de mourir) juste avant l'instant  $Z_{(j)}$ . L'estimateur  $S_n(t)$  de  $S(t)$  est alors donné par

$$\widehat{S}_{KM}(t) = \prod_{Z_{(j)} \leq t} \left(1 - \frac{M(Z_{(j)})}{R(Z_{(j)})}\right).$$

**Remarque 1.3.** Si  $Z_{(j)}$  est l'instant où ne se produisent que des censures alors  $M(Z_{(j)}) = 0$ , donc  $S_n$  ne varie qu'aux points où il y a effectivement une mort (donnée complète). S'il n'y a pas d'ex-aequo alors  $M(Z_{(j)}) = 0$  ou  $M(Z_{(j)}) = 1$  et dans ce dernier cas  $R(Z_{(j)}) = n - j + 1$ , on obtient alors

$$\widehat{S}_{KM}(t) = \prod_{Z_{(j)} \leq t} \left(1 - \frac{1}{n - j + 1}\right)^{\delta_{(j)}}.$$

Donc

$$\widehat{S}_{KM}(t) =: \bar{F}_n(t) = \begin{cases} \prod_{j: Z_{(j)} \leq t} \left(\frac{n-j}{n-j+1}\right)^{\delta_{(j)}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)}, \end{cases}$$

où  $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$  représentent les statistiques d'ordre associées à  $Z_i$ .

L'estimateur de Kaplan-Meier (L'EKM) pour la survie de la variable de censure est défini de la même façon par :

$$\bar{G}_n(t) = \begin{cases} \prod_{j: Z_{(j)} \leq t} \left(\frac{n-j}{n-j+1}\right)^{1-\delta_{(j)}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)}. \end{cases}$$

**Remarque 1.4.** L'estimateur de Kaplan Meier peut aussi se mettre sous la forme suivante

$$\widehat{S}_{KM}(t) = \begin{cases} \prod_{j=1}^n \left(1 - \frac{\delta_{(j)}}{n-j+1}\right)^{\mathbb{1}_{\{Z_{(j)} \leq t\}}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)}, \end{cases}$$

et

$$\bar{G}_n(t) = \begin{cases} \prod_{j=1}^n \left(1 - \frac{1-\delta_{(j)}}{n-j+1}\right)^{\mathbb{1}_{\{Z_{(j)} \leq t\}}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)}, \end{cases}$$

### 1.3 Estimation de la fonction de régression pour des données complètes

Dans la statistique non paramétrique la méthode la plus communément utilisée pour étudier la relation entre deux variables  $X$  et  $Y$  est la régression, tels que  $Y$  est une variable de réponse et  $X$  un

vecteur aléatoire de prédicteurs (co-variables) prenant ses valeurs dans  $\mathbb{R}^d$ . La fonction de régression à un point  $x \in \mathbb{R}^d$  est l'espérance conditionnelle de  $Y$  sachant  $X = x$

$$r(x) := \mathbb{E}[Y|X = x].$$

Il existe plusieurs méthodes pour estimer une fonction de régression, dans ce travail, nous nous intéressons uniquement à l'estimateur par la méthode du noyau.

Nadaraya(1964) et Watson (1964) furent les premiers à s'intéresser à l'estimation de la fonction de régression par la méthode du noyau noté [NW], il est construit à partir d'une fonction noyau  $K(\cdot)$  et d'une fenêtre  $h$ , de manière analogue à l'estimateur à noyau de la fonction de densité  $\ell(x)$ . Nous pouvons alors définir proprement la fonction de régression ou l'espérance conditionnelle de  $Y$  sachant  $X = x$ , par :

$$r(x) = \mathbb{E}(Y|X = x) = \int_{-\infty}^{+\infty} y f_{Y|X=x}(y) dy = \int_{-\infty}^{+\infty} \frac{y f_{X,Y}(x, y)}{\ell(x)} dy =: \frac{r_1(x)}{\ell(x)}.$$

Comme les densités  $f_{X,Y}(x, y)$  et  $\ell(x)$  sont inconnus on peut les estimer en suivant les même étapes que pour l'estimateur à noyau. On considère donc

$$\forall (x, y) \in \mathbb{R}^{d+1} f_n(x, y) = \frac{1}{nh^{d+1}} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) K_d \left( \frac{y - Y_i}{h} \right),$$

et

$$\ell_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right),$$

Un estimateur naturel de  $r(x)$  est alors donné par

$$\begin{aligned} r_n(x) &= \frac{\int_{-\infty}^{+\infty} y f_n(x, y) dy}{\ell_n(x)} \cdot \mathbb{1}_{\{\ell_n(x) \neq 0\}} \\ &= \frac{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \frac{1}{h} \int_{-\infty}^{+\infty} y K_d \left( \frac{y - Y_i}{h} \right) dy}{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right)}. \end{aligned}$$

Si on suppose de plus que le noyau  $K$  est symétrique, et en posant  $u = \frac{y - Y_i}{h}$ , il vient

$$\begin{aligned} r_n(x) &= \frac{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \left[ \int_{-\infty}^{+\infty} (uh + Y_i) K_d(u) du \right]}{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right)} \\ &= \frac{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \left[ h \int_{-\infty}^{+\infty} (uK(u) du + \int_{-\infty}^{+\infty} Y_i K(u) du \right]}{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right)}. \end{aligned}$$

Donc un estimateur à noyau de la fonction de régression est donné par :

$$r_n(x) = \frac{\sum_{i=1}^n Y_i K_d \left( \frac{x - X_i}{h} \right)}{\sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right)},$$

Plus généralement, des estimateurs à poids de  $r(x)$  peuvent s'écrire

$$r_n(x) = \sum_{i=1}^n W_{n,i}(x) Y_i, \quad (1.2)$$

où  $W_{n,i}(x)$  sont des poids dépendants de  $x$  et de  $(X_1, \dots, X_n)$ .

## 1.4 Estimation de la fonction de régression pour des données incomplètes

### 1.4.1 censure à droite

Lorsque l'observation d'un couple  $(X, Y)$  est possible, l'estimation est basée sur un échantillon de variables de même lois que ce couple. Cependant, dans certaines applications un tel échantillon n'est pas disponible. Lorsque la variable réponse  $Y$  est censurée à droite, on ne peut qu'observer l'échantillon composé par  $(X, Z, \delta)$ , où  $C$  est une variable de censure et  $\delta$  est l'indicateur de censure qui désigne laquelle des variable  $Y$  ou  $C$  est réellement observée. Une idée, qui a été introduite par Carbonez et al. (1995), et reprise par Köhler et al. (2002) est de remplacer  $r_n$  de l'expression (1.2) dans le cas de données complètes par

$$\tilde{r}_n(x) = \sum_{i=1}^n W_{n,i}(x) (\bar{G}(Z_i))^{-1} \delta_i Z_i, \quad (1.3)$$

où  $\bar{G}$  est la fonction de survie de la v.a  $C$ , et

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

sont les poids de Nadaraya-Watson, ainsi l'expression (1.3) peut s'écrire comme

$$\tilde{r}_n(x) =: \frac{\tilde{r}_{1,n}(x)}{\ell_n(x)}$$

avec

$$\tilde{r}_{1,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}(Z_i))^{-1} \delta_i Z_i K_d\left(\frac{x-X_i}{h}\right),$$

et

$$\ell_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K_d\left(\frac{x-X_i}{h}\right).$$

En pratique, puisque  $\bar{G}$  est généralement inconnue, on la remplace par l'estimateur de Kaplan-Meier  $\bar{G}_n$  correspondant. Alors, un estimateur possible de  $r(x)$  est donné par

$$r_n(x) = \frac{\frac{1}{nh^d} \sum_{i=1}^n (\bar{G}_n(Z_i))^{-1} \delta_i Z_i K_d\left(\frac{x-X_i}{h}\right)}{\frac{1}{nh^d} \sum_{i=1}^n K_d\left(\frac{x-X_i}{h}\right)} =: \frac{r_{1,n}(x)}{\ell_n(x)}.$$

## 1.5 Données Dépendantes

Le fait de supposer que les données étudiées sont toujours indépendantes n'est que rarement justifié, c'est pour cela, depuis quelques années, plusieurs auteurs ont concentré leurs études sur un autre type de données, qui sont les données dépendantes. Il existe plusieurs notions de dépendance, et parmi ces formes, **le mélange** et **l'association**.

### 1.5.1 Notions de mélange

Les différentes notions de mélange sont reliées à des mesures sous-jacentes de la dépendance entre les  $\sigma$ -algèbres. Plus précisément, soit  $(\Omega, F, \mathbb{P})$  un espace de probabilité, et  $\mathfrak{L}$  et  $\mathfrak{L}'$  deux sous  $\sigma$ -algèbres de  $F$ . Plusieurs mesures de dépendance entre  $\mathfrak{L}$  et  $\mathfrak{L}'$  sont définies comme suit :

1. Le coefficient de mélange fort ou d' $\alpha$ -mélange introduit par Rosenblatt (1956)

$$\alpha(\mathfrak{L}, \mathfrak{L}') := \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| ; A \in \mathfrak{L}, B \in \mathfrak{L}' \right\} ;$$

Les coefficient d' $\alpha$ -mélange sont dits arithmétiques d'ordre  $a > 0$  s'il existe deux constantes  $c \in \mathbb{R}_+^*$  et  $a \in \mathbb{R}_+^*$  telles que les coefficients de mélange vérifient

$$\alpha(n) \leq cn^{-a} = e^{-a \log n} = (e^{\log n})^{-a} ,$$

et ils sont dits géométriques, s'il existe deux constantes  $s \in \mathbb{R}_+^*$  et  $t \in ]0, 1[$  telles que les coefficients de mélange vérifient

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

2. Le coefficient de régularité absolue ou de  $\beta$ -mélange introduit par Volkonskii et Rozanov (1959)

$$\beta(\mathfrak{L}, \mathfrak{L}') := \sup \left\{ \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)| \right\} ;$$

où le sup est pris sur toutes les partitions  $(A_i)_{i \in I}, (B_j)_{j \in J}$  de  $\Omega$  avec  $A_i \in \mathfrak{L}, B_j \in \mathfrak{L}'$ .

3. Le coefficient de mélange uniforme ou de  $\phi$ -mélange introduit par Ibragimov (1959)

$$\phi(\mathfrak{L}, \mathfrak{L}') := \sup \left\{ |\mathbb{P}(B|A) - \mathbb{P}(B)| ; A \in \mathfrak{L}, B \in \mathfrak{L}' \text{ et } \mathbb{P}(A) > 0 \right\} ;$$

4. Le coefficient de  $\psi$ -mélange introduit par Blum et al. (1963)

$$\psi(\mathfrak{L}, \mathfrak{L}') := \sup \left\{ \left| \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1 \right| ; A \in \mathfrak{L}, B \in \mathfrak{L}' \text{ et } \mathbb{P}(A) > 0, \mathbb{P}(B) > 0 \right\} ;$$

5. Le coefficient de  $\rho$ -mélange introduit par Kolmogrov et Rozanov (1960)

$$\rho(\mathfrak{L}, \mathfrak{L}') := \sup \left\{ |\text{corr}(X, Y)| ; X \in L^2(\mathfrak{L}), Y \in L^2(\mathfrak{L}') \right\} ;$$

où  $L^2(A)$  est l'espace de variables aléatoires de carré-intégrables et  $A$ -mesurables.

**Propriétés 1.1.** *L'échelle de ces coefficients est définie par les inégalités suivantes :*

- $0 \leq \alpha(\mathfrak{L}, \mathfrak{L}') \leq \frac{1}{4}$ ,
- $0 \leq \beta(\mathfrak{L}, \mathfrak{L}') \leq 1$ ,
- $0 \leq \phi(\mathfrak{L}, \mathfrak{L}') \leq 1$ ,
- $0 \leq \psi(\mathfrak{L}, \mathfrak{L}') \leq \infty$ ,
- $0 \leq \rho(\mathfrak{L}, \mathfrak{L}') \leq 1$ .

### 1.5.1.1 Les conditions de mélange fort

Supposons que  $c(., .)$  désigne un coefficient de mélange arbitrairement pris parmi ceux précédemment définis ( $\alpha, \beta, \phi, \psi, \rho$ ). Soit  $\{T_i, i \geq 1\}$  une suite de v.a et soit  $\mathfrak{F}_i^k$  la  $\sigma$ -algèbres engendrée par  $\{T_j, -\infty \leq i \leq j \leq k \leq +\infty\}$ . Etant donné un entier positif  $l$ , posons pour tout  $k$  fixé,  $k \geq 1$ ,

$$c(l) = c(A, B) \text{ où } A \in \mathfrak{F}_{-\infty}^k, B \in \mathfrak{F}_{k+l}^{+\infty}$$

La suite de v.a.  $\{T_i, i \geq 1\}$  est dite  $c$ -mélangeante si  $\lim_{l \rightarrow +\infty} c(l) = 0$ .

Cette définition, mise en relation avec celles des différents coefficients de mélange, permet donc de considérer le  $c$ -mélange comme une indépendance asymptotique entre le passé et le future. Ces notions de mélange ne sont bien évidemment pas équivalentes mais il existe différentes implications entre les formes de mélanges :

$$\psi - \text{mélange} \implies \phi - \text{mélange} \implies \left\{ \begin{array}{l} \beta - \text{mélange} \\ \rho - \text{mélange} \end{array} \right\} \implies \alpha - \text{mélange}.$$

Dans le cas de la stricte stationnarité de  $T$  ; on a simplement

$$\alpha(n) := \alpha(\mathfrak{F}_{-\infty}^0, \mathfrak{F}_n^{+\infty}).$$

Parmi tous ces types de dépendances, l' $\alpha$ -mélange est la plus faible et la moins restrictive de toutes ces formes de mélanges, autrement dit, toute suite de variable aléatoire sera alors forcément  $\alpha$ -mélangeante. Les processus  $\alpha$ -mélangeants sont très présents dans la pratique statistique, en particulier les processus  $AR$  et  $ARMA$ , qui sont largement utilisés dans l'analyse des séries temporelles, sont  $\alpha$ -mélangeante avec des coefficients géométrique et exponentiel respectivement. Le champ d'application de telles données est très large et il est important de pouvoir faire de l'estimation dans ce cadre de données dépendantes. Dans la littérature, on trouve un nombre conséquent de travaux consacrés à l'étude des variables  $\alpha$ -mélangeante réelles dont les monographies de Doukhan (1994), Bosq (1996, 1998), Rio (2000) et Yoshihara (2004) donnent un point de vue global.

## 1.5.2 Association

Dans cette thèse, nous intéressons au concept d'association introduit indépendamment par Esary, Proschan et Walkup(1967) et Fortuyn, Fastelyn et Ginibre (FKG, 1971).

Leur objectif était de trouver des applications dans la fiabilité des systèmes et dans la statistique mécanique, en se basant sur l'inégalité de FKG. Ce concept est une généralisation de la dépendance positive introduite par Lehmann (1966) comme suit :

**Définition 1.5.** *Un vecteur  $(X, Y)$  de variables aléatoire est dit positivement dépendant si, pour tous réels  $x$  et  $y$ ,*

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

**Définition 1.6.** *Une famille finie de variables aléatoires réelles  $Y = \{Y_1, Y_2, \dots, Y_n\}$ , est dite associée si pour toute paire de fonctions  $\phi_1$  et  $\phi_2$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , non décroissantes (composante par composante), on a :*

$$\text{cov}(\phi_1(Y_i, 1 \leq i \leq n), \phi_2(Y_j, 1 \leq j \leq n)) \geq 0,$$

*lorsque cette covariance existe.*

*Une famille infinie de variables aléatoires est associée si toute sous famille finie est associée.*

**Propriétés 1.2.** *Les Propriétés principales de l'association présentées par Esary et al. (1967) sont les suivantes :*

**(P1).** *Tout sous-ensemble d'un ensemble fini de variables aléatoires associées est associé.*

**(P2).** *L'union de deux sous-ensembles indépendants de variables aléatoires associées est associée.*

**(P3).** *Tout singleton formé d'une variable aléatoire réelle est associé.*

*Cette propriété entraîne avec la précédente l'association des variables aléatoires indépendantes.*

**(P4).** *Si  $(Y_1, Y_2, \dots, Y_n)$  est un vecteur de variables associées et  $(f_1, f_2, \dots, f_k)$  des fonctions réelles croissantes de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , alors le vecteur  $(f_1(Y), f_2(Y), \dots, f_k(Y))$  est associé.*

**(P5).** *Si pour tout  $k \geq 1$ ,  $(Y_1^{(k)}, Y_2^{(k)}, \dots, Y_n^{(k)})$  est un vecteur associé et si  $Y_i^{(k)} \rightarrow Y_i$  en loi pour tout  $i = 1, 2, \dots, n$ , alors est  $(Y_1, Y_2, \dots, Y_n)$  associé.*

**Exemple 1.1.****a) Statistique d'ordre**

Si  $Y := (Y_1, Y_2, \dots, Y_n)$  est un vecteur associé, alors le vecteur  $(Y_{n1}, Y_{n2}, \dots, Y_{nn})$  de la statistique d'ordre engendré par  $Y$  est aussi associé (d'après **(P4)**).

**b) Processus gaussien**(Pitt (1982))

Tout vecteur gaussien  $(Y_1, Y_2, \dots, Y_n)$  est associé si et seulement si  $\text{cov}(Y_i, Y_j) \geq 0$ , pour tous  $(i, j) \in \{1, \dots, n\}^2$ .

**c) Variables aléatoires binaires**

Une variable aléatoire réelle  $Y$  est dite binaire si elle prend les valeurs 0 ou 1. Un vecteur  $Y = (Y_1, Y_2)$  de variables binaires est associé si et seulement si  $\text{cov}(Y_1, Y_2) \geq 0$ .

En particulier, ce type de variables est fréquemment utilisé en théorie de la fiabilité ( voir le livre de Barlow et Proschan (1975)).

**d) Famille de variables indépendantes**

Tout vecteur aléatoire dont les composantes sont indépendantes est associé.

**e) Processus linéaire**

Soit  $(\varepsilon_i)_{i \in \mathbb{Z}}$  une suite de variables aléatoires indépendantes ou associées et  $(a_i)_{i \in \mathbb{Z}}$  une suite de réels. Pour tout  $n \in \mathbb{Z}$  et  $N \geq 1$ , on pose  $Y_{n,N} = \sum_{|i| \leq N} a_i \varepsilon_{n-i}$ . Supposons qu'il existe une variable aléatoire  $Y_n$  telle que

$$\lim_{N \rightarrow \infty} Y_{n,N} = Y_n \text{ p.s. } |Y_n| < +\infty \text{ p.s. } \forall n \in \mathbb{Z}.$$

$(Y_n)_{n \in \mathbb{Z}}$  est un processus linéaire défini pour tout  $n \in \mathbb{Z}$  par

$$Y_n = \sum_{i \in \mathbb{Z}} a_i \varepsilon_{n-i}.$$

Si les termes de la suite  $(a_i)_{i \in \mathbb{Z}}$  sont positifs, alors le processus linéaire  $(Y_n)_{n \in \mathbb{Z}}$  est associé. En effet, pour tout  $N \geq 1$ , la suite  $(Y_{n,N})_{n \in \mathbb{Z}}$  est associée (d'après **(P4)**). Par conséquent, l'association de la suite  $(Y_n)_{n \in \mathbb{Z}}$  se déduit de la propriété **(P5)**.

**1.5.2.1 Un bref historique sur l'association**

Les premières études sur les données associées ont été faites au début des années 60 par Harris pour des processus de percolation, puis par Lehmann (1966) pour des données dépendantes.

Divers propriétés asymptotiques pour des sommes variables aléatoires associées ont été étudiées

par Newman(1980, 1984), Birkel (1988) et plusieurs autres auteurs, ils ont observé que, dans toute propriété asymptotique des variables aléatoires associés, la structure de covariance joue un rôle fondamental.

Roussas (1991) a établi sous des conditions de régularité, la convergence ponctuelle et uniforme de l'estimateur à noyau de la densité sous données vectorielles associées, il a donné également une caractérisation de la vitesse de convergence.

Bagai et Prakasa Rao (1995) ont étudié l'estimateur de la densité pour un processus associé stationnaire, ils ont étudié la convergence uniforme de l'estimateur à noyau, sans vitesse de convergence.

Des résultats importants sur le Théorème limite pour des données associées, en particulier négativement associées ont été obtenus par Bozorgnia et al.(1996).

Cai et Roussas (1998) établissent des propriétés asymptotiques (convergence uniforme et normalité asymptotique) de l'estimateur de Kaplan-Meier de la fonction de répartition sous association.

Des rappels sur l'association (positive ou négative) et ses applications en statistique peuvent être trouvés dans Roussas (1999, 2000, 2001), Cai et Roussas (1999).

Doukhan et Louhichi(1999) ont défini une notion de dépendance faible qui permet de traiter le mélange et l'association dans une approche unifiée.

Masry (2001) a prouvé la convergence uniforme et la normalité asymptotique avec vitesse de convergence de la densité d'une variable vectorielle. Masry(2002) prolonge l'étude de Roussas au cas multivarié.

Douge (2007) établit une nouvelle inégalité exponentielle, qui conduit à un taux de convergence uniforme presque sûr de  $\ell_n$  sur des ensembles compacts, sous une condition de décroissance géométrique des covariances, améliore ainsi sensiblement les résultats de Doukhan et Louhichi (2001) et de Masry (2002).

Récemment, Guessoum et al. (2012) ont établi la consistance uniforme forte de l'estimateur de Lynden-Bell sous l'association, plus récemment, Ferrani et al (2014) ont établi la consistance uniforme de la densité et du mode sous l'association.

### 1.5.2.2 Quelques formes d'associations

L'usage fréquent de ce type de dépendance a permis de dégager quelques types d'association notamment l'association positive (ou faible), l'association négative et quasi-association.

#### a) Association positive (PA)

Burton, Dabrowski et Dehling (1986) définissent une classe strictement plus large de variables aléatoires incluant l'association : l'association positive ou faible association.

**Définition 1.7.** *La famille finie  $\{Y_1, Y_2, \dots, Y_n\}$  est dite positivement ( ou faiblement) associée (PA)*

si pour tout sous ensembles disjoints  $I$  et  $J \in \{1, 2, \dots, n\}$

$$\text{cov}(\phi_1(Y_i, i \in I), \phi_2(Y_j, j \in J)) \geq 0,$$

pour toutes fonctions non décroissantes  $\phi_1 : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ ,  $\phi_2 : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ , pour lesquelles la covariance existe.

Une famille infinie de variables aléatoires est positivement associée (PA) si toute sous famille finie est positivement associée.

### Remarque 1.5.

- La définition de l'association et de la PA sont des définitions semblables, mais pas équivalentes.
- La définition de la PA est strictement plus faible que celle de l'association, dans le sens où l'association implique la PA mais la réciproque est en général fausse.

### b) Association négative (NA)

La notion d'association négative a été introduite par Alam et Saxena(1981) et développée par Joag-Dev et Proschan(1983).

**Définition 1.8.** La famille finie  $\{Y_1, Y_2, \dots, Y_n\}$  est dite négativement associée (NA) si pour tout sous ensembles disjoints  $I$  et  $J \in \{1, 2, \dots, n\}$

$$\text{cov}(\phi_1(Y_i, i \in I), \phi_2(Y_j, j \in J)) \leq 0,$$

pour toutes fonctions non décroissantes  $\phi_1 : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ ,  $\phi_2 : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ , lorsque cette covariance existe. Ce concept a été utilisé dans de nombreuses applications en statistique et en fiabilité. De plus, certaines loi multivariées connues possèdent cette propriété d'association négative.

### c. Quasi-association (QA) et dépendance faible

Bulinski et Shabanovich (1998) ont introduit un concept, appelé "quasi-association", pour montré que toute collection de variables aléatoires positivement ou négativement associées admettant un second moment fini.

**Définition 1.9. (Bulinski et Shabanovich)** Une suite  $(Y_n)_{n \in \mathbb{N}}$  de variables aléatoires réelles est dite quasi-associée si, pour sous ensembles finis disjoints  $I$  et  $J$  de  $\mathbb{N}$  et toutes fonctions lipschitziennes  $\phi_1 : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ ,  $\phi_2 : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ , on a :

$$|\text{cov}(\phi_1(Y_i, i \in I), \phi_2(Y_j, j \in J))| \leq \sum_{i \in I} \sum_{j \in J} \text{Lip}_i(\phi_1) \text{Lip}_j(\phi_2) |\text{cov}(Y_i, Y_j)| \quad (1.4)$$

où les constantes de lipschitz  $\text{Lip}_i(\phi_1)$  sont telles que pour  $x = (x_i, i \in I)$ ,  $y = (y_i, i \in I)$  dans  $\mathbb{R}^{|I|}$

$$|\phi_1(x) - \phi_1(y)| \leq \sum_{i \in I} \text{Lip}_i(\phi_1) |x_i - y_i|,$$

avec

$$\text{Lip}_i(\phi_1) = \sup_{x_i \neq y_i} \frac{|\phi_1(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{|I|}) - \phi_1(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_{|I|})|}{|x_i - y_i|}. \quad (1.5)$$

Le sup étant pris pour  $x_1, x_2, \dots, x_{|I|}, y_i \in \mathbb{R}$ .

On peut trouver cette définition dans l'ouvrage ( Bulinski et Shashkin (2007)).

**Remarque 1.6.** *L'association positive (PA) ou l'association négative (NA) impliquent la quasi-association. Ce qui fait de la quasi-association la plus faible forme de dépendance dans l'association.*

Doukhan et Louhichi (1999) ont défini une notion de dépendance faible qui permet de traiter le mélange et l'association dans une approche unifiée.

On note  $L$  l'ensemble de toutes les fonctions lipschitziennes définies sur  $\mathbb{R}^n, n \in \mathbb{N}^*$ , et  $L_1$  le sous-espace défini par

$$L_1 = \{h \in L; \|h\|_\infty \leq 1\},$$

où  $\|\cdot\|_\infty$  désigne la norme infinie

**Définition 1.10.** (*Doukhan et Louhichi*) Une suite  $(X_t)_{t \in \mathbb{Z}}$  est dite  $(\theta, \psi)$  faiblement dépendante, s'il existe une suite  $\theta = (\theta_r)_{r \in \mathbb{N}}$  décroissante vers 0 à l'infini et une fonction  $\psi : \mathbb{R}_+^2 \times \mathbb{N}^2 \rightarrow \mathbb{R}_+$  telles que, pour tous  $u$ -uplets  $(s_1, \dots, s_u)$  et  $v$ -uplets  $(t_1, \dots, t_v)$  avec  $s_1 \leq \dots \leq s_u + r \leq t_1 \leq \dots \leq t_v$ ,

$$|\text{cov}(f(X_{s_1}, \dots, X_{s_u}), g(X_{t_1}, \dots, X_{t_v}))| \leq \psi(\text{Lip}(f), \text{Lip}(g), u, v) \theta_r,$$

pour toutes les fonctions  $f, g \in L_1$  qui sont définies respectivement sur  $\mathbb{R}^u$  et  $\mathbb{R}^v$ ,

où

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

représente le module de Lipschitz de la fonction  $f : \mathbb{R}^u \rightarrow \mathbb{R}$ .

**Remarque 1.7.** Si  $(X_t)_{t \in \mathbb{Z}}$  est une suite de variables aléatoires quasi-associées, alors  $(X_t)$  est  $(\theta, \psi)$  faiblement dépendante avec  $\theta_r = \sup_i \sum_{j: |i-j| \geq r} |\text{cov}(X_i, X_j)|$  et  $\psi(a, b, u, v) = \min(u, v)ab$ .

## 1.6 Présentation des résultats existants

Nous donnons dans cette partie une brève présentation des résultats établis sur les propriétés asymptotiques ( la convergence uniforme, la normalité) de la fonction de régression, pour des données complètes dans le cas i.i.d.,  $\alpha$ -mélangeant et associées ainsi que pour le cas de données incomplètes.

### 1.6.1 Cas des données complètes

#### 1.6.1.1 Le cas i.i.d.

Les premières études effectuées en ce sens sont celles de Nadaraya(1964) sur l'estimateur de la fonction de régression, ainsi que celles de Watson(1964) portant sur l'analyse de régression lisse.

#### 1.6.1.2 Le cas $\alpha$ -mélange

Roussas, G. G. (1990) a établi d'un point de vue général le comportement de l'estimateur non paramétrique de la fonction de régression.

#### 1.6.1.3 Le cas associé

A notre connaissance, le seul résultat qui existe dans la littérature sur la convergence presque sûre de l'estimateur  $r_n$  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  $r_n$  qui est établi par Ango Nze, Bühlmann et Doukhan (2002) pour des variables faiblement dépendantes qui incluent les variables associées dans le cas univarié.

### 1.6.2 Cas des données incomplètes

Nous présenterons brièvement les travaux élaborés dans le cas de données tronquées à gauche. Nous procéderons par la suite à la description des travaux réalisés dans le cadre de données censurées à droite.

#### 1.6.2.1 Données tronquées à gauche

Ould saïd et Lamdani (2006) ont établi les propriétés asymptotiques de l'estimateur non paramétrique de la fonction de régression dans un modèle de données tronquées dans le cas i.i.d. Dans le cadre des données  $\alpha$ -mélangeantes tronquées à gauche, nous citons les travaux de Liang et al. (2009) sur la convergence forte de l'estimateur de la fonction de régression. Enfin, Hamrani et Gues-soum (2017) ont traité la consistance de l'estimateur de la fonction de régression pour des données dépendantes de type associé, tronquées à gauche.

### 1.6.2.2 Données censurées à droite

#### Cas i.i.d.

Dans cette partie nous allons reprendre les résultats obtenus par Guessoum et Ould saïd (2008) sur les propriétés asymptotiques de l'estimateur à noyau de la fonction de régression.

#### Hypothèses et résultats principaux :

Soit  $\Omega$  un sous ensemble compact de  $\Omega_0 = \{x \in \mathbb{R} / \ell(x) > 0\}$ . Les hypothèses que nous devons imposer pour établir ce résultat sont les suivantes :

- A1.** La fenêtre  $h$  satisfie :  $\lim_{n \rightarrow +\infty} h = 0$ ,  $\lim_{n \rightarrow +\infty} \frac{nh}{\log n} = +\infty$  et  $\log \log n = o\left(\frac{1}{h^\mu}\right)$  où  $0 < \mu < 1$ .
- A2.** Le noyau  $K$  est borné, symétrique et à support compact, il est aussi Hölderien de l'ordre  $\gamma > 0$ , satisfaisant  $\int_{\mathbb{R}} t|K(t)|dt < +\infty$  et  $\int_{\mathbb{R}} t^2 K(t)dt < +\infty$
- A3.** La fonction  $r_1(x)$  est deux fois différentiable avec  $\sup_{x \in \Omega} |r_1''(x)| < +\infty$ .
- A4.** La densité marginale  $\ell(\cdot)$  est deux fois différentiable et satisfait une condition de lipschitzienne, tel que  $\ell(x) > \Gamma$  pour tout  $x \in \Omega$  et  $\Gamma > 0$ .
- A5.** La fonction  $r_2(x) := \int_{\mathbb{R}} (\bar{G}(y))^{-1} y^2 f_{X,Y}(x, y) dy$  est deux fois différentiable avec  $|r_2''(x)| < +\infty$ .

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

- B1.** La fenêtre  $h$  satisfie  $\lim_{n \rightarrow +\infty} nh^5 = 0$ .
- B2.** Il existe  $\beta > 2$  tel que  $\int_{\mathbb{R}} \int_{\mathbb{R}} y^\beta f_{X,Y}(x, y) dy dx < +\infty$ .

Le premier résultat concernant la convergence uniforme presque sûre de l'estimateur de la fonction de régression est donné dans le Théorème suivant

**Théorème 1.1.** (*Guessoum et Ould saïd (2008)*) : *Sous les hypothèses (A1)-(A5), on a*

$$\sup_{x \in \Omega} |r_n(x) - r(x)| = O\left(\max\left\{\left(\sqrt{\frac{\log n}{nh}}\right), h^2\right\}\right) \text{ p.s. lorsque } n \rightarrow +\infty.$$

La démonstration du Théorème repose sur les Lemmes suivants :

#### Lemme 1.1.

$$\sup_{x \in \Omega} |r_{1,n}(x) - \tilde{r}_{1,n}(x)| = O\left(\frac{1}{\sqrt{nh^\mu}}\right), \text{ p.s. lorsque } n \rightarrow +\infty.$$

#### Lemme 1.2.

$$\sup_{x \in \Omega} |\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| = O(h^2), \text{ p.s. lorsque } n \rightarrow +\infty.$$

#### Lemme 1.3.

$$\sup_{x \in \Omega} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| = O\left(\sqrt{\frac{\log n}{nh}}\right), \text{ p.s. lorsque } n \rightarrow +\infty.$$

Nous énonçons dans le prochain Théorème, la normalité asymptotique de l'estimateur de la fonction de régression

**Théorème 1.2.** (*Guessoum et Ould saïd (2008)*) : Sous les hypothèses (A1)-(A5) et (B1)-(B2), on a

$$\sqrt{nh} [r_n(x) - r(x)] \xrightarrow[n \rightarrow +\infty]{\mathbb{L}} \mathcal{N}(0, \sigma^2(x)),$$

où  $\xrightarrow{\mathbb{L}}$  désigne la convergence en loi,

$$\sigma^2(x) = \kappa \frac{r_2(x)\ell^2(x) - r_1^2(x)\ell(x)}{\ell^4(x)},$$

avec  $\kappa = \int_{\mathbb{R}} K_d^2(t) dt$ .

• **Le cas  $\alpha$ -mélange**

Guessoum et Ould saïd (2010) ont étendu les travaux réalisés dans le cas i.i.d. au cas des données  $\alpha$ -mélangeantes. Ils montrent la convergence uniforme presque sûre de l'estimateur à noyau de la fonction de régression. Nous avons besoin des hypothèses suivantes :

**Hypothèses**

Soit  $\Omega$  un sous ensemble compact de  $\Omega_0 = \{x \in \mathbb{R} / \ell(x) > 0\}$ .

**A'1.** La fenêtre  $h$  satisfie :

- i).  $\lim_{n \rightarrow +\infty} \frac{nh^d}{\log n} = +\infty$  ;
- ii).  $\sqrt{\frac{\log \log n}{n}} = o(h)$  ;
- iii).  $\lim_{n \rightarrow +\infty} h^{d(\nu-2)} \log n = 0$ .

**A'2.** Le noyau  $K_d$  est bornée et satisfait :

- i).  $\int_{\mathbb{R}^d} \|t\| K_d(t) dt = 0$  ;
- ii).  $\int_{\mathbb{R}^d} |t_1 + t_2 + \dots + t_d| K_d^2(t) dt < +\infty$ , et  $\int_{\mathbb{R}^d} K_d^2(t) dt < +\infty$  ;
- iii).  $\forall (t, s) \in \Omega^2 |K_d(t) - K_d(s)| \leq \|t - s\|$  for  $\gamma > 0$ .

**A'3.** La fonction  $r_1(\cdot)$  est continûment différentiable et  $\sup_{x \in \Omega} \left| \frac{\partial r_1}{\partial x_i} \right| < +\infty$  pour  $i = 1, \dots, d$  ;

**A'4.** La fonction  $r_2(x) := \int_{\mathbb{R}} (\bar{G}(y))^{-1} y^2 f_{X,Y}(x, y) dy$  est continûment différentiable et  $\sup_{x \in \Omega} \left| \frac{\partial r_2}{\partial x_i}(x) \right| < \infty$  pour  $i = 1, \dots, d$  ;

**A'5.**  $\exists D_1 > 0$  et  $\exists D_2 > 0$  tels que  $\sup_{u,v \in \Omega} |\ell_{i,j}(u, v)| < D_1$  et  $\sup_{u \in \Omega} |\ell(u)| < D_2$  où  $\ell_{i,j}$  est la densité conjointe de  $(X_i, X_j)$ .

**A'6.**  $\exists \theta > 0$  et  $\exists c_1 > 0$  tel que  $c_1 n^{\frac{\gamma(3-\nu)}{d[\gamma(\nu+1)+2\gamma+1]} + \theta d} \leq h^d$

**A'7.** La densité marginale  $\ell(\cdot)$  est continûment différentiable et  $\sup_{x \in \Omega} \left| \frac{\partial \ell}{\partial x_i}(x) \right| < +\infty$  pour  $i = 1, \dots, d$ . Il existe  $\xi > 0$ , tel que  $\ell > \xi, \forall x \in \Omega$ .

Nous donnons dans le Théorème suivant la consistance pour la fonction de régression

**Théorème 1.3.** (*Guessoum et Ould saïd (2010)*) : Sous les hypothèses **(A'1)**-**(A'7)**, on a

$$\sup_{x \in \Omega} |r_n(x) - r(x)| = O \left( \sqrt{\frac{\log n}{nh^d}} + \sqrt{h^{d(\nu-2)} \log n} \right) + O(h). \text{ p.s. lorsque } n \rightarrow +\infty.$$

La preuve du Théorème (1.3) est basée sur des inégalités de type exponentielle, Fuk-Nageav (Rio 2000) et les techniques de recouvrement du compact par des petites boules.

Pour la normalité asymptotique de l'estimateur  $r_n$ , nous avons besoin des hypothèses **A'2-A'5** et **A'7** précédentes et des hypothèses suivantes :

### Hypothèses

**H.** La fenêtre  $h$  satisfait :

- i).  $\lim_{n \rightarrow +\infty} nh^{d+2} = +\infty$ ;
- ii).  $\sqrt{\frac{\log \log n}{n}} = o(h^{-\mu})$  pour  $0 < \mu < d$ ;
- iii).  $\lim_{n \rightarrow +\infty} nh^{d+4} = 0$ .

**B.** Le coefficient de mélange  $\alpha(\cdot)$  est tel que :

- i).  $\alpha(n) = \alpha(n^{-\nu})$  pour certains  $\nu > 3$ ; item[**ii**].] Soit  $v_n$  une suite de nombres entiers tendant vers  $\infty$  tel que  $\frac{1}{h^{d\delta}} \sum_{k \geq v_n} \alpha^\delta(k) < +\infty$  pour  $0 < \delta < 1$ .

**Théorème 1.4.** : (*Guessoum et Ould saïd (2012)*) : Sous les hypothèses **H**, **B**, **A'2-A'5** et **A'7**, on a

$$\sqrt{nh} [r_n(x) - r(x)] \xrightarrow[n \rightarrow +\infty]{\mathbb{L}} \mathcal{N}(0, \sigma^2(x)),$$

où  $\xrightarrow{\mathbb{L}}$  désigne la convergence en loi,

$$\sigma^2(x) = \kappa \frac{r_2(x)\ell(x) - r_1^2(x)}{\ell^3(x)},$$

avec  $\kappa = \int_{\mathbb{R}} K_d^2(t) dt$  déjà défini.

La preuve du Théorème (1.4) est basée sur la procédure des grands et petits blocs ainsi que sur le critère de Lindeberg-Feller.

• **Le cas associé**

La littérature dans ce domaine reste restreinte, le but de notre étude étant d'étudier les propriétés asymptotiques d'un estimateur à noyau de la fonction de régression, nous nous inspirerons des travaux établis par Ferrani et al. (2014) et qui portent sur l'estimateur à noyau de la densité et du mode pour des données associées et censurées.

Dans un premier temps, ils établissent la convergence uniforme de l'estimateur à noyau de la densité et du mode

**Théorème 1.5.** (Ferrani et al. (2014)) : *Sous les hypothèses de régularité ainsi que des conditions standards sur  $K$ ,  $h$  et une condition de décroissance géométrique des covariances, on a*

$$\sup_{y \in \mathcal{D}} |\tilde{f}_n(y) - \mathbb{E}(\tilde{f}_n(y))| = O\left(\sqrt{\frac{\log n}{nh}}\right), \text{ p.s. lorsque } n \rightarrow +\infty.$$

**Théorème 1.6.** *Sous les hypothèses du Théorème (1.5), on a*

$$\sup_{y \in \mathcal{D}} \left| \hat{f}_n(y) - f(y) \right| = O\left\{ \max\left(\sqrt{\frac{\log n}{nh}}, h^2\right) \right\}. \text{ p.s. lorsque } n \rightarrow +\infty.$$

avec

$$\tilde{f}_n(y) = \frac{1}{nh} \sum_{i=1}^n (\bar{G}(Z_i))^{-1} \delta_i K\left(\frac{y - Z_i}{h}\right),$$

et

$$\hat{f}_n(y) = \frac{1}{nh} \sum_{i=1}^n (\bar{G}_n(Z_i))^{-1} \delta_i K\left(\frac{y - Z_i}{h}\right),$$

**Corollaire 1.1.** (Ferrani et al. (2014)) : *Sous les hypothèses de régularité ainsi que des conditions standards sur  $K$ ,  $h$  et une condition de décroissance géométrique des covariances, on a*

$$|\hat{\theta}_n - \theta| = O\left\{ \max\left(\left(\sqrt{\frac{\log n}{nh}}\right)^{\frac{1}{4}}, h\right) \right\}. \text{ p.s. lorsque } n \rightarrow +\infty,$$

où

$$\hat{\theta}_n := \arg \max_{y \in \mathbb{R}} \hat{f}_n(y),$$

et

$$\theta := \arg \max_{y \in \mathbb{R}} f(y).$$

Dans une second partie, ils étendent les travaux réalisés par Roussas (2000) dans la cadre de la normalité asymptotique de l'estimateur de la densité dans le cas censuré.

**Théorème 1.7.** *Sous les hypothèses du Théorème (1.5), on a :*

$$\sqrt{nh} \left[ \tilde{f}_n(y) - \mathbb{E}(\tilde{f}_n(y)) \right] \xrightarrow[n \rightarrow +\infty]{\mathbb{L}} \mathcal{N}(0, \sigma^2(y)),$$

où

$$\sigma^2(y) := \sigma^2 = \lim_{n \rightarrow \infty} nh \text{var}(\tilde{f}_n(y)) = \frac{f(y)}{\bar{G}(y)} \int_{\mathbb{R}} K_d^2(u) du,$$

**Corollaire 1.2.** *Sous les mêmes hypothèses du Théorème (1.7), on a*

$$\sqrt{nh} \left[ \hat{f}_n(y) - f(y) \right] \xrightarrow[n \rightarrow +\infty]{\mathbb{L}} \mathcal{N}(0, \sigma^2(y)).$$

## 1.7 Apport de la thèse

Dans ce travail on s'intéresse à l'estimation non paramétrique de la fonction de régression par la méthode du noyau, pour des données incomplètes (censurées à droite) et sous une forme de dépendance (association).

Cette thèse est divisée en quatre chapitres.

Le premier chapitre est une introduction générale qui contient des concepts utiles pour les chapitres qui suivent et un rappel de certains résultats établis sur le comportement asymptotique de l'estimateur non paramétrique à noyau de fonction de régression pour des données complètement observées dans le cas i.i.d.,  $\alpha$ -mélange et associées, aussi pour des données incomplètes cas i.i.d. et  $\alpha$ -mélange. Le cas associé fait l'objet de notre étude.

Dans le second chapitre, nous établissons un résultat de convergence uniforme presque sûre sur des compacts avec une vitesse de convergence, en utilisant des inégalités exponentielles, plus précisément l'inégalité de Doukhan-Neumann(2007). Nous illustrons les résultats obtenus par des simulations pour montrer les performances des estimateurs proposés à un échantillon fini.

Le troisième chapitre traite le problème de la normalité asymptotique de notre estimateur et des intervalles de confiance pour la fonction de régression  $r(x)$ , illustrés par des simulations.

Enfin, le chapitre 4 est une conclusion suivie de quelques perspectives de recherche permettant d'étendre et aussi de généraliser les résultats de cette thèse.

### 1.7.1 Présentation des résultats obtenus

Nous donnons dans cette partie une brève présentation des résultats obtenus dans la thèse

#### 1. La consistance :

**Théorème 1.8.** *Sous les hypothèses de régularité ainsi que des conditions standards sur le noyau, la fenêtre et une condition de décroissance géométrique des covariances, on a*

$$\sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| = O \left( \sqrt{\frac{\log n}{nh^d}} \right), \text{ p.s. lorsque } n \rightarrow +\infty.$$

**Théorème 1.9.** *Sous les hypothèses du Théorème (1.8), on a*

$$\sup_{x \in \mathcal{D}} |r_n(x) - r(x)| = O(h^2) + O_{a.s.}(n^{-\theta}) + O_{a.s.}\left(\sqrt{\frac{\log n}{nh^d}}\right),$$

où  $0 < \theta < \frac{\gamma}{2\gamma+9+\frac{3}{2}\rho}$ , pour tout  $\rho > 0$ .

La démonstration et le détail des conditions imposées font l'objet d'un article qui est présenté dans le chapitre 2 de cette thèse.

## 2. La normalité asymptotique :

**Théorème 1.10.** *Sous les conditions standards sur le noyau, la fenêtre, les densités, ainsi d'autres hypothèses de régularités, on a*

$$r_n(x) \xrightarrow{\mathbb{P}} r(x),$$

**Théorème 1.11.** *Sous les mêmes hypothèses du Théorème (1.10), ainsi que des hypothèses sur les grands blocs, petits blocs et une condition de décroissance géométrique des covariances, on a*

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

où

$$\sigma^2(x) = \kappa \frac{r_2(x)\ell(x) - r_1^2(x)}{\ell^3(x)}, \quad (1.6)$$

$$\kappa = \int_{\mathbb{R}^d} K_d^2(t) dt$$

La démonstration de ces résultats ainsi que les hypothèses utilisées sont détaillées au chapitre 3 de cette thèse.

## Chapitre 2

# Convergence uniforme presque sûre

Ce chapitre traite la convergence uniforme presque sûre d'un estimateur à noyau de la fonction de régression pour le modèle censuré à droite avec des données associées. Ce travail a fait l'objet d'une publication dans le Journal : Statistical Papers.

# Strong uniform consistency for the regression function estimator under censoring and association

## 2.1 Introduction

Let  $\{(X_i, Y_i), 1 \leq i \leq n\}$  be a sequence of strictly stationary random vectors with the same distribution as  $(X, Y) \in \mathbb{R}^d \times \mathbb{R}$ , for which there exists a joint probability density function (pdf)  $f_{X,Y}$ . We will denote by  $F_{X,Y}$  the joint distribution function (df) of  $(X, Y)$ , and by  $\ell$  and  $F_Y$  the marginal pdf corresponding to the covariate  $X$  and the marginal df corresponding to the random variable (rv) of interest  $Y$ , respectively. In the sequel, we will suppose  $Y$  bounded and non-negative.

Survival analysis is the part of statistics, in which the variable of interest may often be interpreted as the time elapsed between two events. Such variables (lifetimes) typically appear in a medical or an engineering context. A typical feature of such lifetime data analysis is that due to time limitations,  $Y$  may not always be observable. Instead, one gets only the minimum of the survival time  $Y_i$  and a non-informative censoring time  $C_i$  together with the information  $\delta_i$  as to whether the survival time of an individual is less than the censoring time or not. This means that the observable rv's are given by  $X_i, Z_i = \min(Y_i, C_i)$  and the censoring indicator  $\delta_i = \mathbb{1}_{Y_i \leq C_i}$ . Hence the data are randomly right censored and the available data only provide partial information and, as a consequence, standard statistical procedures are not applicable.

In this paper, the rv's  $\{C_i, 1 \leq i \leq n\}$  are non-negative, assumed to be independent and identically distributed (iid) and having a common Lipschitz continuous marginal df  $G$ .

It is well known that under this model, the observable random vector  $(X, Y, \delta)$  does not identify the conditional distribution of  $Y$  given  $X$ . To overcome the identifiability problem, we will assume that  $C$  is independent of  $(X, Y)$ .

Since Nadaraya (1964) and Watson (1964) proposed their kernel method to estimate an unknown regression function

$$r(x) = \mathbb{E}[Y|X = x], \quad (2.1)$$

from a sample  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  of independent replicates of  $(X, Y)$  there have been numerous activities to study their well known estimator

$$m_n(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}. \quad (2.2)$$

Here,  $K$  denotes a smoothing kernel integrating to one (we refer to Wied and Weißbach (2012) for a good survey) and  $h_n =: h$  is a sequence of bandwidths tending to zero as  $n$  goes to infinity. For more details on the estimator in (2.2) we refer to Härdle (1990) and Wand and Jones (1995). However,

as above-mentioned, the latter estimator does not estimate consistently the regression function in presence of a number of events which potentially censor the event of interest. Now let us define

$$r_1(x) = \int_{\mathbb{R}} y f_{X,Y}(x, y) dy. \quad (2.3)$$

Hence, for the kind of censored data, the regression function

$$r(x) = \frac{r_1(x)}{\ell(x)},$$

must be estimated from the  $n$  available triplets  $(Z_i, \delta_i, X_i)$ ,  $i = 1, \dots, n$ . Based on these observations, let us define

$$r_{1,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}_n(Z_i))^{-1} \delta_i Z_i K\left(\frac{x - X_i}{h}\right)$$

and

$$\ell_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right),$$

as estimators for  $r_1(x)$  and  $\ell(x)$ , respectively. Recall that  $\bar{G}_n(\cdot)$  is the celebrated Kaplan-Meier (1958) estimate of  $\bar{G}(\cdot) := 1 - G(\cdot)$ , defined as

$$\bar{G}_n(t) = \begin{cases} \prod_{i: Z_{(i)} \leq t} \left(\frac{n-i}{n-i+1}\right)^{1-\delta_{(i)}} & \text{if } t < Z_{(n)} \\ 0 & \text{otherwise,} \end{cases} \quad (2.4)$$

where  $(Z_{(i)}, \delta_{(i)}; 1 \leq i \leq n)$  are the  $n$  pairs of observed  $(Z_i, \delta_i)$  ordered on the  $Z_{(i)}$ .

Then we use the following estimator for  $r(x)$ , defined as

$$r_n(x) = \frac{r_{1,n}(x)}{\ell_n(x)}. \quad (2.5)$$

Note that the estimate in (3.5) and its version for  $F_Y$  have been deeply studied in the by many authors. To cite only a work related to our model, Cai and Roussas (1998) stated its strong uniform consistency and asymptotic normality under association hypothesis. More references can be found therein.

The regression function estimation under censoring has gained a great interest and to quote only a few of works dealing with this topic, we cite Koul et al. (1981), Stute (1993) and the monograph by Györfi et al. (2002). Recently, the strong uniform consistency and asymptotic normality for the regression function estimate in (3.6) have been established in the iid case by Guessoum and Ould Saïd (2008), while under  $\alpha$ -mixing condition, El Ghouch and Van Keilegom (2008), and Guessoum and Ould Saïd (2010,2012) stated asymptotic results for (3.6). In absence of covariables, Ferrani et al. (2016) established strong consistencies for kernel density and mode estimates under associated and censored model.

In the spirit to extend the existing results to association dependence, we aim in the present note to state a strong uniform consistency rate of  $r_n(x)$ . For this purpose, we need the following definition.

**Définition 2.1.** A finite collection  $Y=(Y_1, Y_2, \dots, Y_n)$  of real-valued rv's  $Y_k$ ,  $k = 1, 2, \dots, n$  is called associated if  $\text{cov}(\phi_1(Y), \phi_2(Y)) \geq 0$  for any coordinate-wise non-decreasing functions  $\phi_1, \phi_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ , whenever the covariance exists. An infinite family of rv's is associated if this is valid for every finite sub-family.

Associated rv's form an important class of dependent systems and were first introduced by Esary et al. (1967). These systems are encountered in mathematical statistics, reliability and statistical physics. In the latter area, the concept of association is known as FKG-inequalities (see Fortuin et al., 1971). As examples of associated rv's, we recall that most often in reliability studies, the rv's which are generally lifetimes of components, are not independent but are associated. If independent components of a system are subject to the same stress, then their lifetimes are associated. A variety of relevant examples and ample bibliographical references can be found in Bulinski and Shashkin (2007).

Let us now recall that a strong mixing condition refers more to  $\sigma$ -algebra than to rv's. On the one hand, a main inconvenience of mixing assumptions is the difficulty of checking them (see, e.g. Doukhan, 1994). On the other hand, an important property of associated random rv's is that zero correlation implies independence. Also, large classes of examples of associated processes are deduced from the fact that any independent sequence is associated and that monotonic functions of independent sequences remain associated (see Esary et al., 1967). So, the main advantage of the concept of association compared to mixing is that the conditions of limit theorems are easier to verify : indeed, a covariance is much easier to compute than a mixing coefficient. However, it is of interest to note that the classes of  $\alpha$ -mixing processes and associated processes are distinct but may overlap (see, e.g. Louhichi, 2000) for examples of sequences that are associated but not mixing, associated and mixing, and mixing but not associated ones.

As far as we know, the problem of drawing nonparametric inference about the regression function under associated-censored model is not available and this motivates the study we consider here.

The rest of the paper is organized as the following. The next section introduces notations and main results. A simulation study is presented in Section 3 while the last section is devoted to the proofs.

## 2.2 Notations and main results

Before listing the required conditions, let us define

$$\tilde{r}_{1,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}(Z_i))^{-1} \delta_i Z_i K \left( \frac{x - X_i}{h} \right). \quad (2.6)$$

Then to formulate our main results, we introduce another interesting feature of the above adapted  $r_n(\cdot)$ , namely  $\tilde{r}_n(\cdot)$  defined as

$$\tilde{r}_n(x) = \frac{\tilde{r}_{1,n}(x)}{\ell_n(x)}. \quad (2.7)$$

Note that in a real data situation or in simulation studies we shall, however, not dwell on (2.7) since  $\bar{G}$  is unknown. This pseudo-estimator will be needed as a technical tool in the proofs.

Let us now define  $\tau_F$ ,  $\tau_G$  and  $\tau_H$  as the upper points of the right-continuous df's  $F_Y(y) = \mathbb{P}(Y \leq y)$ ,  $G(y) = \mathbb{P}(C \leq y)$  and  $H(y) = \mathbb{P}(Z \leq y)$ , respectively, with  $\tau_F = \sup\{y, \bar{F}_Y(y) > 0\}$ ,  $\tau_G = \sup\{y, \bar{G}(y) > 0\}$  and  $\tau_H = \tau_F \wedge \tau_G$ . In what follows, we assume that  $\tau_H = \tau_F$ .

And let  $[0, \tau_F]$  and  $\mathcal{D}$  be compact sets such that  $\mathcal{D} \subset \{x \in \mathbb{R}^d / \ell(x) > \Gamma > 0; \Gamma := \inf \ell(x)\} \subset \mathbb{R}^d$ . Furthermore, we set  $\Delta_{i,j} := \sum_{k=1}^d \sum_{l=1}^d \text{cov}(X_i^k, X_j^l) + 2 \sum_{k=1}^d \text{cov}(X_i^k, Y_j) + \text{cov}(Y_i, Y_j)$ . Finally,  $c$  will denote a positive constant which may take different values.

### 2.2.1 Assumptions

**A.**  $(Y, C) \in \mathbb{R}^+ \times \mathbb{R}^+$ ,  $\tau_F < \infty$ ,  $G$  is continuous, and  $G(\tau_F) < 1$ ;

**H.** The bandwidth  $h$  satisfies :  $h \rightarrow 0$ ,  $nh^d \rightarrow +\infty$  and  $\frac{\log^5 n}{nh^d} = o(1)$ , as  $n \rightarrow +\infty$ .

**K.** The kernel function  $K_d$  is bounded ( $\|K_d\|_\infty := \sup |K_d|$ ), Hölder continuous with component  $\alpha$ , and satisfies :

**a).**  $\int_{\mathbb{R}^d} t_i K_d(t) dt = 0$ , for all  $i = 1, 2, \dots, d; t = (t_1, t_2, \dots, t_d)^\top$ ;

**b).**  $\int_{\mathbb{R}^d} |t_1 + t_2 + \dots + t_d| K_d^2(t) dt < \infty$ ;

**c).**  $\int_{\mathbb{R}^d} |t_1^{i_1} \times t_2^{i_2} \times \dots \times t_d^{i_d}| K_d(t) dt < \infty$ ,  $i_1 + i_2 + \dots + i_d = 2$ .

**M.**  $\rho(s) := \sup_{j: |i-j| \geq s} \Delta_{i,j} = O(e^{-\gamma s})$ , for all  $i \geq 1$ ,  $\gamma > 0$  and  $s > 0$ .

**D1.** The function  $r_1(\cdot)$  defined in (3.1) is bounded, twice differentiable and  $\sup_{x \in \mathcal{D}} \left| \frac{\partial^k r_1}{\partial x_i \partial x_j^{k-1}}(x) \right| < \infty$  for  $i, j = 1, \dots, d$  and  $k = 1, 2$ ;

**D2.** The marginal density  $\ell(\cdot)$  is bounded, twice differentiable and  $\sup_{x \in \mathcal{D}} \left| \frac{\partial^k \ell}{\partial x_i \partial x_j^{k-1}}(x) \right| < \infty$  for  $i, j = 1, \dots, d$  and  $k = 1, 2$ ;

**D3.** The function  $r_2(x) := \int_{\mathbb{R}} (\bar{G}(y))^{-1} y^2 f_{X,Y}(x, y) dy$  is continuously differentiable and  $\sup_{x \in \mathcal{D}} \left| \frac{\partial r_2}{\partial x_i}(x) \right| < \infty$  for  $i = 1, \dots, d$ ;

**D4.** The joint pdf  $\ell_{i,j}(\cdot, \cdot)$  of  $(X_i, X_j)$  satisfies  $\sup_{u,v \in \mathcal{D}} |\ell_{i,j}(u, v) - \ell(u)\ell(v)| < \infty$ .

### 2.2.2 Comments on the assumptions

The condition  $G(\tau_F) < 1$  in **A** implies  $\tau_F < \tau_G$  and will be needed in our proof of the result stated in Theorem 2.2.

The conditions **H** on the bandwidth are standard except the last one which is needed in applying an exponential inequality (see Theorem 1, inequality (4) in Doukhan and Neumann, 2007). The conditions **K** on the kernel was supposed by many authors in studying regression function under  $\alpha$ -mixing condition. The condition **M** quantifies the asymptotic independence when the distance between the "past" and the "future" is sufficiently large, while the conditions **D1-D4** are encountered in many papers when one deals with models under dependence.

### 2.2.3 Main results

The asymptotic behavior of the estimates  $\tilde{r}_{1,n}(\cdot)$  and  $r_n(\cdot)$  defined in (3.4) and (3.6) respectively, is described in the next theorems.

**Theorem 2.1.** *If the assumptions **A**, **H**, **K**, **M**, **D1**, **D3** and **D4** hold, then for  $n$  large enough we have*

$$\sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| = O\left(\sqrt{\frac{\log n}{nh^d}}\right), \text{ a.s.}$$

As it was mentioned above, there are processes which are associated but not mixing. In such cases, it would be interesting to have at our disposal the modest contribution of this paper.

It is noteworthy that for the proof of Theorem 2.1 we use similar tools used in the  $\alpha$ -mixing frameworks. The main difference here is that functional of associated rv's are not associated in general, which is the case when dealing with kernel functions for instance. This is due to the fact that the rv's  $K\left(\frac{z-Z_i}{h}\right)$  are, in general not associated while these variables remain  $\alpha$ -mixing if the  $Z_i$ 's are, since  $K$  is a measurable function in general. To keep the association, we should apply only monotone transformations to the original variables, which is not the case with a general kernel. This may be resolved by assuming the kernel  $K$  to be of bounded variation. This condition implies that we can write  $K = K_1 - K_2$ , where  $K_1$  and  $K_2$  are monotone functions. In this paper, we do not follow this idea but we use results stated for weakly dependent models in the sense of Doukhan and Louhichi (1999), since associated models are  $\kappa$ -weakly dependent.

Note also that to treat the fluctuation part we use bounds for covariances in applying an exponential inequality stated by Doukhan and Neumann (2007) for weakly dependent rv's. To do so, we use Theorem 5.3 in Bulinski and Shashkin (2007), and Proposition 8 in Doukhan and Neumann (2007).

**Theorem 2.2.** *Under the assumptions of Theorem 2.1, **M** and **D2**, for  $n$  large enough we have*

$$\sup_{x \in \mathcal{D}} |r_n(x) - r(x)| = O(h^2) + O_{a.s.}(n^{-\theta}) + O_{a.s.}\left(\sqrt{\frac{\log n}{nh^d}}\right),$$

where  $0 < \theta < \frac{\gamma}{2\gamma+9+\frac{3}{2}\rho}$ , for any real  $\rho > 0$ .

The proof of Theorem 2.2 is based on a classical decomposition of  $r_n(x) - r(x)$ . For this, we split the main result into three lemmas : Lemma 1 permits to get rates of convergence of order  $o(n^{-\theta})$  a.s., while Lemma 2 states a classical bound of order  $O(h^2)$  for the bias term of  $\tilde{r}_{1,n}$ . The last lemma states the strong uniform consistency rate for the estimator  $\ell_n$ , which is similar to those obtained in iid or  $\alpha$ -mixing cases. Then, to conclude the proof we use Theorem 2.1.

## 2.3 Simulation results

To support Theorem 2.2 above, a simulation was carried out to illustrate the behavior of the estimator  $r_n(\cdot)$  for different combinations of sample sizes  $n$  and censoring rates ( $CR$ ) with a one dimensional ( $d = 1$ ) covariate. For this, we restricted our self to the two cases described hereafter.

### 2.3.1 Models and procedure

The models are :

- Linear case :  $Y_i = 3X_i - 2 + \varepsilon_i$  for the target regression function  $r(x) = 3x - 2$ ,
- Nonlinear case :  $Y_i = X_i^2/(X_i + 1) + \varepsilon_i$  for that of  $r(x) = \frac{x^2}{x+1}$ .

Note that the association property is conserved by any monotonic transformation.

The sequence  $\{X_k, k = 1, \dots, n\}$  of associated rv's is generated as follows :

For  $i = 1, 2, \dots, n$ , we generate  $X_i = (W_{i-1} + W_{i-2})/2$  (see Chaubey and al. (2011)), where the sequences  $\{W_l, l = -1, 0, \dots, n-1\}$  and  $\{\varepsilon_i, i = 1, 2, \dots, n\}$  are iid rv's drawn from  $\exp(1)$  and  $\mathcal{N}(0, 0.02)$ , respectively. Then, we simulate a sequence  $\{C_i, i = 1, 2, \dots, n\}$  of iid rv's distributed as  $\exp(\lambda)$ . Hence, we keep the observed sequence  $\{Z_i = \min(Y_i, C_i), \delta_i, X_i, i = 1, \dots, n\}$ . We thus compute  $r_n(\cdot)$  using a gaussian kernel and bandwidths  $h = O\left((n^{-1} \log n)^{\frac{1}{3}}\right)$ .

On the one hand, we plotted the target curves together with  $r_n(\cdot)$  to display the quality of fit. For this, we first considered the combinations  $n = 150, CR = 0\%, 10\%, 30\%$  and  $CR = 30\%, n = 50, 150, 300$  for the linear model (Figure 3.1 and Figure 3.2) and then  $CR = 20\%, n = 100, 300, 500$  for the nonlinear case (Figure 3.3).

On the other hand, for each model we repeated  $B = 200$  times the above procedure for every fixed combination of  $n = 50, 150, 300$  and  $CR = 0\%, 10\%, 30\%$ . The estimating errors were quantified by using the mean square error (MSE) and the mean integrated square error (MISE) criteria, that is

$$MSE(x) = \frac{1}{B} \sum_{k=1}^B (r_{n,k}(x) - r(x))^2 \quad \text{and} \quad MISE = \int_0^{3/2} MSE(x) dx,$$

where  $r_{n,k}(\cdot)$  is the value of  $r_n(\cdot)$  computed at iteration  $k$ .

In Table 2.1, we considered the median values of the MSEs instead of those of the MSEs. Note that the values reported in Table 2.1 and Table 2.2 were computed over the range  $[0, \frac{3}{2}]$ .

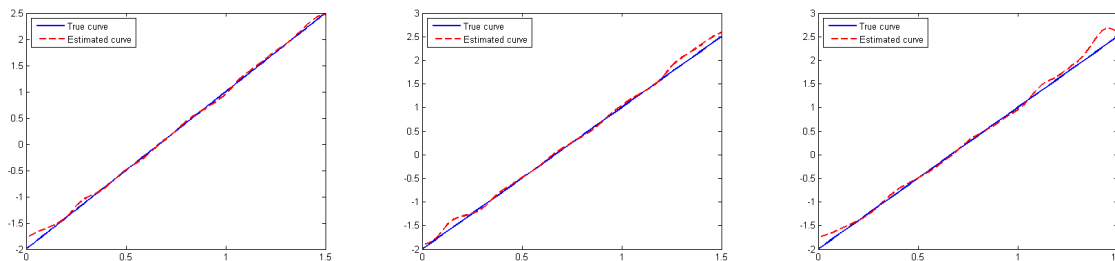


FIGURE 2.1 – Linear case with  $n = 150$  and  $CR = 0\%, 10\%, 30\%$ .

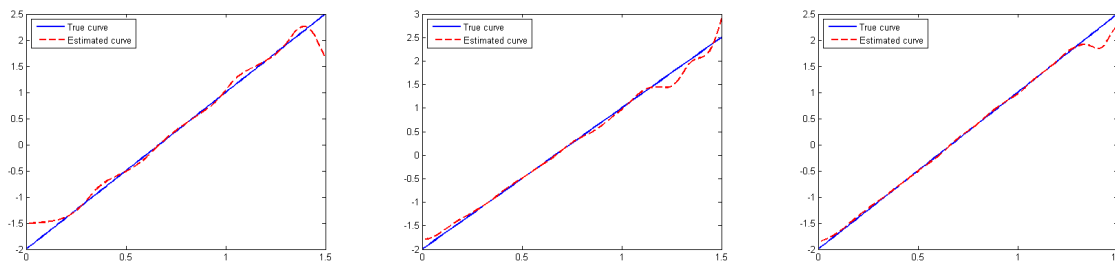


FIGURE 2.2 – Linear case with  $CR = 30\%$  and  $n = 50, 150, 300$ .

### 2.3.2 Comments on the simulation results

As it can be seen from the tables and figures, the errors increase along with the censoring rate (CR) values and decrease when the sample sizes increase. This means that the estimation is sensitive to the censoring effect. Nevertheless, the errors approach zero in each case when  $n$  increases. Likewise, the quality of fit tends to be damaged for sufficiently high ratio of censoring but, this quality increases when  $n$  increases and becomes better in any cases.

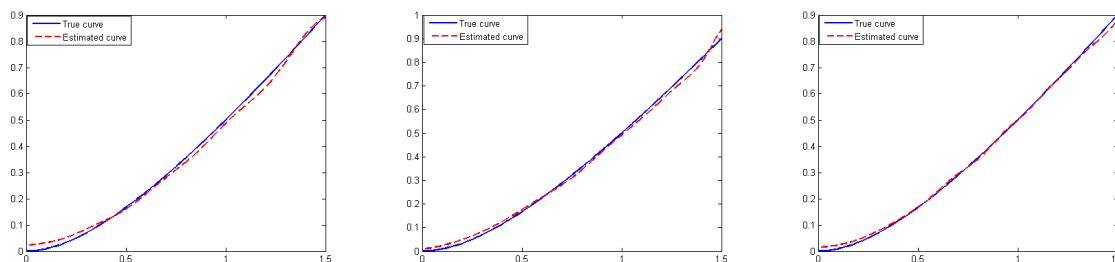


FIGURE 2.3 – Nonlinear case with  $CR = 20\%$  and  $n = 100, 300, 500$ .

CR	n=50	n=150	n=300
0%	0.0132	0.0053	0.0024
10%	0.0172	0.0057	0.0031
30%	0.0259	0.0078	0.0050

TABLE 2.1 – Median values of the  $MSEs$  for  $r_n$ .

CR	n=50	n=150	n=300
0%	0.0497	0.0206	0.0129
10%	0.0727	0.0290	0.0166
30%	0.1197	0.0450	0.0342

TABLE 2.2 –  $MISE$  values for  $r_n$ .

## 2.4 Proofs

For notational convenience we set

$$W_i(x, h) := \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) - \mathbb{E} \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right]. \quad (2.8)$$

Then we have

$$\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x)) = \frac{1}{nh^d} \sum_{i=1}^n W_i(x, h). \quad (2.9)$$

The proof of Theorem 2.1 needs the following results listed in Proposition 2.1 and Proposition 2.2.

**Proposition 2.1.** *Let  $W_1(x, h), \dots, W_n(x, h)$  be rv's defined as in (3.29). Under the assumptions of Theorem 2.1, and for any  $u$ -tuples  $(s_1, s_2, \dots, s_u)$  and any  $v$ -tuples  $(w_1, w_2, \dots, w_v)$  such that  $1 \leq s_1 \leq \dots \leq s_u \leq w_1 \leq \dots \leq w_v \leq n$ , we have*

- i)  $\text{cov} \left( \prod_{i=s_1}^{s_u} W_i(x, h), \prod_{j=w_1}^{w_v} W_j(x, h) \right) =: C_1 \leq c^{u+v} h^{-2} uv \rho(w_1 - s_u)$ ,
- ii)  $\text{cov} \left( \prod_{i=s_1}^{s_u} W_i(x, h), \prod_{j=w_1}^{w_v} W_j(x, h) \right) =: C_2 \leq c^{u+v} h^{2d}$ .

**Proof of Proposition 2.1** To prove item i), we first recall that  $L : \mathbb{R}^m \rightarrow \mathbb{R}$  is a Lipschitz function if

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

where  $\|z\|_1 = |z_1| + |z_2| + \dots + |z_n|$ . For such  $L$ , the partial Lipschitz constants (see Definition 5.1, p.88, in Bulinski and Shashkin, 2007) are

$$\text{Lip}_i(L) = \sup_{\substack{z_1, \dots, z_m, z'_i \in \mathbb{R} \\ z_i \neq z'_i}} \frac{|L(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_m) - L(z_1, \dots, z_{i-1}, z'_i, z_{i+1}, \dots, z_m)|}{|z_i - z'_i|}.$$

Then by setting  $\prod_{i=t_1}^{t_\zeta} W_i =: L_\zeta : \mathbb{R}^{\zeta(d+1)} \rightarrow \mathbb{R}; \zeta \in \mathbb{N} - \{0\}$ , and using Theorem 5.3 in Bulinski and Shashkin (2007, p.89) we get

$$C_1 \leq \sum_{i=s_1}^{s_u} \sum_{j=w_1}^{w_v} \text{Lip}_i \left( \prod_{l=s_1}^{s_u} W_l \right) \text{Lip}_j \left( \prod_{m=w_1}^{w_v} W_m \right) \Delta_{i,j} \leq \frac{c^2}{h^2} \left( \frac{2\tau_F \|K_d\|_\infty}{\bar{G}(\tau_F)} \right)^{u+v-2} \sum_{i=s_1}^{s_u} \sum_{j=w_1}^{w_v} \Delta_{i,j},$$

with  $c = \tau_F (\bar{G}(\tau_F))^{-1} \text{Lip}(K_d) \vee \|K_d\|_\infty \left( (\bar{G}(\tau_F))^{-1} + \tau_F (\bar{G}^2(\tau_F))^{-1} \text{Lip}(G) \right)$ .

Next, by stationarity and Assumption **M** we obtain

$$C_1 \leq \frac{c^2}{h^2} \left( \frac{2\tau_F \|K_d\|_\infty}{\bar{G}(\tau_F)} \right)^{u+v-2} uv\rho(w_1 - s_u) = c^{u+v} h^{-2} uv\rho(w_1 - s_u)$$

Thus result *i*) holds. The result in item *ii*) follows by simple algebra using assumptions **K** and **D2**. The proof of Proposition 2.1 is now finished.  $\square$

**Proposition 2.2.** *There exist constants  $m_0, L_1, L_2 < +\infty, \mu, \lambda \geq 0$  and a non-increasing sequence of real numbers  $(\phi(n))_{n \geq 1}$  such that for any  $u$ -tuples  $(s_1, s_2, \dots, s_u)$  and any  $v$ -tuples  $(w_1, w_2, \dots, w_v)$  satisfying  $1 \leq s_1 \leq \dots \leq s_u \leq w_1 \leq \dots \leq w_v \leq n$ , we have*

- a).  $\text{cov} \left( \prod_{i=s_1}^{s_u} W_i(x, h), \prod_{j=w_1}^{w_v} W_j(x, h) \right) =: \text{Cov} \leq c^{u+v} h^d uv\phi(w_1 - s_u)$ ,
- b).  $\sum_{s=0}^{\infty} (s+1)^{k_0} \phi(s) \leq L_1 L_2^{k_0} (k_0!)^\mu, \forall k_0 \geq 0$ ,
- c).  $\mathbb{E}|W_i(x, h)|^{k_0} \leq (k_0!)^\lambda m_0^{k_0}$ .

**Proof of Proposition 2.2** From Proposition 2.1 we have

$$C_1^{\frac{d}{2d+2}} \leq (c^{u+v} h^{-2} uv\rho(w_1 - s_u))^{\frac{d}{2d+2}}, \quad (2.10)$$

$$C_2^{\frac{d+2}{2d+2}} \leq (c^{u+v} h^{2d})^{\frac{d+2}{2d+2}}. \quad (2.11)$$

Then combining (2.10) and (2.11) we get

$$C_1^{\frac{d}{2d+2}} C_2^{\frac{d+2}{2d+2}} \leq c^{u+v} h^d (uv)^{\frac{d}{2d+2}} \rho^{\frac{d}{2d+2}}(w_1 - s_u) \leq c^{u+v} h^d uv\rho^{\frac{d}{2d+2}}(w_1 - s_u). \quad (2.12)$$

Consequently, by setting  $\phi(\cdot) = \rho^{\frac{d}{2d+2}}(\cdot)$ , the inequality in (2.12) implies item a), that is

$$\text{Cov} \leq c^{u+v} h^d uv\phi(w_1 - s_u).$$

The proofs of b) and c) follow nearly the same lines as in the proof of Proposition 8 in Doukhan and Neumann (2007) by choosing  $k_0 = 0, \mu = 1$  and  $L_1 = L_2 = \frac{1}{1 - \exp(-\gamma d/(2d+2))}$ . Then we omit them.  $\square$

**Proof of Theorem 2.1** The proof uses a classical technique to cover the compact  $\mathcal{D}$ . Let  $p_n$  denote a number of balls  $B_k(x_k, a_n^d)$  centered at  $x_k \in \mathbb{R}^d$ . Then for all  $x \in \mathcal{D}$  there exists an integer

$k$ ,  $1 \leq k \leq p_n$  such that  $\|x - x_k\| \leq a_n^d$ , with  $a_n^d = (n^{-1}h^{d+2\alpha})^{1/2\alpha}$ . And, as  $\mathcal{D}$  is bounded, then there exists a constant  $M_0 > 0$  such that  $p_n \leq M_0 a_n^{-d}$ . Furthermore observe that we have

$$\begin{aligned} \sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| &\leq \max_{1 \leq k \leq p_n} \sup_{x \in B_k} |\tilde{r}_{1,n}(x) - \tilde{r}_{1,n}(x_k)| + \max_{1 \leq k \leq p_n} \sup_{x \in B_k} |\mathbb{E}(\tilde{r}_{1,n}(x)) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| \\ &\quad + \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| \\ &=: J_1 + J_2 + J_3. \end{aligned}$$

Firstly, under Assumption **K** and using the SLLN for associated rv's (see Birkel, 1988b), we have

$$\begin{aligned} J_1 &\leq \max_{1 \leq k \leq p_n} \sup_{x \in B_k} \frac{1}{nh^d} \sum_{i=1}^n \left| \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) - \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x_k - X_i}{h} \right) \right| \\ &\leq \max_{1 \leq k \leq p_n} \sup_{x \in B_k} \frac{1}{nh^d} \sum_{i=1}^n \left| \frac{\delta_i Y_i}{\bar{G}(Y_i)} K_d \left( \frac{x - X_i}{h} \right) - \frac{\delta_i Y_i}{\bar{G}(Y_i)} K_d \left( \frac{x_k - X_i}{h} \right) \right| \\ &\leq \max_{1 \leq k \leq p_n} \sup_{x \in B_k} \left\{ \frac{1}{h^d \bar{G}(\tau_F)} \left\| \frac{x - x_k}{h} \right\|^\alpha \frac{1}{n} \sum_{i=1}^n Y_i \right\} \\ &\leq \frac{c}{\bar{G}(\tau_F)} \frac{1}{\sqrt{nh^d}} = O \left( \frac{1}{\sqrt{nh^d}} \right). \end{aligned} \tag{2.13}$$

The term  $J_2$  is treated using similar arguments as for  $J_1$  and admits the same upper bound in (2.13). At this stage and further, let  $\varepsilon > 0$ . Then under Assumption **H** and (2.13), for  $n$  sufficiently large we have

$$\sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| \leq 2\varepsilon + \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))|. \tag{2.14}$$

Now in order to study the term  $J_3$ , first and foremost observe that from (2.9) one has

$$J_3 = \max_{1 \leq k \leq p_n} \frac{1}{nh^d} \left| \sum_{i=1}^n W_i(x_k, h) \right|.$$

Then to upper bound the fluctuation part  $J_3$ , we first use a Bernstein-type inequality (see Doukhan and Neumann, 2007, Theorem 1, inequality (4), p. 880) which is applicable to associated sequences, since associated process is  $\kappa$ -weakly dependent process (see Doukhan and Neumann, 2008, Section 3, Example 1, p. 154). Then we apply the Borel-Cantelli lemma. We have

$$\mathbb{P} \left( \sum_{i=1}^n W_i(x_k, h) \geq \varepsilon \right) \leq \exp \left( -\frac{1}{2} \frac{\varepsilon^2}{A_n + B_n^{1/(\mu+\lambda+2)} \varepsilon^{(2\eta+2\lambda+3)/(\eta+\lambda+2)}} \right), \tag{2.15}$$

where  $A_n$  can be chosen such that

$$A_n \leq \sigma_n^2 \text{ with } \sigma_n^2 := \text{var} \left( \sum_{i=1}^n W_i(x_k, h) \right)$$

and

$$B_n = 2cL_2 \left( (2^{4+\mu+\lambda} cnh^d L_1 / A_n) \vee 1 \right).$$

For this purpose, we first evaluate  $\sigma_n^2 = (nh^d)^2 \text{var}(\tilde{r}_{1,n}(x_k))$ . We have

$$\begin{aligned} (nh^d)^2 \text{var}(\tilde{r}_{1,n}(x_k)) &= n \text{var} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x_k - X_1}{h} \right) \right) \\ &\quad + \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x_k - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x_k - X_j}{h} \right) \right) \\ &=: V_1 + V_2. \end{aligned}$$

On the one hand, we have

$$\begin{aligned} V_1 &= n \left\{ \mathbb{E} \left[ \frac{\delta_1^2 Z_1^2}{\bar{G}^2(Z_1)} K_d^2 \left( \frac{x_k - X_1}{h} \right) \right] - \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x_k - X_1}{h} \right) \right] \right\} \\ &=: n \{V_{11} - V_{12}\}, \end{aligned}$$

with

$$V_{11} = \mathbb{E} \left[ K_d^2 \left( \frac{x_k - X_1}{h} \right) \mathbb{E} \left( \frac{\delta_1^2 Z_1^2}{\bar{G}^2(Z_1)} | X_1 \right) \right] = \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h} \right) \mathbb{E} \left( \frac{\mathbb{1}_{\{Y_1 \leq C_1\}} Y_1^2}{\bar{G}^2(Y_1)} | X_1 = u \right) \ell(u) du.$$

Again, conditioning with respect to  $Y_1$  and using a change of variables we get

$$\begin{aligned} V_{11} &= \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h} \right) \mathbb{E} \left( \frac{Y_1^2}{\bar{G}(Y_1)} | X_1 = u \right) \ell(u) du = \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h} \right) \left( \int_{\mathbb{R}} \frac{y^2}{\bar{G}(y)} f_{X,Y}(u, y) dy \right) du \\ &= \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h} \right) r_2(u) du = h^d \int_{\mathbb{R}^d} K_d^2(t) r_2(x_k - th) dt, \quad t = (t_1, t_2, \dots, t_d)^\top. \end{aligned}$$

Then using **D3** and Taylor expansion around  $x_k = (x_{k,1}, x_{k,2}, \dots, x_{k,d})^\top$ , we obtain

$$V_{11} = h^d \int r_2(x_k) K_d^2(t) dt - h^{d+1} \int K_d^2(t) \left[ t_1 \frac{\partial r_2}{\partial x_{k,1}}(x^*) + \dots + t_d \frac{\partial r_2}{\partial x_{k,d}}(x^*) \right] = O(h^d),$$

where  $x^*$  is between  $x - th$  and  $x$ . Now, to end with  $V_1$ , we evaluate  $V_{12}$ .

$$\begin{aligned} V_{12} &= \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x_k - X_1}{h} \right) \right] = \mathbb{E}^2 \left[ K_d \left( \frac{x_k - X_1}{h} \right) \mathbb{E} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} | X_1 \right) \right] \\ &= \left[ \int_{\mathbb{R}^d} K_d \left( \frac{x_k - u}{h} \right) r(u) \ell(u) du \right]^2 = \left[ \int_{\mathbb{R}^d} K_d \left( \frac{x_k - u}{h} \right) r_1(u) du \right]^2. \end{aligned}$$

Next, using a change of variables, a Taylor expansion and assumptions **H,K c)** and **D1** we get  $V_{12} = O(h^{2d})$ . Thus  $V_1 = O(nh^d)$ .

On the other hand we have

$$\begin{aligned} V_2 &= \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x_k - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x_k - X_j}{h} \right) \right) \\ &=: \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i,k}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j,k} \right). \end{aligned}$$

The computation of the covariance term is done using conditional expectation techniques. Indeed, under assumptions **H,K**, and **D4** we have

$$\begin{aligned}
\text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i,k}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j,k} \right) &= \mathbb{E} \left( \mathbb{E} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,i,k} K_{d,j,k} \mid X_i, X_j \right) \right) \\
&- \mathbb{E} \left( \mathbb{E} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i,k} \mid X_i \right) \right) \mathbb{E} \left( \mathbb{E} \left( \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j,k} \mid X_j \right) \right) \\
&= \mathbb{E} \left( K_{d,i,k} K_{d,j,k} \mathbb{E} \left( \frac{\delta_i \delta_j Z_i Z_j}{\bar{G}(Z_i) \bar{G}(Z_j)} \mid X_i, X_j \right) \right) \\
&- \mathbb{E} \left( K_{d,i,k} \mathbb{E} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} \mid X_i \right) \right) \mathbb{E} \left( K_{d,j,k} \mathbb{E} \left( \frac{\delta_j Z_j}{\bar{G}(Z_j)} \mid X_j \right) \right) \\
&= \mathbb{E} \left[ K_{d,i,k} K_{d,j,k} \mathbb{E} \left[ \mathbb{E} \left( \frac{\mathbb{1}_{\{Y_i \leq c_i\}} \mathbb{1}_{\{Y_j \leq c_j\}} Y_i Y_j}{\bar{G}(Y_i) \bar{G}(Y_j)} \mid Y_i, Y_j \right) \mid X_i, X_j \right] \right] \\
&- \mathbb{E} \left[ K_{d,i,k} \mathbb{E} \left[ \mathbb{E} \left( \frac{\mathbb{1}_{\{Y_i \leq c_i\}} Y_i}{\bar{G}(Y_i)} \mid Y_i \right) \mid X_i \right] \right] \\
&\times \mathbb{E} \left[ K_{d,j,k} \mathbb{E} \left[ \mathbb{E} \left( \frac{\mathbb{1}_{\{Y_j \leq c_j\}} Y_j}{\bar{G}(Y_j)} \mid Y_j \right) \mid X_j \right] \right] \\
&\leq ch^{2d} \int_{\mathbb{R}^d \times \mathbb{R}^d} K_d(t) K_d(z) \tag{2.16} \\
&\times [\ell_{ij}(x_k - th, x_k - zh) - \ell(x_k - th) \ell(x_k - zh)] dt dz
\end{aligned}$$

$$= O(h^{2d}). \tag{2.17}$$

Now to deal with  $V_2$ , let us define

$$\mathcal{B}_1 = \{(i, j); 1 \leq |i - j| \leq \eta_n\} \text{ and } \mathcal{B}_2 = \{(i, j); \eta_n + 1 \leq |i - j| \leq n - 1\}, \tag{2.18}$$

where  $\eta_n = o(n)$ . Then we write

$$\begin{aligned}
V_2 &= \sum_{i=1}^n \sum_{j \in \mathcal{B}_1} \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i,k}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j,k} \right) + \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i,k}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j,k} \right) \\
&=: V_{21} + V_{22}.
\end{aligned}$$

Firstly, from (2.17) and (3.19) we obtain

$$V_{21} = O(\eta_n n h^{2d}). \tag{2.19}$$

Secondly, by Assumption **M** and Proposition 2.2, result a) we get

$$\begin{aligned}
V_{22} &\leq \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} ch^d \phi(|i - j|) \\
&\leq cnh^d \sum_{j \in \mathcal{B}_2} \exp \left( -\frac{\gamma d}{2d+2} |i - j| \right) \\
&\leq cnh^d \int_{\eta_n}^n \exp \left( -\frac{\gamma d}{2d+2} u \right) du \\
&= O \left( nh^d \exp \left( -\frac{\gamma d}{2d+2} \eta_n \right) \right). \tag{2.20}
\end{aligned}$$

Now choosing  $\eta_n = O(h^{\nu-d})$  with  $0 < \nu < d$ , the statements in (2.19) and (2.20) yield  $V_{21} = o(nh^d)$  and  $V_{22} = o(nh^d)$ . Consequently, we get  $\sigma_n^2 = V_1 + V_2 = O(nh^d) + o(nh^d) = O(nh^d)$ .

Finally, in the right-hand side of (2.15) we choose  $A_n = O(nh^d)$ ,  $\lambda = 0$ , and the above values for  $\mu$ ,  $L_1$  and  $L_2$  (see the proof of *b*) and *c*) which permit to get  $B_n = O(1)$ . Then from (2.15) we have

$$\begin{aligned}
\mathbb{P} \left\{ \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| > \varepsilon \right\} &= \mathbb{P} \left\{ \max_{1 \leq k \leq p_n} \left| \sum_{i=1}^n W_i(x_k, h) \right| > nh^d \varepsilon \right\} \\
&\leq \sum_{i=1}^{p_n} \mathbb{P} \left\{ \left| \sum_{i=1}^n W_i(x_k, h) \right| > nh^d \varepsilon \right\} \\
&\leq 2p_n \exp \left( -\frac{\varepsilon^2 n^2 h^{2d} / 2}{A_n + B_n^{1/(\mu+\lambda+2)} (\varepsilon nh^d)^{(2\eta+2\lambda+3)/(\eta+\lambda+2)}} \right) \\
&\leq 2p_n \exp \left( -\frac{\varepsilon^2 n^2 h^{2d} / 2}{c nh^d + (\varepsilon nh^d)^{5/3}} \right). \tag{2.21}
\end{aligned}$$

If we replace  $\varepsilon$  by  $\varepsilon_0 \sqrt{\frac{\log n}{nh^d}}$  in statement (2.21) we get

$$\begin{aligned}
\mathbb{P} \left\{ \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| > \varepsilon_0 \sqrt{\frac{\log n}{nh^d}} \right\} &\leq 2p_n \exp \left( -\frac{\varepsilon_0^2 n^2 h^{2d} \log n / 2 nh^d}{c nh^d + \varepsilon_0^{5/3} (nh^d)^{5/3} \left(\frac{\log n}{nh^d}\right)^{5/6}} \right) \\
&\leq 2p_n \exp \left( -\frac{\varepsilon_0^2 \log n / 2}{c + \varepsilon_0^{5/3} (nh^d)^{2/3} (\log n / nh^d)^{5/6}} \right) \\
&\leq 2p_n n^{-\left(\varepsilon_0^2/2\right) \left(c + \varepsilon_0^{5/3} (\log^5 n / nh^d)^{1/6}\right)^{-1}} \\
&\leq 2M_0 a_n^{-d} n^{-c\varepsilon_0^2} = M_0 (n^{-1} h^{-d-2\alpha})^{\frac{1}{2\alpha}} n^{-c\varepsilon_0^2} \\
&\leq 2 \frac{M_0}{\sqrt{(nh)^{\frac{2\alpha+d}{\alpha}}}} n^{-c\varepsilon_0^2 + 1 + \frac{d+1}{2\alpha}}. \tag{2.22}
\end{aligned}$$

Then, for an appropriate choice of  $\varepsilon_0$  and Assumption **H**, the term in (2.22) becomes an  $O(n^{-\zeta})$  with  $\zeta > 1$ . So, we obtain

$$\mathbb{P} \left\{ \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| > \varepsilon_0 \sqrt{\frac{\log n}{nh^d}} \right\} = O(n^{-\zeta}). \tag{2.23}$$

Then from (2.14) and by setting  $\varepsilon'_0 = 3\varepsilon_0$  we can conclude that

$$\mathbb{P} \left\{ \sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| > \varepsilon'_0 \sqrt{\frac{\log n}{nh^d}} \right\} \leq \mathbb{P} \left\{ \max_{1 \leq k \leq p_n} |\tilde{r}_{1,n}(x_k) - \mathbb{E}(\tilde{r}_{1,n}(x_k))| > \varepsilon_0 \sqrt{\frac{\log n}{nh^d}} \right\}.$$

Now, as the right-hand side in (2.23) can be made as a general term of a convergent series, then

$$\sum_{n \geq 1} \mathbb{P} \left\{ \sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| > \varepsilon'_0 \sqrt{\frac{\log n}{nh^d}} \right\} < \infty,$$

which, by the Borel-Cantelli lemma implies that

$$\sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| = O\left(\sqrt{\frac{\log n}{nh^d}}\right), \text{ a.s. as } n \rightarrow +\infty.$$

This ends the proof of Theorem 2.1.  $\square$

**Proof of Theorem 2.2** The proof is mainly based on the following inequality.

$$\begin{aligned} \sup_{x \in \mathcal{D}} |r_n(x) - r(x)| &\leq \sup_{x \in \mathcal{D}} \left| \frac{r_{1,n}(x)}{\ell_n(x)} - \frac{\tilde{r}_{1,n}(x)}{\ell_n(x)} \right| + \sup_{x \in \mathcal{D}} \left| \frac{\tilde{r}_{1,n}(x)}{\ell_n(x)} - \frac{\mathbb{E}(\tilde{r}_{1,n}(x))}{\ell_n(x)} \right| \\ &\quad + \sup_{x \in \mathcal{D}} \left| \frac{\mathbb{E}(\tilde{r}_{1,n}(x))}{\ell_n(x)} - \frac{r_1(x)}{\ell_n(x)} \right| + \sup_{x \in \mathcal{D}} \left| \frac{r_1(x)}{\ell_n(x)} - \frac{r_1(x)}{\ell(x)} \right| \\ &\leq \frac{1}{\inf_{x \in \mathcal{D}} (\ell_n(x))} \left\{ \sup_{x \in \mathcal{D}} |r_{1,n}(x) - \tilde{r}_{1,n}(x)| + \sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| \right. \\ &\quad \left. + \sup_{x \in \mathcal{D}} |\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| + \sup_{x \in \mathcal{D}} \frac{|r_1(x)|}{\ell(x)} \sup_{x \in \mathcal{D}} |\ell(x) - \ell_n(x)| \right\} \\ &\leq \frac{1}{\Gamma - \sup_{x \in \mathcal{D}} |\ell_n(x) - \ell(x)|} \left\{ \sup_{x \in \mathcal{D}} |r_{1,n}(x) - \tilde{r}_{1,n}(x)| + \sup_{x \in \mathcal{D}} |\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))| \right. \\ &\quad \left. + \sup_{x \in \mathcal{D}} |\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| + \Gamma^{-1} \sup_{x \in \mathcal{D}} |r_1(x)| \sup_{x \in \mathcal{D}} |\ell(x) - \ell_n(x)| \right\} \\ &=: \frac{1}{\Gamma - \sup_{x \in \mathcal{D}} |\ell_n(x) - \ell(x)|} \left\{ \kappa_1 + \kappa_2 + \kappa_3 + \Gamma^{-1} \kappa_4 \sup_{x \in \mathcal{D}} |r_1(x)| \right\}. \end{aligned} \quad (2.24)$$

As  $\kappa_2$  is bounded by Theorem 2.1, it remains to bound  $\kappa_1$ ,  $\kappa_3$  and  $\kappa_4$ . For this we prove the following lemmas.

**Lemma 2.1.** *Under assumptions **A**, **M**, **K** and **D2** we have*

$$\sup_{x \in \mathcal{D}} |r_{1,n}(x) - \tilde{r}_{1,n}(x)| = o(n^{-\theta}), \text{ a.s. as } n \rightarrow +\infty.$$

**Proof of Lemma 2.1** The proof uses essentially the results of strong uniform consistency rate for  $G_n$  stated by Cai and Roussas (1998), and the fact  $0 \leq Y \leq \tau_F < \infty$  a.s. Indeed we have

$$\begin{aligned} \sup_{x \in \mathcal{D}} |r_{1,n}(x) - \tilde{r}_{1,n}(x)| &= \sup_{x \in \mathcal{D}} \left| \frac{1}{nh^d} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}_n(Z_i)} K_d\left(\frac{x - X_i}{h}\right) - \frac{1}{nh^d} \sup_{x \in \mathcal{D}} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d\left(\frac{x - X_i}{h}\right) \right| \\ &= \sup_{x \in \mathcal{D}} \left| \frac{1}{nh^d} \sum_{i=1}^n \frac{Y_i \mathbf{1}_{\{Y_i \leq C_i\}}}{\bar{G}_n(Y_i)} K_d\left(\frac{x - X_i}{h}\right) - \frac{1}{nh^d} \sum_{i=1}^n \frac{Y_i \mathbf{1}_{\{Y_i \leq C_i\}}}{\bar{G}(Y_i)} K_d\left(\frac{x - X_i}{h}\right) \right| \\ &\leq \sup_{x \in \mathcal{D}} \frac{1}{nh^d} \sum_{i=1}^n \left| Y_i K_d\left(\frac{x - X_i}{h}\right) \left[ \frac{1}{\bar{G}_n(Y_i)} - \frac{1}{\bar{G}(Y_i)} \right] \right| \\ &\leq \frac{\tau_F}{\bar{G}_n(\tau_F) \bar{G}(\tau_F)} \sup_{t \leq \tau_F} |\bar{G}_n(t) - \bar{G}(t)| \times \sup_{x \in \mathcal{D}} \frac{1}{nh^d} \sum_{i=1}^n K_d\left(\frac{x - X_i}{h}\right), \text{ a.s.} \end{aligned}$$

Then using Markov's inequality, a change of variables, assumptions **K,D2** and the Dominated Convergence Theorem we get, for any  $\varepsilon > 0$ ,

$$\mathbb{P} \left( \frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \geq \varepsilon \right) \leq \frac{1}{\varepsilon h^d} \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) \ell(u) du = \varepsilon^{-1} \int_{\mathbb{R}^d} K_d(t) \ell(x - th) dt = O(1).$$

Therefore, by using Theorem 1.4 in Cai and Roussas (1998), we obtain

$$\sup_{t \leq \tau_F} |\bar{G}_n(t) - \bar{G}(t)| = o(n^{-\theta}) \text{ a.s. as } n \rightarrow +\infty.$$

where  $0 < \theta < \frac{\gamma}{2\gamma + 9 + \frac{3}{2}\rho}$ , for any real  $\rho > 0$ , and  $\gamma$  is given in assumption **M**.

Finally,  $\bar{G}_n(\tau_F) \rightarrow \bar{G}(\tau_F)$  and  $\frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) = O(1)$ , complete the proof of Lemma 2.1.  $\square$

**Lemma 2.2.** *Under assumptions **H,K** and **D1** we have*

$$\sup_{x \in \mathcal{D}} |\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| = O(h^2), \text{ a.s. as } n \rightarrow +\infty.$$

**Proof of Lemma 2.2.** The proof is based on conditional expectation techniques and the fact that  $\mathbb{E} \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} | X_1 = u \right] = \mathbb{E} \left[ \frac{\delta_1 Y_1}{\bar{G}(Y_1)} | X_1 = u \right] = r(u)$ . Indeed we have

$$\kappa_3 = \sup_{x \in \mathcal{D}} \left| \mathbb{E} \left( \frac{1}{nh^d} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) \right) - r_1(x) \right| = \sup_{x \in \mathcal{D}} \left| \mathbb{E} \left( \frac{1}{h^d} \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right) - r_1(x) \right|.$$

Firstly we have

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{h^d} \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right] &= \int_{\mathbb{R}^d} \frac{1}{h^d} K_d \left( \frac{x - X_1}{h} \right) \mathbb{E} \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} | X_1 = u \right] \ell(u) du \\ &= \int_{\mathbb{R}^d} \frac{1}{h^d} K_d \left( \frac{x - u}{h} \right) r(u) \ell(u) du. \end{aligned}$$

Then under assumptions **H,K a),K c),D1**, a change of variables and a Taylor expansion give

$$\begin{aligned} \kappa_3 &= \sup_{x \in \mathcal{D}} \left| \int_{\mathbb{R}^d} \frac{1}{h^d} K_d \left( \frac{x - u}{h} \right) r_1(u) du - r_1(x) \right| = \sup_{x \in \mathcal{D}} \left| \int_{\mathbb{R}^d} K_d(t) [r_1(x - th) - r_1(x)] dt \right| \\ &= O(h^2). \end{aligned}$$

The proof of Lemma 2.2 is finished.  $\square$

**Lemma 2.3.** *Under assumptions **H,K,D2,D3** and **D4** we have*

$$\sup_{x \in \mathcal{D}} |\ell_n(x) - \ell(x)| = O \left( \max \left\{ h^2, \sqrt{\frac{\log n}{nh^d}} \right\} \right), \text{ a.s. as } n \rightarrow +\infty.$$

**Proof of Lemma 2.3.** First of all, observe that this lemma can be derived from Theorem 2.1 with  $\bar{G}(Z_i) = 1$ ,  $\delta_i = 1$  and  $Z_i = 1$ , for any  $i = 1, \dots, n$ . Then under assumptions **H,K,D2,D3** and **D4**, the following inequality

$$\sup_{x \in \mathcal{D}} |\ell_n(x) - \ell(x)| \leq \sup_{x \in \mathcal{D}} |\ell_n(x) - \mathbb{E}(\ell_n(x))| + \sup_{x \in \mathcal{D}} |\mathbb{E}(\ell_n(x)) - \ell(x)|, \quad (2.25)$$

suggests to treat the first term in the right-hand side by following the same steps and similar arguments used in proving Theorem 2.1. Hence we get

$$\sup_{x \in \mathcal{D}} |\ell_n(x) - \mathbb{E}(\ell_n(x))| = O\left(\sqrt{\frac{\log n}{nh^d}}\right), \text{ a.s. as } n \rightarrow +\infty.$$

Furthermore, the bias term in (2.25) does not depend upon the dependency structure of the data. Then under **K** and **D2** and using a Taylor expansion, we get

$$\sup_{x \in \mathcal{D}} |\mathbb{E}(\ell_n(x)) - \ell(x)| = O(h^2), \text{ a.s. as } n \rightarrow +\infty,$$

which permits to end the proof of Lemma 2.3.  $\square$

Thus, the statement in (2.24), together with lemmas 2.1, 2.2, 2.3 and Theorem 2.1 yield the result in Theorem 2.2.  $\square$

# Chapitre 3

## Normalité asymptotique

### Normalité asymptotique de l'estimateur à noyau de la fonction de régression pour des données censurées et associées

#### 3.1 Introduction

Nous reprenons dans ce chapitre une partie de l'introduction du chapitre précédent ainsi que certains résultats pour des raisons de notations et pour rendre la lecture de ce chapitre indépendante. Soit  $\{Y_i, i \geq 1\}$  une suite strictement stationnaire de variables aléatoires (v.a.) associées, de fonction de répartition (f.r)  $F$  et  $X$  un vecteur aléatoire (covariable) prenant ses valeurs dans  $\mathbb{R}^d$  de densité  $\ell$ . Dans la suite, nous supposons  $Y$  bornée et non négative.

L'analyse de survie est une partie des statistiques, dans laquelle la variable d'intérêt peut souvent être interprétée comme le temps écoulé entre deux événements. De telles variables (durée de vie) apparaissent généralement dans un contexte médical ou d'ingénierie.

Une caractéristique typique de cette analyse des données de survie est que, en raison des limitations de temps,  $Y$  peut ne pas toujours être observable. Au lieu de cela, on obtient seulement le minimum du temps de survie  $Y_i$  et un temps de censure non informatif  $C_i$  avec l'information  $\delta_i$  pour savoir si le temps de survie d'un individu est inférieur au temps de censure ou non. Dans ce chapitre, les v.a.  $\{C_i, 1 \leq i \leq n\}$  sont non négatives, supposées être indépendantes et identiquement distribuées (i.i.d.) et de fonction de répartition  $G$  inconnue.

Dans le modèle de censure à droite, on n'observe que les  $n$  paires  $(Z_i, \delta_i)$  avec  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ , où  $\mathbb{1}_A$  est la fonction indicatrice de l'évènement  $A$ . Il est bien connu que, dans ce modèle le vecteur aléatoire observé  $(X, Y, \delta)$  n'identifie pas la distribution conditionnelle de  $Y$  sachant  $X$ . Pour surmonter le problème d'identifiabilité, on suppose que  $C$  est indépendante de  $(X, Y)$ .

La fonction de régression à un point  $x \in \mathbb{R}^d$  est l'espérance conditionnelle de  $Y$  sachant  $X = x$

$$r(x) := \mathbb{E}[Y|X = x]. \quad (3.1)$$

Dans ce travail, nous nous intéressons à l'estimateur à noyau de Nadaraya-Watson(1964). Nous pouvons alors définir la fonction de régression par :

$$r(x) = \mathbb{E}[Y|X = x] = \int_{\mathbb{R}} \frac{yf_{X,Y}(x, y)}{\ell(x)} dy =: \frac{r_1(x)}{\ell(x)}, \quad (3.2)$$

où  $f_{X,Y}(\cdot, \cdot)$  est une densité conjointe de  $X, Y$  et  $\ell(\cdot)$  est une densité marginale de la covariable  $X$ . Dans le cas des données complètes, il existe une vaste littérature consacrée à l'étude de l'estimateur non paramétrique à noyau de  $r(\cdot)$  pour les variables aléatoires i.i.d., nous pouvons citer Gonzalez-Manteiza et al. (2002).

Pour le modèle de censure, dans le cas i.i.d, Guessoum et Ould Saïd (2008) ont établi la normalité asymptotique de l'estimateur à noyau de la fonction de régression. Le but de ce chapitre est d'étendre leurs résultats au cas des données dépendantes. Deux types de dépendance sont largement utilisés dans la littérature : le mélange et l'association.

Pour les données censurées, sous la condition  $\alpha$ -mélange, nous citons Guessoum et Ould Saïd (2012) qui ont établi la normalité asymptotique du même estimateur.

Dans ce travail, nous nous concentrons sur le concept d'association qui a été introduit et défini par Esary et al. (1967) avec l'objectif de trouver des applications en fiabilité et en statistique.

Une famille finie de variables aléatoires réelles  $Y = \{Y_1, Y_2, \dots, Y_n\}$ , est dite associée si pour toute paire de fonctions  $\phi_1$  et  $\phi_2$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , non décroissantes (composante par composante), on a :

$$\text{cov}(\phi_1(Y_i, 1 \leq i \leq n), \phi_2(Y_j, 1 \leq j \leq n)) \geq 0,$$

lorsque cette covariance existe.

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

Pour plus de détails sur le concept d'association, nous renvoyons le lecteur à Bulinski et Shashkin (2007). Dans ce livre, on peut trouver de nombreux de nouveaux résultats et exemples, aussi bien qu'une revue sur l'association et les concepts relatifs sur plus de 30 années.

Dans le cas de données complètes associées, l'estimateur à noyau de densité a fait l'objet de plusieurs travaux, comme exemple ceux de Bagai et Prakasa Rao (1995), Roussas (1991, 2000). Douge (2007) ont établi une nouvelle inégalité exponentielle, qui a un taux de convergence uniforme presque sûr sur des ensembles compacts pour l'estimateur de la densité.

Dans le cadre de données incomplètes (censurées), Cai et Roussas (1998) ont établi la convergence uniforme et la normalité asymptotique de l'estimateur de Kaplan-Meier. Ferrani et al. (2014) ont traité la consistance de l'estimateur à noyau de la densité et du mode.

Notre objectif principal est d'établir la normalité asymptotique de l'estimateur à noyau de la fonction de régression, sous un modèle de données censurées à droite, vérifiant une structure de dépendance

de type associé. Ce chapitre est organisé comme suit. La section 2, est consacrée aux notations et définitions des estimateurs utilisés dans notre modèle. Dans la section 3, nous donnons les hypothèses et les résultats principaux. La section 4, est consacré à l'application des résultats asymptotiques pour construire un intervalle de confiance. Quelques simulations sont données dans la section 5. Enfin, les preuves sont reportées à la section 6.

## 3.2 Définitions et Notations

Tout au long de ce chapitre, nous définissons  $\tau_F$ ,  $\tau_G$  et  $\tau_H$  comme les extremums supérieurs de fonction de répartition respectivement  $F_Y(y) = \mathbb{P}(Y \leq y)$ ,  $G(y) = \mathbb{P}(C \leq y)$  et  $H(y) = \mathbb{P}(Z \leq y)$ , avec  $\tau_F = \sup\{y, \bar{F}_Y(y) > 0\}$ ,  $\tau_G = \sup\{y, \bar{G}(y) > 0\}$  et  $\tau_H = \tau_F \wedge \tau_G$ .

En absence de censure, il est connu que l'estimateur à noyau de la fonction de régression est donnée par

$$r_n(x) = \frac{\sum_{i=1}^n Y_i K_d\left(\frac{x-X_i}{h}\right)}{\sum_{i=1}^n K_d\left(\frac{x-X_i}{h}\right)},$$

où  $K$  est une densité de probabilité (appelée noyau),  $h_n =: h > 0$  est une suite de nombres réels positifs tendant vers zéro lorsque  $n$  tend vers l'infini (appelés fenêtres).

Plus généralement, des estimateurs à poids de  $r(x)$  peuvent s'écrire

$$r_n(x) = \sum_{i=1}^n W_{n,i}(x) Y_i,$$

où  $W_{n,i}(\cdot)$  sont des poids dépendants de  $x$  et de  $(X_1, \dots, X_n)$ .

Dans le modèle censuré, suivant Carbonez et al. (1995), et Köhler et al.(2002), nous définissons  $\tilde{r}_n(x)$  comme un estimateur de  $r(x)$  par

$$\begin{aligned} \tilde{r}_n(x) &= \sum_{i=1}^n W_{ni}(x) \frac{\delta_i Z_i}{\bar{G}(Z_i)}, \\ &=: \frac{\tilde{r}_{1,n}(x)}{\ell_n(x)}, \end{aligned} \quad (3.3)$$

où

$$\tilde{r}_{1,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}(Z_i))^{-1} \delta_i Z_i K\left(\frac{x-X_i}{h}\right), \quad (3.4)$$

et  $\ell_n$  est l'estimateur à noyau de la densité marginale de  $\ell$  est donné par

$$\ell_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right).$$

En pratique, puisque  $G$  est généralement inconnu d'où il est impossible d'utiliser l'estimateur (3.3). Alors, pour construire un estimateur pratiquement faisable, on utilise l'estimateur de Kaplan-Meier

(1958)  $G_n$  donné par :

$$\bar{G}_n(t) = \begin{cases} \prod_{i:Z_{(i)} \leq t} \left(\frac{n-i}{n-i+1}\right)^{1-\delta_{(i)}} & \text{si } t < Z_{(n)} \\ 0 & \text{sinon,} \end{cases} \quad (3.5)$$

où  $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$  représentent les statistiques d'ordre associées à  $Z_i$  et  $\delta_{(i)}$  le concomitant de  $Z_{(i)}$ . Ainsi, l'estimateur de  $r(\cdot)$  sera défini par :

$$\begin{aligned} r_n(x) &= \frac{\sum_{i=1}^n (\bar{G}_n(Z_i))^{-1} \delta_i Z_i K\left(\frac{x-X_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)} \\ &=: \frac{r_{1,n}(x)}{\ell_n(x)}, \end{aligned} \quad (3.6)$$

où

$$r_{1,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}_n(Z_i))^{-1} \delta_i Z_i K\left(\frac{x-X_i}{h}\right). \quad (3.7)$$

Tout au long de ce chapitre, nous utilisons  $c$  pour désigner une constante positive qui peut prendre différentes valeurs.

### 3.3 Hypothèses et résultats

On considère des compacts  $[0, \tau_F]$  et  $\mathcal{D}$  tel que  $\mathcal{D} \subset \{x \in \mathbb{R}^d / \ell(x) > \Gamma > 0; \Gamma := \inf \ell(x)\} \subset \mathbb{R}^d$ , on définit  $\Delta_{i,j} := \sum_{k=1}^d \sum_{l=1}^d \text{cov}(X_i^k, X_j^l) + 2 \sum_{k=1}^d \text{cov}(X_i^k, Y_j) + \text{cov}(Y_i, Y_j)$ , où  $X_i^k$  est la  $k^{\text{ème}}$  composante du vecteur  $X_i$ .

Les hypothèses pour établir les résultats sont les suivantes :

**A.**  $(Y, C) \in \mathbb{R}^+ \times \mathbb{R}^+$ ,  $\tau_F < \infty$ ,  $G$  est continue, et  $G(\tau_F) < 1$ ;

**H.** La fenêtre  $h$  satisfait :

**i).**  $\lim_{n \rightarrow +\infty} nh^{d+4} = 0$ ;

**ii).**  $h \simeq o\left(\frac{1}{n}\right)^\alpha$  pour tout  $\frac{1}{d+4} < \alpha < \frac{1}{d+3}$ .

**K1.** Le noyau  $K_d$  est bornée ( $\|K_d\|_\infty := \sup |K_d| < +\infty$ ), à support compact et à premières dérivées partielles bornées, il satisfait :

**a).**  $\int_{\mathbb{R}^d} t_i K_d(t) dt = 0$ , pour tous  $i = 1, 2, \dots, d$ ;  $t = (t_1, t_2, \dots, t_d)^\top$ ;

**b).**  $\int_{\mathbb{R}^d} |t_1 + t_2 + \dots + t_d| K_d^2(t) dt < \infty$ ;

**c).** pour tous  $\{i_1, i_2, \dots, i_d\}$  tel que  $i_1 + i_2 + \dots + i_d = 2$  on a  $\int_{\mathbb{R}^d} |t_1^{i_1} t_2^{i_2} \dots t_d^{i_d}| K_d(t) dt < \infty$ ;

**M.** Le terme de covariance vérifie :  $\rho(s) := \sup_{j:|i-j| \geq s} \Delta_{i,j} = O(e^{-\gamma s})$ , pour tous  $i \geq 1$ ,  $\gamma > 0$  et  $s > 0$ ;

**D1.** La fonction  $r_1(\cdot)$  définie dans (3.1) est bornée, deux fois différentiable et  $\sup_{x \in \mathcal{D}} \left| \frac{\partial^k r_1}{\partial x_i \partial x_j^{k-1}}(x) \right| < \infty$   
pour  $i, j = 1, \dots, d$  et  $k = 1, 2$ ;

- D2.** La densité marginale  $\ell(\cdot)$  est bornée, deux fois différentiable et  $\sup_{x \in \mathcal{D}} \left| \frac{\partial^k \ell}{\partial x_i \partial x_j^{k-1}}(x) \right| < \infty$   
 $i, j = 1, \dots, d$  et  $k = 1, 2$ ;
- D3.** La fonction  $r_2(x) := \int_{\mathbb{R}} (\bar{G}(y))^{-1} y^2 f_{X,Y}(x, y) dy$  est continûment différentiable et  
 $\sup_{x \in \mathcal{D}} \left| \frac{\partial r_2}{\partial x_i}(x) \right| < \infty$  pour  $i = 1, \dots, d$ ;
- D4.** La densité conjointe  $\ell_{i,j}(\cdot, \cdot)$  de  $(X_i, X_j)$  satisfait  $\sup_{u, v \in \mathcal{D}} |\ell_{i,j}(u, v) - \ell(u)\ell(v)| < \infty$ .
- B**; Il existe des suites de nombres entiers  $(p_n)_n$ ,  $(q_n)_n$  et  $(k_n)_n$  définie par  $k_n := \lfloor \frac{n}{p+q} \rfloor$ , (où  $\lfloor x \rfloor$  désigne la partie entière de  $x$ ), tendant vers l'infini avec  $n$ , de sorte que  $k(p+q) \leq n$  et  $\frac{k(p+q)}{n} \rightarrow 1$  lorsque  $n \rightarrow +\infty$ , telles que :
- i).  $\frac{pk}{n} \xrightarrow{n \rightarrow +\infty} 1$  ;
- ii).  $ph^d \xrightarrow{n \rightarrow +\infty} 0$  et  $\frac{p^2}{nh^d} \xrightarrow{n \rightarrow +\infty} 0$  ;
- iii).  $\frac{1}{h^{d+2}} \sum_{j=q}^{\infty} |\Delta_{1,j+1}| \rightarrow 0$ ,  
 où  $p, q$  et  $k$  désignent respectivement  $p_n, q_n$  et  $k_n$ .

### 3.3.1 Discussion sur les hypothèses

1. Les hypothèses **(K1)** sont courantes dans l'estimation à noyau.
2. Les hypothèses **(D1)** et **(D3)** sont des conditions de régularité qui sont nécessaires pour établir la convergence de certaines grandeurs qui interviennent dans la preuve du résultat principal.
3. L'hypothèse **(D2)** intervient dans la convergence de l'estimateur à noyau de la densité.
4. L'hypothèse **(D4)** est technique, elle est employée pour simplifier les calculs de covariances dans les preuves.
5. l'hypothèse **(B)** permet d'employer la technique de Bernstein des grands blocs et petits blocs pour prouver la normalité asymptotique pour les associées.
6. Comme  $\frac{qk}{n} = \frac{k(q+p)}{n} - \frac{pk}{n}$ , l'hypothèse **(B)**(i) implique que  $\frac{qk}{n} \rightarrow 0$ . Aussi,  $\frac{q}{p} = \frac{(\frac{qk}{n})}{(\frac{pk}{n})} \rightarrow 0$ , pour que  $q < p$ .
7. La première hypothèse **(B)**(ii) implique que  $qh^d \rightarrow 0$ , ce qui implique également  $h^d \rightarrow 0$ . La seconde hypothèse de **(B)**(ii) implique que  $nh^d \rightarrow +\infty$ .

### 3.3.2 Principaux résultats

**Théorème 3.1.** *Sous les hypothèses **A, H, K, M, D3** et **D4**, nous avons*

$$r_n(x) \xrightarrow{\mathbb{P}} r(x).$$

**Théorème 3.2.** *Sous les hypothèses **A**, **H**, **K**, **M**, **D1-D4** et **B**, nous avons*

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

où

$$\sigma^2(x) = \kappa \frac{r_2(x)\ell(x) - r_1^2(x)}{\ell^3(x)}, \quad (3.8)$$

$\kappa = \int_{\mathbb{R}^d} K_d^2(t) dt$  et  $\xrightarrow{\mathcal{L}}$  désigne la convergence en loi.

### 3.4 Application aux intervalles de confiance

En utilisant une méthode plug-in en remplaçant  $\ell$ ,  $\bar{G}$  et  $r_1$  par leurs estimations  $\ell_n$ ,  $\bar{G}_n$  et  $r_{1,n}$  et donné par (3.3), (3.6) et (3.5) respectivement, et  $r_2$  par

$$r_{2,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n (\bar{G}_n(Z_i))^{-2} \delta_i Z_i^2 K\left(\frac{x - X_i}{h}\right), \quad (3.9)$$

Nous permet d'obtenir une estimation  $\sigma_n$  de  $\sigma$  à partir (3.8). Ensuite, à partir du Théorème 3.2 nous obtenons le résultat suivant.

**Corollaire 3.1.** *Sous les hypothèses du Théorème 3.2, nous avons*

$$\left(\frac{nh^d}{\sigma_n^2(x)}\right)^{1/2} [r_n(x) - r(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1). \quad (3.10)$$

A partir du corollaire 3.1, on obtient pour chaque  $\xi \in (0, 1)$  fixé, l'intervalle de confiance à  $(1-\xi)\%$  suivant

$$\left[ r_n(x) - \frac{U_{1-\xi/2} \sigma_n(x)}{\sqrt{nh^d}}, r_n(x) + \frac{U_{1-\xi/2} \sigma_n(x)}{\sqrt{nh^d}} \right]$$

où  $U_{1-\xi/2}$  représente le quantile d'ordre  $1 - \xi/2$  de la distribution normale standard.

### 3.5 Simulations

Dans cette section, nous effectuons les simulations pour étudier la performance de l'échantillon fini de l'estimateur  $r_n(x)$  de  $r(x)$  dans le cas  $d = 1$ . En particulier, la simulation consiste à examiner comment fonctionne l'approximation asymptotique. Pour cela, nous générons les données observées comme suit

*Etape 1.*

Nous générons la suite associée  $\{X_t, t = 1, \dots, n\}$  par  $X_t = (W_{t-1} + W_{t-2})/2$  ( voir Chaubey et al.(2011)), où  $\{W_t, t = -1, 0, \dots, n-1\}$  sont  $n+1$  variables aléatoires i.i.d. suivant une loi exponentielle de paramètre 1.

*Etape 2.*

Soit  $Y_i$  du modèle

$$Y_i = r(X_i) + \varepsilon_i,$$

où  $r(\cdot)$  est la fonction de régression et  $\{\varepsilon_t, t = 1, \dots, n\}$  sont  $n$  variables aléatoires i.i.d. de loi normale  $\mathcal{N}(0, 0.8)$ .

*Etape 3.*

Nous simulons une suite  $\{C_t, t = 1, \dots, n\}$  de v.a. i.i.d distribuée comme une  $\exp(\lambda)$ . Ensuite, nous calculons notre estimateur avec les données observées  $(X_i, Z_i, \delta_i)$ , où  $Z_i = Y_i \wedge C_i$  et  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ . Nous utilisons le noyau Gaussien standard. La fenêtre a été prise pour être  $h = O\left((n^{-1} \log n)^{\frac{1}{3}}\right)$  qui satisfait l'hypothèse **(H)**.

### 3.5.1 Normalité asymptotique

Nous considérons la propriété de la normalité asymptotique. Sur la base d'un  $n$ -échantillon, nous comparons la forme de la densité estimée (des écarts normalisés) à celle de la densité normale dans le cas du modèle linéaire.

$$Y_i = 2X_i - 3 + \varepsilon_i, i = 1, \dots, n. \quad (3.11)$$

Les données proviennent des mêmes distributions que précédemment. Pour une taille d'échantillon  $n$  donnée, nous estimons la fonction de régression  $r(x)$  par  $r_n(x)$ , et nous calculons l'écart normalisé entre cette estimation et la fonction de régression théorique donnée dans l'équation (3.11)

pour  $x = 1$  (*i.e.*,  $\bar{r}_n = \sigma_n^{-1}(1) \sqrt{nh} \{r_n(1) - r(1)\} = \sigma_n^{-1}(1) \sqrt{nh} \{r_n(1) + 1\}$ ).

Nous traçons en faisant appel à ce schéma,  $B = 1000$   $n$ -échantillons indépendants, nous faisons le choix de la fenêtre classique (voir, e.g. Sliverman, 1986 p.40)  $h'_B = C.B^{-1/5}$ , où la constante  $C$  est choisie de manière appropriée.

Nous comparons la forme de la densité estimée avec l'écart normalisé (courbes en pointillés) à celle (courbes continues) de la densité normale standard. Nous traçons également les histogrammes correspondants par rapport à la densité normale standard et les tracés Q-Q- plots pour différentes valeurs de  $n$  et un pourcentage de censure ( $PC$ ). Nous ne donnons que la courbe pour  $n = 100$ ,  $n = 300$ , et  $n = 500$ , avec  $PC = 10\%$  et  $PC = 30\%$ .

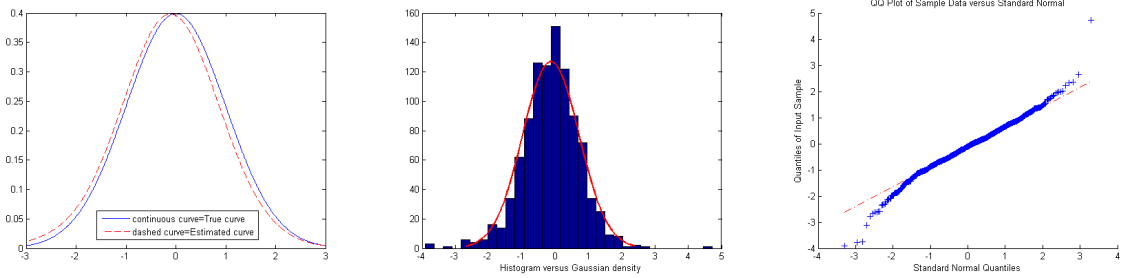


FIGURE 3.1 – Les graphes de la normalité pour  $n = 100$  et  $B = 1000$  avec  $PC = 10\%$ .

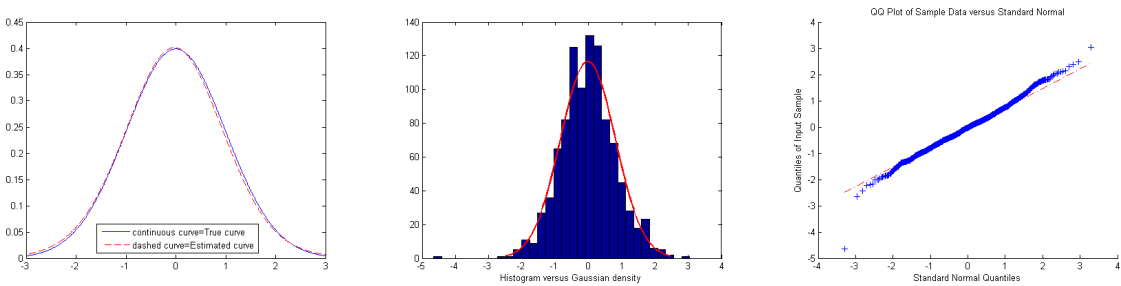


FIGURE 3.2 – Les graphes de la normalité pour  $n = 300$  et  $B = 1000$  avec  $PC = 10\%$ .

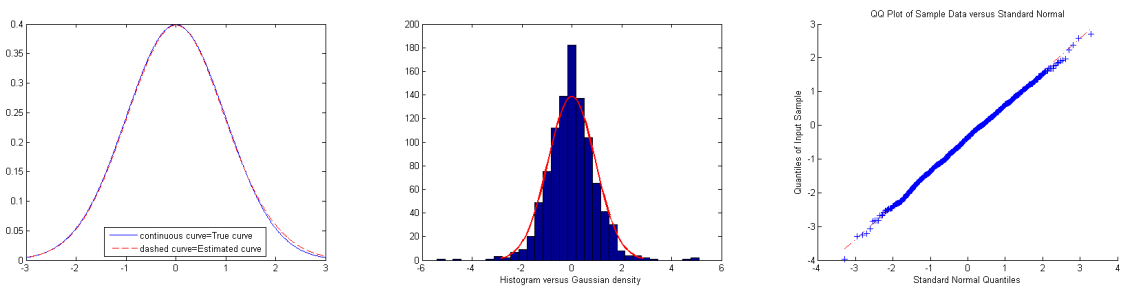


FIGURE 3.3 – Les graphes de la normalité pour  $n = 500$  et  $B = 1000$  avec  $PC = 10\%$ .

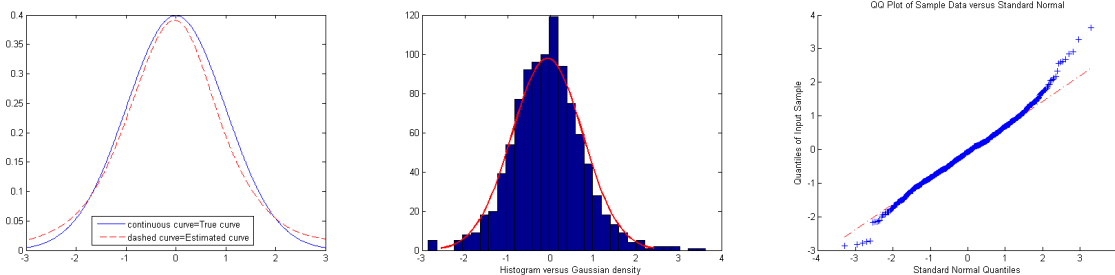


FIGURE 3.4 – Les graphes de la normalité pour  $n = 100$  et  $B = 1000$  avec  $PC = 30\%$ .

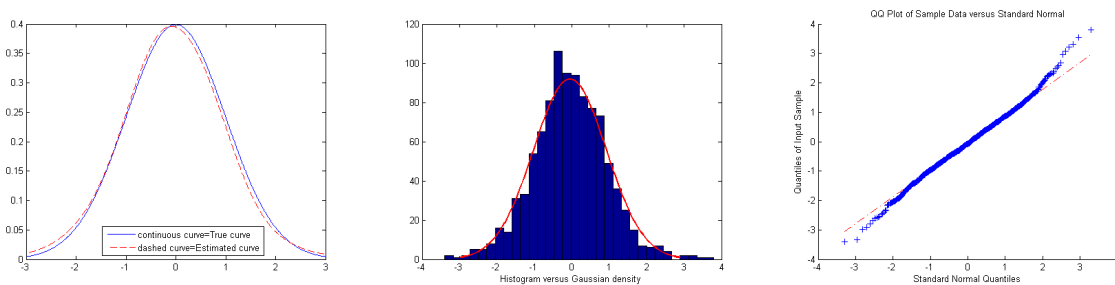


FIGURE 3.5 – Les graphes de la normalité pour  $n = 300$  et  $B = 1000$  avec  $PC = 30\%$ .

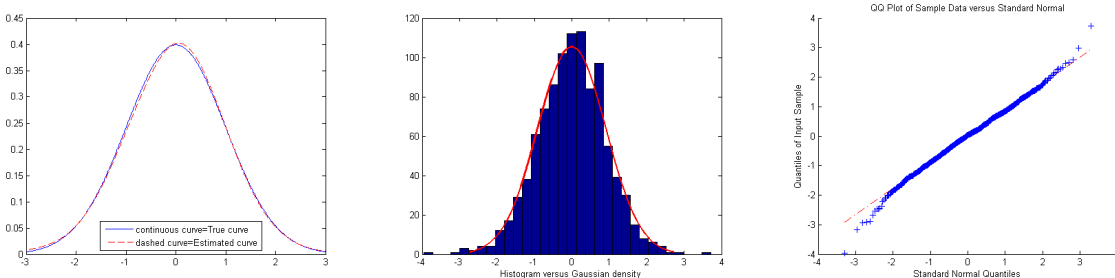


FIGURE 3.6 – Les graphes de la normalité pour  $n = 500$  et  $B = 1000$  avec  $PC = 30\%$ .

D’après les figures 1 à 6, la qualité de l’ajustement augmente proportionnellement avec la taille de l’échantillon et inversement au pourcentage de censure.

### 3.5.2 Intervalles de confiance

Nous avons construit les intervalles de confiance à  $(1 - \xi)\%$  obtenus à partir du corollaire 3.1, c'est-à-dire

$$r_n(x) \pm \frac{1.96 \times \sigma_n(x)}{\sqrt{nh}} \tag{3.12}$$

Pour plusieurs valeurs de  $n$ , nous dessinons dans le même graphe deux courbes correspondant aux bornes inférieur et supérieur de l'intervalle de confiance à  $(1 - \xi/2)\%$ , pour  $x \in [0, 2]$  et  $\xi = 5\%$ . Nous donnons le graphe pour  $n = 50$ ,  $n = 100$ ,  $n = 300$ , et  $n = 500$ , et nous prenons différentes valeurs du PC :  $PC = 10\%$  et  $PC = 30\%$ .

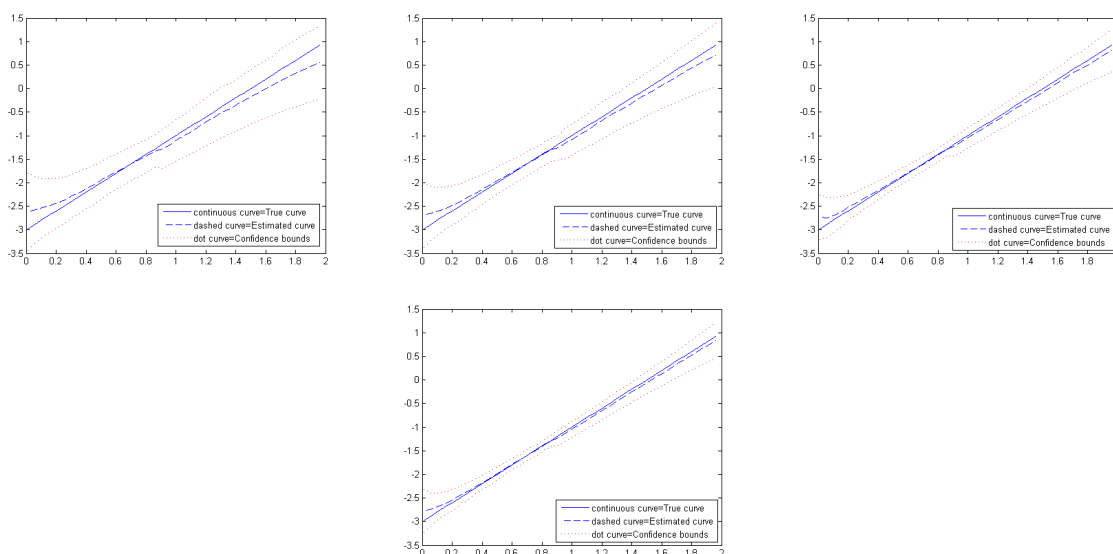


FIGURE 3.7 – Intervalles de confiance avec  $PC = 10\%$  et  $n = 50, 100, 300, 500$ .

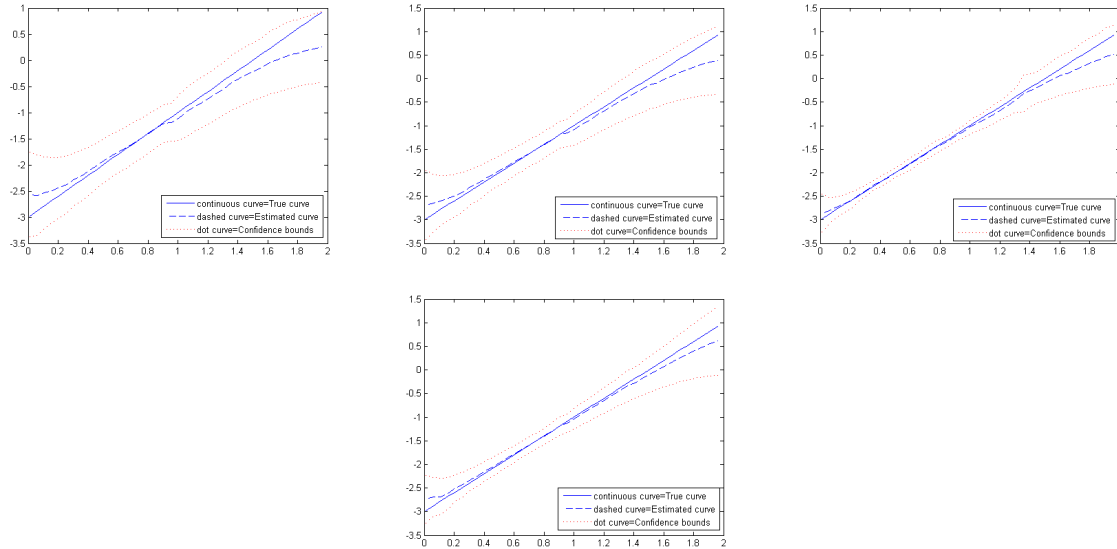


FIGURE 3.8 – Intervalles de confiance. avec  $PC = 30\%$  et  $n = 50, 100, 300, 500$ .

Il est clair qu'à partir des figures 7 et 8, l'intervalle de confiance devient plus précis quand  $n$  augmente et le PC diminue.

Ensuite, nous calculons la probabilité de couverture d'un niveau de confiance  $1 - \xi = 0.95$  et  $0.99$ . On prend  $x \in [0, 2]$  et nous choisissons la taille de l'échantillon  $n$  pour être 50, 100, 300, et 500, pour 1000 répétitions. Le Tableau 1 contient les probabilités de couverture comme le pourcentage de couverture de la vraie courbe en tout point de données par les bandes de confiance dans l'équation 3.12.

D'après le tableau 1, il est évident qu'une taille d'échantillon plus large et un faible PC, garantit

	conf.level	n=50	n=100	n=300	n=500
$PC = 0\%$	0.95	0.8921	0.9167	0.9341	0.9550
	0.99	0.9533	0.9724	0.9856	0.9901
$PC = 15\%$	0.95	0.8712	0.8866	0.9161	0.9389
	0.99	0.9607	0.9697	0.9708	0.9772
$PC = 30\%$	0.95	0.7937	0.8356	0.8675	0.9062
	0.99	0.8559	0.8929	0.9282	0.9433

TABLE 3.1 – Les probabilités de couverture de l'intervalle de confiance pour 1000 répétitions.

une meilleure couverture de la bande de confiance.

### 3.6 Preuves

Nous avons besoin de résultats et de notations auxiliaires pour prouver nos résultats. En utilisant (3.1) et (3.6), nous écrivons

$$\begin{aligned}
r_n(x) - r(x) &= \frac{r_{1,n}(x)}{\ell_n(x)} - \frac{r_1(x)}{\ell(x)} \\
&= \frac{\ell(x)r_{1,n}(x) - \ell_n(x)r_1(x)}{\ell_n(x)\ell(x)} \\
&= \frac{1}{\ell_n(x)\ell(x)} [\ell(x)r_{1,n}(x) - \ell_n(x)r_1(x)] \\
&= \frac{1}{\ell_n(x)\ell(x)} [\ell(x)r_{1,n}(x) - \ell_n(x)r_1(x) + \ell(x)r_1(x) - \ell(x)r_1(x)] \\
&= \frac{1}{\ell_n(x)\ell(x)} [\ell(x)(r_{1,n}(x) - r_1(x)) - r_1(x)(\ell(x) - \ell_n(x))]. \tag{3.13}
\end{aligned}$$

Notons que

$$\begin{aligned}
r_{1,n}(x) - r_1(x) &= [r_{1,n}(x) + \tilde{r}_{1,n}(x) - \tilde{r}_{1,n}(x) + \mathbb{E}(\tilde{r}_{1,n}(x)) - \mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)] \\
&= [(r_{1,n}(x) - \tilde{r}_{1,n}(x)) + (\tilde{r}_{1,n}(x) - \mathbb{E}(\tilde{r}_{1,n}(x))) + (\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x))] \\
&=: \alpha_{1,n} + \alpha_{2,n} + \alpha_{3,n}. \tag{3.14}
\end{aligned}$$

et

$$\begin{aligned}
\ell_n(x) - \ell(x) &= \ell_n(x) + \mathbb{E}(\ell_n(x)) - \mathbb{E}(\ell_n(x)) - \ell(x) \\
&= (\ell_n(x) - \mathbb{E}(\ell_n(x))) + (\mathbb{E}(\ell_n(x)) - \ell(x)) \\
&=: \beta_{1,n}(x) + \beta_{2,n}(x). \tag{3.15}
\end{aligned}$$

#### Preuve du Théorème 3.1

La convergence en probabilité de l'estimateur de la densité  $\ell_n(x)$  a été traitée par Liang et Peng (2010). L'étape suivante consiste à prouver la convergence en probabilité de  $r_{1,n}(x)$  vers  $r_1(x)$ . Donc, il faut montrer que chaque terme de (3.14) tend vers zéro en probabilité. Pour cela, nous utilisons l'inégalité de Markov sur les termes  $\alpha_{1,n}$  et  $\alpha_{3,n}$  et l'inégalité de Tchebychev pour  $\alpha_{2,n}$ . Le Théorème de Slutsky et (3.13) donnent le résultat.

- Pour  $\alpha_{1,n}$ , nous montrons que

$$\mathbb{E} [|r_{1,n}(x) - \tilde{r}_{1,n}(x)|] \rightarrow 0,$$

à partir de (3.4) et (3.7), nous avons

$$\begin{aligned}
 \mathbb{E} [|r_{1,n}(x) - \tilde{r}_{1,n}(x)|] &= \mathbb{E} \left[ \left| \frac{1}{nh^d} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}_n(Z_i)} K_d \left( \frac{x - X_i}{h} \right) - \frac{1}{nh^d} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) \right| \right] \\
 &= \mathbb{E} \left[ \left| \frac{1}{nh^d} \sum_{i=1}^n \frac{Y_i \mathbb{1}_{\{Y_i \leq C_i\}}}{\bar{G}_n(Y_i)} K_d \left( \frac{x - X_i}{h} \right) - \frac{1}{nh^d} \sum_{i=1}^n \frac{Y_i \mathbb{1}_{\{Y_i \leq C_i\}}}{\bar{G}(Y_i)} K_d \left( \frac{x - X_i}{h} \right) \right| \right] \\
 &\leq \mathbb{E} \left[ \frac{1}{nh^d} \sum_{i=1}^n \left| \frac{Y_i}{\bar{G}_n(Y_i)} K_d \left( \frac{x - X_i}{h} \right) - \frac{Y_i}{\bar{G}(Y_i)} K_d \left( \frac{x - X_i}{h} \right) \right| \right] \\
 &\leq \mathbb{E} \left[ \frac{1}{nh^d} \sum_{i=1}^n \left| Y_i K_d \left( \frac{x - X_i}{h} \right) \left[ \frac{1}{\bar{G}_n(Y_i)} - \frac{1}{\bar{G}(Y_i)} \right] \right| \right] \\
 &\leq \frac{\tau_F}{\bar{G}_n(\tau_F) \bar{G}(\tau_F)} \sup_{t \leq \tau_F} |\bar{G}_n(t) - \bar{G}(t)| \times \mathbb{E} \left[ \frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \right],
 \end{aligned}$$

sous les hypothèses **K1** et **D2**, on a

$$\begin{aligned}
 \mathbb{E} \left( \frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \right) &= \frac{1}{h^d} \mathbb{E} \left( K_d \left( \frac{x - X_i}{h} \right) \right) \\
 &= \frac{1}{\varepsilon h^d} \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) \ell(u) du \\
 &= \int_{\mathbb{R}^d} K_d(t) \ell(x - th) dt = O(1).
 \end{aligned}$$

De ce fait, en prenant le Théorème 1.4 dans Cai et Roussas (page 323 (1998)) et l'hypothèse **M**, on obtient

$$\sup_{t \leq \tau_F} |\bar{G}_n(t) - \bar{G}(t)| = o(n^{-\theta}) \text{ p.s. lorsque } n \rightarrow +\infty,$$

où  $0 < \theta < \frac{\gamma}{2\gamma+9+\frac{3}{2}\rho}$ , pour tout  $\rho < \frac{1}{3}\gamma d - 6$ , et  $\gamma > \frac{18}{d}$ .

Comme  $\bar{G}_n(\tau_F) \rightarrow \bar{G}(\tau_F)$  et  $\mathbb{E} \left( \frac{1}{nh^d} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h} \right) \right) = O(1)$ , on obtient  $\mathbb{E} [|r_{1,n}(x) - \tilde{r}_{1,n}(x)|] = o(1)$ .

L'inégalité de Markov conclut la preuve.

- Pour  $\alpha_{3,n}$ , le Lemme 2 dans Menni et Tatachak (2016), donne  $|\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| = o(1)$ . Une fois de plus, l'inégalité de Markov conclut la preuve.

• Pour  $\alpha_{2,n}(x)$ , on montre que  $\text{var}(\tilde{r}_{1,n}(x)) \rightarrow 0$ . D'après (3.4), nous avons

$$\begin{aligned}
\text{var}(\alpha_{2,n}(x)) &= \text{var}(\tilde{r}_{1,n}(x)) \\
&= \text{var} \left( \frac{1}{nh^d} \sum_{i=1}^n \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) \right) \\
&= \frac{n}{n^2 h^{2d}} \text{var} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right) \\
&\quad + \frac{1}{n^2 h^{2d}} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) \quad (3.16) \\
&=: I_1 + I_2.
\end{aligned}$$

D'une part, nous avons

$$\begin{aligned}
I_1 &= \frac{n}{n^2 h^{2d}} \left\{ \mathbb{E} \left[ \frac{\delta_1^2 Z_1^2}{\bar{G}^2(Z_1)} K_d^2 \left( \frac{x - X_1}{h} \right) \right] - \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right] \right\} \\
&=: \frac{1}{nh^{2d}} \{I_{11} - I_{12}\},
\end{aligned}$$

avec

$$I_{11} = \mathbb{E} \left[ K_d^2 \left( \frac{x - X_1}{h} \right) \mathbb{E} \left( \frac{\delta_1^2 Z_1^2}{\bar{G}^2(Z_1)} | X_1 \right) \right] = \int_{\mathbb{R}^d} K_d^2 \left( \frac{x - u}{h} \right) \mathbb{E} \left( \frac{\mathbb{1}_{\{Y_1 \leq C_1\}} Y_1^2}{\bar{G}^2(Y_1)} | X_1 = u \right) \ell(u) du.$$

De plus, conditionnellement à  $Y_1$  et en utilisant un changement de variable, on obtient

$$\begin{aligned}
I_{11} &= \int_{\mathbb{R}^d} K_d^2 \left( \frac{x - u}{h} \right) \mathbb{E} \left( \frac{Y_1^2}{\bar{G}(Y_1)} | X_1 = u \right) \ell(u) du = \int_{\mathbb{R}^d} K_d^2 \left( \frac{x - u}{h} \right) \left( \int_{\mathbb{R}} \frac{y^2}{\bar{G}(y)} f_{X,Y}(u, y) dy \right) du \\
&= \int_{\mathbb{R}^d} K_d^2 \left( \frac{x - u}{h} \right) r_2(u) du = h^d \int_{\mathbb{R}^d} K_d^2(t) r_2(x - th) dt.
\end{aligned}$$

Ainsi, en utilisant **D1** et un développement de Taylor au voisinage de  $x = (x_1, x_2, \dots, x_d)^\top$ , nous obtenons

$$I_{11} = h^d \int r_2(x) K_d^2(t) dt - h^{d+1} \int K_d^2(t) \left[ t_1 \frac{\partial r_2}{\partial x_1}(x^*) + \dots + t_d \frac{\partial r_2}{\partial x_d}(x^*) \right] = O(h^d),$$

où  $x^*$  est entre  $x - th$  et  $x$ .

Maintenant, on évalue  $I_{12}$

$$\begin{aligned}
I_{12} &= \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right] = \mathbb{E}^2 \left[ K_d \left( \frac{x - X_1}{h} \right) \mathbb{E} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} | X_1 \right) \right] \\
&= \left[ \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) r(u) \ell(u) du \right]^2 \\
&= \left[ \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) r_1(u) du \right]^2.
\end{aligned}$$

Ensuite, en utilisant un changement de variables, un développement de Taylor et les hypothèses **K(c)** et **D1**

$$I_{12} = \left[ h^d \int_{\mathbb{R}^d} K_d(t) r_1(x - th) dt \right]^2 = h^{2d} \left[ \int_{\mathbb{R}^d} K_d(t) r_1(x - th) dt \right]^2.$$

Alors

$$\begin{aligned}
I_{12} &= h^{2d} \left\{ \int_{\mathbb{R}^d} K_d(t) \left[ \frac{h^2}{2} \left( t_1^2 \frac{\partial^2 r_1}{\partial x_1^2}(x^*) + \cdots + t_d^2 \frac{\partial^2 r_1}{\partial x_d^2}(x^*) \right) + 2 \sum_{i=1}^n \sum_{j=1: j \neq i}^n t_i t_j \frac{\partial^2 r_1}{\partial x_i \partial x_j}(x^*) \right] dt \right. \\
&\quad \left. + \int_{\mathbb{R}^d} K_d(t) |r_1(x)| dt \right\}^2 \\
&= O(h^{2d+4}).
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
I_1 &= \frac{1}{nh^{2d}} \{c \cdot h^d - c \cdot h^{2d+4}\} \\
&= O\left(\frac{1}{nh^d}\right) = o(1).
\end{aligned} \tag{3.17}$$

D'autre part, par les hypothèses **H**, **K** et **D4** on obtient

$$\text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) = O(h^{2d}) \tag{3.18}$$

Pour évaluer  $I_2$ , on utilise la technique de Masry (1986), on définit

$$\mathcal{B}_1 = \{(i, j); 1 \leq |i - j| \leq \vartheta_n\} \text{ et } \mathcal{B}_2 = \{(i, j); \vartheta_n + 1 \leq |i - j| \leq n - 1\}, \tag{3.19}$$

où  $\vartheta_n = o(n)$ . On décompose  $I_2$  comme suit :

$$\begin{aligned}
I_2 &= \frac{1}{n^2 h^{2d}} \sum_{i=1}^n \sum_{j \in \mathcal{B}_1} \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) \\
&\quad + \frac{1}{n^2 h^{2d}} \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) \\
&=: I_{21} + I_{22}.
\end{aligned}$$

Ainsi, de (3.18) on obtient

$$I_{21} = O \left( \vartheta_n \cdot n \cdot \frac{h^{2d}}{n^2 h^{2d}} \right) = O \left( \frac{\vartheta_n}{n} \right). \tag{3.20}$$

Pour  $\vartheta_n \leq |i - j| \leq n$ , en utilisant la Proposition 2, le résultat (a) dans Menni et Tatachak (2016) et l'hypothèse **M**, nous obtenons

$$\begin{aligned}
I_{22} &\leq \frac{1}{n^2 h^{2d}} \sum_{i=1}^n \sum_{\vartheta_n \leq |i-j| \leq n} c h^d \phi(|i - j|) \\
&\leq \frac{c}{n h^d} \sum_{\vartheta_n \leq |i-j| \leq n} \exp \left( -\frac{\gamma d}{2d+2} |i - j| \right) \\
&\leq \frac{c}{n h^d} \int_{\vartheta_n}^n \exp \left( -\frac{\gamma d}{2d+2} u \right) du \\
&= O \left( \frac{1}{n h^d} \exp \left( -\frac{\gamma d}{2d+2} \vartheta_n \right) \right).
\end{aligned} \tag{3.21}$$

En choisissant  $\vartheta_n = O(h^{\nu-d})$  avec  $0 < \nu < d$ , (3.20) et (3.21) deviennent

$$I_{21} = O\left(\frac{\vartheta_n}{n}\right) = O\left(\frac{h^{\nu-d}}{n}\right) = O\left(\frac{h^\nu}{nh^d}\right) = o(1) \quad (3.22)$$

$$I_{22} = O\left(\frac{1}{nh^d} \exp\left(-\frac{\gamma dh^{\nu-d}}{2d+2}\right)\right) = o\left(\frac{1}{nh^d}\right) = o(1) \quad (3.23)$$

Finalement, d'après (3.16), (3.17), (3.22), et (3.23) nous obtenons  $\text{var}(\tilde{r}_{1,n}(x)) = o(1)$ , en appliquant l'inégalité de Tchebychev, on arrive au résultat.  $\square$

### Preuve du Théorème 3.2

La démonstration est basée sur quelques Lemmes et propositions préliminaires. Le Lemme 3.1 traite les termes négligeables  $\alpha_{1,n}(x)$ ,  $\alpha_{3,n}(x)$  et  $\beta_{2,n}(x)$  apparaissant dans (3.14) et (3.15). Les autres Lemmes considèrent les termes  $\alpha_{2,n}(x)$  et  $\beta_{1,n}(x)$  en calculant leur variance et covariance.

**Lemme 3.1.** *Sous les hypothèses **A**, **D1**, **D2**, **K** et **H**, on a*

$$(a) \sqrt{nh^d} |\alpha_{1,n}(x)| = o(1),$$

$$(b) \sqrt{nh^d} |\alpha_{3,n}(x)| = o(1),$$

$$(c) \sqrt{nh^d} |\beta_{2,n}(x)| = o(1).$$

### Preuve du Lemme 3.1.

Suivant les preuves des Lemmes 1, 2 et 3 de Menni et Tatachak (2016), nous obtenons

a) Sous les hypothèses **A**, **K**, **M**, et **D2**, on a

$$|\alpha_{1,n}(x)| = |r_{1,n}(x) - \tilde{r}_{1,n}(x)| = o(n^{-\theta}).$$

Ainsi, en utilisant l'hypothèse **H(ii)**, nous obtenons

$$\begin{aligned} \sqrt{nh^d} |\alpha_{1,n}(x)| &= \sqrt{nh^d} o(n^{-\theta}) \\ &= o(1). \end{aligned}$$

b) Sous les hypothèses **K(a)**, **K(c)**, **D1**, et un développement de Taylor, on obtient

$$|\alpha_{3,n}(x)| = |\mathbb{E}(\tilde{r}_{1,n}(x)) - r_1(x)| = O(h^2).$$

Ensuite, en utilisant l'hypothèse **H(i)**, nous obtenons

$$\begin{aligned} \sqrt{nh^d} |\alpha_{3,n}(x)| &= \sqrt{nh^d} (O(h^2)) \\ &= O(\sqrt{nh^{d+4}}) = o(1). \end{aligned}$$

c) Sous les hypothèses **K(a)**, **D2** et un développement de Taylor, on obtient

$$|\beta_{2,n}(x)| = |\mathbb{E}(\ell_n(x)) - \ell(x)| = O(h^2).$$

Ainsi, en utilisant l'hypothèses **H(i)** on obtient

$$\begin{aligned}\sqrt{nh^d} |\beta_{2,n}(x)| &= \sqrt{nh^d} (O(h^2)) \\ &= O(\sqrt{nh^{d+4}}) = o(1).\end{aligned}$$

Ce qui nous permet de terminer la preuve de Lemme 3.1.  $\square$

**Lemme 3.2.** *Sous les hypothèses **M**, **H**, **K**, **D1** et **D4**, on a*

$$\text{var} \left( \sqrt{nh^d} \alpha_{2,n}(x) \right) \longrightarrow \sigma_1^2(x) := r_2(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \text{ as } n \rightarrow +\infty.$$

**Preuve du Lemme 3.2.**

Dans toute la suite et pour des commodités de notation, nous mettons

$$W_i(x, h) := \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right) - \mathbb{E} \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right], \quad (3.24)$$

donc

$$\alpha_{2,n}(x) = \frac{1}{nh^d} \sum_{i=1}^n W_i(x, h).$$

$$\begin{aligned}\text{var} \left( \sqrt{nh^d} \alpha_{2,n}(x) \right) &= nh^d \text{var} (\alpha_{2,n}(x)) \\ &= \frac{1}{nh^d} \text{var} \left( \sum_{i=1}^n W_i(x, h) \right) \\ &= \frac{1}{h^d} \text{var} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right) \\ &\quad + \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) \\ &=: S_1 + S_2.\end{aligned}$$

D'une part, nous avons

$$\begin{aligned}S_1 &= \frac{1}{h^d} \left\{ \mathbb{E} \left[ \frac{\delta_1^2 Z_1^2}{\bar{G}^2(Z_1)} K_d^2 \left( \frac{x - X_1}{h} \right) \right] - \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right] \right\} \\ &=: \frac{1}{h^d} \{S_{11} - S_{12}\},\end{aligned}$$

De plus, en utilisant le conditionnement par rapport à  $Y_1$  et un changement de variables on obtient

$$S_{11} = h^d \int_{\mathbb{R}^d} K_d^2(t) r_2(x - th) dt.$$

Maintenant, nous évaluons  $S_{12}$

$$\begin{aligned}S_{12} &= \mathbb{E}^2 \left[ \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right] \\ &= \left[ \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) r_1(u) du \right]^2.\end{aligned}$$

Ensuite, en utilisant un changement de variables, on obtient

$$S_{12} = \left[ h^d \int_{\mathbb{R}^d} K_d(t) r_1(x - th) dt \right]^2 = h^{2d} \left[ \int_{\mathbb{R}^d} K_d(t) r_1(x - th) dt \right]^2$$

Par conséquent

$$S_1 = \int_{\mathbb{R}^d} K_d^2(t) r_2(x - th) dt - h^d \left[ \int_{\mathbb{R}^d} K_d(t) r_1(x - th) dt \right]^2$$

Ainsi, en utilisant le Théorème de convergence dominée (T.C.D), nous obtenons

$$S_1 \longrightarrow \int_{\mathbb{R}^d} K_d^2(t) r_2(x) dt.$$

D'autre part, nous avons

$$\begin{aligned} S_2 &= \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \operatorname{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_d \left( \frac{x - X_i}{h} \right), \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_d \left( \frac{x - X_j}{h} \right) \right) \\ &=: \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \operatorname{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j} \right). \end{aligned}$$

Le calcul du terme de covariance est effectué en utilisant les techniques d'espérance par conditionnement, et sous les hypothèses **H**, **K**, et **D4**, nous obtenons

$$\operatorname{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j} \right) = O(h^{2d}). \quad (3.25)$$

Maintenant, pour traiter  $S_2$ , nous définissons

$$\mathcal{B}_1 = \{(i, j); 1 \leq |i - j| \leq \vartheta_n\} \text{ et } \mathcal{B}_2 = \{(i, j); \vartheta_n + 1 \leq |i - j| \leq n - 1\}, \quad (3.26)$$

où  $\vartheta_n = o(n)$ .

Alors, nous écrivons

$$\begin{aligned} S_2 &= \frac{1}{nh^d} \sum_{i=1}^n \sum_{j \in \mathcal{B}_1} \operatorname{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j} \right) + \frac{1}{nh^d} \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} \operatorname{cov} \left( \frac{\delta_i Z_i}{\bar{G}(Z_i)} K_{d,i}, \frac{\delta_j Z_j}{\bar{G}(Z_j)} K_{d,j} \right) \\ &=: S_{21} + S_{22}. \end{aligned}$$

Premièrement, à partir de (3.25) et (3.26), on a

$$S_{21} = O \left( \frac{\vartheta_n n h^{2d}}{n h^d} \right) = O(\vartheta_n h^d). \quad (3.27)$$

Ensuite, en utilisant la Proposition 2, le résultat (a) dans Menni et Tatachak (2016) et l'hypothèse

**M**, on obtient

$$\begin{aligned}
S_{22} &\leq \frac{1}{nh^d} \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} ch^d \phi(|i-j|) \\
&\leq \frac{cnh^d}{nh^d} \sum_{j \in \mathcal{B}_2} \exp\left(-\frac{\gamma^d}{2d+2}|i-j|\right) \\
&\leq c \int_{\vartheta_n}^n \exp\left(-\frac{\gamma^d}{2d+2}u\right) du \\
&= O\left(\exp\left(-\frac{\gamma^d}{2d+2}\vartheta_n\right)\right). \tag{3.28}
\end{aligned}$$

En choisissant  $\vartheta_n = O(h^{\nu-d})$  avec  $0 < \nu < d$ , les expressions dans (3.27) et (3.28) donnent  $S_{21} = O(h^\nu)$  et  $S_{22} = \exp\left(-\frac{\gamma^d}{2d+2}h^{\nu-d}\right)$ . Puis par T.C.D nous obtenons  $S_2 \rightarrow o(1)$ .  $\square$

**Lemme 3.3.** *Sous les hypothèses **M**, **H**, **K**, **D2** et **D4**, on a*

$$\text{var}\left(\sqrt{nh^d}\beta_{1,n}(x)\right) \rightarrow \sigma_2^2(x) := \ell(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \text{ as } n \rightarrow +\infty.$$

**Preuve du Lemme 3.3.**

En utilisant les mêmes étapes de la preuve du Lemme 3.2, nous obtenons le résultat.  $\square$

**Lemme 3.4.** *Sous les hypothèses **K**, **M** et **D4**, on a*

$$\text{cov}\left(\sqrt{nh^d}\alpha_{2,n}(x), \sqrt{nh^d}\beta_{1,n}(x)\right) \rightarrow \sigma_{12}^2(x) := r_1(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \text{ as } n \rightarrow +\infty.$$

**Preuve du Lemme 3.4.**

Soit

$$\eta_i(x, h) := K_d\left(\frac{x - X_i}{h}\right) - \mathbb{E}\left(K_d\left(\frac{x - X_1}{h}\right)\right), \tag{3.29}$$

à partir de (3.24) et (3.29), on peut écrire :

$$\begin{aligned}
\text{cov}\left(\sqrt{nh^d}\alpha_{2,n}(x), \sqrt{nh^d}\beta_{1,n}(x)\right) &= \text{cov}\left(\sqrt{nh^d}\frac{1}{nh^d} \sum_{i=1}^n W_i(x, h), \sqrt{nh^d}\frac{1}{nh^d} \sum_{j=1}^n \eta_j(x, h)\right) \\
&= \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1}^n \text{cov}(W_i(x, h), \eta_j(x, h)) \\
&= \frac{n}{nh^d} \text{cov}(W_1(x, h), \eta_1(x, h)) + \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \text{cov}(W_i(x, h), \eta_j(x, h)) \\
&= \frac{1}{h^d} \mathbb{E}(W_1(x, h)\eta_1(x, h)) + \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \mathbb{E}(W_i(x, h)\eta_j(x, h)).
\end{aligned}$$

D'une part, en utilisant un changement de variables, un développement de Taylor et sous les hypothèses **K(a)**, **D1** et **D2**, on obtient

$$\begin{aligned}
 \frac{1}{h^d} \mathbb{E}(W_1(x, h)\eta_1(x, h)) &= \frac{1}{h^d} \left\{ \mathbb{E} \left( \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d^2 \left( \frac{x - X_1}{h} \right) \right) \right) - \mathbb{E} \left( \frac{\delta_1 Z_1}{\bar{G}(Z_1)} K_d \left( \frac{x - X_1}{h} \right) \right) \right\} \\
 &\quad \times \mathbb{E} \left( K_d \left( \frac{x - X_1}{h} \right) \right) \\
 &= \frac{1}{h^d} \left\{ \left( \int_{\mathbb{R}^d} K_d^2 \left( \frac{x - u}{h} \right) r_1(u) du \right) - \left( \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) r_1(u) du \right) \right. \\
 &\quad \left. \times \left( \int_{\mathbb{R}^d} K_d \left( \frac{x - u}{h} \right) \ell(u) du \right) \right\} \\
 &= \frac{1}{h^d} \left\{ \left( \int_{\mathbb{R}^d} K_d^2(t) r_1(x - ht) (h^d) dt \right) - \left( \int_{\mathbb{R}^d} K_d(t) r_1(x - ht) (h^d) dt \right) \right. \\
 &\quad \left. \times \left( \int_{\mathbb{R}^d} K_d(t) \ell(x - ht) (h^d) dt \right) \right\} \\
 &= r_1(x) \int_{\mathbb{R}^d} K_d^2(t) dt + o(1).
 \end{aligned}$$

D'autre part, comme dans la preuve de (3.25) et par l'hypothèse **D4**, on peut montrer que

$$\mathbb{E}(W_i(x, h)\eta_j(x, h)) = O(h^{2d}).$$

Ainsi, en adaptant les preuves de (3.27) et (3.28), nous obtenons

$$\lim_{n \rightarrow \infty} \frac{1}{nh^d} \sum_{i=1}^n \sum_{j=1: j \neq i}^n \mathbb{E}(W_i(x, h)\eta_j(x, h)) = 0.$$

La preuve est maintenant achevée.  $\square$

Nous allons prouver à présent la normalité asymptotique de  $\sqrt{nh^d}\alpha_{2,n}(x), \sqrt{nh^d}\beta_{1,n}(x)$ , et de toute combinaison linéaire entre eux.

Par conséquent, étant donné un couple de nombres réels  $(b_1, b_2)$ , nous avons :

$$\psi_n(x) := b_1 \sqrt{nh^d} \alpha_{2,n}(x) + b_2 \sqrt{nh^d} \beta_{1,n}(x). \quad (3.30)$$

**Proposition 3.1.** *Sous les hypothèses **K1**, **D1** -**D4** et **M**, on a*

$$\psi_n(x) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_3^2(x)), \quad (3.31)$$

où  $\sigma_3^2(x) = \kappa b^t \Sigma(x) b$  et  $\Sigma = \begin{pmatrix} r_2(x) & r_1(x) \\ r_1(x) & \ell(x) \end{pmatrix}$ , et  $b = (b_1, b_2)^t$ .

**Preuve de la Proposition 3.1**

En utilisant les Lemmes 3.2 – 3.4, nous avons

$$\begin{aligned}
\text{var}(\psi_n(x)) &= \text{var}\left(b_1\sqrt{nh^d}\alpha_{2,n}(x) + b_2\sqrt{nh^d}\beta_{1,n}(x)\right) \\
&= b_1^2\text{var}\left(\sqrt{nh^d}\alpha_{2,n}(x)\right) + b_2^2\text{var}\left(\sqrt{nh^d}\beta_{1,n}(x)\right) + 2b_1b_2\text{cov}\left(\sqrt{nh^d}\alpha_{2,n}(x), \sqrt{nh^d}\beta_{1,n}(x)\right) \\
&= b_1^2\sigma_1^2(x) + b_2^2\sigma_2^2(x) + 2b_1b_2\sigma_{12}(x) \\
&= \sigma_3^2(x).
\end{aligned}$$

Soit

$$V_i(x) := b_1W_i(x) + b_2\eta_i(x),$$

alors

$$\begin{aligned}
\psi_n(x) &= \left(\frac{1}{\sqrt{n}}\right) \sum_{i=1}^n \left(\frac{V_i(x)}{\sqrt{h^d}}\right), \\
&= : \frac{1}{\sqrt{n}} \sum_{i=1}^n U_i.
\end{aligned}$$

Le but est d'établir la convergence en loi comme suit :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n U_i \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_3^2(x)). \quad (3.32)$$

Nous employons la procédure de Bernstein des petits blocs et grands blocs qui consiste à partager l'ensemble  $\{1, 2, \dots, n\}$  en  $k$  grands  $p$ -blocs  $I_m$  et  $k$  petits  $q$ -blocs  $J_m$ , où :

$I_m = \{i, i = (m-1)(p+q) + 1, \dots, (m-1)(p+q) + p\}$ , contient  $p$  éléments pour chaque  $m = 1, \dots, k$ ,

$J_m = \{j, j = (m-1)(p+q) + p + 1, \dots, m(p+q)\}$ , contient  $q$  éléments pour chaque  $m = 1, \dots, k$ .

Les points restants vont constituer l'ensemble  $\{l, k(p+q) + 1 \leq l \leq n\}$  qui peut être vide.

Pour  $m = 1, \dots, k$ , on définit les variables aléatoires suivantes :

$$y_m = \sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} U_i, \quad y'_m = \sum_{j=(m-1)(p+q)+p+1}^{m(p+q)} U_j, \quad y''_k = \sum_{l=k(p+q)+1}^n U_l.$$

Alors,

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{i=1}^n U_i &= n^{-\frac{1}{2}} \left\{ \sum_{m=1}^k y_m + \sum_{m=1}^k y'_m + y''_k \right\}, \\
&=: n^{-\frac{1}{2}} \left\{ T_n + T'_n + T''_n \right\}.
\end{aligned}$$

Donc, pour prouver la normalité asymptotique de (3.32), il suffit de prouver que :

$$n^{-\frac{1}{2}} T_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_3^2(x)), \quad (3.33)$$

et

$$\frac{1}{n} \left\{ \mathbb{E}(T_n'^2) + \mathbb{E}(T_n''^2) \right\} \xrightarrow{n \rightarrow +\infty} 0. \quad (3.34)$$

Pour montrer (3.33) et (3.34) on a besoin des Lemmes suivants :

**Lemme 3.5.** *Sous les hypothèses **K1**, **M**, **D4** et **B**, on a*

i)  $\frac{k}{n} \text{var}(y'_1) \rightarrow 0$ , as  $n \rightarrow +\infty$ ,

ii)  $\frac{1}{n} \left| \text{cov}(y'_1, y'_{l+1}) \right| \leq q \cdot c \frac{1}{nh^{d+2}} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |\Delta_{1,r+1}|$ ,

iii)  $\frac{1}{n} \sum_{1 \leq i < j \leq k} \left| \text{cov}(y'_i, y'_j) \right| \rightarrow 0$ .

**Preuve.**

Pour le point (i), on a

$$\begin{aligned} \frac{k}{n} \text{var}(y'_1) &= \frac{k}{n} \text{var} \left( \sum_{j=p+1}^{p+q} U_j \right) \\ &= \frac{k \cdot q}{n} \text{var}(U_1) + \frac{2 \cdot k}{n} \sum_{1 \leq i < j \leq q} |\text{cov}(U_i, U_j)| \\ &=: J_1 + J_2. \end{aligned}$$

Nous montrons premièrement que

$$\begin{aligned} J_1 &= \frac{k \cdot q}{n} \text{var}(U_1) \\ &= \frac{k \cdot q}{n} \text{var} \left( \frac{1}{\sqrt{h^d}} (b_1 W_1(x) + b_2 \eta_1(x)) \right), \\ &= \frac{k \cdot q}{n} \frac{1}{h^d} \text{var}(b_1 W_1(x) + b_2 \eta_1(x)), \\ &= \frac{k \cdot q}{n} \left[ \frac{1}{h^d} \text{var}(b_1 W_1(x)) + \frac{1}{h^d} \text{var}(b_2 \eta_1(x)) + \frac{2}{h^d} \text{cov}(b_1 W_1(x), b_2 \eta_1(x)) \right], \\ &= \frac{k \cdot q}{n} \left[ b_1^2 \frac{1}{h^d} \text{var}(W_1(x)) + b_2^2 \frac{1}{h^d} \text{var}(\eta_1(x)) + 2b_1 b_2 \frac{1}{h^d} \text{cov}(W_1(x), \eta_1(x)) \right], \\ &= \frac{k \cdot q}{n} \sigma_3^2(x) = c \frac{k \cdot q}{n} = o(1), \end{aligned}$$

d'après l'hypothèse **B**(i).

On constate que

$$\begin{aligned} J_2 &= \frac{2 \cdot k}{n} \sum_{1 \leq i < j \leq q} |\text{cov}(U_i, U_j)| \\ &= \frac{2 \cdot k}{n} \sum_{l=1}^{q-1} (q-l) |\text{cov}(U_1, U_{l+1})| \quad (\text{par stationnarité}) \\ &\leq \frac{2kq}{n} \sum_{l=1}^{q-1} |\text{cov}(U_1, U_{l+1})| \end{aligned}$$

$$\begin{aligned}
|\text{cov}(U_1, U_{l+1})| &= \left| \text{cov} \left( \frac{b_1 W_1(x) + b_2 \eta_1(x)}{\sqrt{h^d}}, \frac{b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x)}{\sqrt{h^d}} \right) \right| \\
&= \left| \frac{1}{h^d} \text{cov}(b_1 W_1(x) + b_2 \eta_1(x), b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x)) \right| \\
&\leq \frac{1}{h^d} [b_1^2 |\text{cov}(W_1(x), W_{l+1}(x))| + b_1 b_2 |\text{cov}(W_1(x), \eta_{l+1}(x))| + b_1 b_2 |\text{cov}(\eta_1(x), W_{l+1}(x))| \\
&\quad + |b_2^2 \text{cov}(\eta_1(x), \eta_{l+1}(x))|] = O(h^d).
\end{aligned}$$

Alors,

$$\begin{aligned}
J_2 &\leq \frac{2kq}{n} \sum_{l=1}^{q-1} c.h^d = \frac{2kq}{n} (q-1)c.h^d \\
&\leq c \frac{2kq}{n} qh^d \\
&\leq c \frac{kq}{n} qh^d,
\end{aligned}$$

sous les hypothèses **B**(i) et (ii) on obtient  $J_2 = o(1)$ . Donc  $\frac{k}{n} \text{var}(y'_1) \rightarrow 0$ .

Pour le point (ii), on a

$$\begin{aligned}
 \frac{1}{n} \left| \text{cov} \left( y'_1, y'_{l+1} \right) \right| &= \frac{1}{n} \left| \text{cov} \left( \sum_{i=p+1}^{p+q} U_i, \sum_{j=l(p+q)+p+1}^{(l+1)(p+q)} U_j \right) \right| \\
 &= \frac{1}{n} \left| \sum_{i=p+1}^{p+q} \sum_{j=l(p+q)+p+1}^{(l+1)(p+q)} \text{cov}(U_i, U_j) \right| \\
 &= \frac{1}{n} \left| \sum_{r=1}^q (q-r+1) \text{cov}(U_1, U_{l(p+q)+r}) + \sum_{r=1}^{q-1} (q-r) \text{cov}(U_{r+1}, U_{l(p+q)+1}) \right| \\
 &= \frac{1}{n} \left| \sum_{r=1}^q (q-r+1) \text{cov}(U_1, U_{l(p+q)+r}) + \sum_{r=1}^{q-1} (q-r) \text{cov}(U_1, U_{l(p+q)-r+1}) \right| \\
 &\leq \frac{q}{n} \sum_{r=1}^q \left| \text{cov}(U_1, U_{l(p+q)+r}) \right| + \frac{q}{n} \sum_{r=1}^{q-1} \left| \text{cov}(U_1, U_{l(p+q)-r+1}) \right| \\
 &= \frac{q}{n} \sum_{r=l(p+q)+1}^{l(p+q)+q} \left| \text{cov}(U_1, U_r) \right| + \frac{q}{n} \sum_{r=l(p+q)-(q-2)}^{l(p+q)} \left| \text{cov}(U_1, U_r) \right| \\
 &= \frac{q}{n} \sum_{r=l(p+q)-(q-2)}^{l(p+q)+q} \left| \text{cov}(U_1, U_r) \right| \\
 &= \frac{q}{n} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} \left| \text{cov}(U_1, U_{r+1}) \right| \\
 &= \frac{q}{n} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} \left| \text{cov} \left( \frac{1}{\sqrt{h^d}} V_1, \frac{1}{\sqrt{h^d}} V_{r+1} \right) \right| \\
 &= \frac{q}{nh^d} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} \left| \text{cov}(V_1, V_{r+1}) \right| \\
 &= \frac{q}{nh^d} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} \left| \left[ b_1^2 \text{cov}(W_1, W_{r+1}) + b_2^2 \text{cov}(\eta_1, \eta_{r+1}) + 2b_1 b_2 \text{cov}(W_1, \eta_{r+1}) \right] \right|
 \end{aligned}$$

En utilisant le Théorème 5.3 dans Bulinski et Shashkin(2007), on peut montrer que

$$\text{cov}(W_1, W_{r+1}) \leq c_1^2 \frac{1}{h^2} \Delta_{1,r+1},$$

avec

$$c_1 = \tau_F \left( \bar{G}(\tau_F) \right)^{-1} \text{Lip}(K_d) \vee \|K_d\|_\infty \left( \left( \bar{G}(\tau_F) \right)^{-1} + \tau_F \left( \bar{G}^2(\tau_F) \right)^{-1} \text{Lip}(G) \right),$$

et

$$\text{cov}(\eta_1, \eta_{r+1}) \leq c_2^2 \frac{1}{h^2} \Delta_{1,r+1},$$

avec

$$c_2 = \left( \bar{G}(\tau_F) \right)^{-1} \text{Lip}(K_d) \vee \|K_d\|_\infty \left( \bar{G}^2(\tau_F) \right)^{-1} \text{Lip}(G),$$

et

$$\text{cov}(W_1, \eta_{r+1}) \leq c_1 c_2 \frac{1}{h^2} \Delta_{1,r+1}.$$

Donc

$$\frac{1}{n} \left| \text{cov}(y'_1, y'_{l+1}) \right| \leq \frac{q \cdot c}{n} \frac{1}{h^{d+2}} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |\Delta_{1,r+1}|.$$

Enfin, pour le point (iii), par stationnarité et les hypothèses **B**(i) et (iii) on obtient

$$\begin{aligned} \frac{1}{n} \sum_{1 \leq i < j \leq k} \left| \text{cov}(y'_i, y'_j) \right| &= \frac{1}{n} \sum_{l=1}^{k-1} (k-l) \left| \text{cov}(y'_1, y'_{l+1}) \right| \\ &\leq \frac{k}{n} \sum_{l=1}^{k-1} \left| \text{cov}(y'_1, y'_{l+1}) \right| \\ &\leq \frac{qk}{nh^{d+2}} c \sum_{l=1}^{k-1} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |\Delta_{1,r+1}| \\ &\leq \frac{qk}{n} \frac{1}{nh^{d+2}} c \sum_{r=p}^{\infty} |\Delta_{1,r+1}| \rightarrow 0. \end{aligned}$$

La preuve du Lemme 3.5 est terminée.  $\square$

**Lemme 3.6.** *Sous les mêmes hypothèses que celles du Lemme 3.5, nous avons*

i)  $\frac{k}{n} \text{var}(y_1) \rightarrow \sigma_3^2(x),$

ii)  $\frac{1}{n} |\text{cov}(y_1, y_{l+1})| \leq p \frac{1}{nh^{d+2}} \sum_{r=l(p+q)-p}^{l(p+q)+p} |\Delta_{1,r+1}|,$

iii)  $\frac{1}{n} \sum_{1 \leq i < j \leq k} |\text{cov}(y_i, y_j)| \rightarrow 0,$

iv)  $\text{var}(n^{-\frac{1}{2}} T_n) \rightarrow \sigma_3^2(x).$

**Preuve.** Pour le point (i), on a

$$\begin{aligned} \frac{k}{n} \text{var}(y_1) &= \frac{k}{n} \text{var} \left( \sum_{i=1}^p U_i \right) \\ &= \frac{kp}{n} \text{var}(U_1) + \frac{2k}{n} \sum_{1 \leq i < j \leq p} |\text{cov}(U_i, U_j)| \\ &=: J'_1 + J'_2 \end{aligned}$$

Nous montrons premièrement que

$$\begin{aligned}
J'_1 &= \frac{kp}{n} \text{var}(U_1) \\
&= \frac{kp}{n} \text{var}\left(\frac{b_1 W_1(x) + b_2 \eta_1(x)}{\sqrt{h^d}}\right) \\
&= \frac{kp}{n} \frac{1}{h^d} \text{var}(b_1 W_1(x) + b_2 \eta_1(x)) \\
&= \frac{kp}{n} \left[ \frac{1}{h^d} \text{var}(b_1 W_1(x)) + \frac{1}{h^d} \text{var}(b_2 \eta_1(x)) + \frac{2}{h^d} \text{cov}(b_1 W_1(x), b_2 \eta_1(x)) \right] \\
&= \frac{kp}{n} \left[ b_1^2 \frac{1}{h^d} \text{var}(W_1(x)) + b_2^2 \frac{1}{h^d} \text{var}(\eta_1(x)) + 2b_1 b_2 \frac{1}{h^d} \text{cov}(W_1(x), \eta_1(x)) \right] \\
&= \frac{kp}{n} \sigma_3^2(x),
\end{aligned}$$

en utilisant **B(i)**, on obtient

$$J'_1 = \sigma_3^2(x).$$

On a

$$\begin{aligned}
J'_2 &= \frac{2k}{n} \sum_{1 \leq i < j \leq p} |\text{cov}(U_i, U_j)| \\
&= \frac{2k}{n} \sum_{l=1}^{p-1} (p-l) |\text{cov}(U_l, U_{l+1})| \\
&\leq \frac{2kp}{n} \sum_{l=1}^{p-1} |\text{cov}(U_l, U_{l+1})| \\
&= \frac{2kp^2}{n} \left[ \left| \text{cov}\left(\frac{b_1 W_1(x) + b_2 \eta_1(x)}{\sqrt{h^d}}, \frac{b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x)}{\sqrt{h^d}}\right) \right| \right] \\
&= \frac{2kp^2}{nh^d} [|\text{cov}(b_1 W_1(x) + b_2 \eta_1(x), b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x))|] \\
&\leq \frac{2kp^2}{nh^d} [b_1^2 |\text{cov}(W_1(x), W_{l+1}(x))| + b_1 b_2 |\text{cov}(W_1(x), \eta_{l+1}(x))| \\
&\quad + b_1 b_2 |\text{cov}(\eta_1(x), W_{l+1}(x))| + b_2^2 |\text{cov}(\eta_1(x), \eta_{l+1}(x))|] \\
&= \frac{2kp^2}{nh^d} ch^{2d} \\
&= \frac{2kp^2}{n} ch^d = c \frac{kp}{n} .ph^d
\end{aligned}$$

Sous les hypothèses **B(i)** et **B(ii)**, nous obtenons

$$J'_2 = o(1).$$

Donc,  $\frac{k}{n} \text{var}(y_1) \rightarrow \sigma_3^2(x)$ .

Concernant le point (ii), la preuve est obtenue de manière analogue à celle du Lemme 3.5(ii), nous avons

$$\begin{aligned}
 \frac{1}{n} |\text{cov}(y_1, y_{l+1})| &= \frac{1}{n} \left| \text{cov} \left( \sum_{i=1}^p U_i, \sum_{j=l(p+q)+1}^{l(p+q)+p} U_j \right) \right| \\
 &= \frac{1}{n} \left| \sum_{r=1}^p (p-r+1) \text{cov}(U_1, U_{l(p+q)+r}) + \sum_{r=1}^{p-1} (p-r) \text{cov}(U_{r+1}, U_{l(p+q)+1}) \right| \\
 &= \frac{1}{n} \left| \sum_{r=1}^p (p-r+1) \text{cov}(U_1, U_{l(p+q)+r}) + \sum_{r=1}^{p-1} (p-r) \text{cov}(U_1, U_{l(p+q)-r+1}) \right| \\
 &\leq \frac{p}{n} \left| \sum_{r=1}^p \text{cov}(U_1, U_{l(p+q)+r}) + \sum_{r=1}^{p-1} \text{cov}(U_1, U_{l(p+q)-r+1}) \right| \\
 &\leq \frac{p}{n} \sum_{r=l(p+q)-(p-2)}^{l(p+q)+p} |\text{cov}(U_1, U_r)| = \frac{p}{n} \sum_{r=l(p+q)-(p-1)}^{l(p+q)+(p-1)} |\text{cov}(U_1, U_{r+1})| \\
 &\leq \frac{p}{nh^d} \sum_{r=l(p+q)-(p-1)}^{l(p+q)+(p-1)} |\text{cov}(V_1, V_{r+1})| \\
 &\leq \frac{p}{nh^d} \sum_{r=l(p+q)-p}^{l(p+q)+p} \left| [b_1^2 \text{cov}(W_1, W_{r+1}) + b_2^2 \text{cov}(\eta_1, \eta_{r+1}) + 2b_1 b_2 \text{cov}(W_1, \eta_{r+1})] \right| \\
 &\leq c \frac{p}{n} \frac{1}{h^{d+2}} \sum_{r=l(p+q)-p}^{l(p+q)+p} |\Delta_{1,r+1}|.
 \end{aligned}$$

La preuve du point (iii) s'obtient par le même raisonnement que pour Lemme 3.5 (iii), par stationnarité et les hypothèses **B(i)** et (iii), nous obtenons

$$\begin{aligned}
 \frac{1}{n} \sum_{1 \leq i < j \leq k} |\text{cov}(y_i, y_j)| &= \frac{1}{n} \sum_{l=1}^{k-1} (k-l) |\text{cov}(y_1, y_{l+1})| \\
 &\leq \frac{k}{n} \sum_{l=1}^{k-1} |\text{cov}(y_1, y_{l+1})| \\
 &\leq \frac{c.pk}{n} \frac{1}{h^{d+2}} \sum_{l=1}^{k-1} \sum_{r=l(p+q)-p}^{l(p+q)+p} |\Delta_{1,r+1}| \\
 &\leq \frac{c.pk}{n} \frac{1}{h^{d+2}} \sum_{r=q}^{\infty} |\Delta_{1,r+1}| \longrightarrow 0.
 \end{aligned}$$

Enfin, pour le point (iv) en suivant les mêmes étapes que celles du Lemme 3.6 (ii) et (iii), nous obtenons

$$\text{var} \left( n^{-\frac{1}{2}} T_n \right) = \frac{k}{n} \text{var}(y_1) + \frac{2}{n} \sum_{1 \leq i < j \leq k} |\text{cov}(y_i, y_j)| \longrightarrow \sigma_3^2(x).$$

La preuve du Lemme 3.6 est établie.  $\square$

**Lemme 3.7.** *Sous les hypothèses **D4** et **B**, on a*

$$\text{i) } \text{var} \left( n^{-\frac{1}{2}} T'_n \right) \rightarrow 0$$

$$\text{ii) } \text{var} \left( n^{-\frac{1}{2}} T''_n \right) \rightarrow 0$$

**Preuve.** La preuve du point (i) se fait sous les hypothèses **D4**, **B**(i) et (ii), on a

$$\begin{aligned} \text{var} \left( n^{-\frac{1}{2}} T'_n \right) &= \frac{1}{n} \text{var} \left( T'_n \right) \\ &= \frac{1}{n} \text{var} \left( \sum_{m=1}^k y'_m \right) \\ &= \frac{k}{n} \text{var} \left( y'_1 \right) + \frac{2}{n} \sum_{1 \leq i < j \leq k} \left| \text{cov} \left( y'_i, y'_j \right) \right| \rightarrow 0. \end{aligned}$$

Le résultat est donné par le Lemme 3.5 (i) et (iii).

Pour le point (ii), on a

$$\begin{aligned} \text{var} \left( n^{-\frac{1}{2}} T''_n \right) &= \frac{1}{n} \text{var} \left( y''_k \right) \\ &= \frac{1}{n} \text{var} \left( \sum_{l=k(p+q)+1}^n U_l \right) \\ &= \frac{n - k(p+q)}{n} \text{var} (U_1) + \frac{2}{n} \sum_{k(p+q)+1 \leq i < j \leq n} \left| \text{cov} (U_i, U_j) \right| \\ &= \frac{n - k(p+q)}{n} \text{var} (U_1) + \frac{2}{n} \sum_{1 \leq i < j \leq n-k(p+q)} \left| \text{cov} (U_i, U_j) \right| \quad (\text{par stationnarité}) \\ &=: J''_1 + J''_2. \end{aligned}$$

On montre d'abord que

$$\begin{aligned} J''_1 &= \frac{n - k(p+q)}{n} \text{var} (U_1) \\ &= \frac{n - k(p+q)}{n} \text{var} \left( \frac{b_1 W_1(x) + b_2 \eta_1(x)}{\sqrt{h^d}} \right) \\ &= \frac{n - k(p+q)}{n} \left[ \frac{1}{h^d} \text{var} (b_1 W_1(x)) + \frac{1}{h^d} \text{var} (b_2 \eta_1(x)) + \frac{2}{h^d} \text{cov} (b_1 W_1(x), b_2 \eta_1(x)) \right] \\ &= \frac{n - k(p+q)}{n} \left[ b_1^2 \frac{1}{h^d} \text{var} (W_1(x)) + b_2^2 \frac{1}{h^d} \text{var} (\eta_1(x)) + 2b_1 b_2 \frac{1}{h^d} \text{cov} (W_1(x), \eta_1(x)) \right] \\ &= \frac{n - k(p+q)}{n} \sigma_3^2(x) \\ &\leq c \frac{p}{n}, \end{aligned}$$

( $n-k(p+q)$  étant inférieur à  $p$ ), donc  $J''_1 = o(1)$ .

On constate que

$$\begin{aligned}
J_2'' &= \frac{2}{n} \sum_{1 \leq i < j \leq n-k(p+q)} |\text{cov}(U_i, U_j)| \\
&\leq \frac{2}{n} \sum_{1 \leq i < j \leq p} |\text{cov}(U_i, U_j)| \\
&= \frac{2}{n} \sum_{l=1}^{p-1} (p-l) |\text{cov}(U_1, U_{l+1})| \\
&\leq \frac{2p}{n} \sum_{l=1}^{p-1} |\text{cov}(U_1, U_{l+1})| \\
&= \frac{2p}{n} \sum_{l=1}^{p-1} \left[ \left| \text{cov} \left( \frac{b_1 W_1(x) + b_2 \eta_1(x)}{\sqrt{h^d}}, \frac{b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x)}{\sqrt{h^d}} \right) \right| \right] \\
&= \frac{2p}{nh^d} \sum_{l=1}^{p-1} [|\text{cov}(b_1 W_1(x) + b_2 \eta_1(x), b_1 W_{l+1}(x) + b_2 \eta_{l+1}(x))|] \\
&\leq \frac{2p}{nh^d} \sum_{l=1}^{p-1} [b_1^2 |\text{cov}(W_1(x), W_{l+1}(x))| + b_1 b_2 |\text{cov}(W_1(x), \eta_{l+1}(x))| \\
&\quad + b_1 b_2 |\text{cov}(\eta_1(x), W_{l+1}(x))| + b_2^2 |\text{cov}(\eta_1(x), \eta_{l+1}(x))|] \\
&= \frac{2p}{nh^d} (p-1) ch^{2d} \\
&\leq c \frac{p}{n} ph^d,
\end{aligned}$$

d'après l'hypothèse **B(ii)**, on obtient  $J_2'' = o(1)$ . Donc,  $\text{var} \left( n^{-\frac{1}{2}} T_n'' \right) \rightarrow 0$ . Le Lemme 3.7 est maintenant achevé.  $\square$

D'après les Lemmes 3.7 et 3.5, ainsi que l'inégalité de Tchebychev, on a (3.34).

La démonstration de (3.33) se fait en deux étapes. Tout d'abord, nous devons établir que les fonctions caractéristiques de  $\sum_{m=1}^k n^{-\frac{1}{2}} y_m$  moins le produit de la fonction caractéristique de  $n^{-\frac{1}{2}} y_m$ , converge vers 0, c'est-à-dire

$$\left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^k \mathbb{E} \left( \exp \left( it n^{-\frac{1}{2}} y_m \right) \right) \right| \xrightarrow{n \rightarrow +\infty} 0. \quad (3.35)$$

Cela prouve que les v.a.  $n^{-\frac{1}{2}} y_m$ , sont asymptotiquement indépendantes, deuxièmement, il est prouvé que la distribution déterminée par le produit de la fonction caractéristique de  $n^{-\frac{1}{2}} y_m, m = 1, \dots, k$  est asymptotiquement la distribution  $\mathcal{N}(0, \sigma_3^2(x))$ . Nous devons donc vérifier la condition standard de Lindeberg-Feller (voir loève 1963, p. 286 ).

$$k \mathbb{E} \left( n^{-1} y_1^2 \mathbb{1}_{\{|n^{-\frac{1}{2}} y_1| > \varepsilon \sigma_3(x)\}} \right) \xrightarrow{n \rightarrow +\infty} 0, \text{ pour tout } \varepsilon > 0. \quad (3.36)$$

Ci-après les détails Concernant la première étape, on a

$$\begin{aligned}
 I_k(t) &:= \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^k \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \right| \\
 &= \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^{k-1} \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \cdot \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \right| \\
 &= \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) \right. \\
 &\quad \left. + \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^{k-1} \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \cdot \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \right|.
 \end{aligned}$$

Par l'inégalité triangulaire, on a :

$$\begin{aligned}
 I_k(t) &\leq \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) \right| \\
 &\quad + \left| \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \right| \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^{k-1} \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \right|,
 \end{aligned}$$

ainsi, on obtient

$$\begin{aligned}
 I_k(t) &\leq \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) \right| \\
 &\quad + \left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^{k-1} \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \right|,
 \end{aligned}$$

ce qui donne

$$I_k(t) \leq \left| \text{cov} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right), \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \right| + I_{k-1}(t). \quad (3.37)$$

En répétant le même argument (3.37) devient

$$\left| \mathbb{E} \left( \exp \left( it \sum_{m=1}^k n^{-\frac{1}{2}} y_m \right) \right) - \prod_{m=1}^k \mathbb{E} \left( \exp \left( itn^{-\frac{1}{2}} y_m \right) \right) \right| \quad (3.38)$$

$$\begin{aligned}
 &\leq \left| \text{cov} \left( \exp \left( it \sum_{m=1}^{k-1} n^{-\frac{1}{2}} y_m \right), \exp \left( itn^{-\frac{1}{2}} y_k \right) \right) \right| \\
 &\quad + \left| \text{cov} \left( \exp \left( it \sum_{m=1}^{k-2} n^{-\frac{1}{2}} y_m \right), \exp \left( itn^{-\frac{1}{2}} y_{k-1} \right) \right) \right| \\
 &\quad + \dots + \left| \text{cov} \left( \exp \left( itn^{-\frac{1}{2}} y_1 \right), \exp \left( itn^{-\frac{1}{2}} y_2 \right) \right) \right|. \quad (3.39)
 \end{aligned}$$

Par conséquent, nous appliquons le Lemme 1 dans Bulinski (1996) à chacun des termes à droite de (3.38). Pour cela, il faut déterminer les dérivées partielles du premier ordre de la fonction

$\phi_m : \mathbb{R}^{p(d+1)} \rightarrow \mathbb{R}$ ,  $m = 1, \dots, k$ . défini par

$$\begin{aligned} \phi_m \left( x_{(m-1)(p+q)+1}, \dots, x_{(m-1)(p+q)+p}, z_{(m-1)(p+q)+1}, \dots, z_{(m-1)(p+q)+p} \right) &= e^{it \frac{y_m}{\sqrt{n}}} \\ &= e^{it} \left\{ \sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} \frac{1}{\sqrt{nh^d}} [b_1 W_1(x_l, z_l) + b_2 \eta_1(x_l, z_l)] \right\}, \end{aligned}$$

où  $x_l \in \mathbb{R}^d$  et  $z_l \in \mathbb{R}$  pour  $l = (m-1)(p+q)+1, \dots, (m-1)(p+q)+p$ , nous évaluons  $\frac{\partial \phi_m}{\partial z_l}(x_l, z_l)$  et  $\frac{\partial \phi_m}{\partial x_l^j}(x_l, z_l)$  avec  $j = 1, \dots, d$ .

$$\begin{aligned} \frac{\partial \phi_m}{\partial z_l}(x_l, z_l) &= \frac{\partial}{\partial z_l} e^{it \frac{y_m}{\sqrt{n}}} \\ &= it e^{it \frac{y_m}{\sqrt{n}}} \cdot \frac{\partial}{\partial z_l} \left( n^{-\frac{1}{2}} \sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} U_i \right) \\ &= it e^{it \frac{y_m}{\sqrt{n}}} \cdot \frac{\partial}{\partial z_l} \left[ \frac{1}{\sqrt{nh^d}} b_1 \left\{ \frac{\delta_l z_l}{\bar{G}(z_l)} K_d \left( \frac{x - x_l}{h} \right) - \mathbb{E} \left( \frac{\delta_l z_l}{\bar{G}(z_l)} K_d \left( \frac{x - x_1}{h} \right) \right) \right\} \right. \\ &\quad \left. + \frac{1}{\sqrt{nh^d}} b_2 \left\{ K_d \left( \frac{x - x_l}{h} \right) - \mathbb{E} \left( K_d \left( \frac{x - x_1}{h} \right) \right) \right\} \right] \\ &= \frac{it e^{it \frac{y_m}{\sqrt{n}}}}{\sqrt{nh^d}} \cdot \left[ b_1 \left\{ \frac{\bar{G}(z_l) - g(z_l) z_l}{\bar{G}^2(z_l)} \cdot \delta_l \cdot K_d \left( \frac{x - x_l}{h} \right) \right\} \right]. \end{aligned}$$

$$\begin{aligned} \frac{\partial \phi_m}{\partial x_l^j}(x_l, z_l) &= \frac{\partial}{\partial x_l^j} e^{it \frac{y_m}{\sqrt{n}}}, 1 \leq j \leq d \\ &= it e^{it \frac{y_m}{\sqrt{n}}} \cdot \frac{\partial}{\partial x_l^j} \left( n^{-\frac{1}{2}} \sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} U_i \right) \\ &= \frac{it e^{it \frac{y_m}{\sqrt{n}}}}{\sqrt{nh^d}} \cdot \frac{\partial}{\partial x_l^j} \left[ b_1 \left\{ \frac{\delta_l z_l}{\bar{G}(z_l)} K_d \left( \frac{x - x_l}{h} \right) - \mathbb{E} \left( \frac{\delta_l z_l}{\bar{G}(z_l)} K_d \left( \frac{x - x_1}{h} \right) \right) \right\} \right. \\ &\quad \left. + b_2 \left\{ K_d \left( \frac{x - x_l}{h} \right) - \mathbb{E} \left( K_d \left( \frac{x - x_1}{h} \right) \right) \right\} \right] \\ &= \frac{it e^{it \frac{y_m}{\sqrt{n}}}}{\sqrt{nh^d}} \cdot \left[ b_1 \left\{ \frac{\delta_l z_l}{\bar{G}(z_l)} \cdot \frac{\partial}{\partial x_l^j} K_d \left( \frac{x - x_l}{h} \right) \right\} + b_2 \cdot \left\{ \frac{\partial}{\partial x_l^j} K_d \left( \frac{x - x_l}{h} \right) \right\} \right], \end{aligned}$$

avec,

$$K_d \left( \frac{x - x_l}{h} \right) = \prod_{j=1}^d K \left( \frac{x^j - x_l^j}{h} \right).$$

$g$  étant bornée, en utilisant l'hypothèse **K1**, nous avons

$$\left| \frac{\partial \phi_m}{\partial z_l}(x_l, z_l) \right| \leq \frac{c_1 \cdot t}{\sqrt{nh^d}},$$

et

$$\left| \frac{\partial \phi_m}{\partial x_l^j}(x_l, z_l) \right| \leq \frac{c_2 \cdot t}{\sqrt{nh^{d+2}}}.$$

En appliquant le Lemme 1 dans Bulinski (1996), on obtient

$$\left| \text{cov} \left( e^{itn^{-\frac{1}{2}}y_1}, e^{itn^{-\frac{1}{2}}y_2} \right) \right| \leq c \cdot \frac{t^2}{nh^{d+2}} \sum_{j \in I_1} \sum_{l \in I_2} |\Delta_{j,l}|, \quad (3.40)$$

et en généralisant (3.40), nous obtenons

$$\left| \text{cov} \left( e^{it \sum_{m=1}^{k-1} n^{-\frac{1}{2}}y_m}, e^{itn^{-\frac{1}{2}}y_k} \right) \right| \leq c \cdot \frac{t^2}{nh^{d+2}} \sum_{j \in I_1+I_2+\dots+I_{k-1}} \sum_{l \in I_k} |\Delta_{j,l}|. \quad (3.41)$$

Ainsi,

$$\begin{aligned} \left| \mathbb{E} \left( e^{it \sum_{m=1}^k n^{-\frac{1}{2}}y_m} \right) - \prod_{m=1}^k \mathbb{E} \left( e^{itn^{-\frac{1}{2}}y_m} \right) \right| &\leq \frac{c^2 t^2}{nh^{d+2}} \left[ \sum_{j \in I_1} \sum_{l \in I_2} |\Delta_{j,l}| + \sum_{j \in I_1+I_2} \sum_{l \in I_3} |\Delta_{j,l}| \right. \\ &\quad \left. + \dots + \sum_{j \in I_1+I_2+\dots+I_{k-1}} \sum_{l \in I_k} |\Delta_{j,l}| \right] \\ &=: \frac{c^2 t^2}{nh^{d+2}} \cdot \Upsilon, \end{aligned} \quad (3.42)$$

par stationnarité,

$$\Upsilon = (k-1) \sum_{j \in I_1} \sum_{l \in I_2} |\Delta_{j,l}| + (k-2) \sum_{j \in I_1} \sum_{l \in I_3} |\Delta_{j,l}| + \dots + \sum_{j \in I_1} \sum_{l \in I_k} |\Delta_{j,l}|.$$

Donc (3.42) devient

$$\begin{aligned} \left| \mathbb{E} \left( e^{it \sum_{m=1}^k n^{-\frac{1}{2}}y_m} \right) - \prod_{m=1}^k \mathbb{E} \left( e^{itn^{-\frac{1}{2}}y_m} \right) \right| &\leq \frac{c^2 t^2}{nh^{d+2}} \left[ (k-1) \sum_{j \in I_1} \sum_{l \in I_2} |\Delta_{j,l}| + (k-2) \sum_{j \in I_1} \sum_{l \in I_3} |\Delta_{j,l}| \right. \\ &\quad \left. + \dots + \sum_{j \in I_1} \sum_{l \in I_k} |\Delta_{j,l}| \right]. \end{aligned} \quad (3.43)$$

D'autre part, on a aussi par stationnarité

$$\begin{aligned} \sum_{j \in I_1} \sum_{l \in I_2} |\Delta_{j,l}| &= \sum_{j=1}^p \sum_{l=(p+q)+1}^{(p+q)+p} |\Delta_{j,l}| \\ &= |\Delta_{1,(p+q)-(p-2)}| + 2 |\Delta_{1,(p+q)-(p-3)}| + \dots + (p-1) |\Delta_{1,p+q}| \\ &\quad + p |\Delta_{1,(p+q)+1}| + (p-1) |\Delta_{1,(p+q)+2}| + \dots + 2 |\Delta_{1,(p+q)+(p-1)}| + |\Delta_{1,(p+q)+p}| \\ &\leq p \cdot \sum_{j=(p+q)-(p-2)}^{(p+q)+p} |\Delta_{1,j}|, \end{aligned}$$

$$\begin{aligned}
 \sum_{j \in I_1} \sum_{l \in I_3} |\Delta_{j,l}| &= \sum_{j=1}^p \sum_{l=2(p+q)+1}^{2(p+q)+p} |\Delta_{j,l}| \\
 &= |\Delta_{1,2(p+q)-(p-2)}| + 2 |\Delta_{1,2(p+q)-(p-3)}| + \cdots + (p-1) |\Delta_{1,2(p+q)}| \\
 &\quad + p |\Delta_{1,2(p+q)+1}| + (p-1) |\Delta_{1,2(p+q)+2}| + \cdots + 2 |\Delta_{1,2(p+q)+(p-1)}| + |\Delta_{1,2(p+q)+p}| \\
 &\leq p \cdot \sum_{j=2(p+q)-(p-2)}^{2(p+q)+p} |\Delta_{1,j}|,
 \end{aligned}$$

et ainsi de suite, jusqu'à

$$\begin{aligned}
 \sum_{j \in I_1} \sum_{l \in I_k} |\Delta_{j,l}| &= \sum_{j=1}^p \sum_{l=(k-1)(p+q)+1}^{(k-1)(p+q)+p} |\Delta_{j,l}| \\
 &= |\Delta_{1,(k-1)(p+q)-(p-2)}| + 2 |\Delta_{1,(k-1)(p+q)-(p-3)}| + \cdots + (p-1) |\Delta_{1,(k-1)(p+q)}| \\
 &\quad + p |\Delta_{1,(k-1)(p+q)+1}| + (p-1) |\Delta_{1,(k-1)(p+q)+2}| + \cdots + 2 |\Delta_{1,(k-1)(p+q)+(p-1)}| \\
 &\quad + |\Delta_{1,(k-1)(p+q)+p}| \\
 &\leq p \cdot \sum_{j=(k-1)(p+q)-(p-2)}^{(k-1)(p+q)+p} |\Delta_{1,j}|.
 \end{aligned}$$

En utilisant ces expressions, l'inégalité (3.43) devient

$$\begin{aligned}
 \left| \mathbb{E} \left( e^{it \sum_{m=1}^k n^{-\frac{1}{2}} y_m} \right) - \prod_{m=1}^k \mathbb{E} \left( e^{it n^{-\frac{1}{2}} y_m} \right) \right| &\leq \frac{c^2 t^2}{n h^{d+2}} \cdot p \cdot k \sum_{j=(k-1)(p+q)-(p-2)}^{(k-1)(p+q)+p} |\Delta_{1,j}| \\
 &= \frac{c^2 t^2}{n h^{d+2}} \cdot p \cdot k \sum_{j=(p+q)-(p-1)}^{(k-1)(p+q)+(p-1)} |\Delta_{1,j+1}| \\
 &\leq c^2 t^2 \cdot \frac{pk}{n} \cdot \frac{1}{h^{d+2}} \sum_{j=q}^{\infty} |\Delta_{1,j+1}| \longrightarrow 0.
 \end{aligned}$$

d'après les hypothèses **B(i)** et **B(iii)**.  $\square$

Pour la deuxième étape, sous l'hypothèse **K1** on procède de la manière suivante :

$$\begin{aligned}
 |y_1| &= \left| \sum_{i=1}^p U_i \right| \\
 &= \left| \sum_{i=1}^p \frac{V_i(x)}{\sqrt{h^d}} \right| \\
 &= \left| \sum_{i=1}^p \left( \frac{b_1 W_i(x) + b_2 \eta_i(x)}{\sqrt{h_n^d}} \right) \right| \\
 &\leq \frac{1}{\sqrt{h^d}} \sum_{i=1}^p \left[ \left| b_1 \left\{ \frac{\delta_i z_i}{\bar{G}(z_i)} K_d \left( \frac{x - X_i}{h} \right) - \mathbb{E} \left( \frac{\delta_1 z_1}{\bar{G}(z_1)} K_d \left( \frac{x - X_1}{h} \right) \right) \right\} \right| \right. \\
 &\quad \left. + \left| b_2 \left\{ K_d \left( \frac{x - X_i}{h} \right) - \mathbb{E} \left( K_d \left( \frac{x - X_1}{h} \right) \right) \right\} \right| \right] \\
 &\leq \frac{p}{\sqrt{h^d}} \left[ \frac{2b_1 \tau_F}{\bar{G}(\tau_F)} \|K_d\|_\infty + 2b_2 \|K_d\|_\infty \right].
 \end{aligned}$$

En utilisant l'inégalité de Tchebychev, on obtient

$$\begin{aligned}
 k \mathbb{E} \left( n^{-1} y_1^2 \mathbb{1}_{\{|n^{-\frac{1}{2}} y_1| > \varepsilon \sigma_3(x)\}} \right) &\leq \frac{k p^2}{n h^d} \left[ \frac{2b_1 \tau_F}{\bar{G}(\tau_F)} \|K_d\|_\infty + 2b_2 \|K_d\|_\infty \right]^2 \mathbb{P} \left( \left| n^{-\frac{1}{2}} y_1 \right| > \varepsilon \sigma_3(x) \right) \\
 &\leq \frac{p^2}{n h^d} \left[ \frac{2b_1 \tau_F}{\bar{G}(\tau_F)} \|K_d\|_\infty + 2b_2 \|K_d\|_\infty \right]^2 k \frac{\text{var} \left( n^{-\frac{1}{2}} y_1 \right)}{\varepsilon^2 \sigma_3^2(x)} \\
 &= \frac{p^2}{n h^d} \left[ \frac{2b_1 \tau_F}{\bar{G}(\tau_F)} \|K_d\|_\infty + 2b_2 \|K_d\|_\infty \right]^2 \frac{k \text{var} (y_1)}{\varepsilon^2 \sigma_3^2(x)},
 \end{aligned}$$

sous l'hypothèse **B(ii)**, nous avons

$$k \mathbb{E} \left( n^{-1} y_1^2 \mathbb{1}_{\{|n^{-\frac{1}{2}} y_1| > \varepsilon \sigma_3(x)\}} \right) \longrightarrow 0, \text{ as } n \longrightarrow \infty.$$

Ceci termine la preuve de la Proposition 3.1.  $\square$

**Remarque 3.1.** Si on prend  $b_1 = 1$  et  $b_2 = 0$  dans (3.30), en utilisant la Proposition 3.1 et le Lemme 3.2, nous avons

$$\sqrt{nh_n^d} \alpha_{2,n}(x) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \sigma_1^2(x) \right).$$

**Remarque 3.2.** Si on prend  $b_1 = 0$  et  $b_2 = 1$  dans (3.30), en utilisant la Proposition 3.1 et le Lemme 3.3, on obtient

$$\sqrt{nh_n^d} \beta_{1,n}(x) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \sigma_2^2(x) \right).$$

**Proposition 3.2.** sous les hypothèses **K**, **D1 -D4** et **M**, nous avons

$$\sqrt{nh_n^d} (\alpha_{2,n}(x), \beta_{1,n}(x)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \kappa \Sigma(x) \right),$$

où  $\kappa$  et  $\Sigma(x)$  sont définis dans la Proposition 3.1.

**Preuve de la Proposition 3.2** Il suffit d'appliquer la Proposition 3.1, Remarque 3.1 et 3.2.  $\square$

### Preuve du Théorème 3.2

Considérons l'application définie par  $\Theta(x, y) = \frac{x}{y}$  pour  $y \neq 0$ . Nous remarquons que

$$r_n(x) - r(x) = \frac{r_{1,n}(x)}{\ell_n(x)} - \frac{r_1(x)}{\ell(x)} = \Theta(r_{1,n}(x), \ell_n(x)) - \Theta(r_1(x), \ell(x)).$$

D'après le Lemme 3.1, la Proposition 3.2, le Théorème de Mann-wold (voir Rao 1965, p.321) et la méthode delta (voir Van der Vaart et Wellner 1996, p.374) nous déduisons que

$\sqrt{nh_n^d}(r_n(x) - r(x))$  Converge en loi vers  $\mathcal{N}(0, \kappa \nabla \Theta^t \Sigma(x) \nabla \Theta)$ , où le gradient  $\nabla \Theta$  est évalué en  $(r_1(x), \ell(x))$ ,

c'est-à-dire que

$$\nabla \Theta(r_1(x), \ell(x)) = \begin{pmatrix} \frac{1}{\ell(x)} \\ -\frac{r_1(x)}{\ell^2(x)} \end{pmatrix}$$

donc

$$\begin{aligned} \sigma^2(x) &= \kappa \nabla \Theta^t \Sigma(x) \nabla \Theta \\ &= \kappa \begin{pmatrix} \frac{1}{\ell(x)} & -\frac{r_1(x)}{\ell^2(x)} \end{pmatrix} \cdot \begin{pmatrix} r_2(x) & r_1(x) \\ r_1(x) & \ell(x) \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\ell(x)} \\ -\frac{r_1(x)}{\ell^2(x)} \end{pmatrix} \\ &= \kappa \begin{pmatrix} \frac{r_2(x)}{\ell(x)} - \frac{r_1(x)r_1(x)}{\ell^2(x)} & \frac{r_1(x)}{\ell(x)} - \frac{r_1(x)\ell(x)}{\ell^2(x)} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\ell(x)} \\ -\frac{r_1(x)}{\ell^2(x)} \end{pmatrix} \\ &= \kappa \left[ \frac{r_2(x)}{\ell^2(x)} - \frac{r_1^2(x)}{\ell^3(x)} - \frac{r_1(x)r_1(x)}{\ell^3(x)} + \frac{r_1^2(x)\ell(x)}{\ell^4(x)} \right] \\ &= \kappa \left[ \frac{r_2(x)\ell^2(x) - r_1^2(x)\ell(x) - r_1^2(x)\ell(x) + r_1^2(x)\ell(x)}{\ell^4(x)} \right] \\ &= \kappa \left[ \frac{r_2(x)\ell^2(x) - r_1^2(x)\ell(x)}{\ell^4(x)} \right] \\ &= \kappa \left[ \frac{r_2(x)\ell(x) - r_1^2(x)}{\ell^3(x)} \right]. \end{aligned}$$

La preuve du Théorème 3.2 est établie.  $\square$

# Chapitre 4

## Conclusion et perspectives

### 4.1 Conclusion

Les travaux réalisés dans cette thèse constituent un prolongement des travaux réalisés par Guessoum et al. (2008, 2010, 2012) sur le comportement asymptotique de l'estimateur non paramétrique de la fonction de régression. Les travaux antérieurs à ceux de cette thèse traitent le cas de données censurées à droite indépendantes et identiquement distribuées ainsi qu'alpha-mélangeantes. Certains processus dans la pratique ne sont ni indépendants ni fortement mélangeants mais associés, comme c'est le cas dans le domaine de la finance ou la fiabilité des systèmes.

La structure de dépendance étudiée dans cette thèse, appelée association, est caractérisée par la positivité des covariances. L'estimateur utilisé est celui de Nadaraya-Watson, légèrement modifié pour être adapté au modèle étudié.

Après avoir présenté quelques rappels et un bref historique sur les travaux réalisés dans ce domaine, nous avons établi la convergence uniforme presque sûre de l'estimateur tout en précisant sa vitesse, ainsi que la normalité asymptotique.

Des simulations ont été réalisées pour illustrer les performances de l'estimateur à taille finie ainsi que la construction des intervalles de confiance.

Les travaux réalisés dans le cadre de cette thèse ont fait l'objet de trois communications dans des congrès, et une publication parue en 2016 dans le journal *Statistical Papers*, (voir référence [51]) , possédant un *impact factor* égal à 0.781 et cité dans *Web of Science*.

## 4.2 Perspectives

Le travail que nous avons présenté dans cette thèse, nous permet d'envisager plusieurs pistes de recherche, nous en citons quelques-uns :

- Les résultats établis au deuxième et troisième chapitre pour la consistance et la normalité asymptotique peuvent être étendus au cas de données fonctionnelles ( lorsque la covariable est dans un espace de dimension infinie)

- Il serait intéressant d'établir un résultat de type Berry-Esseen pour la fonction de régression dans un modèle de données incomplètes et associées.

- Derrar et al.(2015) ont étudié la convergence presque sûre et la normalité asymptotique d'un estimateur robuste de régression fonctionnelle pour des données i.i.d. tronquées à gauche. On pourrait étudier cet estimateur pour des données censurées à droite et associées.

## Bibliographie

- [1] Alam, K., Saxena, K. M. L. (1981). Positive dependence in multivariate distributions. *Comm. Statist. A* 10 :1183–1196.
- [2] Ango Nze, P., Bühlmann, P., Doukhan, P. (2002). Weak dependence beyond mixing and asymptotics for nonparametric regression. *Ann. Statist.*, 30(2) :397–430.
- [3] Bagai, I., Prakasa Rao, B. L. S. (1995). Kernel-type density and failure rate estimation for associated sequences. *Ann. Inst. Statist. Math.* 47 : 253–266.
- [4] Barlow, R. E., Proschan, F. (1975). Statistical theory of reliability and life testing. Holt, Rinehart and Winston, Inc., New York. Probability models, International Series in Decision Processes, Series in Quantitative Methods for Decision Making.
- [5] Birkel, T. (1988a). Moment bounds for associated sequences. *Ann. Probab.* 16 (3) : 1184–1193.
- [6] Birkel, T. (1988b). A note on the strong law of large numbers for positively dependent random variables. *Statist. Probab. Lett.*, 7(1) : 17–20.
- [7] Birkel, T. (1988c). On the convergence rate in the central limit theorem for associated processes. *Ann. Probab.* 16 (4) : 1685–1698.
- [8] Bosq, D. (1996). Limit theorems for Banach-valued autoregressive processes. Applications to real continuous time processes. *Bull. Belg. Math. Soc. Simon Stevin*, 3(5) :537–555.
- [9] Bosq, D. (1998). Nonparametric statistics for stochastic processes, volume 110 of Lecture Notes in Statistics. *Springer-Verlag, New York, second edition. Estimation and prediction.*
- [10] Bozorgnia, A., Patterson, R. F., Taylor, R.L.(1996). Limit theorems for negatively dependent random variables. *Proceedings of the First world congress of Nonlinear Analysis, Tampa, Florida, August, 19–26.*
- [11] Blum, J.R., Hanson, D.L., Koopmans, L.H.(1963). On the strong law of large numbers for a class of stochastic processes. *Z. Wahrsch. verw. Gebiete* 2 :1–11.
- [12] Bulinski, A. V. (1996). On the convergence rates in the central limit theorem for positively and negatively dependent random fields. In : bragimov, I. A. and Zaitsev, A. Yu. (eds.), *Probab. Theory and Math. Statist., Gordon and Breach*, pp. 3–14.
- [13] Bulinski, A., Shabanovich, E. (1998) Asymptotical behavior of some functionals of positively and negatively dependent random fields. *Fund. Appl. Math.* 4, 479–492.
- [14] Bulinski, A., Suquet, C. (2001). Normal approximation for quasi-associated random fields. *Statist. Probab. Lett.*, 54(2) :215–226.
- [15] Bulinski, A., Shashkin, A. (2007) Limit theorems for associated random fields and related systems. Vol. 10. Advanced series on statistical science and applied probability. World Scientific, Singapore.

- [16] Burton, R.M., Dabrowski, A. R., Dehling, H. (1986). An invariance principle for weakly associated random vectors. *Stochastic Processes and their applications* 23 :301–306.
- [17] Cai, Z., Roussas, G.G. (1998) Kaplan-Meier estimator under association. *J Multivariate Anal* 67 :318–348.
- [18] Cai, Z., Roussas, G.G. (1999a). Weak convergence for a smooth estimator of a distribution function under association. *Stochastic Anal. Appl.* 17 : 145–168.
- [19] Cai, Z., Roussas, G.G. (1999b). Berry-Esseen bounds for smooth estimator of a distribution function under association. *J. Nonparametric Statist.* 10 : 79–106.
- [20] Carbonez, A., Györfi, L., Van der Meulin, E.C. (1995). Partition-estimate of a regression function under random censoring. *Statist. Decisions*, 1321–37.
- [21] Chaubey, Y.P., Dewan, I., 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., Lo, S-H. (1997) On the rate of uniform convergence of the product-limit-estimator : strong and weak laws. *Ann. Stat.* 25 :1050–108.
- [23] Deheuvels, P., Einmahl J (2000) Functional limit laws for the increments of Kaplan-Meier product limit processes and applications. *Ann. Probab.* 28 :1301–1335.
- [24] Devroye, L. P. (1978). The uniform convergence of the Nadaraya-Watson regression function estimate. *The Canadian J. of Statist.*, 62 :179–191.
- [25] 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.
- [26] Doukhan, P. (1994) Mixing : Properties and examples. Lecture Notes in Statistics, 85, *Springer-Verlag, New York*.
- [27] Doukhan, P., Louhichi S (1999) A new weak dependence condition and applications to moment inequalities. *Stochastic Process Appl* 84 :313–342.
- [28] Doukhan, P., Louhichi, S. (2001). Functional estimation of a density under a new weak dependence condition. *Scandinavian Journal of Statist.* 28 : 325–341.
- [29] Doukhan, P., Neumann, M. (2007) Probability and moment inequalities for sums of weakly dependent random variables, with applications. *Stochastic Process Appl* 117 :878–903.
- [30] Doukhan, P., Neumann, M. (2008) The notion of  $\psi$ -weak dependence and its applications to bootstrapping time series, *Prob Surveys* 5 :146–168
- [31] El-Ghouch A, Van Keilegom I (2008) Non-Parametric Regression with Dependent Censored Data, *Scand J Stat* 35 :228–247

- [32] Esary J, Proschan F, Walkup D (1967) Association of random variables with applications., *Ann Math Stat* 38 :1466–1476
- [33] Ferrani Y, Ould Saïd E, Tatachak A (2013) On kernel density and mode estimates for associated and censored data. *Commun Stat Theory Methods* 45 :1853–1862
- [34] Fortuin, C.M., Kasteleyn, P.W. and Ginibre, J. (1971) Correlation inequalities on some partially ordered sets. *Comm Math Phys* 22 :89–103
- [35] Gu, M., Lai, T.L. (1990) Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann Probab* 18 :160–189
- [36] Guessoum, Z., Ould Saïd, E. (2008) On the nonparametric estimation of the regression function under censorship model. *Statist and Decisions* 26 :159–177
- [37] Guessoum, Z., Ould Saïd, E. (2010) Kernel regression uniform rate estimation for censored data under  $\alpha$ - mixing condition. *Electronic J Stat* 4 :117–132
- [38] Guessoum, Z., 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
- [39] Guessoum, Z., Ould Saïd, E., Sadki, O., Tatachak, A. (2012). A note on the Lynden-Bell estimator under association. *Statist. probab. Lett.* 82 : 1994–2000.
- [40] Györfi, L., Kohler, M., Krzyzak, A., Walk, H. (2002) A Distribution-Free Theory of Nonparametric Regression. *Springer-Verlag, New York*
- [41] Härdle, W. (1990) Applied Nonparametric Regression. *Cambridge University Press, Cambridge.*
- [42] Ibragimov, I.A. (1959). Some limit theorems for stochastic processes stationary in the strict sense. *Dokl. Akad. Nauk SSSR* 125 :711–714.
- [43] Joag-Dev, K., Proschan, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* 11 : 286–295.
- [44] Kaplan, E.L., Meier P (1958) Nonparametric estimation from incomplete observations. *J Amer Statist Assoc* 53 :457–481.
- [45] Koul, H., Susarla, V., van Ryzin, J. (1981) Regression analysis with randomly right censored data. *Ann Stat* 9 :1276–1288.
- [46] Kolmogorov, A.N., Rozanov, Yu.A. (1960). On strong mixing conditions for stationary Gaussian processes. *Theor. Probab. Appl.* 5 :204–208.
- [47] Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* 37 : 1137–1153.
- [48] Liang, H.Y., Li, D. L., Qi, Y.C. (2009), Strong convergence in nonparametric regression with truncated dependent data, *J. Multivariate Anal.*, 100, 162–174.
- [49] Louhichi, S. (2000) Weak convergence for empirical processes of associated sequences. *Ann Inst Henri Poincaré, Probab Stat* 36 :547–567.

- [50] Masry, E. (2002). Multivariate probability density estimation for associated processes : strong consistency and rates. *Statist. Probab. Lett.* 58 : 205–219.
- [51] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality, *Stoch. Process. Appl.* 115 155–177.
- [52] Menni, N., 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.
- [53] Nadaraya, E.A. (1964) On estimating regression. *Theor Probab Appl* 9 :141–142.
- [54] Newman, C. M. (1980). Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.* 74 : 119–128.
- [55] Newman, C. M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables, in :Tong, Y. L. (Ed.) Inequalities in statistics and probability, IMS Lecture Notes-Monograph Ser., vol. 5 *Inst. Math. Statist., Hayward, CA* : 127–140.
- [56] Oliveira, P.E. (2012). Asymptotics for associated random variables. *Springer-Verlag Berlin Heidelberg*. Heidelberg.
- [57] Ould-Saïd, E., Lemdani, M. (2006). Asymptotic properties of a nonparametric regression function estimator with randomly truncated data. *Ann. Institut. Statist. Math.*, 58, 357–378.
- [58] Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33 : 1065–1076.
- [59] Peterson, A. V. (1977). Expressing the Kaplan-Meier estimator as a function of subsurvival function. *J. Amer. statist. Assoc.*, 72 : 854–858.
- [60] Pitt, L. D. (1982). Positively correlated normal variables are associated. *Ann. Probab.* 10 (3) : 496–499.
- [61] Rio, E. (2000). Théorie asymptotique des processus aléatoires faiblement dépendants, *Springer ISMAI*.
- [62] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* 42 : 43–47.
- [63] Rao, C. R. (1965), *Linear Statistical Inference and its Applications*, New York : John Wiley and Sons.
- [64] Roussas, G.G. (1988). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann. Inference* 18 : 135–149.
- [65] Roussas, G.G. (1990). Nonparametric regression estimation under mixing conditions. *Stochst. Process. Applicat.* 36 : 107–116.
- [66] Roussas, G.G. (1991). Kernel estimates under association : strong uniform consistency. *Statist. Probab. Lett.* 12 : 393–403.

- [67] Roussas, G.G. (1999). Positive and negative dependence with some statistical applications. In *Asymptotics, Nonparametrics and Time Series* (S. Ghosh, Ed.), 757–788. *Marcell Dekker, Inc., New York*.
- [68] Roussas, G.G. (2000). Asymptotic normality of the kernel estimate of a probability density function under association. *Statist. Probab. Lett.* 50 : 1–12.
- [69] Roussas, G.G. (2001). Fixed design regression under association, *Lecture Notes-Monograph Series.* 71–89.
- [70] Schimek, M. G. (2000). *Smoothing and Regression : Approaches, computation and Application.* John Wiley & Son, Inc, *New York*.
- [71] Shashkin, A. P. (2002). Quasi-associatedness of a gaussian system of random vectors. *Russ. Math. Surv.* 57 : 1243–1244.
- [72] Shorack, G. R., Wellner, J.A. (1986). *Empirical processes with applications to statistics.* Wiley, *New York*.
- [73] Silverman, B.W. (1986). *Estimation for statistics and data analysis.* Monographs on Statist. and Appl. Probab. Chapman and Hall. London. MR0848134
- [74] Stone, J. C. (1977). Consistence nonparametric regression. *Ann. Statist.* 5.4, 595–645.
- [75] Stute, W. (1993) Consistent Estimation Under Random Censorship When Covariables are Present. *J of Multivar Anal* 1 :89–103.
- [76] Tsybakov, A. B. (2003). Introduction à l'estimation non paramétrique. *Mathématiques et applications*, 41, *Springer*.
- [77] Van Der Vaart, A. W ., Wellner, J.A. (1996), *Weak Convergence and Empirical Process,* *New York : Springer*.
- [78] Volkonskii, V. A., Rozanov, Y. A. (1959). Some limit theorems for random functions. I. *Theor. Probability Appl.* 4, 178–197.
- [79] Wand, M.P., Jones, M.C. (1995) *Kernel Smoothing.* *Chapman & Hall, London*.
- [80] Watson, G.S. (1964) Smooth regression analysis. *Sankhya Series A* 26 :359–372.
- [81] Wied, D., Weißbach, R. (2012) Consistency of the kernel density estimator : asurvey. *Stat Papers* 53 :1–21.