

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE



UNIVERSITÉ LES FRÈRES MENTOURI CONSTANTINE  
FACULTÉ DES SCIENCES EXACTES  
DÉPARTEMENT DE MATHÉMATIQUES

N° d'ordre : .....

N° de série : .....

## THÈSE

PRÉSENTÉE POUR L'OBTENTION DU DIPLÔME DE DOCTORAT EN SCIENCES  
EN MATHÉMATIQUES

« Estimation non paramétrique du mode conditionnel  
dans un modèle de censure »

Par

Lamia AOUICHA

OPTION

Probabilités et Statistique

Devant le jury :

|                     |                 |               |        |   |
|---------------------|-----------------|---------------|--------|---|
| Président           | M <sup>r</sup>  | Z. MOHDEB     | Prof.  | École polytechnique de Constantine        |
| Directrice de thèse | M <sup>me</sup> | F. MESSACI    | Prof.  | Université Salah Boubnider, Constantine 3 |
| Examinatrice        | M <sup>me</sup> | N. NEMOUCHI   | Prof.  | Université frères Mentouri, Constantine 1 |
| Examinatrice        | M <sup>me</sup> | O. SADKI      | Prof.  | U.S.T.H.B., Alger                         |
| Examinatrice        | M <sup>me</sup> | M. CHIKHI     | M.C.A. | Université frères Mentouri, Constantine 1 |
| Examinatrice        | M <sup>me</sup> | S. KHARFOUCHI | M.C.A. | Université Salah Boubnider, Constantine 3 |

Soutenue le : 13 juin 2019

À mes chères parents  
À la mémoire de mon beau père  
À ma petite famille  
À tous ceux qui me sont proches

# Remerciements

Au terme de ce travail, je tiens à remercier vivement ma directrice de thèse madame Fatiha Messaci, pour leur soutien constant durant ces années de travail.

Vous avez madame, réussie à me transmettre le plaisir de la recherche et à mener à bien ce travail et je me considère comme particulièrement chanceuse d'avoir pu travailler sous votre direction.

J'adresse mes vifs remerciements à monsieur le professeur Z. Mohdeb d'avoir accepté de présider le jury. Je tiens à remercier les professeures N. Nemouchi et O. Sadki qui ont bien accepté d'examiner et de juger ce travail.

Je veux exprimer ma reconnaissance à madame M. Chikhi, qui a acceptée sans hésitation d'examiner ce travail bien qu'elle soit occupée.

Je suis très heureuse que Madame S. Kharfouchi ait accepté de faire partie de mon jury. Je tiens sincèrement à la remercier.

Mes remerciements vont à tous mes amies ainsi qu'à l'ensemble de mes camarades du Laboratoire LAMASD.

Je tiens enfin à remercier tous les membres de ma famille et toutes les personnes proches qui m'ont apporté leur soutien durant ces années.

## Article et communications

1. Aouicha, L., Messaci, F., 2019. Kernel estimation of the conditional density under a censorship model, *Statistics & Probability Letters* 145, 173–180.
2. Aouicha, L. et Messaci, F. Taux de convergence presque complète du mode conditionnel dans un modèle de censure mixte. CISAA2014. Constantine, 30 novembre et 1 décembre 2014.
3. Aouicha, L. et Messaci, F. Taux de convergence en moyenne quadratique d'un estimateur du mode conditionnel dans un modèle de censure mixte. Congrès des Mathématiciens Algériens, CMA 2016, Batna, 08-09 novembre 2016.

# Table des matières

|  |           |
|--|-----------|
| <b>Introduction</b>  | <b>4</b>  |
| <b>1 L'estimation à noyau des fonctions de densité et du mode conditionnels</b>  | <b>10</b> |
| 1.1 Cas des données complètes . . . . .  | 10        |
| 1.1.1 Estimateur à noyau de la densité . . . . .   | 10        |
| 1.1.2 Estimation de la densité et du mode conditionnels . .  | 13        |
| 1.2 Cas des données censurées . . . . .  | 14        |
| 1.2.1 Exemples de censure . . . . .  | 15        |
| 1.2.2 Cas de la censure à droite . . . . .   | 19        |
| 1.2.3 Cas de la censure mixte . . . . .  | 20        |
| <b>2 Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle dans un modèle de censure mixte</b> | <b>24</b> |
| 2.1 Estimation de la densité conditionnelle . . . . .  | 24        |
| 2.2 Hypothèses . . . . .   | 28        |
| 2.3 Résultats et preuve . . . . .  | 29        |
| <b>3 Estimation du mode conditionnel dans un modèle de censure mixte</b>   | <b>42</b> |
| 3.1 Convergence presque complète de l'estimateur à noyau de la densité conditionnelle . . . . .                                  | 42        |

|          |   |           |
|----------|---|-----------|
| 3.2      | Taux de convergence presque complète de l'estimateur du mode conditionnel . . . . . | 52        |
| <b>4</b> | <b>Simulation</b>   | <b>55</b> |
| 4.1      | Estimateur de la densité conditionnelle . . . . .                                   | 55        |
| 4.2      | Estimateur du mode conditionnel . . . . .   | 60        |
|          | <b>Perspectives</b>   | <b>74</b> |
|          | <b>Bibliographie</b>  | <b>75</b> |

# Introduction

L'estimation non paramétrique connaît un grand essor chez de nombreux auteurs et dans différents domaines. Celle-ci ne fait pas d'hypothèse sur la forme de la loi de probabilité qui régit le phénomène sous étude lorsque nos connaissances sur le modèle ne sont pas précises, ce qui est souvent le cas dans la pratique. Dans cette situation, il est procédé à l'estimation d'une des fonctions décrivant le modèle ou le lien entre des variables aléatoires comme la fonction de répartition, la densité, la densité conditionnelle ou le mode conditionnel.

Dans ce travail, nous considérons d'abord l'estimation non paramétrique de la fonction de la densité conditionnelle.

L'importance d'une telle étude est mise en évidence par l'application de tels estimateurs à des données réelles, par exemple, dans Hyndman et al. (1996) et récemment dans Arora et Taylor (2016); Wen and Wu (2017).

En dimension finie, en ce qui concerne l'estimateur à noyau de la fonction densité conditionnelle, Roussas (1968) fut le premier à établir des propriétés asymptotiques pour des données markoviennes, ainsi que sa convergence en probabilité.

Des propriétés de convergence sur des estimateurs de la densité conditionnelle ont été obtenues entre autres par Bosq (1971), Bosq (1973), Delecroix (1975) et Chahboun (1984). Pour des références plus récentes sur le sujet, Youndje (1993) s'est intéressée à l'étude de la densité conditionnelle pour des données

complètes indépendantes, Laksaci et Yousfate (2002) ont établi, pour un processus markovien stationnaire, la convergence en norme  $L^p$  pour l'estimateur à noyau de la densité conditionnelle. Hyndman et al. (1996) ont étudié les propriétés quadratiques de l'estimateur à noyau. Dans Hyndman et Yao, (2002) sont considérés les méthodes basées sur le modèle loglinéaire local et les polynômes locaux sous contraintes, un test de symétrie pour la densité conditionnelle y est proposé et étudié. Hall et al. (2004) ont montré que la densité conditionnelle joue un rôle clef en statistique appliquée et particulièrement en économie. Pour Fan et Yim (2004), une densité conditionnelle offre le résumé le plus informatif de la relation entre variable dépendante et indépendante, enfin Efromovich (2007), expose la théorie minimax d'une densité conditionnelle, il montre en outre que l'estimateur des séries orthogonales atteint la vitesse minimax et que ce résultat reste vrai même si  $X$  et  $Y$  sont indépendantes.

Par ailleurs, la densité conditionnelle va permettre de prévoir la variable aléatoire  $Y$  à l'aide du mode conditionnel. Le mode conditionnel, de par son importance dans le domaine de prévision non paramétrique, était la motivation de beaucoup d'auteurs, les propriétés de convergence et de normalité asymptotiques ont été établies par Samanta et Thavasncswaran (1990) dans le cadre de données indépendantes et identiquement distribuées, alors que des conditions de convergence dans le cas de données  $\phi$ -mélangeantes ont été établies par Collomb et al. (1987), dans le cas de données  $\alpha$ -mélangeantes par Ould Saïd (1993) et dans le cas de données ergodiques par Rosa (1993); Ould Saïd (1997). De leur coté Quintela-Del-Rio et Vieu (1997) ont estimé le mode conditionnel comme étant le point annulant la dérivée d'ordre un de l'estimateur de la densité conditionnelle et établi la convergence presque complète de cet estimateur. Berlinet et al. (1998) ont donné des théorèmes généraux de normalité asymptotique des estimateurs du mode conditionnel,



indépendamment de la structure de dépendance des données et ils l'ont appliqué au cas d'un processus stationnaire  $\alpha$ -mélangeant du mode conditionnel. Louani et Ould Saïd (1999) ont établi la normalité asymptotique dans le cas de données fortement mélangeantes. Mehra et al. (2000) ont établi la loi du logarithme itéré, la convergence uniforme presque sûre sur un ensemble compact, et la normalité asymptotique de l'estimateur du voisinage le plus proche lissé de la fonction du mode conditionnel, Rossi (2004) a utilisé le mode conditionnel, dans le domaine des hautes technologies, pour décrire un procédé de dépollution biologique. Enfin, citons aussi les travaux de Einbeck et Tutz (2006) et Matzner (1996) pour avoir plus d'informations sur l'estimateur du mode conditionnel comme étant un type de prévision.

En dimension infinie, le mode conditionnel connaît, depuis peu, un intérêt croissant, malgré le peu de résultats existants. Dans ce contexte des données fonctionnelles, les premiers travaux ont été réalisés par Ferraty et al. (2006). Ils ont montré, sous des conditions de régularité de la densité conditionnelle, la convergence presque complète de l'estimateur à noyau qu'ils ont introduit pour la densité conditionnelle, ainsi que celle du mode conditionnel, et établi la vitesse de convergence. Une application de leurs résultats aux données issues de l'industrie agro-alimentaire est présentée.

Depuis cet article, une littérature abondante s'est développée sur l'estimation de la densité conditionnelle et ses dérivées, notamment afin de l'utiliser pour estimer le mode conditionnel. En effet, en considérant des observations  $\alpha$ -mélangeantes, Ferraty et al. (2005) ont établi la convergence presque complète d'un estimateur à noyau du mode conditionnel défini par la variable aléatoire maximisant la densité conditionnelle. Alternativement, Ezzahrioui et Ould Saïd (2005, 2006) ont estimé le mode conditionnel par le point qui annule la dérivée de l'estimateur à noyau de la densité conditionnelle. Ces derniers se sont concentrés sur la normalité asymptotique de l'estimateur

proposé dans les deux cas (i.i.d. et mélangeantes). La précision des termes dominants de l'erreur quadratique de l'estimateur à noyau de la densité conditionnelle a été obtenue par Laksaci (2007).

Mais, il est bien connu que dans plusieurs situations on ne peut pas observer la variable d'intérêt. C'est le cas lorsque  $Y$  est censurée à droite, ce qui signifie que l'observation ne correspond pas à la vraie réalisation de  $Y$  mais qu'elle lui est inférieure.

Ce type de censure est souvent rencontré dans la pratique comme dans le suivi médical ou dans les tests de vie d'ingénierie. Parmi les travaux sur la densité conditionnelle et le mode conditionnel en présence de données censurées, nous citons dans le cas des données indépendantes et identiquement distribuées, Ould Saïd et Cai (2005) qui ont défini un nouveau estimateur de la densité conditionnelle et ont établi la convergence presque sûre uniforme avec taux de l'estimateur proposé dans un modèle de censure à droite.

Salha et El Shekh Ahmed (2009) ont proposé un estimateur modifié du mode conditionnel, ils ont établi la normalité et la consistance asymptotique de l'estimateur proposé dont ils ont testé l'efficacité par une étude de simulation et une application à des données réelles.

Dans ce contexte de censure, Khardani et al. (2010) ont défini un nouveau estimateur du mode conditionnel et ils ont établi sa convergence presque sûre et sa normalité asymptotique.

Khardani et Semmar (2014) ont employé une version récursive de l'estimateur de la densité conditionnelle pour laquelle ils ont étudié la consistance uniforme et la normalité asymptotique.

Malheureusement, le mécanisme de censure complique le calcul de la variance asymptotique et nous ne connaissons aucun travail donnant l'erreur

moyenne quadratique asymptotique de l'estimateur de la densité conditionnelle à noyau dans un modèle de censure à droite.

Cependant, ce résultat peut facilement être déduit du théorème 2 établi au chapitre 2. Ce dernier théorème est obtenu sous le modèle plus général de censure mixte. Ce qui signifie que l'observation correspond non pas à une réalisation de la variable d'intérêt  $Y$  mais qu'il existe deux variables aléatoires  $R$  et  $L$  telles que la donnée correspond à l'observation de  $\max(\min(Y, R), L)$  et d'un indicateur de censure exprimant laquelle des trois variables latentes précédentes est réellement observée. Des exemples de l'application de ce modèle peuvent être trouvés dans Patilea et Rolin (2006) et Morales et al. (1991). Sous ce modèle de censure, Messaci (2010); Kebabi et al. (2011); Kebabi et Messaci (2012) ont étudié des estimateurs non paramétriques de la régression, tandis que Boukeloua (2015) a établi la convergence en moyenne quadratique, avec taux, pour l'estimateur à noyau de la densité (inconditionnelle) dont nous nous sommes inspirées pour montrer le théorème 2.

Dans cette thèse, notre contribution consiste à introduire et étudier des propriétés asymptotiques d'estimateurs non paramétriques de la densité et du mode conditionnels dans le cas d'une variable réponse soumise à la censure mixte. La méthode d'estimation choisie est celle bien connue et usuelle du noyau. Ce travail est mené comme suit.

## Organisation de la thèse

**Chapitre 1 :** Nous y rappelons les différents modèles de censure ainsi que les estimateurs à noyau de la densité et du mode conditionnels pour une variable réponse totalement observée. Puis nous présentons des estimateurs non paramétriques de la fonction de survie pour le

modèle d'une censure unique (à gauche ou à droite) et pour celui de la censure mixte tout en rappelant quelques uns de leurs résultats de convergence que nous utilisons dans les chapitres ultérieurs.

**Chapitre 2 :** Nous y montrons la convergence en moyenne quadratique, avec taux, de l'estimateur à noyau de la densité conditionnelle que nous construisons dans le cas de la censure mixte.

**Chapitre 3 :** Nous y traitons la convergence presque complète, avec taux, de l'estimateur de la densité conditionnelle, introduit au chapitre précédent, ainsi que celle de l'estimateur du mode conditionnel que nous en déduisons.

**Chapitre 4 :** Nous y présentons une étude de simulation afin d'illustrer les performances, à taille finie, des estimateurs introduits et étudiés théoriquement aux chapitres antérieurs.

**Annexe :** Nous y présentons quelques définitions et résultats utiles dans la thèse.

# Chapitre 1

## L'estimation à noyau des fonctions de densité et du mode conditionnels

### 1.1 Cas des données complètes

#### 1.1.1 Estimateur à noyau de la densité

L'estimateur à noyau de la densité a été introduit par Rosenblatt (1956), il s'écrit comme un produit de convolution entre un noyau  $K$  convenablement rééchelonné et la fonction de répartition empirique. Il a été introduit comme suit.

Soit l'échantillon de  $n$  variables aléatoires réelles (v.a.r.) indépendantes et identiquement distribuées (i.i.d.)  $X_1, \dots, X_n$  de même loi que  $X$ . Un estimateur naturel de la fonction de répartition  $F$ , de la variable aléatoire  $X$  est la fonction de répartition empirique, définie pour tout  $x \in \mathbb{R}$  par :

$$F_n(x) := \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq x\}}.$$

Bien que la fonction de répartition détermine entièrement la loi d'une variable aléatoire, elle ne donne pas d'information visuelle concernant la symétrie, la multimodalité ou l'aplatissement de la loi. C'est la fonction de densité  $f_X$ , quand elle existe, qui permet de mieux le faire. Une solution intuitive pour estimer cette dernière a été proposée par Rosenblatt (1956).

Pour  $h > 0$  assez petit on a :

$$f_X(x) \approx \frac{F(x+h) - F(x-h)}{2h},$$

en remplaçant  $F$  par son estimateur  $F_n$ , nous obtenons l'estimateur suivant de  $f_X$  :

$$\begin{aligned} f_n(x) &:= \frac{F_n(x+h) - F_n(x-h)}{2h} \\ &= \frac{1}{2nh} \sum_{i=1}^n 1_{\{x-h \leq X_i \leq x+h\}} \\ &= \frac{1}{nh} \sum_{i=1}^n K_0\left(\frac{x - X_i}{h}\right), \end{aligned}$$

où  $K_0(u) = \frac{1}{2}1_{\{|u| \leq 1\}}$ .

Ceci mène à généraliser cet estimateur comme suit

$$f_n(x) := \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \quad (1.1)$$

où  $K$  est une fonction intégrable de  $\mathbb{R} \rightarrow \mathbb{R}$ , telle que  $\int K(u)du = 1$ , c'est l'estimateur à noyau de  $f_X$ .

La fonction  $K$  est le noyau et le paramètre  $h$  est la fenêtre.

Voici quelques exemples de noyaux usuels :

- Noyau rectangulaire :  $K(u) = \begin{cases} \frac{1}{2}, & \text{si } \{|u| \leq 1\}, \\ 0, & \text{si non.} \end{cases}$

- Noyau triangulaire :  $K(u) = \begin{cases} (1 - |u|), & \text{si } \{|u| \leq 1\}, \\ 0, & \text{si non.} \end{cases}$
  - Noyau d'Epanechnikov ou parabolique :  $K(u) = \begin{cases} \frac{3}{4}(1 - u^2), & \text{si } \{|u| \leq 1\}, \\ 0, & \text{si non.} \end{cases}$
  - Noyau quadratique :  $K(u) = \begin{cases} \frac{15}{16}(1 - u^2)^2, & \text{si } \{|u| \leq 1\}, \\ 0, & \text{si non.} \end{cases}$
  - Noyau gaussien :
- $$K(u) = \frac{1}{\sqrt{2\pi}} \exp(-u^2/2).$$

- Noyau de Silverman :

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-|u|}{\sqrt{2}}\right) \sin\left(\frac{|u|}{\sqrt{2}} + \frac{\pi}{4}\right).$$

En général, la fenêtre  $h$  est prise comme une suite  $(h_n)_{n \geq 1}$  qui tend vers zéro lorsque  $n$  tend vers l'infini.

L'estimateur  $f_n$ , de la fonction densité a été largement étudié dans la littérature. Rosenblatt (1956) donna dans son article l'erreur quadratique moyenne relative à l'estimation de la densité dans le cas d'observations univariées indépendantes et identiquement distribuées et où le noyau uniforme  $K = \frac{1}{2}1_{[-1,1]}$  est utilisé (1 désigne la fonction indicatrice). Parzen (1962) généralisa ce résultat en considérant une classe très vaste de noyaux et établit aussi la normalité asymptotique, puis le cas multivarié fut traité par Cacoullos (1966).

Rosenblatt (1971) a donné des conditions de convergence en moyenne quadratique de l'estimateur de la densité. Deheuvels (1974) a, en outre, étudié les convergences ponctuelle et uniforme presque sûre.

Sa convergence uniforme faible et forte a été considérée aussi par plusieurs auteurs comme Schuster (1969), Van Ryzin (1969), Rosenblatt (1971) et Silverman (1978). La loi du logarithme itéré a été établie par Deheuvels (1991).

Rappelons que cette méthode d'estimation (à noyau) se généralise au cas de  $\mathbb{R}^p$ . Ainsi, si  $X_1, \dots, X_n$ , sont  $n$  vecteurs aléatoires i.i.d. de  $\mathbb{R}^p$ , de même densité inconnue, alors on peut l'estimer par  $f_n(x) = \frac{1}{nh_n^p} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)$ ,  $x \in \mathbb{R}^p$ .

Nous pouvons maintenant introduire l'estimateur de la densité conditionnelle à partir de l'estimation précédente et en déduire un estimateur du mode conditionnel.

### 1.1.2 Estimation de la densité et du mode conditionnels

Soit  $(X, Y)$  un couple aléatoire à valeurs dans  $\mathbb{R}^2$ , admettant une densité  $f$ . La densité marginale  $f_X$  de  $X$  est obtenue par la formule

$$f_X(x) = \int f(x, y) dy.$$

La densité conditionnelle de  $Y$  sachant  $X = x$  est définie par

$$f^x(y) = \frac{f(x, y)}{f_X(x)}.$$

Soit  $(X_1, Y_1), \dots, (X_n, Y_n)$  un  $n$  échantillon de v.a. i.i.d. de même loi que  $(X, Y)$ . L'estimateur à noyau de  $f$  est défini par

$$f_n(x, y) := \frac{1}{nhg} \sum_{i=1}^n K' \left( \frac{x - X_i}{h}, \frac{y - Y_i}{g} \right), \quad (1.2)$$

où  $K'$  est un noyau de  $\mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $h$  et  $g$  sont les fenêtres ou les paramètres de lissage.



Cette écriture peut être simplifiée en supposant qu'il existe deux noyaux  $K : \mathbb{R} \rightarrow \mathbb{R}$  et  $K_0 : \mathbb{R} \rightarrow \mathbb{R}$  tels que  $K'(x, y) = K(x).K_0(y)$  pour tout  $(x, y) \in \mathbb{R}^2$ .

Pour  $h = g$ , on obtient alors l'estimateur suivant de  $f^x$  :

$$\begin{aligned}\hat{f}_n^x(y) &:= \frac{\frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) K_0\left(\frac{y-Y_i}{h}\right)}{\frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)} \\ &= \sum_{i=1}^n W_{n,i}(x) \frac{1}{h} K_0\left(\frac{y-Y_i}{h}\right),\end{aligned}$$

où

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

est le poids de Nadaraya-Watson.

On suppose que le mode conditionnel est unique sur un compact  $S$ . On note  $\theta(x)$  ce mode, il est donné par

$$f^x(\theta(x)) = \sup_{y \in S} f^x(y).$$

Un estimateur naturel du mode conditionnel  $\theta(x)$  est alors défini comme la v.a.  $\theta_n(x)$  maximisant l'estimateur de la densité conditionnelle, autrement dit c'est la variable aléatoire  $\theta_n(x)$  solution du problème suivant :

$$\hat{f}_n^x(\theta_n(x)) = \sup_{y \in S} \hat{f}_n^x(y).$$

## 1.2 Cas des données censurées

Le problème des données incomplètes est très vaste et a suscité beaucoup d'intérêt parmi les statisticiens ces dernières années. Dans le domaine des durées de survie, les données sont souvent incomplètes à cause de phénomène de la censure.

Par définition, le temps de survie d'un individu est dit censuré, lorsque sa

valeur exacte n'est pas observée, seules des bornes supérieures et (ou) inférieures pour cette valeur sont disponibles.

La censure peut se manifester pour différentes raisons : l'événement d'intérêt n'est pas survenu au moment de l'analyse, un sujet peut être perdu de vue avant d'avoir expérimenté l'événement d'intérêt, un événement concurrent peut être survenu avant l'événement d'intérêt, un sujet peut être exclu de l'étude sans avoir expérimenté l'événement d'intérêt, etc.

Il existe plusieurs types de censure dont la censure à droite, la censure à gauche et la censure par intervalle. La censure à droite est la forme de censure la plus commune dans les études médicales.

Un ouvrage qui fait autorité sur le sujet est le livre de Andersen et al. (1992). On peut citer aussi Survival Analysis écrit par Klein et Goel (1991).

### 1.2.1 Exemples de censure

**Censure à droite** Une durée de vie aléatoire  $X$  est dite censurée par une variable aléatoire de censure  $C$  si on observe parfois  $C$  au lieu de  $X$ . L'information donnée par  $C$  sur  $X$  est  $X > C$ .

Un exemple classique de censure droite est celui où l'étude porte sur la durée de survie  $X$  de patients atteints d'une certaine maladie. Pour les patients perdus de vue au bout du temps  $C$  alors qu'ils étaient encore vivants,  $C$  censure  $X$  à droite puisque, pour eux,  $X$  est inconnue mais supérieur à  $C$  :  $X > C$ .

*Exemple 1.* Un exemple typique est celui où l'événement d'intérêt est le décès d'un patient malade et la durée d'observation est une durée totale d'hospitalisation (on n'a plus de nouvelles après l'hospitalisation) .

On peut aussi observer ce genre de phénomène dans des études de fiabilité quand la panne d'un appareil ne permet pas de poursuivre l'observation du

composant objet de notre étude. Pour ce type de censure, tout ce que l'on sait est que la vraie durée est supérieure à la durée observée.

**Censure à gauche** Il y a censure à gauche lorsque la durée de survie est inférieure à la durée observée.

*Exemple 2.* En fiabilité, l'exemple d'une telle situation est celui d'un composant électronique monté en parallèle avec un ou plusieurs autres composants. Une panne de ce composant n'entraîne pas nécessairement l'arrêt du système : le système peut continuer à fonctionner (bien que pouvant présenter certaines anomalies) jusqu'à ce que cette panne soit détectée (par exemple lors d'un contrôle ou en cas de l'arrêt du système). La durée observée pour ce composant est alors censurée à gauche.

**Censure double** Il y a des situations où les données d'un même échantillon peuvent être censurées soit à droite, soit à gauche.

*Exemple 3.* Leiderman et al. (1973) ont étudié l'âge auquel les enfants d'une communauté africaine apprennent à accomplir certaines tâches. Au début de l'étude, certains enfants savaient déjà effectuer les tâches étudiées, on sait seulement alors que l'âge où ils ont appris est inférieur à leur âge à la date du début de l'étude (censure à gauche). A la fin de l'étude, certains enfants ne savaient pas encore accomplir ces tâches et on sait alors seulement que l'âge auquel ils ont appris est supérieur à leur âge à la fin de l'étude (censure à droite).

Dans cet exemple, on trouve dans un même échantillon des données censurées à gauche aussi bien que des données censurées à droite.

Selon le processus qui génère la censure, on est en présence de censure double Turnbull (1974) ou mixte Patilea et Rolin (2006).

**Censure mixte** Nous disons qu'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. Dans le modèle I de Patilea et Rolin (2006), au lieu d'observer un échantillon de  $Y$  on observe un échantillon du couple  $(Z, A)$  où  $Z = \max(\min(Y, R), L)$  et

$$A = \begin{cases} 0, & \text{si } L < Y \leq R, \\ 1, & \text{si } L < R < Y, \\ 2, & \text{si } \min(Y, R) \leq L. \end{cases}$$

où  $L$  et  $R$  sont des variables de censure et  $A$  est l'indicateur de censure.

*Exemple 4.* Un exemple d'application de ce modèle donné par Patilea et Rolin (2006) consiste à considérer un système de trois composants  $C1$ ,  $C2$  et  $C3$ , avec  $C1$  et  $C2$  en série et  $C3$  en parallèle avec le système composé de  $C1$  et  $C2$ . Les variables  $Y$ ,  $R$  et  $L$  représentent les temps de survie de  $C1$ ,  $C2$  et  $C3$  respectivement, et nous pouvons déterminer quel composant tombe en panne en même temps que le système. Donc au lieu d'observer  $Y$ , on peut seulement observer le couple  $(Z, A)$ .

*Remarque 1.* Le modèle de Patilea et Rolin (2006) et celui de Turnbull (1974), bien que similaires, ne sont pas identiques. Dans le modèle de Turnbull (1974), on suppose que  $Y$  est indépendante du couple  $(R, L)$  et que  $P(L < R) = 1$  alors que dans le modèle de Patilea,  $Y$ ,  $R$  et  $L$  sont indépendantes, si on appelle  $Y$  la variable d'intérêt,  $L$  la variable de censure à gauche et  $R$  la variable de censure à droite. La similarité vient du fait qu'on observe  $Z = \max(\min(Y, R), L)$ .

**Censure par intervalle** Dans le cas de la censure par intervalle, on observe à la fois une borne inférieure et une borne supérieure de la durée d'intérêt, autrement dit ; au lieu d'observer  $Y$ , on observe  $C_1 < C_2$  tels que  $C_1 < Y < C_2$ . En particulier, la censure gauche peut être considérée comme une censure par un intervalle tel que  $C_1 = -\infty$ , et la censure à droite comme une censure par intervalle tel que  $C_2 = +\infty$ .

Ceci arrive dans des études de suivi médical où les patients sont contrôlés périodiquement, si un patient ne se présente pas à un ou plusieurs contrôles et se présente ensuite après que l'événement d'intérêt se soit produit. On a aussi pour ce genre d'expériences des données qui sont censurées à droite ou, plus rarement, à gauche. Un avantage de ce type est qu'il permet de représenter les données censurées à droite ou à gauche par des intervalles du type  $[a, +\infty[$  et  $[0, a]$  respectivement, ce qui permet de considérer ce modèle comme étant plus générique. Turnbull (1976) présente ce genre de censure avec plus de détails.

*Exemple 5.* Pour détecter les composants défectueux d'un processus de production industriel, on effectue des contrôles selon des dates aléatoires. Lorsqu'on constate qu'un composant est à changer, on sait seulement qu'il est tombé en panne entre les dates de deux contrôles successifs.

Un problème important en statistique est celui de l'estimation de la fonction de répartition qui décrit complètement la loi de probabilité des observations. Ce qui revient à estimer la fonction de survie qui est le complément à 1 de la fonction de répartition. L'étude d'un tel estimateur est aussi justifiée par le fait qu'il intervient explicitement dans l'expression de l'estimateur de la densité conditionnelle que nous nous proposons d'étudier avec l'estimateur du mode conditionnel.

## 1.2.2 Cas de la censure à droite

### Estimateur de Kaplan-Meier

Dans la pratique, un échantillon est souvent censuré. Dans le cas de la censure à droite, Kaplan et Meier (1958) ont proposé un estimateur de la fonction de survie  $(1 - F(x))$ , qui se réduit au complément à 1 de la fonction de répartition empirique lorsque les observations sont complètes. Soit  $X_1, \dots, X_n$  un échantillon représentant les durées d'intérêt (ces variables sont donc supposées positives), de fonction de répartition  $F$ , et  $R_1, \dots, R_n$  un échantillon représentant les temps de censure, que l'on suppose indépendants des durées d'intérêt, de fonction de répartition  $F_R$ .

Dans le modèle de censure aléatoire à droite, on observe non pas la durée d'intérêt  $X_i$  mais plutôt la plus petite des deux valeurs  $Z_i = \min(X_i, R_i)$ , ainsi que l'indicateur de censure  $\delta_i$  qui vaut 1 si la durée d'intérêt est observée, et 0 si elle est censurée, i.e.  $\delta_i = 1_{\{X_i \leq R_i\}}$ .

Pour ce genre de données la fonction de répartition  $F$  est estimée par l'estimateur introduit par Kaplan et Meier (1958), donné pour  $z < Z_{(n)}$  où  $Z_{(n)} = \max\{Z_1, \dots, Z_n\}$  par

$$F_n(z) = 1 - \prod_{i: Z_i \leq z} \left( \frac{N_n(Z_i) - 1}{N_n(Z_i)} \right)^{\delta_i}$$

avec

$$N_n(x) = \sum_{i=1}^n 1_{\{Z_i \geq x\}}.$$

Pour  $z \geq Z_{(n)}$ , il y a plusieurs conventions pour définir  $F_n(z)$  :

Soit on le définit par  $F_n(Z_{(n)})$ , ce qui fait que  $F_n$  peut ne pas être une fonction de répartition si  $Z_{(n)}$  est une donnée censurée, soit on le définit par 0, soit on le laisse non défini.

Cet estimateur coïncide avec la fonction de répartition empirique quand il n'y a pas de données censurées. Il est donc naturel que les statisticiens se

soient intéressés à étendre les résultats connus pour la fonction de répartition empirique au cas de l'estimateur de Kaplan-Meier.

La convergence presque sure uniforme et la loi du logarithme itéré ont été montrées respectivement par Winter et al. (1978) et Földes et Rejtő (1981a). Ensuite, Stute et Wang (1993) ont montré la loi forte des grands nombres dans le cas de la censure à droite. Quant à la convergence presque complète, elle a été montrée par Földes et al. (1980), avec un taux de convergence de l'ordre de  $\sqrt{\log n}/\sqrt{n}$ .

Puis en imposant la continuité de  $F$  et de  $F_R$ , Földes et Rejtő (1981b) ont amélioré le taux de convergence qui est passé à l'ordre de  $\sqrt{\log n/n}$ . Ensuite, Kitouni et al. (2015) ont signalé qu'on peut se passer de cette hypothèse de continuité en utilisant la borne exponentielle suivante.

*Théorème 1. Il existe une constante absolue  $D$  telle que, pour tout réel positif  $u$ ,*

$$P\left(\sqrt{n} \sup_{x \in \mathbb{R}} S_R(x) |F_n(x) - F(x)| > u\right) \leq 2.5 \exp(-2u^2 + Du).$$

*Démonstration.*

Voir Bitouzé et al. (1999). □

### 1.2.3 Cas de la censure mixte

Le but de cette thèse est justement d'introduire et d'étudier des estimateurs de la densité conditionnelle et du mode conditionnel quand la variable réponse est soumise à une censure mixte.

Ces estimateurs faisant intervenir l'estimateur de Patilea et Rolin, nous commençons par introduire et rappeler quelques une de ses propriétés, notamment celles qui nous seront nécessaires dans la suite.

Considérons trois variables aléatoires positives indépendantes  $Y$ ,  $L$  et  $R$  de fonctions de répartition respectives  $F_Y$ ,  $F_L$  et  $F_R$ , et de fonctions de survie respectives  $S_Y$ ,  $S_L$  et  $S_R$ , où  $Y$  représente la durée d'intérêt et  $L$  et  $R$  sont

les durées de censure à gauche et à droite respectivement.

De plus, pour toute variable aléatoire réelle  $U$ ,  $F_U$ ,  $S_U = 1 - F_U$ ,  $T_U := \sup\{t : F_U(t) < 1\}$  et  $I_U := \inf\{t : F_U(t) > 0\}$ , représente la fonction de distribution de  $U$ , la fonction de survie de  $U$ , le point terminal et le point initial du support de  $U$  respectivement.

Dans le modèle I de Patilea et Rolin (2006), au lieu d'observer un échantillon de  $Y$  on observe un échantillon du couple  $(Z, A)$  où  $Z = \max(\min(Y, R), L)$  et

$$A = \begin{cases} 0, & \text{si } L < Y \leq R, \\ 1, & \text{si } L < R < Y, \\ 2, & \text{si } \min(Y, R) \leq L. \end{cases}$$

Ce modèle considère la censure à droite et la censure à gauche comme deux phénomènes qui agissent indépendamment l'un de l'autre mais que l'un peut censurer l'autre.

### Estimation de la fonction de survie

Patilea et Rolin (2006) ont proposé d'estimer  $S_Y$ , fonction de survie de la variable d'intérêt  $Y$  comme suit :

Considérons  $H$  la fonction de répartition de  $Z$ , elle peut s'écrire comme  $\sum_{k=0}^2 H_k(t)$  où

$$H_k(t) = P(Z \leq t, A = k), \quad \text{pour } k = 0, 1, 2.$$

En notant pour toute application  $R$  de  $\mathbb{R} \rightarrow \mathbb{R}$ ,  $R(t_-)$  la limite de  $R$  à gauche



de  $t$ , lorsque cette limite existe, ces fonctions s'écrivent

$$\begin{aligned} H_0(t) &= \int_0^t F_L(u_-) S_R(u_-) dF_Y(u), \\ H_1(t) &= \int_0^t F_L(u_-) S_Y(u) dF_R(u), \\ H_2(t) &= \int_0^t \{1 - S_Y(u) S_R(u)\} dF_L(u), \end{aligned}$$

et c'est à partir de ces équations que l'estimateur est obtenu.

L'idée est de considérer dans un premier temps  $X' = \min(Y, R)$  et  $L$  dans un modèle de censure à gauche (c'est-à-dire que l'on considère une donnée complète si  $A = 0$  ou  $A = 1$  et censurée à gauche si  $A = 2$ ), et d'estimer la fonction de répartition de  $X'$ , puis l'utiliser pour estimer la fonction de répartition de la variable d'intérêt  $Y$  en considérant un modèle de censure à droite.

L'estimateur de la fonction de survie  $S_Y$  ainsi obtenu, en remplaçant à la fin les fonctions  $H_0$ ,  $H_1$  et  $H_2$  par leurs estimateurs empiriques, obtenus à partir d'un échantillon  $(Z_i, A_i)_{1 \leq i \leq n}$ , est donné par :

$$S_n(t) = \prod_{\{j: Z'_j \leq t\}} \left\{ 1 - \frac{D_{1j}}{U_{j-1} - N_{j-1}} \right\},$$

où  $(Z'_j)_{1 \leq j \leq M}$ , sont les valeurs distinctes de  $(Z_i)_{1 \leq i \leq n}$  considérées dans l'ordre croissant, et

$$\begin{cases} D_{kj} = \sum_{1 \leq i \leq n} 1_{\{Z_i = Z'_j, A_i = k\}}, \\ N_j = \sum_{1 \leq i \leq n} 1_{\{Z_i \leq Z'_j\}}, \\ U_{j-1} = n \prod_{j \leq l \leq M} \left\{ 1 - \frac{D_{2l}}{N_l} \right\} \end{cases}$$

pour  $0 \leq l \leq 2$  et  $1 \leq j \leq M$ .

*Remarque 2.* Si  $L \equiv 0$  (pas de censure à gauche),  $S_n$  se réduit à l'estimateur de Kaplan-Meier qui lui même se réduit au complément à 1 de la fonction de répartition empirique si  $R \equiv \infty$ .

En adaptant le Théorème 1 au cas de la censure à gauche et en notant par  $F_L^{(n)}$  l'estimateur de  $F_L$ , nous obtenons pour tout  $c > 0$  et  $\tau > \min(I_Y, I_R)$

$$\begin{aligned} P \left( \sup_{t \geq \tau} |(1 - S_Y(t)S_R(t))(F_L^{(n)}(t) - F_L(t))| > \sqrt{\frac{(c+1) \log n}{2n}} \right) \\ \leq 2.5 \exp \left\{ -(c+1) \log n + k \sqrt{\frac{(c+1) \log n}{2}} \right\}, \end{aligned}$$

où  $k$  est une constante.

En remarquant que

$$\exp \left\{ k \sqrt{\frac{(c+1) \log n}{2}} \right\} \leq n$$

pour  $n$  assez grand, nous obtenons

$$P \left( \sup_{t \geq \tau} |(1 - S_Y(t)S_R(t))(F_L^{(n)}(t) - F_L(t))| > \sqrt{\frac{(c+1) \log n}{2n}} \right) = O(n^{-c})$$

et puisque

$$\sup_{t \geq \tau} |(1 - S_Y(t)S_R(t))(F_L^{(n)}(t) - F_L(t))| \geq (1 - S_Y(\tau)S_R(\tau)) \sup_{t \geq \tau} |F_L^{(n)}(t) - F_L(t)|,$$

il s'ensuit que

$$P \left( \sup_{t \geq \tau} |F_L^{(n)}(t) - F_L(t)| > \frac{1}{1 - S_Y(\tau)S_R(\tau)} \sqrt{\frac{(c+1) \log n}{2n}} \right) = O(n^{-c}), \quad (1.3)$$

qui correspond à la formule (3) dans Boukeloua (2015).

Par ailleurs, en vertu du Lemme 1 dans Boukeloua (2015), nous avons :

si  $\max(I_L, I_R) < I_Y$ , alors pour tout  $c > 0$ ,  $\theta < \min(T_X, T_R)$ , il existe  $a > 0$  tel que

$$P \left( \sup_{t \leq \theta} |S_n(t) - S_Y(t)| > a \sqrt{\frac{\log n}{n}} \right) = O(n^{-c}). \quad (1.4)$$

## Chapitre 2

# Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle dans un modèle de censure mixte

Dans ce chapitre, notre but principal est d'étudier l'erreur quadratique moyenne de l'estimateur à noyau de la densité conditionnelle dans un modèle de censure mixte et de préciser sa vitesse de convergence. Les résultats de ce chapitre ont fait l'objet d'une publication dans la revue *Statistics & Probability Letters*, voir Aouicha et Messaci (2019).

### 2.1 Estimation de la densité conditionnelle

Soit  $X$  et  $Y$  deux variables aléatoires réelles. Tout au long de cette thèse, nous supposons l'existence de la densité de probabilité conjointe (par rapport à la mesure de Lebesgue)  $f$  du couple  $(X, Y)$ , nous notons  $f_X$  la densité de  $X$  et par  $f^x$  la densité de  $Y$  sachant  $X = x$ .

Nous observons un échantillon de variables aléatoires indépendantes et identiquement distribuées (i.i.d.)  $(X_i, Z_i = \max(\min(Y_i, R_i), L_i), A_i)_{1 \leq i \leq n}$  de même loi que  $(X, Z = \max(\min(Y, R), L), A)$  où  $R$  et  $L$  sont respectivement une variable de censure à droite et à gauche et sont telles que  $R, L$  et  $Y$  sont positives et indépendantes et

$$A = 1_{\{L < R \leq Y\}} + 2 \times 1_{\{\min(Y, R) \leq L\}}$$

est l'indicateur de censure.

Dans ce contexte de censure, rapellons que Patilea et Rolin (2006) ont proposé un estimateur de la fonction de survie  $S_R$  (voir Section 1.2.3) donné par

$$S_n(t) = \prod_{\{j: Z'_j \leq t\}} \left\{ 1 - \frac{D_{1j}}{U_{j-1} - N_{j-1}} \right\},$$

où  $(Z'_j)_{1 \leq j \leq M}$ , sont les valeurs distinctes de  $(Z_i)_{1 \leq i \leq n}$  considérées dans l'ordre croissant,  $\forall 0 \leq k \leq 2$ ,

$$D_{kj} = \sum_{1 \leq i \leq n} 1_{\{Z_i = Z'_j, A_i = k\}},$$

$$N_j = \sum_{1 \leq i \leq n} 1_{\{Z_i \leq Z'_j\}},$$

et

$$U_{j-1} = n \prod_{j \leq l \leq M} \left\{ 1 - \frac{D_{2l}}{N_l} \right\}.$$

Kebabi et Messaci (2012) ont étudié l'estimateur à noyau de la régression de  $Y$  sachant  $X$ .

Dans ce travail, nous proposons d'introduire et d'étudier un estimateur à noyau de la densité conditionnelle  $f^x$ .

D'une part, si  $Y$  n'est pas censurée, l'estimateur à noyau de la densité conditionnelle est défini comme

$$\hat{f}_n^x(y) = \sum_{i=1}^n W_{n,i}(x) \frac{1}{g_n} K_0\left(\frac{y - Y_i}{g_n}\right)$$

où

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

est le poids de Nadaraya-Watson.  $K$  et  $K_0$  sont des noyaux réels,  $(h_n)_{n \in \mathbb{N}}$  et  $(g_n)_{n \in \mathbb{N}}$  sont deux séquences de nombres réels strictement positifs convergent vers zéro (voir Collomb et al. (1987)).

D'autre part, Soit  $h'$  une fonction de  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , suivant Messaci (2010), un estimateur sans biais de  $Eh'(X, Y)$  dans notre contexte de modèle de censure est

$$\frac{1}{n} \sum_{i=1}^n 1_{\{A_i=0\}} \frac{h'(X_i, Z_i)}{S_R(Z_i)F_L(Z_i)}.$$

En effet

$$\begin{aligned} E\left(\frac{1_{\{A=0\}}h'(X, Z)}{S_R(Z)F_L(Z)}\right) &= E\left(E\left(\frac{1_{\{A=0\}}h'(X, Y)}{S_R(Z)F_L(Z)}\right) \middle/ (X, Y)\right) \\ &= E\left(\frac{h'(X, Y)}{S_R(Y)F_L(Y)} E(1_{\{A=0\}} \middle/ (X, Y))\right) \\ &= E(h'(X, Y)) \end{aligned}$$

selon l'hypothèse  $\mathbf{H}_{3,1}$ , donnÃ©e à la section 2.2, on a

$$E(1_{\{A=0\}} \middle/ X, Y) = S_R(Y)F_L(Y). \quad (2.1)$$

Car pour tout  $B$  dans  $\sigma(X, Y)$  (tribu engendrée par le couple  $(X, Y)$ ) il existe un borélien  $C$  tel que  $B = (X, Y)^{-1}(C)$ .

L'indépendance de  $(X, Y)$  et  $(L, R)$  permet d'écrire

$$\begin{aligned} \int_B (1_{A=0}) dP &= \int_B (1_{L<Y<R}) dP \\ &= \int_{C \times \mathbb{R}_+^2} (1_{l<y<r}) dP_{(X,Y,L,R)} \\ &= \int_{C \times \mathbb{R}_+^2} (1_{l<y<r}) dP_{(X,Y)} \otimes dP_{(L,R)}. \end{aligned}$$

Maintenant par le théorème de Fubini et l'indépendance de  $R$  et  $L$ , nous obtenons

$$\begin{aligned}
\int_B (1_{A=0}) dP &= \int_C \left( \int_{\mathbb{R}_+^2} (1_{l < y < r}) dP_{(L,R)} \right) dP_{(X,Y)} \\
&= \int_C \left( \int_{\mathbb{R}_+} (1_{l < y}) dP_L \times \int_{\mathbb{R}_+} (1_{y < r}) dP_R \right) dP_{(X,Y)} \\
&= \int_C (F_L(y) S_R(y)) dP_{(X,Y)} \\
&= \int_B F_L(Y) S_R(Y) dP,
\end{aligned}$$

car  $F$  est continue.

De plus,  $F_L(Y) S_R(Y)$  étant clairement mesurable par rapport à  $\sigma(X, Y)$ , le résultat en découle.

Donc, après avoir estimé  $S_R$  et  $F_L$ , nous proposons comme estimateur de  $f^x(y)$

$$\hat{f}_n^x(y) = \sum_{i=1}^n W_{n,i}(x) 1_{\{A_i=0\}} \frac{\frac{1}{g_n} K_0\left(\frac{y-Z_i}{g_n}\right)}{S_n(Z_i) F_n(Z_i) + U_n}, \quad (2.2)$$

où  $F_n$  est l'estimateur de Kaplan-Meier de la fonction de distribution  $F_L$  (pour les données censurées à gauche) et est donné par

$$F_n(t) = \prod_{\{j: Z'_j > t\}} \left\{ 1 - \frac{1_{\{A_j=2\}}}{j} \right\}.$$

Malgré que  $S_n F_n$  ne s'annule pas lorsque  $A_i = 0$ , voir dernière ligne de la page 1509 Messaci (2010), nous avons introduit le terme  $U_n = \frac{1}{n}$  pour obtenir une borne inférieure non aléatoire pour le dénominateur de  $\hat{f}_n^x$ .

Remarquez que toute suite  $(U_n)$ , telle qu'il existe  $\beta > 0$  satisfaisant  $U_n^2 n^{\beta-2} \rightarrow \infty$  et  $U_n < \sqrt{\frac{\log n}{n}}$ , peut être choisie. En effet, dans ce cas l'étape iii) de la démonstration du Lemme 1 et (2.16) restent toutes les deux valides.

De plus, notez que le mécanisme de censure intervient à travers le dénominateur de  $\hat{f}_n^x$ , qui n'est autre en vue de (2.1), qu'un estimateur consistant du taux de non censure conditionnellement à  $(X, Y)$ .

## 2.2 Hypothèses

Soit  $S \subset ]\max(I_Y, I_L), \min(T_Y, T_R)[$  un compact sur  $\mathbb{R}$ . Nous pouvons déterminer le taux de l'erreur quadratique moyenne de  $\hat{f}_n^x$  sur  $S$  grâce aux hypothèses suivantes.

$$\mathbf{H}_1 : \max(I_L, I_Y) < I_R.$$

$$\mathbf{H}_{2,1} : f_X \text{ est 2-fois continûment différentiable autour de } x.$$

$$\mathbf{H}_{2,2} : f_X(x) > 0.$$

$$\mathbf{H}_{2,3} : \lim_{n \rightarrow \infty} h_n = 0, \lim_{n \rightarrow \infty} g_n = 0, \lim_{n \rightarrow \infty} nh_n g_n = +\infty, \lim_{n \rightarrow \infty} \frac{\log n}{nh_n} = 0 \text{ et } h_n = O\left(\frac{1}{g_n \log n}\right).$$

$$\mathbf{H}_{2,4} : \text{La densité conjointe } f(., .) \text{ du couple } (X, Y) \text{ est deux fois continûment différentiable.}$$

$$\mathbf{H}_{2,5} : K \text{ est une densité bornée à support compact et } \int tK(t)dt = 0.$$

$$\mathbf{H}_{2,6} : K_0 \text{ est une densité bornée à support compact et } \int tK_0(t)dt = 0.$$

$$\mathbf{H}_{3,1} : (R, L) \text{ et } (X, Y) \text{ sont indépendants et } F_L \text{ est continue sur } ]0, \infty[.$$

*Remarque 3.*  $\mathbf{H}_1$  et  $\mathbf{H}_{3,1}$  sont couramment utilisés dans le modèle de censure mixte.  $\mathbf{H}_1$  assure l'identifiabilité du modèle étudié comme dans Boukeloua (2015).

La condition  $\mathbf{H}_{3,1}$  joue un rôle crucial pour obtenir (2.1) ci-dessus et a déjà été utilisé, par exemple, dans Messaci (2010); Kebabi et Messaci (2012).

Les hypothèses  $\mathbf{H}_{2,1}$  et  $\mathbf{H}_{2,4} - \mathbf{H}_{2,6}$  sont généralement utilisés dans le cadre de l'estimation à noyau pour obtenir les taux de convergence des estimateurs. De plus, on peut facilement vérifier que les deux exemples suivants

$$h_n = n^{-s}, g_n = n^{-t} \text{ avec } s > 0, t > 0 \text{ et } s + t < 1$$

et

$$h_n = \frac{1}{\log n}, g_n = n^{-s} \text{ avec } 0 < s < 1$$

satisfont la condition **H<sub>2,3</sub>**.

*Remarque 4.* Comme  $K_0$  a un support compact,  $y \in S$  et  $\lim_{n \rightarrow +\infty} g_n = 0$ , on a :

$$\exists n_0 \in \mathbb{N}, \exists T \in \mathbb{R} \text{ et } \exists I \in \mathbb{R} \text{ tel que } [I, T] \subset ]\max(I_Y, I_L), \min(T_Y, T_R)[$$

et nous pouvons supposer dans (2.2) et (2.4) que

$$\forall n \geq n_0, \forall i \in \{1 \cdots n\}, I \leq Z_i \leq T. \quad (2.3)$$

Soulignons le fait que  $S_R(T) > 0$  et  $F_L(I) > 0$ .

## 2.3 Résultats et preuve

Maintenant, nous pouvons énoncer notre résultat principal.

*Théorème 2.* Sous les conditions **H<sub>1</sub>**, **H<sub>2,1</sub>** – **H<sub>2,6</sub>** et **H<sub>3,1</sub>**, on a

$$\sup_{y \in S} E(\hat{f}_n^x(y) - f^x(y))^2 = O(h_n^4 + h_n^2 g_n^2 + g_n^4) + O\left(\frac{1}{nh_n g_n}\right).$$

*Remarque 5.* Posons  $L \equiv 0$ , nous déduisons le cas commun des données censurées à droite qui, à notre connaissance, est obtenu pour la première fois.

*Remarque 6.* Notez que sous un renforcement standard des hypothèses **H<sub>2,1</sub>**, **H<sub>2,2</sub>** et **H<sub>2,5</sub>**, nous obtenons l'uniformité du résultat précédent à la fois sur  $y \in S$  et  $x$  appartenant à un ensemble compact. Il suffit de suivre des étapes similaires à celles suivies dans la preuve du Théorème 2.



*Démonstration.* du Théorème 2

Dans toute cette preuve, C désigne une constante générique. On a

$$\begin{aligned}\hat{f}_n^x(y) &= \frac{\hat{f}_{n,N}(x, y)}{f_n(x)} \\ &= \frac{\frac{1}{nh_n g_n} \sum_{i=1}^n 1_{\{A_i=0\}} \frac{K(\frac{x-X_i}{h_n})K_0(\frac{y-Z_i}{g_n})}{S_n(Z_i)F_n(Z_i)+U_n}}{\frac{1}{nh_n} \sum_{i=1}^n K(\frac{x-X_i}{h_n})}\end{aligned}$$

où

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

est l'estimateur à noyau de  $f_X(x)$ .

Cela suggère d'introduire la quantité suivante

$$\tilde{f}_n^x(y) = \frac{\tilde{f}_{n,N}(x, y)}{f_n(x)}$$

où

$$\tilde{f}_{n,N}(x, y) = \frac{1}{nh_n g_n} \sum_{i=1}^n 1_{\{A_i=0\}} \frac{K(\frac{x-X_i}{h_n})K_0(\frac{y-Z_i}{g_n})}{S_R(Z_i)F_L(Z_i)}. \quad (2.4)$$

Nous devons d'abord traiter les points suivants et prouver le Lemme.

- L'indépendance et l'équidistribution de  $(X_i, Y_i)_{1 \leq i \leq n}$  combiné avec la relation (2.3) et la bornitude de  $K$  et  $K_0$  permettent d'écrire

$$\begin{aligned}\sup_{y \in S} \text{Var}(\tilde{f}_{n,N}(x, y)) &\leq \frac{1}{n^2 h_n^2 g_n^2} \sup_{y \in S} \sum_{i=1}^n E \left[ \frac{K(\frac{x-X_i}{h_n})K_0(\frac{y-Z_i}{g_n})}{S_R(Z_i)F_L(Z_i)} 1_{\{A_i=0\}} \right]^2 \\ &\leq \frac{C}{nh_n g_n S_R^2(T)F_L^2(I)} \sup_{y \in S} E \left[ \frac{1}{h_n g_n} K\left(\frac{x-X}{h_n}\right) K_0\left(\frac{y-Y}{g_n}\right) \right].\end{aligned}$$

De plus

$$\sup_{y \in S} E \left[ \frac{1}{h_n g_n} K\left(\frac{x-X}{h_n}\right) K_0\left(\frac{y-Y}{g_n}\right) \right] \leq \sup_{y \in S} \int \int K(s)K_0(v)f(x-sh_n, y-vg_n)dsdv.$$

Grace à l'hypothèse **H<sub>2,4</sub>**, on peut utiliser le développement de Taylor autour de  $(x, y)$ . Ceci, associé avec les hypothèses **H<sub>2,5</sub>** et **H<sub>2,6</sub>**,

permet de conclure qu'il existe une constante  $C$  telle que

$$\sup_{y \in S} E \left[ \frac{1}{h_n g_n} K\left(\frac{x-X}{h_n}\right) K_0\left(\frac{y-Y}{g_n}\right) \right] \leq C. \quad (2.5)$$

Nous obtenons alors

$$\sup_{y \in S} \text{Var}(\tilde{f}_{n,N}(x, y)) = O\left(\frac{1}{nh_n g_n}\right). \quad (2.6)$$

- L'indépendance et l'équidistribution de  $(X_i)_{1 \leq i \leq n}$  avec l'hypothèse  $\mathbf{H}_{2,5}$ , conduit à

$$\begin{aligned} \text{Var}(f_n(x)) &= \frac{1}{n^2 h_n^2} \sum_{i=1}^n \text{Var}\left(K\left(\frac{x-X_i}{h_n}\right)\right) \\ &\leq \frac{1}{nh_n^2} E\left(K^2\left(\frac{x-X}{h_n}\right)\right) \\ &\leq \frac{C}{nh_n} \left(\frac{1}{h_n} EK\left(\frac{x-X}{h_n}\right)\right). \end{aligned}$$

De plus, sous les hypothèses  $\mathbf{H}_{2,1}$ ,  $\mathbf{H}_{2,3}$  et  $\mathbf{H}_{2,5}$ , le théorème de Bochner (voir Théorème 1A dans Parzen (1962) ) est applicable et assure que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{h_n} E\left(K\left(\frac{x-X}{h_n}\right)\right) &= \lim_{n \rightarrow \infty} E(f_n(x)) \\ &= f_X(x), \end{aligned} \quad (2.7)$$

donc

$$\text{Var}(f_n(x)) = O\left(\frac{1}{nh_n}\right). \quad (2.8)$$

Le Lemme suivant, a une grande importance dans le déroulement de la démonstration du Théorème 2.

*Lemme 1. Sous  $\mathbf{H}_1, \mathbf{H}_{2,1}, \mathbf{H}_{2,3}$  et  $\mathbf{H}_{2,5}$ , on a pour tout  $\beta > 0$ , il existe  $b > 0$  tel que*

$$P \left[ \sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b \sqrt{\frac{\log n}{n}} \right] = O(n^{-\beta}).$$

*Démonstration.* i) Nous commençons par démontrer qu'il existe  $\epsilon_0 > 0$  tel que pour tous  $\beta > 0$ ,

$$P(f_n(x) < \epsilon_0) = O(n^{-\beta}). \quad (2.9)$$

Pour cela, nous avons

- D'une part, sous  $\mathbf{H}_{2,1}$  et  $\mathbf{H}_{2,5}$ ,

$$Ef_n(x) - f_X(x) = O(h_n^2), \quad (2.10)$$

(voir par exemple la preuve du Lemme 4 dans le chapitre suivant), donc nous pouvons écrire

$$\begin{aligned} P[f_n(x) \leq \frac{f_X(x)}{2}] &\leq P[|f_n(x) - f_X(x)| > \frac{f_X(x)}{2}] \\ &\leq P[|f_n(x) - Ef_n(x)| > \frac{f_X(x)}{4}], \end{aligned} \quad (2.11)$$

pour  $n$  suffisamment grand.

- D'autre part, remarquons que

$$f_n(x) - Ef_n(x) = \frac{1}{n} \sum_{i=1}^n Z_i,$$

où

$$Z_i = \frac{1}{h_n} [K(\frac{x - X_i}{h_n}) - EK(\frac{x - X_i}{h_n})]$$

avec

$$|Z_i| \leq \frac{C}{h_n} := M$$

en vue de l'hypothèse  $\mathbf{H}_{2,5}$ .

Sous  $\mathbf{H}_{2,1}$ ,  $\mathbf{H}_{2,3}$  et  $\mathbf{H}_{2,5}$ , nous pouvons appliquer le théorème de Bochner pour obtenir qu'il existe  $C_1 > 0$  tel que,

$$\sigma^2 = EZ_1^2 \leq \frac{C_1}{h_n}.$$

Maintenant en appliquant le corollaire A.9 dans Ferraty et Vieu (2006), voir corollaire 1 dans l'annexe, on obtient pour  $0 < \epsilon < \frac{\sigma^2}{M}$ ,

$$P(|f_n(x) - Ef_n(x)| > \epsilon) \leq 2e^{-\frac{n\epsilon^2 h_n}{4C_1}}.$$

Posons  $\epsilon = \sqrt{\frac{4C_1\beta \log n}{nh_n}}$ , nous pouvons voir que sous **H<sub>2,3</sub>**,

$$\lim_{n \rightarrow \infty} \epsilon = 0$$

et la dernière inégalité devient

$$\forall \beta > 0, P\left(|f_n(x) - Ef_n(x)| > \sqrt{\frac{4C_1\beta \log n}{nh_n}}\right) = O(n^{-\beta}).$$

Cette relation avec (2.11) donne (2.9).

ii) Nous devons également prouver qu'il existe  $\epsilon_1 > 0$  tel que pour tous  $\beta > 0$ ,

$$P(F_n(t)S_n(t) \leq \epsilon_1) = O(n^{-\beta}). \quad (2.12)$$

À cette fin, pour  $0 < \epsilon_1 < S_R(T)F_L(I)$ , posons  $\epsilon = S_R(T)F_L(I) - \epsilon_1$ . Alors

$$\begin{aligned} P(S_n(T)F_n(I) < \epsilon_1) &\leq P\left(\sup_{I \leq t \leq T} |S_R(t)F_L(t) - S_n(t)F_n(t)| > \epsilon\right) \\ &\leq P\left(\sup_{I \leq t \leq T} |S_n(t) - S_R(t)| > \frac{\epsilon}{2}\right) \\ &\quad + P\left(\sup_{I \leq t \leq T} |F_n(t) - F_L(t)| > \frac{\epsilon}{2}\right). \end{aligned} \quad (2.13)$$

De plus, Sous **H<sub>1</sub>**, en appliquant le Lemme 1 dans Boukeloua (2015) ou voir Section 1.2.3, nous avons pour tout  $\beta > 0$ , tout  $\theta < \min(T_R, T_Y)$ , il existe  $a_1 > 0$  tel que

$$P\left(\sup_{t \leq \theta} |S_n(t) - S_R(t)| > a_1 \sqrt{\frac{\log n}{n}}\right) = O(n^{-\beta}). \quad (2.14)$$

Nous dérivons également de la relation (3) dans le même document ou voir Section 1.2.3, pour tout  $\beta > 0$ , tout  $\theta > \min(I_Y, I_R)$ , il existe  $a_2 > 0$  tel que

$$P \left( \sup_{t \geq \theta} |F_n(t) - F_L(t)| > a_2 \sqrt{\frac{\log n}{n}} \right) = O(n^{-\beta}). \quad (2.15)$$

Donc, il est clair que ii) découle de (2.13)–(2.15).

iii) Nous sommes maintenant en mesure de d'obtenir le résultat du Lemme 1.

Comme

$$\begin{aligned} & \sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| \\ & \leq \frac{1}{f_n(x)} \frac{\sup_{I \leq u \leq T} |S_n(u) - S_R(u)| + \sup_{I \leq u \leq T} |F_n(u) - F_L(u)| + U_n}{(S_n(T)F_n(I) + U_n) S_R(T)F_L(I)}. \end{aligned}$$

et en posant  $b = \frac{(a_1 + a_2 + 1)}{S_R(T)F_L(I)\epsilon_0\epsilon_1}$ , nous déduisons de (2.9), (2.12), (2.14), (2.15) et le fait que  $U_n = 1/n$  le résultat que nous cherchons.  $\square$

Maintenant, nous devons utiliser la décomposition suivante

$$E(\hat{f}_n^x(y) - f^x(y))^2 \leq 2E(\hat{f}_n^x(y) - \tilde{f}_n^x(y))^2 + 4\text{Var} \tilde{f}_n^x(y) + 4(E\tilde{f}_n^x(y) - f^x(y))^2.$$

- **Traitement du terme**  $\sup_{y \in S} E(\hat{f}_n^x(y) - \tilde{f}_n^x(y))^2$ .

Nous pouvons écrire

$$\sup_{y \in S} E(\hat{f}_n^x(y) - \tilde{f}_n^x(y))^2 = T_{n,1}(x, y) + T_{n,2}(x, y),$$

où

$$T_{n,1}(x, y) := \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| \leq b \sqrt{\frac{\log n}{n}}\}} \left[ \hat{f}_n^x(y) - \tilde{f}_n^x(y) \right]^2 dP$$

et

$$T_{n,2}(x, y) := \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t)+U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b\sqrt{\frac{\log n}{n}}\}} \left[ \hat{f}_n^x(y) - \tilde{f}_n^x(y) \right]^2 dP.$$

Nous avons sous **H<sub>2,4</sub>**-**H<sub>2,6</sub>**

$$\begin{aligned} T_{n,1}(x, y) &:= \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t)+U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| \leq b\sqrt{\frac{\log n}{n}}\}} \left[ \hat{f}_n^x(y) - \tilde{f}_n^x(y) \right]^2 dP \\ &\leq \frac{b^2}{n^2 h_n^2 g_n^2} \frac{\log n}{n} \sup_{y \in S} E \left[ \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) K_0\left(\frac{y - Y_i}{g_n}\right) \right]^2 \\ &\leq \frac{b^2 \log n}{n^3 h_n^2 g_n^2} \sup_{y \in S} \sum_{i=1}^n E \left[ K\left(\frac{x - X_i}{h_n}\right) K_0\left(\frac{y - Y_i}{g_n}\right) \right]^2 \\ &\quad + \frac{b^2 \log n}{n^3 h_n^2 g_n^2} \sup_{y \in S} \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \left\{ E \left[ K\left(\frac{x - X_j}{h_n}\right) K_0\left(\frac{y - Y_j}{g_n}\right) \right] E \left[ K\left(\frac{x - X_i}{h_n}\right) K_0\left(\frac{y - Y_i}{g_n}\right) \right] \right\} \\ &\leq \frac{C \log n}{n^2 h_n g_n} \sup_{y \in S} E \left[ \frac{1}{h_n g_n} K\left(\frac{x - X_1}{h_n}\right) K_0\left(\frac{y - Y_1}{g_n}\right) \right] \\ &\quad + \frac{C(n-1) \log n}{n^2} \sup_{y \in S} \left\{ E \left[ \frac{1}{h_n g_n} K\left(\frac{x - X_1}{h_n}\right) K_0\left(\frac{y - Y_1}{g_n}\right) \right] \right\}^2 \\ &= O\left(\frac{1}{n h_n g_n}\right), \end{aligned}$$

en raison de (2.5) et en vertu de l'hypothèse **H<sub>2,3</sub>**.

Maintenant, nous examinons le terme  $T_{n,2}$ .

Nous avons sous  $\mathbf{H}_{2,6}$  et selon l'expression de  $f_n(x)$

$$\begin{aligned}
T_{n,2}(x, y) &:= \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b\sqrt{\frac{\log n}{n}}\}} \left[ \hat{f}_n^x(y) - \tilde{f}_n^x(y) \right]^2 dP \\
&= \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b\sqrt{\frac{\log n}{n}}\}} \frac{1}{n^2 h_n^2 g_n^2 f_n^2(x)} \\
&\quad \times \left[ \sum_{i=1}^n 1_{\{A_i=0\}} \left( \frac{K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Z_i}{g_n}\right)}{S_n(Z_i)F_n(Z_i) + U_n} - \frac{K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Z_i}{g_n}\right)}{S_R(Z_i)F_L(Z_i)} \right) \right]^2 dP \\
&\leq \sup_{y \in S} \int_{\{\sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b\sqrt{\frac{\log n}{n}}\}} \frac{1}{g_n^2} \left( \frac{1}{U_n} + \frac{1}{S_R(T)F_L(I)} \right)^2 \\
&\quad \times \left( \frac{1}{nh_n f_n(x)} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Z_i}{g_n}\right) 1_{\{A_i=0\}} \right)^2 dP \\
&\leq P \left[ \sup_{I \leq t \leq T} \left| \frac{1}{f_n(x)} \left( \frac{1}{S_n(t)F_n(t) + U_n} - \frac{1}{S_R(t)F_L(t)} \right) \right| > b\sqrt{\frac{\log n}{n}} \right] \\
&\quad \times \frac{1}{g_n^2} \left( \frac{1}{U_n} + \frac{1}{S_R(T)F_L(I)} \right)^2.
\end{aligned}$$

En utilisant le Lemme 1, en choisissant  $\beta > 4$  et en appliquant l'hypothèse  $\mathbf{H}_{2,3}$ , nous obtenons

$$T_{n,2}(x, y) = O\left(\frac{1}{nh_n g_n}\right). \quad (2.16)$$

- **Traitement du terme**  $\sup_{y \in S} (E\tilde{f}_n^x(y) - f^x(y))^2$ .

En utilisant l'identité usuelle suivante

$$\frac{1}{U} = \frac{1}{C} - \frac{U-C}{C^2} + \frac{(U-C)^2}{UC^2} \quad (C \neq 0, U \neq 0),$$

avec  $U = f_n(x)$  et  $C = Ef_n(x)$ , nous obtenons

$$\begin{aligned}
[E\tilde{f}_n^x(y) - f^x(y)]^2 &= \left\{ \frac{E\tilde{f}_{n,N}(x,y)}{Ef_n(x)} - \frac{E[\tilde{f}_{n,N}(x,y)(f_n(x) - Ef_n(x))]}{[Ef_n(x)]^2} \right. \\
&\quad \left. + \frac{E[\tilde{f}_n^x(y)(f_n(x) - Ef_n(x))^2]}{[Ef_n(x)]^2} - \frac{f(x,y)}{f_X(x)} \right\}^2 \\
&= \left\{ \frac{f_X(x)E\tilde{f}_{n,N}(x,y) - Ef_n(x)f(x,y)}{f_X(x)Ef_n(x)} - \frac{\text{Cov}[\tilde{f}_{n,N}(x,y), f_n(x)]}{(Ef_n(x))^2} \right. \\
&\quad \left. + \frac{E[\tilde{f}_n^x(y)(f_n(x) - Ef_n(x))^2]}{[Ef_n(x)]^2} \right\}^2 \\
&= \left\{ \frac{E\tilde{f}_{n,N}(x,y) - f(x,y)}{Ef_n(x)} - \frac{f^x(y)[Ef_n(x) - f_X(x)]}{Ef_n(x)} \right. \\
&\quad \left. + \frac{E[\tilde{f}_n^x(y)(f_n(x) - Ef_n(x))^2]}{[Ef_n(x)]^2} - \frac{\text{Cov}[\tilde{f}_{n,N}(x,y), f_n(x)]}{[Ef_n(x)]^2} \right\}^2 \\
&\leq 4 \left\{ \frac{E\tilde{f}_{n,N}(x,y) - f(x,y)}{Ef_n(x)} \right\}^2 + 4 \left\{ \frac{f^x(y)[Ef_n(x) - f_X(x)]}{Ef_n(x)} \right\}^2 \\
&\quad + 4 \left\{ \frac{E[\tilde{f}_n^x(y)(f_n(x) - Ef_n(x))^2]}{[Ef_n(x)]^2} \right\}^2 + 4 \left\{ \frac{\text{Cov}[\tilde{f}_{n,N}(x,y), f_n(x)]}{[Ef_n(x)]^2} \right\}^2.
\end{aligned}$$

En appliquant (2.7) et l'hypothèse  $\mathbf{H}_{2,2}$ , nous trouvons que

$$\sup_{y \in S} [E\tilde{f}_n^x(y) - f^x(y)]^2 \leq C(L_{1,n} + L_{2,n} + L_{3,n} + L_{4,n}),$$

où

$$L_{1,n} := \sup_{y \in S} \left( E\tilde{f}_{n,N}(x,y) - f(x,y) \right)^2,$$

$$L_{2,n} := \sup_{y \in S} [f^x(y)(Ef_n(x) - f_X(x))]^2,$$

$$L_{3,n} := \sup_{y \in S} \left\{ E[\tilde{f}_n^x(y)(f_n(x) - Ef_n(x))^2] \right\}^2$$



et

$$L_{4,n} := \sup_{y \in S} \left[ \text{Cov} \left( \tilde{f}_{n,N}(x, y), f_n(x) \right) \right]^2.$$

— En utilisant la formule (2.4) et le fait que  $(X_i, Z_i, A_i)$  sont i.i.d., nous obtenons

$$E \tilde{f}_{n,N}(x, y) - f(x, y) = E \left( \frac{1}{h_n g_n} 1_{\{A=0\}} \frac{K\left(\frac{x-X}{h_n}\right) K_0\left(\frac{y-Y}{g_n}\right)}{S_R(Y) F_L(Y)} \right) - f(x, y).$$

De plus, la relation (2.1) mène à

$$E \tilde{f}_{n,N}(x, y) - f(x, y) = E \left[ \frac{1}{h_n g_n} K\left(\frac{x-X}{h_n}\right) K_0\left(\frac{y-Y}{g_n}\right) \right] - f(x, y).$$

En procédant comme pour prouver (2.5), nous obtenons

$$\sup_{y \in S} \left| E \tilde{f}_{n,N}(x, y) - f(x, y) \right| = O(h_n^2 + g_n^2),$$

qui donne

$$\begin{aligned} L_{1,n} &:= \sup_{y \in S} \left( E \tilde{f}_{n,N}(x, y) - f(x, y) \right)^2 \\ &= O(h_n^4 + g_n^4 + h_n^2 g_n^2). \end{aligned}$$

— Sous les conditions  $\mathbf{H}_{2,2}$  et  $\mathbf{H}_{2,4}$ ,  $\sup_{y \in S} |f^x(y)|$  est borné, donc en utilisant (3.8), nous obtenons

$$\begin{aligned} L_{2,n} &:= \sup_{y \in S} [f^x(y) (E f_n(x) - f_X(x))]^2 \\ &= O(h_n^4). \end{aligned}$$

— Comme

$$\tilde{f}_n^x(y) = \sum_{i=1}^n W_{n,i}(x) 1_{\{A_i=0\}} \frac{\frac{1}{g_n} K_0\left(\frac{y-Z_i}{g_n}\right)}{S_R(Z_i) F_L(Z_i)}$$

et en utilisant le fait que  $\sum_{i=1}^n |W_{n,i}(x)| = 1$ , nous obtenons sous

les conditions  $\mathbf{H}_{2,5}$ ,  $\mathbf{H}_{2,6}$  et la relation (2.3) avec la formule (2.8),

$$\begin{aligned} L_{3,n} &:= \sup_{y \in S} \left\{ E \left[ \tilde{f}_n^x(y) (f_n(x) - E f_n(x))^2 \right] \right\}^2 \\ &\leq \left[ \frac{C}{g_n} \text{Var}(f_n(x)) \right]^2 \\ &= O\left(\frac{1}{nh_n g_n}\right). \end{aligned}$$

— L'inégalité de Cauchy-Schwarz, combinée avec (2.6) et (2.8), nous donne

$$\begin{aligned} L_{4,n} &:= \sup_{y \in S} \left[ \text{Cov} \left( \tilde{f}_{n,N}(x, y), f_n(x) \right) \right]^2 \\ &\leq \sup_{y \in S} \left[ \text{Var} \left( \tilde{f}_{n,N}(x, y) \right) \text{Var} (f_n(x)) \right] \\ &= O\left(\frac{1}{n^2 h_n^2 g_n}\right), \end{aligned} \tag{2.17}$$

qui, sous l'hypothèse  $\mathbf{H}_{2,3}$ , donne

$$L_{4,n} = O\left(\frac{1}{nh_n g_n}\right).$$

- **Traitement du terme**  $\sup_{y \in S} \text{Var}(\tilde{f}_n^x(y))$ .

Nous pouvons calculer la variance de  $\tilde{f}_n^x$  en utilisant l'identité suivante (voir la formule (8.12) dans Scott (1992)).

$$\begin{aligned} \text{Var}(\tilde{f}_n^x(y)) &= \frac{\text{Var} \left( \tilde{f}_{n,N}(x, y) \right)}{(E f_n(x))^2} - 2 E \tilde{f}_{n,N}(x, y) \frac{\text{Cov} \left( \tilde{f}_{n,N}(x, y), f_n(x) \right)}{(E f_n(x))^3} \\ &\quad + \text{Var}(f_n(x)) \frac{\left( E \tilde{f}_{n,N}(x, y) \right)^2}{(E f_n(x))^4}. \end{aligned}$$

Le fait que  $E f_n$  converge et sous l'hypothèse  $\mathbf{H}_{2,2}$ , nous avons

$$\sup_{y \in S} \text{Var}(\tilde{f}_n^x(y)) \leq C(L'_{1,n} + L'_{2,n} + L'_{3,n}),$$

où

$$L'_{1,n} := \sup_{y \in S} \text{Var} \left( \tilde{f}_{n,N}(x, y) \right),$$

$$L'_{2,n} := \sup_{y \in S} \left[ E \tilde{f}_{n,N}(x, y) \text{Cov} \left( \tilde{f}_{n,N}(x, y), f_n(x) \right) \right]$$

et

$$L'_{3,n} := \text{Var} (f_n(x)) \sup_{y \in S} \left( E \tilde{f}_{n,N}(x, y) \right)^2.$$

Il reste à étudier chaque terme  $L'_{j,n}$ , ( $1 \leq j \leq 3$ ).

— La relation (2.6) conduit à

$$L'_{1,n} := \sup_{y \in S} \text{Var} \left( \tilde{f}_{n,N}(x, y) \right)$$

$$= O \left( \frac{1}{nh_n g_n} \right).$$

— En utilisant les relations (2.3), (2.4) et (2.5), nous constatons que  $\sup_{y \in S} E \tilde{f}_{n,N}(x, y)$  est borné. Donc, en procédant comme pour prouver (2.17) avec l'utilisation de l'hypothèse  $\mathbf{H}_{2,3}$ , nous obtenons

$$L'_{2,n} := \sup_{y \in S} \left[ E \tilde{f}_{n,N}(x, y) \text{Cov} \left( \tilde{f}_{n,N}(x, y), f_n(x) \right) \right]$$

$$= O \left( \frac{1}{nh_n g_n} \right).$$

— En utilisant à nouveau le fait que  $\sup_{y \in S} E \tilde{f}_{n,N}(x, y)$  est Borné avec (2.8) et l'hypothèse  $\mathbf{H}_{2,3}$ , nous obtenons

$$L'_{3,n} := \text{Var} (f_n(x)) \sup_{y \in S} \left( E \tilde{f}_{n,N}(x, y) \right)^2$$

$$= O \left( \frac{1}{nh_n g_n} \right).$$

La preuve est alors terminée.



# Chapitre 3

## Estimation du mode conditionnel dans un modèle de censure mixte

L'objectif de ce chapitre est de proposer un estimateur non paramétrique du mode conditionnel lorsque la variable réponse est soumise à une censure mixte et d'établir son taux de convergence presque complète.

A cette fin, nous avons d'abord besoin d'étudier la convergence presque complète uniforme sur un compact  $S$  de l'estimateur à noyau de la densité conditionnelle, que nous avons déjà introduit au chapitre précédent et de donner sa vitesse de convergence.

### 3.1 Convergence presque complète de l'estimateur à noyau de la densité conditionnelle

Remarquons que sous l'hypothèse  $\mathbf{H}_1$ , nous pouvons appliquer la relation (1.4), pour obtenir pour tout  $\theta < \min(T_R, T_Y)$

$$\sup_{t \leq \theta} |S_n(t) - S_R(t)| = O_{p.co} \left( \sqrt{\frac{\log n}{n}} \right). \quad (3.1)$$

Nous utilisons également la relation (1.3), nous obtenons alors pour tout  $\theta > \min(I_Y, I_R)$

$$\sup_{t \geq \theta} |F_n(t) - F_L(t)| = O_{p.co} \left( \sqrt{\frac{\log n}{n}} \right). \quad (3.2)$$

Pour obtenir le taux de convergence presque complète de l'estimateur de la densité conditionnelle, nous avons besoin d'introduire les conditions suivantes.

$$\mathbf{H}'_{2,3} : \lim_{n \rightarrow \infty} h_n = 0, \lim_{n \rightarrow \infty} g_n = 0, \lim_{n \rightarrow \infty} \frac{nh_n g_n}{\log n} = +\infty.$$

$\mathbf{H}'_{2,6} : \mathbf{H}_{2,6}$  est satisfaite de plus  $K_0$  est holdérienne, ce qui s'écrit :

$$\exists \beta > 0, \exists C > 0 \text{ tels que } \forall x, y \in \mathbb{R}, |K_0(x) - K_0(y)| \leq C|x - y|^\beta.$$

*Proposition 1.* Sous les hypothèses suivantes  $\mathbf{H}_1, \mathbf{H}_{2,1} - \mathbf{H}_{2,2}, \mathbf{H}'_{2,3}, \mathbf{H}_{2,4} - \mathbf{H}_{2,5}, \mathbf{H}'_{2,6}$  et  $\mathbf{H}_{3,1}$ , on a

$$\sup_{y \in S} \left| \hat{f}_n^x(y) - f^x(y) \right| = O(h_n^2 + g_n^2) + O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

Pour  $L \equiv 0$ , notre modèle se ramène au modèle classique de censure à droite et nous retrouvons le même taux de convergence que celui donné par Khardani et al. (2010).

*Démonstration.* La preuve de cette proposition est basée sur la décomposition suivante

$$\hat{f}_n^x(y) - f^x(y) = \frac{\hat{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, y)}{f_n(x)} + \frac{\tilde{f}_{n,N}(x, y) - f(x, y)}{f_n(x)} + \left( \frac{f_X(x) - f_n(x)}{f_n(x)} \right) f^x(y),$$

où (rappelons le)

$$\hat{f}_n^x(y) = \frac{\hat{f}_{n,N}(x, y)}{f_n(x)},$$

avec

$$\hat{f}_{n,N}(x, y) = \frac{1}{nh_n g_n} \sum_{i=1}^n 1_{\{A_i=0\}} \frac{K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Z_i}{g_n}\right)}{S_n(Z_i) F_n(Z_i) + U_n},$$

et

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

est l'estimateur à noyau de  $f_X(x)$ . Rappelons aussi que

$$\tilde{f}_n^x(y) = \frac{\tilde{f}_{n,N}(x, y)}{f_n(x)}$$

où

$$\tilde{f}_{n,N}(x, y) = \frac{1}{nh_n g_n} \sum_{i=1}^n 1_{\{A_i=0\}} \frac{K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Z_i}{g_n}\right)}{S_R(Z_i) F_L(Z_i)}.$$

Il suffit donc de démontrer les lemmes suivants . □

*Lemme 2.* Sous les conditions  $\mathbf{H}_1$ ,  $\mathbf{H}'_{2,3}$ ,  $\mathbf{H}_{2,4} - \mathbf{H}_{2,5}$  et  $\mathbf{H}'_{2,6}$ , on a

$$\sup_{y \in S} \left| \hat{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, y) \right| = O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

*Démonstration.* Selon les définitions de  $\hat{f}_{n,N}$ ,  $\tilde{f}_{n,N}$  et grâce à la relation (2.3), nous pouvons écrire

$$\begin{aligned} \sup_{y \in S} \left| \hat{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, y) \right| &\leq \frac{1}{nh_n g_n} \sup_{y \in S} \left[ \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Y_i}{g_n}\right) \right] \\ &\quad \times \frac{\sup_{t \geq I} |F_n(t) - F_L(t)| + \sup_{t \leq T} |S_n(t) - S_R(t)| + U_n}{(S_n(T) F_n(I) + U_n) (S_R(T) F_L(I))} \\ &\leq \left( \frac{\sup_{t \geq I} |F_n(t) - F_L(t)| + \sup_{t \leq T} |S_n(t) - S_R(t)| + U_n}{S_n(T) F_n(I) S_R(T) F_L(I)} \right) \left( \frac{1}{n} \sum_{i=1}^n \psi_i \right) \end{aligned}$$

où

$$\psi_i = \frac{1}{h_n g_n} \sup_{y \in S} \left( K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{y-Y_i}{g_n}\right) \right).$$

- Sous  $\mathbf{H}_1$ , les relations (3.1) et (3.2) impliquent que

$$S_n(T) \xrightarrow{p.co} S_R(T) \neq 0, F_n(I) \xrightarrow{p.co} F_L(I) \neq 0 \quad (3.3)$$

et

$$\begin{aligned} \sup_{t \geq I} |F_n(t) - F_L(t)| + \sup_{t \leq T} |S_n(t) - S_R(t)| &= O_{p.co} \left( \sqrt{\frac{\log n}{n}} \right), \\ U_n = \frac{1}{n} &= O \left( \frac{\log n}{nh_n g_n} \right) \end{aligned} \quad (3.4)$$

- Maintenant, montrons que  $\frac{1}{n} \sum_{i=1}^n \psi_i$  converge presque complètement.

On a

$$\frac{1}{n} \sum_{i=1}^n \psi_i = \left( \frac{1}{n} \sum_{i=1}^n U_i \right) + E\psi_1,$$

où  $U_i = \psi_i - E\psi_i$ .

Nous appliquons l'inégalité de Bernstein (voir Corollaire 1.i dans l'annexe).

Pour cela, en utilisant les hypothèses  $\mathbf{H}_{2,5}$  et  $\mathbf{H}'_{2,6}$ , nous obtenons

$$|U_i| \leq \frac{C}{h_n g_n} := M.$$

Bornons maintenant  $\sigma^2 = \text{Var}U_i$ .

$$\sigma^2 = \text{Var}U_i = E(U_i)^2 \leq E(\psi_i)^2.$$

Comme les noyaux  $K$  et  $K_0$  sont positifs et bornés, nous avons

$$E\psi_1^2 \leq CE \left[ \frac{1}{h_n^2 g_n^2} K\left(\frac{x - X_1}{h_n}\right) \sup_{y \in S} K_0\left(\frac{y - Y_1}{g_n}\right) \right].$$

Ceci combiné avec la relation (2.5) conduit à



$$\sigma^2 \leq \frac{C}{h_n g_n}.$$

Alors pour  $\epsilon < \frac{\sigma^2}{M}$ , nous obtenons

$$P \left[ \frac{1}{n} \left| \sum_{i=1}^n U_i \right| > \epsilon \right] \leq 2 \exp \left\{ -\frac{n\epsilon^2}{2\sigma^2(1 + \epsilon \frac{M}{\sigma^2})} \right\} \leq 2 \exp \left\{ -\frac{n\epsilon^2 h_n g_n}{4C} \right\}. \quad (3.5)$$

Posons  $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{nh_n g_n}}$ , nous avons

$$P \left[ \frac{1}{n} \left| \sum_{i=1}^n U_i \right| > \epsilon_0 \sqrt{\frac{\log n}{nh_n g_n}} \right] \leq 2 \exp \left\{ -\frac{n\epsilon_0^2}{2\sigma^2(1 + \epsilon_0 \frac{M}{\sigma^2})} \right\} \leq 2n^{-\frac{\epsilon_0^2}{4C}}.$$

Pour  $\epsilon_0$  assez grand, le dernier terme de cette inégalité est le terme d'une série convergente et en vue de l'hypothèse  $\mathbf{H}'_{2,3}$ ,

$$\frac{1}{n} \sum_{i=1}^n \psi_i \xrightarrow{p.co} E\psi_1.$$

Ceci avec (3.3) et (3.4) donne par application de la Proposition 3, voir dans l'annexe,

$$\sup_{y \in S} \left| \hat{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, y) \right| = O_{p.co} \left( \sqrt{\frac{\log n}{n}} \right),$$

qui est égal, grâce à l'hypothèse  $\mathbf{H}'_{2,3}$ , à  $O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right)$ .  $\square$

*Lemme 3.* En vertu des hypothèses  $\mathbf{H}_{2,4}$ ,  $\mathbf{H}_{2,5}$ ,  $\mathbf{H}'_{2,6}$  et  $\mathbf{H}_{3,1}$ , nous obtenons

$$\sup_{y \in S} \left| E\tilde{f}_{n,N}(x, y) - f(x, y) \right| = O(h_n^2 + g_n^2).$$

*Démonstration.* Comme  $(X_i, Z_i, A_i)$  sont i.i.d., nous avons

$$E\tilde{f}_{n,N}(x, y) - f(x, y) \leq E \left( \frac{1}{h_n g_n} 1_{\{A_1=0\}} \frac{K\left(\frac{x-X_1}{h_n}\right) K_0\left(\frac{y-Y_1}{g_n}\right)}{S_R(Y_1) F_L(Y_1)} \right) - f(x, y).$$

D'autre part, la relation (2.1) conduit à

$$E\tilde{f}_{n,N}(x, y) - f(x, y) \leq E \left( \frac{1}{h_n g_n} K\left(\frac{x - X_1}{h_n}\right) K_0\left(\frac{y - Y_1}{g_n}\right) \right) - f(x, y). \quad (3.6)$$

D'après l'hypothèse  $\mathbf{H}_{2,4}$  et en utilisant le changement de variable,

$$\begin{aligned} r &= \frac{x - u}{h_n}, \\ s &= \frac{y - v}{g_n} \end{aligned}$$

et le développement de Taylor au voisinage de  $(x, y)$ , nous trouvons

$$\begin{aligned} E\tilde{f}_{n,N}(x, y) - f(x, y) &\leq E \left( \frac{1}{h_n g_n} K\left(\frac{x - X_1}{h_n}\right) K_0\left(\frac{y - Y_1}{g_n}\right) \right) - f(x, y) \\ &\leq \int \int \frac{1}{h_n g_n} K\left(\frac{x - u}{h_n}\right) K_0\left(\frac{y - v}{g_n}\right) f(u, v) dudv - f(x, y) \\ &\leq \int \int K(r) K_0(s) f(x - rh_n, y - sg_n) dr ds - f(x, y) \\ &\leq \int \int K(r) K_0(s) [f(x - rh_n, y - sg_n) - f(x, y)] dr ds. \end{aligned}$$

P ar l'utilisation des hypothèses  $\mathbf{H}_{2,4}$ ,  $\mathbf{H}_{2,5}$  et  $\mathbf{H}'_{2,6}$ , nous obtenons

$$\sup_{y \in S} \left| E\tilde{f}_{n,N}(x, y) - f(x, y) \right| = O(h_n^2 + g_n^2).$$

□

*Lemme 4.*  $\mathbf{H}'_{2,3}$ ,  $\mathbf{H}_{2,4} - \mathbf{H}_{2,5}$  et  $\mathbf{H}'_{2,6}$ , nous avons

$$\sup_{y \in S} \left| E\tilde{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, y) \right| = O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

*Démonstration.* En utilisant la compacité de  $S$ , nous pouvons écrire

$$S \subset \bigcup_{j=1}^{\tau_n} ]t_j - l_n, t_j + l_n[$$

avec

$$l_n = C\tau_n^{-1} \text{ et } \tau_n = Cn^{1+\frac{1}{\beta}}. \quad (3.7)$$

Posons

$$t_y = \arg \min_{t \in \{t_1, \dots, t_{\tau_n}\}} |y - t|.$$

Nous allons utiliser l'inégalité suivante

$$\sup_{y \in S} \left| \tilde{f}_{n,N}(x, y) - E\tilde{f}_{n,N}(x, y) \right| \leq A_1 + A_2 + A_3,$$

où

$$A_1 = \sup_{y \in S} \left| \tilde{f}_{n,N}(x, y) - \tilde{f}_{n,N}(x, t_y) \right|,$$

$$A_2 = \sup_{y \in S} \left| \tilde{f}_{n,N}(x, t_y) - E\tilde{f}_{n,N}(x, t_y) \right|$$

et

$$A_3 = \sup_{y \in S} \left| E\tilde{f}_{n,N}(x, t_y) - E\tilde{f}_{n,N}(x, y) \right|.$$

- Traitement du terme  $A_1$ .

Sous les hypothèses  $\mathbf{H}_{2,5}$ ,  $\mathbf{H}'_{2,6}$  et la relation (2.3), nous avons

$$\begin{aligned} A_1 &= \sup_{y \in S} \frac{1}{nh_n g_n} \left| \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \frac{1_{\{A_i=0\}} K_0\left(\frac{y-Z_i}{g_n}\right)}{S_R(Z_i) F_L(Z_i)} - \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \frac{1_{\{A_i=0\}} K_0\left(\frac{t_y-Z_i}{g_n}\right)}{S_R(Z_i) F_L(Z_i)} \right| \\ &\leq \frac{l_n^\beta}{nh_n g_n^{\beta+1} S_R(T) F_L(I)} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \\ &\leq \frac{Cl_n^\beta}{h_n g_n^{\beta+1}}, \end{aligned}$$

$$A_1 = O\left(\sqrt{\frac{\log n}{nh_n g_n}}\right) \text{ à cause de la relation (3.7) et l'hypothèse } \mathbf{H}'_{2,3}.$$

- Traitement du terme  $A_2$ .

Pour  $\epsilon > 0$ ,

$$\begin{aligned}
P(A_2 > \epsilon) &\leq P\left(\max_{j=1,\dots,\tau_n} \left| \tilde{f}_{n,N}(x, t_j) - E\tilde{f}_{n,N}(x, t_j) \right| > \epsilon\right) \\
&\leq \tau_n \max_{j=1,\dots,\tau_n} P\left(\left| \tilde{f}_{n,N}(x, t_j) - E\tilde{f}_{n,N}(x, t_j) \right| > \epsilon\right) \\
&\leq \tau_n \max_{j=1,\dots,\tau_n} P\left(\frac{1}{n} \left| \sum_{i=1}^n \Gamma_i \right| > \epsilon\right),
\end{aligned}$$

où

$$\Gamma_i = V_i - EV_i \text{ avec } V_i = \frac{1}{h_n g_n} 1_{\{A_i=0\}} \frac{K\left(\frac{x-X_i}{h_n}\right) K_0\left(\frac{t_j-Z_i}{g_n}\right)}{S_R(Z_i) F_L(Z_i)}.$$

En vue de la relation (2.3)

$$|V_i| \leq \frac{|\psi_i|}{S_R(T) F_L(I)},$$

En procédant de la même manière que pour prouver la relation (3.5), nous déduis ons que

$$P(A_2 > \epsilon) \leq 2\tau_n \exp\left\{-\frac{n\epsilon^2 h_n g_n}{4C}\right\}.$$

Posons  $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{nh_n g_n}}$ , en appliquant la relation (3.7) et avec un choix adéquat de  $\epsilon_0$ , nous obtenons

$$A_2 = O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

- Traitement du terme  $A_3$ .

Il est suffisant de remarquer que  $|A_3| \leq EA_1$ .

□

*Lemme 5. Sous les conditions  $\mathbf{H}_{2,1}$ ,  $\mathbf{H}_{2,2}$ ,  $\mathbf{H}'_{2,3}$  et  $\mathbf{H}_{2,5}$  nous obtenons*

$$Ef_n(x) - f_X(x) = O(h_n^2), \quad (3.8)$$

$$Ef_n(x) - f_n(x) = O_{a.co} \left( \sqrt{\frac{\log n}{nh_n}} \right) \quad (3.9)$$

et

$$\exists \eta > 0, \sum_{n=1}^{\infty} P(f_n(x) < \eta) < \infty. \quad (3.10)$$

*Démonstration.* • Comme  $(X_i)_{i \leq n}$  sont i.i.d, la condition  $\mathbf{H}_{2,5}$  nous permet d'écrire

$$\begin{aligned} Ef_n(x) - f_X(x) &= E \frac{1}{h_n} K\left(\frac{x - X_1}{h_n}\right) - f_X(x) \\ &= \frac{1}{h_n} \int K\left(\frac{x - u}{h_n}\right) f_X(u) du - f_X(x) \\ &= \int K(z) [f_X(x - zh_n) - f_X(x)] dz, \end{aligned}$$

où  $z = \frac{x-u}{h_n}$ .

Les hypothèses  $\mathbf{H}_{2,1}$ ,  $\mathbf{H}_{2,5}$  avec le développement de Taylor au voisinage de  $x$  permettent d'arriver

$$|Ef_n(x) - f_X(x)| = O(h_n^2).$$

- Maintenant, nous appliquons l'inégalité de Bernstein (cf Corollaire1). Pour cela, posons

$$\begin{aligned} U_i &= \frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right), \\ \Gamma_i &= U_i - EU_i. \end{aligned}$$

Il est clair que

$$f_n(x) - Ef_n(x) = \frac{1}{n} \sum \Gamma_i.$$

En vertu de l'hypothèse  $\mathbf{H}_{2,5}$ , nous avons

$$|\Gamma_i| \leq \frac{C}{h_n} := M.$$

Nous devons maintenant borner  $\Delta^2 = \text{Var}\Gamma_i$ , pour cela calculons  $EU_1^2$ .

En raison de l'hypothèse **H<sub>2,5</sub>**,

$$\Delta^2 \leq \frac{C}{h_n}.$$

Alors pour  $\epsilon < \frac{\Delta^2}{M}$ , nous avons

$$P \left[ \frac{1}{n} \left| \sum_{i=1}^n \Gamma_i \right| > \epsilon \right] \leq 2 \exp \left\{ -\frac{n\epsilon^2 h_n}{4C} \right\}.$$

Mettons  $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{nh_n}}$ , nous obtenons

$$P \left[ \frac{1}{n} \left| \sum_{i=1}^n \Gamma_i \right| > \epsilon_0 \sqrt{\frac{\log n}{nh_n}} \right] \leq 2n^{-\frac{\epsilon_0^2}{4C}},$$

pour  $\epsilon_0$  assez grand, la quantité dans le membre e droit de la dernière inégalité est le terme d'une série convergente, d'où ,

$$Ef_n(x) - f_n(x) = O_{p.co} \left( \sqrt{\frac{\log n}{nh_n}} \right).$$

- Les résultats (3.8) et (3.9) entraînent la convergence presque complète de  $f_n(x)$  vers  $f_X(x)$ . Ainsi pour tout  $\epsilon > 0$  on a

$$\sum_{n \in \mathbb{N}} P [|f_n(x) - f_X(x)| > \epsilon] < \infty.$$

En remarquant que

$$P \left( |f_n(x)| < \frac{f_X(x)}{2} \right) \leq P \left( |f_n(x) - f_X(x)| > \frac{f_X(x)}{2} \right).$$

Grâce à la condition **H<sub>2,2</sub>**, il suffit de poser  $\eta = \epsilon = \frac{f_X(x)}{2}$  pour obtenir la relation (3.10).

□

## 3.2 Taux de convergence presque complète de l'estimateur du mode conditionnel

Un estimateur naturel du mode conditionnel

$$\theta(x) = \arg \sup_{y \in S} f^x(y)$$

sur un compact  $S$  est donné par

$$\hat{\theta}_n(x) = \arg \sup_{y \in S} \hat{f}_n^x(y).$$

En utilisant la Proposition 1 ci dessus, avec les hypothèses supplémentaires suivantes, nous pouvons obtenir le taux de convergence presque complète de l'estimateur introduit  $\hat{\theta}_n(x)$ .

**H<sub>2,7</sub>** :  $\exists \xi > 0$ ,  $f^x \nearrow$  sur  $[\theta(x) - \xi, \theta(x))$  et  $f^x \searrow$  sur  $(\theta(x), \theta(x) + \xi]$ .

Cela montre que  $\theta(x)$  est le mode conditionnel unique sur  $[\theta(x) - \xi, \theta(x) + \xi]$ .

En outre  $\theta(x)$  doit vérifier

**H<sub>2,8</sub>** :  $f^x$  est 2-fois continûment différentiable autour de  $\theta(x)$ , de dérivée seconde  $f^{x(2)}(\cdot)$  satisfaisant  $f^{x(2)}(\theta(x)) < 0$ .

*Théorème 3. Sous les conditions imposées dans la Proposition 1 et si les hypothèses **H<sub>2,7</sub>** et **H<sub>2,8</sub>** sont également satisfaites, nous avons*

$$\left| \hat{\theta}_n(x) - \theta(x) \right|^2 = O(h_n^2 + g_n^2) + O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

*Remarque 7.* Posons  $L \equiv 0$ , notre modèle de censure se réduit au modèle classique de censure à droite. Dans ce cas, si nous posons  $h_n = g_n$ , nous obtenons le même taux de convergence que celui de l'estimateur du mode conditionnel étudié par Khardani et al. (2010), sauf que notre mode de convergence

(presque complète) est plus fort que celui étudié dans l'article précité (convergence presque sûre). Nous avons donc généralisé le mode de censure tout en améliorant le mode de convergence.

*Démonstration.* Avant d'établir le taux de convergence, nous avons d'abord besoin de montrer la convergence presque complète de  $\hat{\theta}_n(x)$  vers  $\theta(x)$ , ce qui s'écrit

$$\hat{\theta}_n(x) \xrightarrow{p.co} \theta(x). \quad (3.11)$$

La continuité et la monotonie stricte de  $f^x(\cdot)$  sur  $(\theta(x) - \xi, \theta(x))$  et sur  $(\theta(x), \theta(x) + \xi)$  montre que la fonction  $(f^x)^{-1}$  existe donc et est continue au point  $f^x(\theta(x))$ , ce qui implique que

$$\forall \epsilon > 0, \exists \delta(\epsilon) > 0, \forall y \in (\theta(x) - \xi, \theta(x) + \xi), |f^x(y) - f^x(\theta(x))| \leq \delta(\epsilon) \implies |y - \theta(x)| \leq \epsilon.$$

Comme  $\hat{\theta}_n(x) \in [\theta(x) - \xi, \theta(x) + \xi]$  par définition,

$$\forall \epsilon > 0, \exists \delta(\epsilon) > 0, |f^x(\hat{\theta}_n(x)) - f^x(\theta(x))| \leq \delta(\epsilon) \implies |\hat{\theta}_n(x) - \theta(x)| \leq \epsilon.$$

Nous obtenons donc

$$\forall \epsilon > 0, \exists \delta(\epsilon) > 0, P[|\hat{\theta}_n(x) - \theta(x)| > \epsilon] \leq P[|f^x(\hat{\theta}_n(x)) - f^x(\theta(x))| > \delta(\epsilon)].$$

D'autre, part il vient directement des définitions de  $\hat{\theta}_n(x)$  et  $\theta(x)$  que

$$\begin{aligned} |f^x(\theta(x)) - f^x(\hat{\theta}_n(x))| &= f^x(\theta(x)) - f^x(\hat{\theta}_n(x)) \\ &= f^x(\theta(x)) - \hat{f}_n^x(\theta(x)) + \hat{f}_n^x(\theta(x)) - f^x(\hat{\theta}_n(x)) \\ &\leq [f^x(\theta(x)) - \hat{f}_n^x(\theta(x))] + [\hat{f}_n^x(\hat{\theta}_n(x)) - f^x(\hat{\theta}_n(x))] \\ &\leq 2 \sup_{y \in [\theta(x) - \xi, \theta(x) + \xi]} |f^x(y) - \hat{f}_n^x(y)|. \end{aligned}$$

Il reste à appliquer la convergence uniforme presque complète de l'estimateur  $\hat{f}_n^x$  sur le compact  $[\theta(x) - \xi, \theta(x) + \xi]$ , montrée à la proposition 1, pour arriver à

$$\sum_{n=1}^{\infty} P[|\hat{\theta}_n(x) - \theta(x)| > \epsilon] < \infty,$$



qui est le résultat visé.

Établissons maintenant le taux de cette convergence.

L'hypothèse **H<sub>2,8</sub>** combinée avec le développement de Taylor d'ordre deux autour de  $\theta(x)$ , nous mène à

$$f^x(\hat{\theta}_n(x)) = f^x(\theta(x)) + \frac{1}{2!} f^{x(2)}(\theta^*)(\hat{\theta}_n(x) - \theta(x))^2,$$

pour une variable aléatoire  $\theta^*$  comprise entre  $\theta(x)$  et  $\hat{\theta}_n(x)$ .

Mais puisque nous avons montré un peu plus haut que

$$\left| f^x(\hat{\theta}_n(x)) - f^x(\theta(x)) \right| \leq 2 \sup_{y \in (\theta(x) - \zeta, \theta(x) + \zeta)} \left| \hat{f}_n^x(y) - f^x(y) \right|,$$

nous en déduisons que

$$\left| f^{x(2)}(\theta^*) \right| (\hat{\theta}_n(x) - \theta(x))^2 \leq 4 \sup_{y \in (\theta(x) - \zeta, \theta(x) + \zeta)} \left| \hat{f}_n^x(y) - f^x(y) \right|. \quad (3.12)$$

La relation (3.11) avec l'hypothèse **H<sub>2,8</sub>** implique que

$$f^{x(2)}(\theta^*) \xrightarrow{p.co} f^{x(2)}(\theta(x)) \neq 0.$$

Il reste à appliquer la Proposition 1 avec (3.12) et la Proposition 3.ii de l'annexe, pour obtenir

$$\left| \hat{\theta}_n(x) - \theta(x) \right|^2 = O(h_n^2 + g_n^2) + O_{p.co} \left( \sqrt{\frac{\log n}{nh_n g_n}} \right).$$

□

# Chapitre 4

## Simulation

### 4.1 Estimateur de la densité conditionnelle

Dans cette section, nous voulons vérifier la performance de l'estimateur étudié  $\hat{f}_n^x$  lequel, rappelons le, est défini en (2.2)

par,

$$\hat{f}_n^x(y) = \sum_{i=1}^n W_{n,i}(x) 1_{\{A_i=0\}} \frac{\frac{1}{g_n} K_0\left(\frac{y-Z_i}{g_n}\right)}{S_n(Z_i) F_n(Z_i) + U_n},$$

pour des tailles finies de l'échantillon. De plus, nous évaluons l'effet du taux de censure (CR) et de la taille de l'échantillon ( $n$ ) sur la qualité de cet estimateur pour un point fixé  $x$ .

À cette fin, nous considérons les modèles connus suivants .

$$\text{M1 : } Y = X + 0.2\epsilon,$$

$$\text{M2 : } Y = \cos(1.5X) + 0.2\epsilon,$$

$$\text{M3 : } Y = \exp(X - 0.2) + 0.2\epsilon,$$

où  $X$  et  $\epsilon$  sont des variables aléatoires indépendantes suivant respectivement la distribution exponentielle  $\mathcal{E}(1/3)$  et la distribution normale  $\mathcal{N}(0,1)$ . Nous prenons  $x = 3$  pour le modèle M1 et  $x = 0.5$  pour les modèles M2 et M3.

Nous modifions le paramètre CR en considérant différentes variables de cen-

sure. C'est à-dire que la variable de censure à gauche  $L$  est générée par une distribution exponentielle pour laquelle nous faisons varier le paramètre, alors que la variable de censure à droite  $R$  est générée par une distribution exponentielle à laquelle nous ajoutons la constante 0.01 pour assurer la validité de l'hypothèse  $\mathbf{H}_1$ .

Nous choisissons le noyau Epanechnikov pour  $K$  et  $K_0$  et nous prenons  $h_n = g_n$ . Nous utilisons une fenêtre optimale obtenue en minimisant la norme supremum entre les valeurs estimées et les vraies valeurs de la densité conditionnelle, sur un ensemble de valeurs de fenêtre dans  $[0, 1]$ . Le supremum sur  $y$  est pris sur l'intervalle  $[0, 4]$ .

La mesure de la distance entre les valeurs théoriques et les valeurs estimées est calculée en fonction de CR et  $n$ .

Pour chaque valeur de CR et  $n$ , nous répétons 200 fois la simulation et prenons la médiane sur  $y \in [0, 4]$  de l'erreur moyenne quadratique.

Plus précisément, le tableau ci-dessous fournit les valeurs

$$\underset{y \in [0,4]}{\text{med}} \frac{1}{200} \sum_{i=1}^{200} [f_{n,i}^{\hat{x}}(y) - f^x(y)]^2$$

| CR(%) | n=100 |       |       | n=200 |       |       | n=500 |       |        |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
|       | M1    | M2    | M3    | M1    | M2    | M3    | M1    | M2    | M3     |
| 0%    | 0.029 | 0.024 | 0.013 | 0.019 | 0.012 | 0.009 | 0.013 | 0.010 | 0.0087 |
| 30%   | 0.054 | 0.043 | 0.027 | 0.025 | 0.024 | 0.017 | 0.023 | 0.020 | 0.012  |
| 60%   | 0.102 | 0.091 | 0.050 | 0.073 | 0.043 | 0.035 | 0.040 | 0.023 | 0.030  |

Il résulte de ce tableau que la qualité de l'estimation s'améliore lorsque la taille de l'échantillon augmente ou lorsque le taux de censure diminue, ce qui est tout à fait prévisible.

Nous illustrons ces résultats par les figures suivantes qui montrent également que tous les graphes de l'estimateur présentent une forme acceptable par rapport aux véritables courbes.

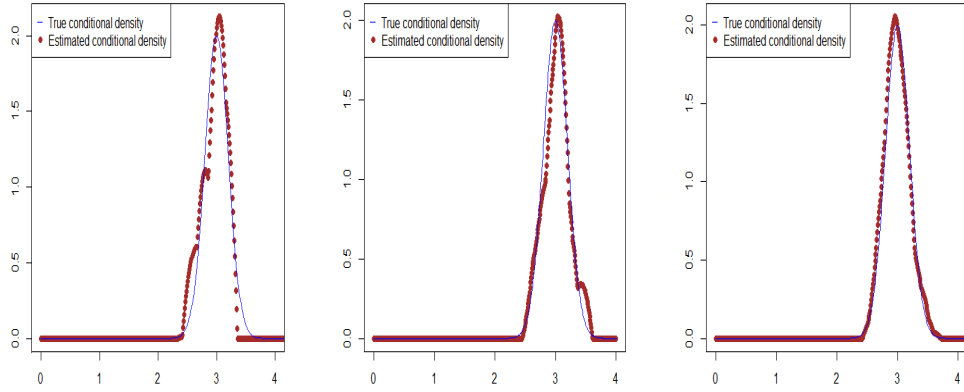


FIGURE 4.1 – Modèle M1, de gauche à droite  $n = 100, 200$  et  $500$  et CR à peu près 30%.

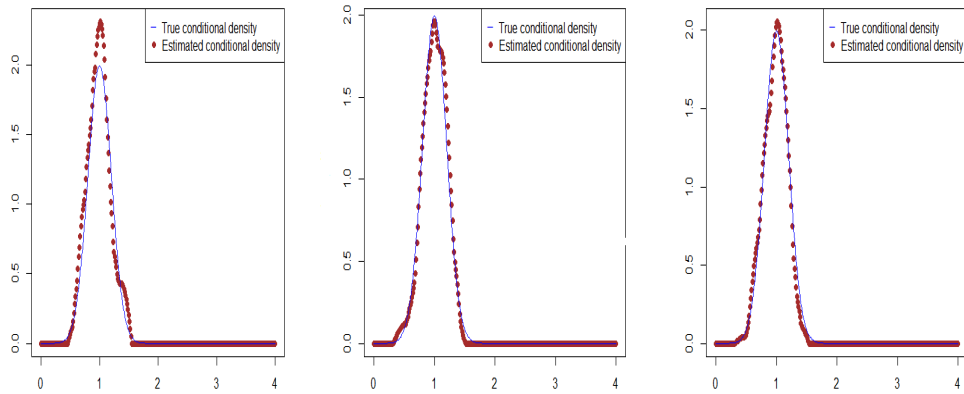


FIGURE 4.2 – Modèle M2, de gauche à droite  $n = 100, 200$  et  $500$  et CR à peu près 30%.

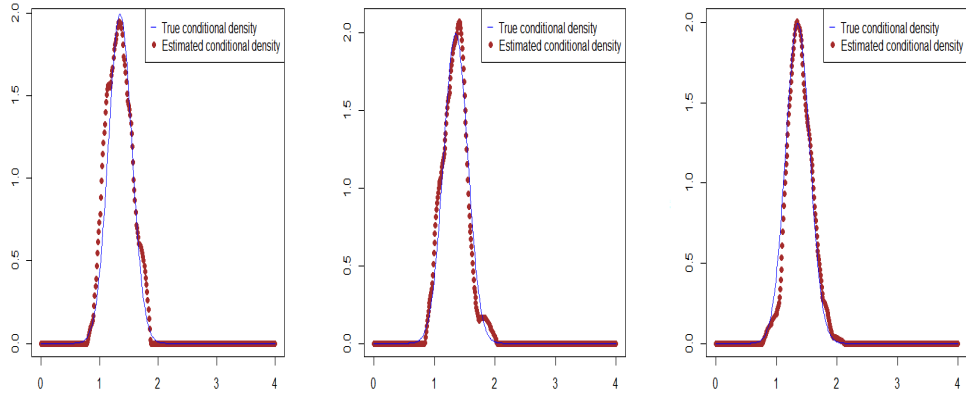


FIGURE 4.3 – Modèle M3, de gauche à droite  $n = 100, 200$  et  $500$  et CR à peu près 30%.

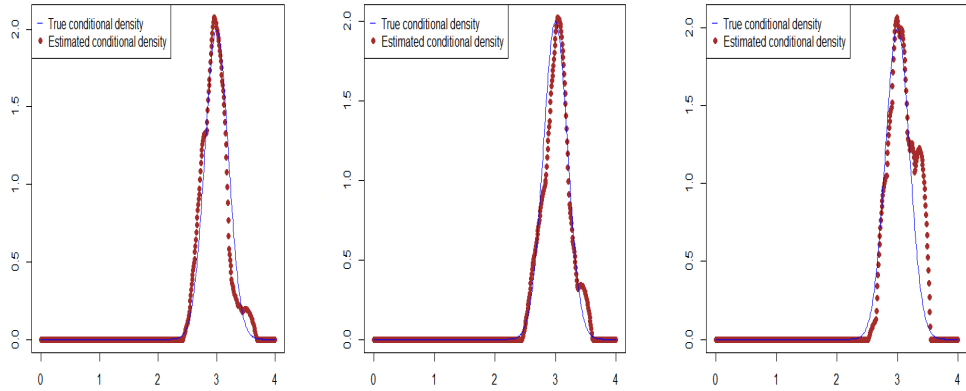


FIGURE 4.4 – Modèle M1, de gauche à droite CR à peu près 00%, 30% et 60% et  $n = 200$ .

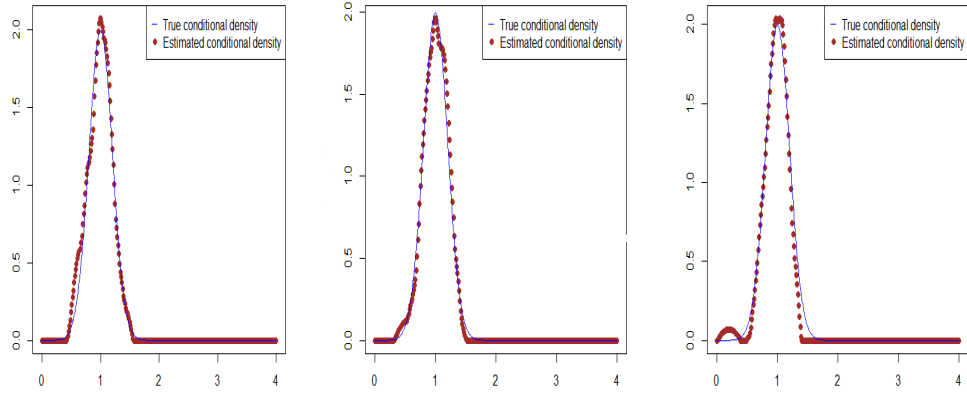


FIGURE 4.5 – Modèle M2, de gauche à droite CR à peu près 00%, 30% et 60% et  $n = 200$ .

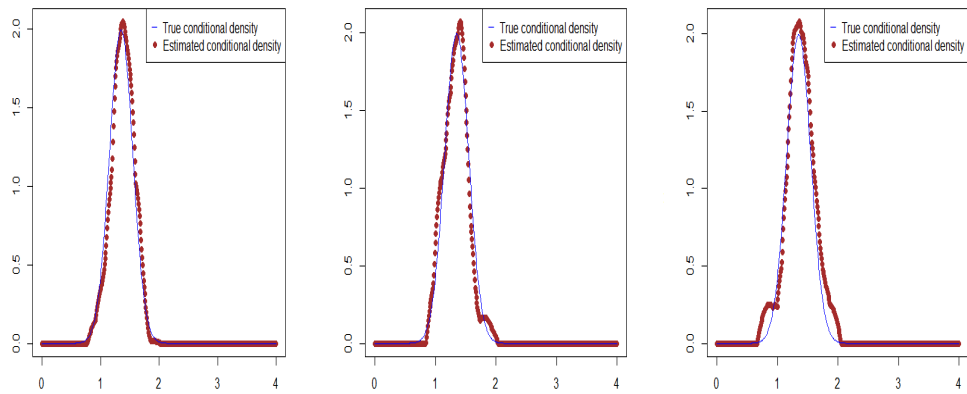


FIGURE 4.6 – Modèle M3, de gauche à droite CR à peu près 00%, 30% et 60% et  $n = 200$ .

## 4.2 Estimateur du mode conditionnel

Dans cette section, nous considérons deux modèles parmi ceux étudiés à la section précédente tout en conservant les mêmes paramètres afin de tester la performance de l'estimateur du mode conditionnel, donné par

$$\hat{\theta}_n(x) = \arg \sup_{y \in \mathcal{S}} \hat{f}_n^x(y).$$

Nous présentons l'effet de la taille de l'échantillon ainsi que celui du taux de censure pour les deux modèles suivants.

M1 :  $Y = X + 0.2\epsilon$

et

M3 :  $Y = \exp(X - 0.2) + 0.2\epsilon$ .

Nous évaluons l'erreur quadratique moyenne de l'estimateur du mode conditionnel étudié.

Pour chaque valeur de CR et  $n$ , nous répétons 200 fois la simulation et prenons la médiane sur  $x \in [0, 4]$  pour le modèle M1 et sur  $x \in [0, 1.5]$  pour le modèle M3, de l'erreur moyenne quadratique.

Plus précisément, le tableau ci-dessous fournit les valeurs

$$\mathit{med}_{x \in [0,4]} \frac{1}{200} \sum_{i=1}^{200} [\hat{\theta}_{n,i}(x) - \theta(x)]^2$$

et

$$\mathit{med}_{x \in [0,1.5]} \frac{1}{200} \sum_{i=1}^{200} [\hat{\theta}_{n,i}(x) - \theta(x)]^2$$

pour les modèles M1 et M3 respectivement.

| CR(%) | n=200 |       | n=300 |       | n=500 |       |
|-------|-------|-------|-------|-------|-------|-------|
|       | M1    | M2    | M1    | M2    | M1    | M2    |
| 00%   | 0.023 | 0.016 | 0.013 | 0.014 | 0.012 | 0.010 |
| 30%   | 0.025 | 0.022 | 0.017 | 0.019 | 0.014 | 0.011 |
| 60%   | 0.070 | 0.027 | 0.040 | 0.025 | 0.016 | 0.015 |

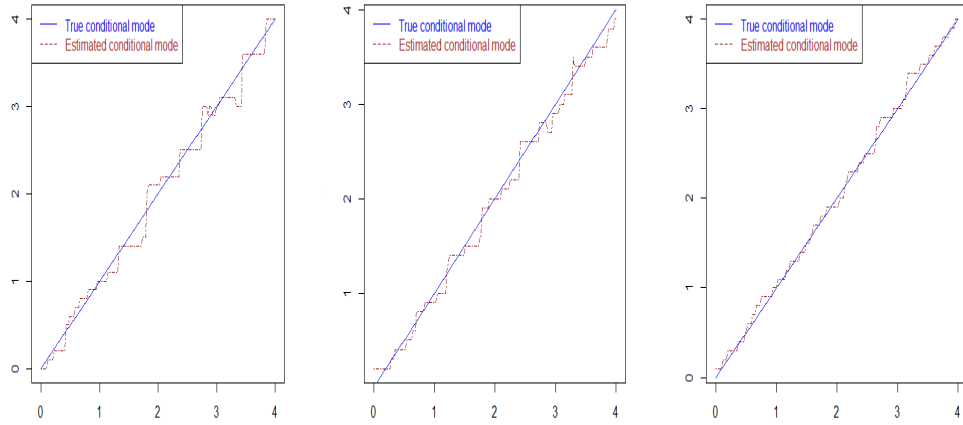


FIGURE 4.7 – Modèle M1, de gauche à droite  $n = 200, 500$  et  $800$  et CR à peu près 30%.

Les figures 4.7, 4.8, 4.9 et 4.10 présentent une bonne qualité d'ajustement au fur et à mesure que les taux de censure diminuent et que les tailles des échantillons augmentent, ce qui est prévisible et est en conformité avec les résultats obtenus au tableau précédent.



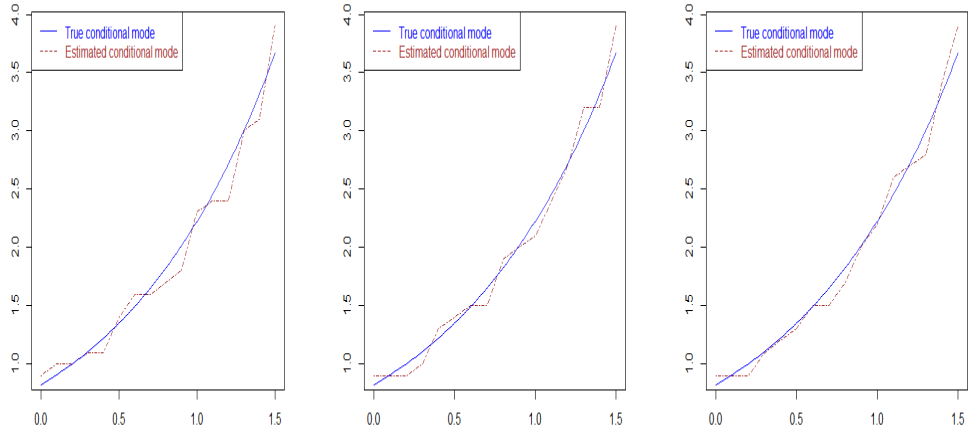


FIGURE 4.8 – Modèle M3, de gauche à droite  $n = 200, 500$  et  $800$  et CR à peu près 30%.

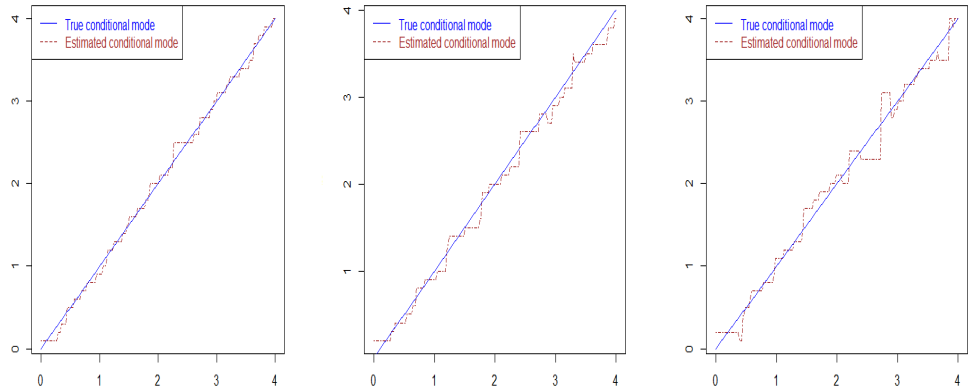


FIGURE 4.9 – Modèle M1, de gauche à droite CR à peu près 00%, 30% et 60% et  $n = 500$ .

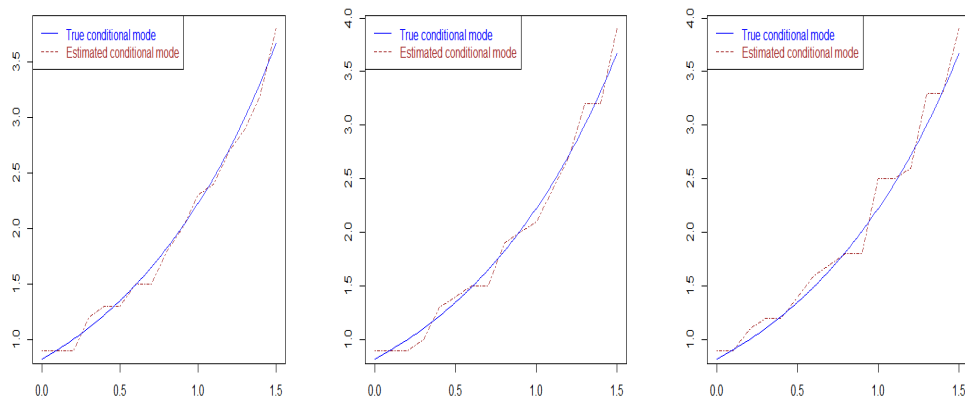


FIGURE 4.10 – Modèle M3, de gauche à droite CR à peu près 00%, 30% et 60% et  $n = 500$ .

# Annexe

Nous rappelons dans cette partie quelques outils et résultats utilisés dans cette thèse.

*Théorème 4. Soient  $g$  et  $K$  deux fonctions intégrables, avec  $K$  bornée et  $|z|K(z) \xrightarrow{|z| \rightarrow \infty} 0$ , posons :*

$$g_n(x) = \frac{1}{h_n} \int K\left(\frac{z}{h_n}\right)g(x-z)dz,$$

où  $(h_n)_{n \in \mathbb{N}}$  est une suite de nombres réels strictement positifs qui converge vers 0.

*Si  $g$  est continue au point  $x \in \mathbb{R}$ , alors  $g_n(x) \xrightarrow{n \rightarrow \infty} g(x) \int K(z)dz$ .*

*Si  $g$  est uniformément continue, alors la convergence de  $g_n$  est uniforme.*

*Démonstration.* . Soit  $\epsilon > 0$ , on a :

$$\begin{aligned} \left| g_n(x) - g(x) \int_{\mathbb{R}} k(z)dz \right| &= \left| \frac{1}{h_n} \int k\left(\frac{z}{h_n}\right)g(x-z)dz - \int_{\mathbb{R}} g(x)k(z)dz \right| \\ &= \left| \int_{\mathbb{R}} k(y)g(x-h_n y)dy - \int_{\mathbb{R}} g(x)k(y)dy \right|. \end{aligned}$$

Comme  $g$  est continue,  $\exists \delta > 0 / |x-z| < \delta$  implique que  $|g(x) - g(z)| \leq \epsilon$ .

Donc

$$\begin{aligned}
\int_{\mathbb{R}} |(g(x - h_n y) - g(x))K(y)| dy &\leq \int_{|y| < \frac{\delta}{h_n}} |g(x - h_n y) - g(x)| K(y) dy + \int_{|y| \geq \frac{\delta}{h_n}} |g(x - h_n y) - g(x)| K(y) dy \\
&\leq \epsilon \int_{|y| < \frac{\delta}{h_n}} K(y) dy + \int_{|y| \geq \frac{\delta}{h_n}} |g(x - h_n y)| K(y) dy \\
&\quad + |g(x)| \int_{|y| \geq \frac{\delta}{h_n}} K(y) dy
\end{aligned}$$

Par l'absolue continuité de l'intégrale et le fait que  $h_n \xrightarrow[n \rightarrow \infty]{} 0 : \exists n_0 \in \mathbb{N} / \forall n \geq n_0$ ,

$$\int_{|y| \geq \frac{\delta}{h_n}} K(y) dy < \epsilon.$$

De plus, sur  $\{|y| \geq \frac{\delta}{h_n}\}$  on a  $|y| \frac{h_n}{\delta} \geq 1$ , d'où

$$\begin{aligned}
\int_{|y| \geq \frac{\delta}{h_n}} |g(x - h_n y)| K(y) dy &\leq \int_{|y| \geq \frac{\delta}{h_n}} |y| \frac{h_n}{\delta} |g(x - h_n y)| K(y) dy \\
&\leq \frac{1}{\delta} \sup_{|y| \geq \frac{\delta}{h_n}} |y| K(y) \int g(z) dz \xrightarrow[n \rightarrow \infty]{} 0.
\end{aligned}$$

□

Soit  $(\Omega, A, P)$  un espace de probabilité,  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles et  $X$  une variable aléatoire réelle définie sur  $\Omega$ . Nous rappelons ci dessous les modes de convergence utilisés dans ce travail .

## Convergence en moyenne quadratique

Si  $X_n \in L^2$  et  $X \in L^2$ , on dit que  $X_n$  converge en moyenne quadratique vers  $X$  si

$$E|X_n - X|^2 = \int |X_n - X|^2 dP \xrightarrow[n \rightarrow \infty]{} 0.$$

On note  $X_n \xrightarrow{m.q.} X$ .

## Convergence presque complète

La suite de v.a.r.  $(X_n)_{n \in \mathbb{N}}$  converge presque complètement vers la v.a.r.  $X$  si

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

On note  $X_n \xrightarrow{p.co.} X$ .

Soulignons le fait que la convergence presque complète n'implique pas la convergence en moyenne quadratique qui, à son tour, n'entraîne pas la convergence presque complète, comme le montre l'exemple suivant

*Exemple 6.* Soit l'espace de probabilité  $([0, 1], B([0, 1]), \lambda)$ , où  $B([0, 1])$  désigne la tribu borélienne sur  $[0, 1]$  et  $\lambda$  est la restriction de la mesure de Lebesgue à cet ensemble, considérons les suites  $(f_n)_{n \in \mathbb{N}} = n1_{[0, \frac{1}{n^2}]}$ ,  $(g_n)_{n \in \mathbb{N}} = 1_{[0, \frac{1}{n}]}$ .

On a

$$\begin{aligned} \sum_{n \in \mathbb{N}} \lambda(|f_n| > \epsilon) &= \sum_{n \in \mathbb{N}} \lambda\left(|n1_{[0, \frac{1}{n^2}]}| > \epsilon\right) \\ &= \sum_{n \in \mathbb{N}} \lambda\left(|1_{[0, \frac{1}{n^2}]}| > \frac{\epsilon}{n}\right) \\ &\leq \sum_{n \in \mathbb{N}} \lambda\left(1_{[0, \frac{1}{n^2}]}\right) \\ &\leq \sum_{n \in \mathbb{N}} \frac{1}{n^2} < +\infty \end{aligned}$$

donc  $(f_n)_{n \in \mathbb{N}}$  converge p.co vers 0,

mais

$$\begin{aligned} \lim_{n \rightarrow \infty} E f_n^2 &= \lim_{n \rightarrow \infty} \int_{\mathbf{R}} n^2 1_{[0, \frac{1}{n^2}]}(x) d\lambda(x) \\ &= \lim_{n \rightarrow \infty} n^2 \lambda\left([0, \frac{1}{n^2}]\right) \\ &= \lim_{n \rightarrow \infty} \frac{n^2}{n^2} = 1 \neq 0, \end{aligned}$$

donc  $(\mathbf{f}_n)_{n \in \mathbb{N}}$  **ne converge pas en m.q vers 0.**

On a

$$\begin{aligned} \sum_{n \in \mathbb{N}} \lambda(|g_n| > \epsilon) &= \sum_{n \in \mathbb{N}} \lambda\left(|1_{[0, \frac{1}{n}]}| > \epsilon\right) \\ &= \sum_{n \in \mathbb{N}} \lambda\left(|1_{[0, \frac{1}{n}]}| > \epsilon\right) \\ &= \sum_{n \in \mathbb{N}} \lambda\left(1_{[0, \frac{1}{n}]}\right) \\ &= \sum_{n \in \mathbb{N}} \frac{1}{n} = +\infty, \end{aligned}$$

donc  $(\mathbf{g}_n)_{n \in \mathbb{N}}$  **ne converge pas p.co vers 0,**

mais

$$\begin{aligned} \lim_{n \rightarrow \infty} E g_n^2 &= \lim_{n \rightarrow \infty} \int_{\mathbf{R}} 1_{[0, \frac{1}{n}]}(x) d\lambda(x) \\ &= \lim_{n \rightarrow \infty} \lambda\left([0, \frac{1}{n}]\right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \\ &= 0, \end{aligned}$$

donc  $(\mathbf{g}_n)_{n \in \mathbb{N}}$  **converge en m.q vers 0.**

Rappelons maintenant quelques propriétés de la convergence presque complète.

## Convergence presque complète et inégalités de Bernstein

Le concept de convergence presque complète a été introduit par Hsu et Robbins (1947). Elle implique la convergence presque sûre et se prête bien aux calculs faisant intervenir des sommes de variables aléatoires. Malgré cela, elle ne commence à devenir populaire dans la communauté statistique que

dans les années 1980 après les travaux de Collomb. Elle est utilisée surtout en statistique non-paramétrique.

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

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

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

$$\exists \epsilon_0 > 0, \quad \sum_{n \in \mathbb{N}} P[|X_n - X| > \epsilon_0 u_n] < \infty.$$

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

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

Dans la partie qui suit, nous rappelons quelques propriétés relatives à la convergence presque complète et nous introduisons quelques inégalité exponentielles pour des sommes de variables aléatoires, utiles à la démonstration des propriétés de convergence presque complète.

Cette convergence et son taux possèdent les propriétés suivantes.

*Proposition 2. (cf Ferraty et Vieu (2006), Proposition A.5.)*

*Soit  $(u_n)$  une suite de nombres réels tels que  $\lim_{n \rightarrow \infty} u_n = 0$ , et soit  $l_x$  et  $l_y$  deux nombres réels telles que  $\lim_{n \rightarrow \infty} X_n = l_X$  p.co. et  $\lim_{n \rightarrow \infty} Y_n = l_Y$  p.co..*

*i) On a*

$$(a) \quad \lim_{n \rightarrow \infty} (X_n + Y_n) = l_X + l_Y, \text{ p.co.,}$$

- (b)  $\lim_{n \rightarrow \infty} (X_n Y_n) = l_X l_Y$ , p.co.,
- (c)  $\lim_{n \rightarrow \infty} \frac{1}{X_n} = \frac{1}{l_X}$  p.co. lorsque  $l_X \neq 0$ .
- ii) Si  $X_n - l_X = O_{a.co.}(u_n)$  et  $Y_n - l_Y = O_{p.co.}(u_n)$ , alors
- (a)  $(X_n + Y_n) - (l_X + l_Y) = O_{p.co.}(u_n)$ ,
- (b)  $(X_n Y_n) - (l_X l_Y) = O_{p.co.}(u_n)$ ,
- (c)  $\frac{1}{X_n} - \frac{1}{l_X} = O_{p.co.}(u_n)$  lorsque  $l_X \neq 0$ .

*Démonstration.* i.a) La preuve découle immédiatement de l'inégalité suivante

$$P(|(X_n + Y_n) - (l_X + l_Y)| > \epsilon) \leq P(|X_n - l_X| > \frac{\epsilon}{2}) + P(|Y_n - l_Y| > \frac{\epsilon}{2}).$$

ii.a) Il suffit d'appliquer ce résultat à  $\epsilon = \epsilon_0 U_n$ .

i.b) Sans perte de généralité, on pose  $l_X = 0$ . La décomposition suivante

$$X_n Y_n = Y_n (X_n - l_X) + Y_n l_X,$$

nous donne

$$\begin{aligned} P(|(X_n Y_n)| > \epsilon) &\leq P\left(|Y_n - l_Y| |X_n| > \frac{\epsilon}{2}\right) + P\left(|l_Y X_n| > \frac{\epsilon}{2}\right) \\ &\leq P\left(|Y_n - l_Y| > \sqrt{\frac{\epsilon}{2}}\right) + P\left(|X_n| > \sqrt{\frac{\epsilon}{2}}\right) + P\left(|X_n l_Y| > \frac{\epsilon}{2}\right) \\ &\leq P\left(|Y_n - l_Y| > \frac{\epsilon}{2}\right) + P\left(|X_n| > \frac{\epsilon}{2}\right) + P\left(|X_n l_Y| > \frac{\epsilon}{2}\right). \end{aligned}$$

L'inégalité précédente et la convergence presque complète de  $X_n$  et  $Y_n$  permettent d'écrire

$$\sum_{n \in \mathbb{N}} P(|X_n Y_n| > \epsilon) < \infty.$$

ii.b) Il suffit d'appliquer le résultat précédent à  $\epsilon = \epsilon_0 U_n$ .



i.c) La convergence presque complète de  $Y_n$  vers  $l_Y$  implique l'existence de  $\delta \geq 0$  ( $\delta = \frac{l_Y}{2}$ ) tel que :

$$\sum_{n \in \mathbb{N}} P(|Y_n| \leq \delta) < \infty, \quad (4.1)$$

de plus nous avons

$$\begin{aligned} P\left(\frac{1}{Y_n} - \frac{1}{l_Y} > \epsilon\right) &= P(|Y_n - l_Y| > \epsilon |l_Y Y_n|) \\ &\leq P(|Y_n - l_Y| > \epsilon |l_Y Y_n|, |Y_n| > \delta) + P(|Y_n| \leq \delta) \\ &\leq P(|Y_n - l_Y| > \epsilon \delta |l_Y|) + P(|Y_n| \leq \delta). \end{aligned}$$

En utilisant la relation (4.1) et la définition de la convergence presque complète de  $Y_n$  vers  $l_Y$ , il vient

$$\forall \epsilon > 0, \quad \sum_{n \in \mathbb{N}} P\left(\left|\frac{1}{Y_n} - \frac{1}{l_Y}\right| \geq \epsilon\right) < \infty.$$

ii.c) On procède de la même manière pour  $\epsilon = \epsilon_0 U_n$ .

□

*Proposition 3. (cf Ferraty et Vieu (2006), Proposition A.6.)*

On suppose que  $\lim_{n \rightarrow \infty} u_n = 0$ , Si  $X_n = O_{p.co.}(u_n)$  et  $\lim_{n \rightarrow \infty} Y_n = l_Y$  p.co., alors

- i)  $X_n Y_n = O_{p.co.}(u_n)$ ,
- ii)  $\frac{X_n}{Y_n} = O_{p.co.}(u_n)$ , lorsque  $l_Y \neq 0$ .

*Démonstration.* i) La définition de la convergence presque complète de  $Y_n$  vers  $l_Y$  implique l'existence de  $\delta > 0$  tel que

$$\sum_{n \in \mathbb{N}} P[|Y_n| > \delta] < \infty.$$

La décomposition suivante

$$\begin{aligned} P(|Y_n X_n| > \epsilon U_n) &= P(|Y_n X_n| > \epsilon U_n, |Y_n| \leq \delta) + P(|X_n Y_n| > \epsilon U_n, |Y_n| > \delta) \\ &\leq P(|X_n| > \epsilon \delta^{-1} U_n) + P(|Y_n| > \delta), \end{aligned}$$

associée à l'inégalité précédente et à l'hypothèse  $X_n = O_{a.co}(U_n)$ , conduisent à  $X_n Y_n = O_{a.co}(U_n)$ .

- ii) La preuve est un résultat direct de i) de cette proposition et de la partie i.c) de la proposition 2.  $\square$

## Inégalités de type Bernstein

Soit  $\{X_n, n \geq 1\}$  une suite de variables aléatoires indépendantes, identiquement distribuées et centrées. Pour démontrer la convergence presque complète, nous avons besoin de trouver des bornes supérieures pour certaines probabilités concernant des sommes de variables aléatoires de la forme

$$P\left(\left|\sum_{i=1}^n Z_i\right| > \epsilon\right)$$

où éventuellement  $\epsilon$  décroît avec  $n$ .

Dans ce contexte, il existe de puissants outils probabilistes appelés inégalités exponentielles.

On en trouve différentes versions dans la littérature. Les inégalités diffèrent selon les hypothèses imposés aux variables aléatoires  $Z_i$ . Nous en présentons ici celles qu'on appelle inégalités de type Bernstein dont la forme convient le plus à notre travail.

Supposons que  $\{X_n, n \geq 1\}$  est une suite de variables aléatoires réelles, indépendantes et centrées.

*Proposition 4. (cf Ferraty et Vieu (2006), Proposition A.7. )*

Si

$$\forall m \geq 2, |E(X_i^m)| \leq \left(\frac{m!}{2}\right) (a_i)^2 b^{m-2},$$

alors

$$\forall \epsilon \geq 0, \quad P \left[ \sum_{i=1}^n |X_i| > \epsilon A_n \right] \leq 2 \exp \left\{ \frac{-\epsilon^2}{2(1 + \frac{b\epsilon}{A_n})} \right\},$$

où  $(a_i)_{1 \leq i \leq n}$  sont des réels positifs,  $b \in \mathbb{R}^+$  et  $A_n^2 = a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2$ .

*Démonstration.* cf. Bernstein (1946); Yurinskii (1976) □

*Corollaire 1. (cf Ferraty et Vieu (2006), Corollaire A.9. )*

1. S'il existe une constante positive  $M < \infty$ , telle que  $|X_1| \leq M$ , alors on a

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

où  $\sigma^2 = EX_i^2$ .

2. Supposons que les  $(X_i)_{1 \leq i \leq n}$  dépendent de  $n$  et que  $\sigma_n^2 = EX_i^2$ , s'il existe  $M = M_n < \infty$  tel que  $|X_1| \leq M$ , si  $\frac{M}{\sigma_n^2} \leq C < \infty$  et si  $u_n = n^{-1} \sigma_n^2 \log n$ , vérifie  $\lim_{n \rightarrow \infty} u_n = 0$ , alors nous avons

$$\frac{1}{n} \sum_{i=1}^n X_i = O_{a.co}(\sqrt{u_n}).$$

*Démonstration.* 1. En appliquant la proposition 4 à  $a_i^2 = \sigma^2$ ,  $A_n^2 = n\sigma^2$  et  $b = M$  nous aboutissons à 1).

2. Comme  $\frac{MU_n}{\sigma_n^2}$  tend vers zéro, il suffit de reprendre le résultat 1) pour  $\epsilon = \epsilon_0 \sqrt{u_n}$ , on arrive donc à l'existence d'une constante  $C'$  telle que

$$\begin{aligned} P \left[ \frac{1}{n} \left| \sum_{i=1}^n X_i \right| > \epsilon_0 U_n \right] &\leq 2 \exp \left\{ \frac{-\epsilon_0^2 \log n}{2 \left( 1 + \epsilon_0 \sqrt{\frac{MU_n}{\sigma_n^2}} \right)} \right\} \\ &\leq 2n^{-C'} \epsilon_0^2. \end{aligned}$$

Pour  $\epsilon_0$  bien choisi le terme de droite est le terme général d'une série convergente. Ainsi s'achève la preuve de ce corollaire.  $\square$

# Perspectives

Nous terminons par quelques perspectives de recherche qui peuvent faire l'objet de futurs travaux.

- Montrer des résultats de convergence en moyenne quadratique similaires à ceux du chapitre 2 dans le cas de la censure double.
- Etablir la normalité asymptotique d'estimateurs de la densité et du mode conditionnels dans un modèle de censure mixte et double.
- On peut aussi étudier le prolongement de nos résultats au cas de données fonctionnelles (lorsque la covariable est dans un espace de dimension infinie).

# Bibliographie

- Andersen P.K., Borgan O., Gill. R.D. and Keiding N., 1992. Statistical Models Based on Counting Processes, *Springer Verlag*.
- Aouicha, L., Messaci, F., 2019. Kernel estimation of the conditional density under a censorship model, *Statistics & Probability Letters* 145, 173–180.
- Arora, S., Taylor, J. W., 2016. Forecasting electricity smart meter data using conditional kernel density estimation. *Omega* 59, 47–59.
- Berlinet, A., Gannoun, A. et Matzner-Lober, E., 1998. Normalité asymptotique d'estimateurs convergent du mode conditionnel. *Canad. J.Statist.*, 26 : 365–380.
- Bernstein, S., 1946. Probability Theory, 4th ed. (in russian). *M. L. Gostechizdat*.
- Bitouzé, D., Laurent, B., et Massart, P., 1999. A Dvoretzky-Kiefer-Wolfowitz type inequality for the Kaplan-Meier estimator. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 35(6), 735–763.
- Bochner, S., 1955. Harmonic analysis and the theory of probability. *Univ. of California Press*.
- Bosq, D., 1971. Contribution à la Théorie de l'Estimation Fonctionnelle. PhD thesis, Paris VI.

- Bosq, D., 1973. Sur l'estimation de la densité d'un processus stationnaire et mélangeant. *C. R. Acad. Sci. Paris Sér. A-B*, 277 :A535–A538.
- Boukeloua, M., 2015. Rates of mean square convergence of density and failure rate estimators under twice censoring. *Statistics & Probability Letters* 106, 121–128.
- Cacoullos, T., 1966. Estimation of a multivariate density. *Ann. Inst. Statist. Math.*, 18, 179–189.
- Cantelli, F. P., 1933. Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari*, 4, 421–424.
- Chahboun, M., 1984. Estimation des densités de la loi et les probabilités de transition par les méthodes du noyau, des  $\delta$ -suites et des fonctions orthogonales pour un processus de Markov stationnaire et mélangeant. PhD thesis, Université de Rouen.
- Chung, K. L., 1949. An estimate concerning the Kolmogoroff limit distribution. *Transactions of American Mathematical Society*, 67(1), 36–50.
- Collomb, G., Hardle, W., Hassani, S., 1987. A note on prediction via estimation of the conditional mode function. *Journal of Statistical Planning and Inference* 15, 227–236.
- Deheuvels, P. 1974. Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité. *C.R. Acad. Sci. Paris Ser. I*, 278 : 1217–1220.
- Deheuvels, P., 1991. Laws of the iterated logarithm for density estimators. In G. Roussas, éditeur : *Nonparametric Functional estimators and Related Topics*, NATO adv.Sci.Ins. C, 19–29.

- Delecroix, M., 1975. Sur l'estimation des densités marginales et de transition d'un processus stationnaire et mélangeant. PhD thesis, Université de Lille.
- Dvoretzky, A., Kiefer, J., et Wolfowitz, J., 1956. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 642–669.
- Efromovich, S., 2007. Optimal nonparametric estimation of the density of regression errors with finite support. *Ann. Inst. Statist. Math.*, 59(4) :617–654.
- Einbeck, J. and Tutz, G., 2006. Modelling beyond regression functions : an application of multimodal regression to speed-flow data. *J. Roy. Statist. Soc. Ser. C*, 55(4) : 461–475.
- Ezzahrioui, M., Ould Saïd, E., (2005). Asymptotic normality of nonparametric estimators of the conditional mode for functional data. Technical report, No.249, LMPA, Univ. Littoral Côte d'Opale.
- Ezzahrioui, M., Ould Saïd, E. (2006). On the asymptotic properties of a nonparametric estimator of the conditional mode for functional dependent data. Preprint, LMPA No 277, Univ. du Littoral Côte d'Opale.
- Fan, J. et Yim, T. H., 2004. A crossvalidation method for estimating conditional densities. *Biometrika*, 91(4) :819–834.
- Ferraty, F., Laksaci, A. et Vieu, P., 2006. Functional time series prediction via conditional mode estimation. *C. R. Math. Acad. Sci. Paris*. 340, 389–392.
- Ferraty, F., Vieu, P., 2006. Nonparametric Functional Data Analysis, Theory and Practice. *Springer Series in Statistics*. Springer, New York.



- Ferraty, F., Laksaci, A. et Vieu, P., 2006. Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. Stoch. Proc.*, 9 : 47–76.
- Földes, A., et Rejtő, L., 1981a. A LIL type result for the product limit estimator. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 56, 75–86.
- Földes, A., et Rejtő, L., 1981b. Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *The Annals of Statistics.*, 9, 122–129.
- Földes, A., Rejtő, L., et Winter, B. B., 1980. Strong consistency properties of nonparametric estimators for randomly censored data, I : The product-limit estimator. *Periodica Mathematica Hungarica*, 11, 233–250.
- Földes, A., Rejtő, L., et Winter, B. B., 1981. Strong consistency properties of nonparametric estimators for randomly censored data, II : Estimation of density and failure rate. *Periodica Mathematica Hungarica*, 12(1), 15–29.
- Gannoun, A. et Saracco, J., 2002. A new proof of strong consistency of kernel estimation of density function and mode under random censorship. *Statist. Probab. Lett.*, 59 : 61–66.
- Glivenko, V., 1933. Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari*, 4, 92–99.
- Hall, P., Racine, J., and Li, Q., 2004. Cross-validation and the estimation of conditional probability densities. *J. Amer. Statist. Assoc.*, 99(468) :1015–1026.
- Han-Ying Liang, a,b., Ai-Ai Liu, a. (2013). Kernel estimation of conditional density with truncated, censored and dependent data. *Journal of Multivariate Analysis* 120 (2013) 40–58.

- Hsu, P. L., Robbins, H., 1947. Complete Convergence and the Law of Large Numbers. *Proceedings of the national academy of sciences* 33 (2).
- Hyndman, R. J., Bashtannyk, D. M., Grunwald, G. K., 1996. Estimating and visualizing conditional densities. *J. Comput. Graph. Statist.* 5 (4), 315–336.
- Hyndman, R. J. and Yao, Q., 2002. Nonparametric estimation and symmetry tests for conditional density functions. *J. Nonparametr. Stat.*, 14(3) :259–278.
- Kaplan, E. L., Meier, P., 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53, 457–481.
- Kebabi, K., Messaci, F., 2012. Rate of the almost complete convergence of a kernel regression estimate with twice censored data. *Statistics & Probability Letters* 82 (11), 1908–1913.
- Kebabi, K., Laroussi, I., Messaci, F., 2011. Least squares estimators of the regression function with twice censored data. *Statistics & Probability Letters* 81, 1588–1593.
- Khardani, S., Lemdani, M., Ould Saïd, E., 2010. Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. *Journal of the Korean Statistical Society* 39 (4), 455–469.
- Khardani, S., Semmar, S., 2014. Nonparametric conditional density estimation for censored data based on a recursive kernel. *Electronic Journal of Statistics* 8, 2541–2556.
- Kiefer, J. (1961). On large deviations of the empiric df of vector chance variables and a law of the iterated logarithm. *Pacific J. Math*, 11(3), 649–660.

- Kitouni, A. and Boukeloua, M. and Messaci, F., 2015. Rate of strong consistency for nonparametric estimators based on twice censored data. *Statistics & Probability Letters* 96, 255–261.
- Klein, J.P. et Goel. P.K., 1991. Survival Analysis : State of the Art, NATO ASI Series, serie E : *Applied Sciences*, Vol. 211.
- Laksaci, A., 2007. Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Ann. I.S.U.P.*, 51(3) :69–80.
- Laksaci, A. et Yousfate, A., 2002. Estimation fonctionnelle de la densité de l'opérateur de transition d'un processus de Markov à temps discret. *C.R. Math., Acad. Sci. Paris, Sér. I*, 334 : 1035–1038.
- Leiderman, P. H., Babu, D., Kagia, J., Kraemer, H. C. et Leiderman, G. F., 1973. African infant precocity and some social influences during the first year. *Nature*, 242, 247–249.
- Louani, D. et Ould Saïd, E., 1999. Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric Statist.*, 11 : 413–442.
- Matzner-Lober, E., Gannoun, A. et De Gooijer. J.G. ( IYYX). Nonparametric forecasting : A comparison of 17 methods
- Mehra, K. L., Ramakrishnaiah, Y. S. and Sashikala, P., 2000. Laws of iterated logarithm and related asymptotics for estimators of conditional density and mode. *Ann. Inst. Math. Stat.*, 52, 630–645.
- Messaci, F., 2010. Local averaging estimates of the regression function with twice censored data. *Statistics & Probability Letters* 80 (9), 1508–1511.

- Messaci, F., et Nemouchi, N., 2011. A law of the iterated logarithm for the product limit estimator with doubly censored data. *Statistics & Probability Letters*, 81(8), 1241–1244.
- Messaci, F., et Nemouchi, N., 2013. Erratum to “A law of the iterated logarithm for the product limit estimator with doubly censored data” [Statist. Probab. Lett. 81 (2011) 1241–1244].
- Morales, D., Pardo, L., Quesada, V., 1991. Bayesian survival estimation for incomplete data when the life distribution is proportionally related to the censoring time distribution. *Communications in Statistics -Theory and Methods* 20, 831–850.
- Ould Saïd, E., 1993. Estimation non paramétrique du mode conditionnel, application à la prévision. *C. R. Acad. Sci. Paris Sér. I Math.*, 316, 043–947.
- Ould Saïd, E. et Cai, Z., 2005. Strong uniform consistency of non parametric estimation of the conditional mode function. *J. Nonparametric. Statist.*, 17 : 797–806.
- Ould Saïd, E., 1997. A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. Statist.*, 24 : 231–239.
- Ould Saïd, E. et Tatachak, A., 2007. Asymptotic properties of the kernel estimator of the conditional mode for the left truncated model. *Statistics/Probability Theory*, 17 : 651–656.
- Parzen, E., 1962. On estimation of a probability density function and mode. *Annals of Mathematical Statistics* 33, 1065–1076.
- Patilea, V., Rolin, J.-M., 2006. Product-limit estimators of the survival function with twice censored data. *The Annals of Statistics* 34 (2), 925–938.

- Quintela-Del-Rio, A. et Vieu, P., 1997. A nonparametric conditional mode estimate. *J. Nonparametric Statist.*, 8 : 253–266.
- Rosa, A.C.M., 1993. Préviation robuste sous une hypothèse ergodique. thèse de Doctorat. Université de Toulouse I.
- Rosenblatt, M., 1956. Remarks on some nonparametric estimates of density function. *The Annals of Mathematical Statistics*, 27, 832–837.
- Rosenblatt, M., 1971. Curve estimates. *The Annals of Mathematical Statistics*, 42(6), 1815–1842.
- Rossi, V., 2004. Filtrage non linéaire par noyaux de convolution Application à un procédé de dépollution biologique. Thèse du Doctorat en Science, école Nationale Supérieure Agronomique de Montpellier.
- Roussas, G., 1968. On some properties of nonparametric estimates of probability density functions. *Bull. Soc. Math. Grèce (N.S.)*, 9 : 29–43.
- Salha, R. et El Shekh Ahmed, H., 2009. On the Kernel Estimation of the Conditional Mode. *Asian Journal of Mathematics and Statistics* 2(1), 1–8.
- Samanta, M. et Thavasncswaran. A., 1990. Non-parametric estimation of the conditional mode. *Comm. Statist. theory methods*, 19, 4515–4524.
- Schuster, E. F., 1969. Estimation of a probability density and its derivatives. *Ann. Math. Statist.*, 40, 1187–1196.
- Scott, D. W., 1992. Multivariate Density Estimation : Theory, Practice and Visualization. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. *A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.*

- Silverman, B. W., 1978. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, 6(1), 177–184.
- Stute W., et Wang J.-L., 1993. The strong law under random censorship. *The Annals of Statistics*, 21(3), 1591–1607.
- Turnbull, B. W., 1974. Nonparametric Estimation of a Survivorship Function with Doubly Censored Data. *Journal of the American Statistical Association* 69 (345), 169–173.
- Turnbull, B. W., 1976. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(3) :290–295.
- Van Ryzin, J., 1969. On strong consistency of density estimates. *Ann. Math. Statist.*, 40, 1765–1772.
- Vieu, P., 1996. A note on density mode estimation. *Statist. Probab. Lett.*, 26, 297–307.
- Wen, K., Wu, X., 2017. Smoothed kernel conditional density estimation. *Economics Letters* 152, 112–116.
- Winter, B. B., Földes, A., et Rejtő, L., 1978. Glivenko-Cantelli theorems for the product limit estimate. *Problems of Control and Information Theory*, 7, 213–225.
- Youndjé, E., 1993. Estimation non paramétrique de la densité conditionnelle par la méthode de noyau. Thèse de Doctorat. Université de Rouen.
- Yurinskii, V., 1976. Exponential inequalities for sums of random vectors. *Journal of Multivariate Analysis* 6, 475–499.

# Abstract

In this thesis, we are interested in the nonparametric estimation of the conditional density and the conditional mode, in a twice censorship model. This means that the response variable  $Y$  is right censored by a variable  $R$  and that  $\min(Y, R)$  is left censored.

On the one hand, we build estimators, by the kernel method, for the conditional density and the conditional mode and establish a rate of the almost complete convergence for them.

On the other hand, we give the rate of the mean square convergence of the conditional density estimator.

Finally, a simulation study is conducted in order to illustrate, for a finite size, the performance of the proposed estimators.

**Keywords** : Censored data, Conditional density, Conditional mode, Mean square error, Almost complete convergence, Nonparametric estimation, Convergence rate, Simulation.

# Résumé

Dans cette thèse, nous nous intéressons à l'estimation non paramétrique de la densité conditionnelle et du mode conditionnel dans un modèle de censure mixte. Ce qui veut dire que la variable réponse  $Y$  est censurée à droite par une variable  $R$  et que  $\min(Y, R)$  est censuré à gauche.

Nous construisons des estimateurs, par la méthode du noyau, pour la densité et le mode conditionnels, pour lesquels nous établissons la vitesse de convergence presque complète.

Par ailleurs, nous étudions la convergence en moyenne quadratique, avec taux, de l'estimateur de la densité conditionnelle introduit.

Finalement, une étude sur des données simulées permet d'illustrer la performance, à taille finie, des estimateurs proposés.

**Mots clefs :** Données censurées, Densité conditionnelle, Mode conditionnel, Erreur quadratique moyenne, convergence presque complète, Estimation non paramétrique, Taux de convergence, Simulation.