

N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT

Présentée par

Abdelkader BENKHALED

Spécialité : Mathématiques

Option : Probabilités - Statistique

Intitulée

**Estimation locale linéaire pour la fonction de densité
et ses dérivées pour des variables censurées**

Soutenue le : 28/10/2020

Devant le jury composé de :

Présidente :

Melle Amina Angelika BOUCHENTOUF Professeur à l'université de S.B.A.

Examineurs :

Mr Abbès RABHI

Professeur. à l'université de S.B.A.

Mr Abdeldjebbar KANDOUCI

Professeur. à l'université de Saida.

Directeur de thèse :

Mr Fethi MADANI

MCA. à l'université de Saida.

Année universitaire : 2019/2020

Dédicace

Je dédie ce travail à mes Parents et ma Famille

Remerciements

Je remercie, en premier lieu, notre Dieu qui nous a donné de la force pour effectuer le présent travail.

En second lieu, je tiens à adresser mes vifs remerciements à mon directeur de thèse **Dr. Madani Fethi** pour son aide, ses conseils et tout le temps qu'il a consacré pour mener à bien ce travail. Je salut en lui son savoir faire, sa patiences et ses connaissances dont il ma transmet.

Je tiens aussi à remercier vivement le **Professeur Amina Angelika Bouchentouf** d'accepter de présider le jury de soutenance de cette thèse ainsi que présence permanente durant la réalisation de ce projet.

Je remercie également les membres de jury examinateurs Mrs : le **Pr. Abbes Rabhi** et le **Pr. Abdeldjebbar Kandouci** de la confiance qu'ils m'accordent et de l'intérêt qu'ils témoignent pour ce travail en acceptant de faire partie du jury.

Mes remerciements vont également vers les membres du laboratoire de Modèles stochastiques, Statistique et Applications (LMSSA) et spécialement son directeur Professeur Abdeldjebbar Kandouci pour leurs soutiens permanent pour achever ce travail. Enfin, j'exprime mes reconnaissances à mes parents, ma famille et mes amis qui m'ont toujours encouragé au cours de la réalisation de ce mémoire

Résumé

Cette thèse est consacrée à l'étude de l'estimation non-paramétrique de la densité conditionnelle lorsque la variable réponse est censurée tandis que la variable explicative est fonctionnelle. Notre objectif est de proposer un estimateur local linéaire inspiré de celui de Madani et al [20]. Dans un premier temps, nous construisons un nouvel estimateur de la densité conditionnelle ainsi le mode conditionnel par l'approche "double noyaux" en minimisant un critère des moindres carrés pondéré par des poids de Kaplan- Meier. Nous établissons la convergence presque sûre de l'estimateur proposé, les propriétés asymptotiques sont obtenues sous des conditions assez générales telles, l'hypothèse de mélange forte et l'hypothèse de concentration de la mesure de probabilité de la variable explicative fonctionnelle. A titre illustratif, nous donnons des exemples d'applications sur des données simulées et réelles. Dans un second temps, sous les mêmes hypothèses que celles du modèle précédent, nous établissons la normalité asymptotique de l'estimateur construit. Nous obtenons des intervalles de confiance en utilisant la méthode du "plug-in" pour les paramètres inconnus.

Mots clés : Densité conditionnelle, Estimateur locale linéaire, Données fonctionnel censuré, Normalité asymptotique, Données mélangeant, Probabilité des petites boules.

Summary

This thesis is devoted to the study of non-parametric estimation of conditional density when the response variable is censored while the explanatory variable is functional. Our objective is to propose a linear local estimator inspired by that of Madani et al [20]. First, we build a new estimator of the conditional density, thus the conditional mode by the "double kernel" approach by minimizing a least squares criterion weighted by Kaplan-Meier weights. We establish the almost sure convergence of the proposed estimator, the asymptotic properties are obtained under fairly general conditions such as the strong mixing hypothesis and the concentration hypothesis of the probability measure of the functional explanatory variable. As an illustration, we give examples of applications on simulated and real data. Secondly, under the same assumptions as those of the previous model, we establish the asymptotic normality of the constructed estimator. We obtain confidence intervals using the "plug-in" method for the unknown parameters.

Keywords : Conditional density, Local linear estimator, Functional censored data, Asymptotic normality, Mixing data, Small balls probability.

ملخص :

هذه الأطروحة مخصصة لدراسة التقدير غير البارامترى للكثافة الشرطية عندما يكون متغير الاستجابة خاضع للرقابة بينما المتغير التوضيحي وظيفي. هدفنا هو اقتراح مقدر محلي خطي مستوحى من مقدر مدني و الاخرين [٢٠].

أولاً ، نقوم ببناء مقدر كثافة شرطية جديد ووضع شرطي من خلال نهج النواة المزدوجة عن طريق تقليل معيار المربعات الصغرى الموزون بأوزان كابلان ماير. نحن نؤسس التقارب المؤكد تقريباً للمقدر المقترح ، ويتم الحصول على الخصائص المقاربة في ظل ظروف عامة إلى حد ما مثل فرضية الخلط القوية وفرضية التركيز لمقياس الاحتمال للمتغير التوضيحي الوظيفي. على سبيل التوضيح ، نقدم أمثلة تطبيقية على البيانات.

ثانياً ، وفقاً لنفس الافتراضات مثل تلك الخاصة بالنموذج السابق ، فإننا نؤسس الحالة الطبيعية المقاربة للمقدر المركب. نحصل على فترات الثقة باستخدام طريقة المكون الإضافي للمعلومات غير المعروفة.

كلمات مفتاحية: الكثافة الشرطية ، المقدر المحلي الخطي ، البيانات الوظيفية الخاضعة للرقابة ، الحالة الطبيعية المقاربة ، بيانات المزج ، احتمالية الكرات الصغيرة.

Liste des travaux

Publications dans des revues à comités de lecture

1. Strong consistency of local linear estimation of a conditional density function under random censorship.(Accepted in : Arabian Journal of Mathematics)
2. Asymptotic normality of the local linear estimation of the conditional density for functional dependent and censored data. (Accepted in : South African Statistical Journal)

Table des figures

5.1	A set of 300 simulated curves.	76
5.2	Distribution of the <i>MSE</i> obtained for different <i>CR</i> for $n = 300$	77
5.3	Estimates of the conditional hazard rate after the transplantation	80

Table des matières

1	Introduction	13
1.1	Généralités	13
1.2	Données Fonctionnelles : Etat de L'art	19
1.3	Brève Présentation des Résultats	21
1.3.1	Présentation de l'estimateur	21
1.3.2	Résultats sur la consistance du modèle	22
1.3.3	Résultats sur la normalité Asymptotique de l'estimateur	23
1.4	Le Plan de la Thèse	23
2	Concepts de Base de l'Analyse de Survie	26
2.1	Durée de Vie	26
2.2	Données Incomplètes	28
2.2.1	Données Tronquées	28
2.2.2	Données Censurées	29
2.3	Les fonctions d'intéret	31
2.3.1	L'estimateur de Kaplan-Meier	32
3	Functional data : Local linear estimation of a conditional density function under random censorship	35
3.1	Introduction	36
3.2	Construction of the conditional density estimator	37
3.3	Assumptions and main results	40
3.3.1	Almost sure consistency of the conditional density estimator .	40

3.3.2	Almost sure consistency of the conditional mode estimator . . .	43
3.4	Proofs	43
4	Local linear estimation of the conditional density for functional dependent and censored data : The Asymptotic Normality	54
4.1	Introduction	55
4.2	The model	57
4.3	Notations and hypotheses	58
4.4	Main results : Asymptotic normality	61
4.5	Appendix	62
5	Computational studies	75
5.1	On simulated data	75
5.2	Real data application	79
	Conclusion et Perspectives	80
	Bibliographie	83

Introduction

Contents

1.1	Généralités	13
1.2	Données Fonctionnelles : Etat de L'art	19
1.3	Brève Présentation des Résultats	21
1.3.1	Présentation de l'estimateur	21
1.3.2	Résultats sur la consistance du modèle	22
1.3.3	Résultats sur la normalité Asymptotique de l'estimateur	23
1.4	Le Plan de la Thèse	23

1.1 Généralités

La statistique est une branche de la théorie de probabilité où le hasard intervient fortement. L'objectif d'une étude statistique est l'estimation de la loi d'une variable aléatoire (v.a) Y qui présente un phénomène aléatoire, sur la base d'un échantillon (Y_1, \dots, Y_n) . La théorie d'estimation joue un rôle important dans nombreuses domaines comme la médecine (la durée entre le déclenchement d'une maladie et la guérison...), la finance (la répartition des revenus d'une population...), l'économie (la distribution des instants de défaillance d'un matériel...), les prévisions météorologiques. Dans la littérature statistique, deux types d'approches sont omniprésentes :

◊ L'approche paramétrique, qui considère des modèles statistiques paramétriques. La loi de la v.a admet une forme générale connue qui dépend d'un ou plusieurs paramètres inconnus à estimer.

◊ L'approche non paramétrique, qui considère des modèles statistiques non paramétriques. Nos connaissances sur la loi de la v.a sont beaucoup plus floues, ce qui correspond d'ailleurs souvent à la réalité. Dans cette situation, il est naturel de vouloir estimer une des fonctions décrivant le modèle, soit généralement la fonction de répartition ou la densité dans le cas continu. Les estimateurs non paramétriques sont introduit par Rosenblatt [66] et Parzen [61] pour estimer une densité de probabilité, et par Nadaraya-Watson [60] pour estimer une fonction de régression, qui sont appelés les estimateurs à noyaux.

L'observation des variables fonctionnelles est devenue habituelle et occupe une place importante dans la recherche en statistique, par exemple grâce au développement d'instruments de mesure permettant d'observer des variables (dépendantes du temps ou de l'espace) à une résolution de plus en plus fine. Il semble alors naturel de supposer que les données sont en réalité des observations d'une v.a prenant des valeurs dans un espace fonctionnel. Il existe actuellement un grand nombre de domaines dans lesquels des données fonctionnelles sont collectées, comme la météorologie, la chimie quantitative, la biométrie, l'économétrie ou l'imagerie médicale. ce qui a conduit à l'extension des techniques statistiques à dimensions finies au données à dimensions infinies.

Historiquement, l'analyse statistique des données fonctionnelles remonte aux années soixante quand plusieurs études, dans différentes disciplines scientifiques, s'intéressaient aux données sous forme de courbes. Il faut dire que les premiers travaux peuvent être conférés à des météorologues et des chimistes (cf. Holmström [38] en climatologie, Deville [24] en économétrie et Kirkpatrick et Heckman [48] en génétique).

Par ailleurs, en statistique fonctionnelle, les premiers résultats sur l'analyse statistique des données fonctionnelles reviennent à la contribution de Ramsay et Silverman [63]. Cette monographie a permis, aux statisticiens, d'avoir une vision globale du traitement des données fonctionnelles en termes des techniques de régression,

de discrimination statistique ainsi que d'analyse factorielle. Plus généralement, il faut noter que la contribution de Ferraty et Vieu [31] peut être considérée comme déterminante dans le cadre nonparamétrique fonctionnel. Ces auteurs ont étudié les propriétés asymptotiques de plusieurs modèles nonparamétriques tels que l'opérateur de régression, la fonction de répartition conditionnelle, la densité conditionnelle, etc.

L'estimation nonparamétrique de la densité conditionnelle est un sujet qui a donné lieu à un grand nombre de travaux. Son champ d'application est très vaste et couvre divers domaines, comme l'analyse de la régression, des séries chronologiques et la théorie de la fiabilité. Il faut dire qu'en dimension finie, il existe une littérature abondante pour la densité conditionnels. Roussas [67] fut le premier à établir ses propriétés asymptotiques pour des données markoviennes, ainsi que sa convergence en probabilité. Youndjé [77] s'est intéressé à l'étude de la densité conditionnelle pour des données complètes indépendantes. Laksasi et Yousfate [54] ont établi, pour un processus markovien stationnaire, la convergence en norme L^p pour l'estimateur à noyau de la densité conditionnelle. Pour Hall et al. [35], la densité conditionnelle joue un rôle clef en statistique appliquée et particulièrement en économie, pour Fan and Yim [28], une densité conditionnelle offre le résumé le plus informatif de la relation entre variable dépendante et indépendante, enfin Efromovich, [27] dit que la densité conditionnelle de la variable dépendante sachant le prédicteur décrit l'association ultime entre le prédicteur et la variable dépendante.

Signalons, qu'en dimension infinie, la densité conditionnel a connu tout récemment un intérêt croissant, malgré le peu de résultats disponibles dans la littérature. Dans ce contexte, les premiers travaux ont été réalisés par Ferraty et al. [32]. Ils ont montré, sous des conditions de régularité, la convergence presque complète des estimateurs à noyau de la densité conditionnelle et du mode conditionnel et ont établi leurs vitesses de convergence. Notons aussi qu'une application de leurs résultats aux données issues de l'industrie agro-alimentaire a été présentée. Dans le même contexte, Dabo-Niang et al. [15] ont étudié un estimateur non paramétrique du mode de la densité d'une variable explicative à valeurs dans un espace vectoriel semi-normé, de dimension éventuellement finie. Ils ont établi la convergence

presque sûre avec une application de ce résultat au cas où la mesure de probabilité de la variable explicative vérifie une condition de concentration. On trouve aussi dans Dabo-Niang et Laksaci [14] l'étude d'un estimateur à noyau du mode de la distribution d'une variable réelle Y conditionnée par une variable explicative X , à valeurs dans un espace semi-métrique. Ils ont établi la convergence en norme L^p de l'estimateur et ils ont montré que les résultats asymptotiques établis sont liés aux probabilités des petites boules de la loi de la variable explicative ainsi que la régularité de la densité conditionnelle. Laksaci et al.[53] ont résolu la question du choix du paramètre de lissage dans l'estimation de la densité conditionnelle à variable explicative fonctionnelle. La détermination des termes dominants de l'erreur quadratique de l'estimateur à noyau de la densité conditionnelle a été obtenue par Laksaci [52].

Dans la statistique, il arrive que les données collectées possèdent une certaine forme de dépendance, Il en est ainsi pour certains processus comme les processus autorégressifs à moyenne mobile (ARMA), les processus (GAREH) ou les chaînes de Markov qui sont soumis à une certaine forme de dépendance mais sont asymptotiquement indépendants. Il existe plusieurs formes de dépendances, exprimées par des coefficients et définies comme suit : Le coefficient de mélange fort ou d' α -mélange introduit par Rosenblatt [66], le coefficient de régularité absolue ou de β -mélange introduit par Kolmogorov [51], le coefficient de mélange uniforme ou de Φ -mélange introduit par Ibragimov [43], le coefficient de Ψ -mélange introduit par Blum et al. [7] et le coefficient de ρ -mélange introduit par Hirschfeld [37] et Gebelein [34].

Parmi tous ces types de dépendances, l' α -mélange est la plus faible et la moins restrictive. Ce qui implique que toute suite de v.a's mélangeantes sera alors forcément α -mélange. Ceci explique donc que l'estimation pour des données mélangeantes est souvent restreinte dans la littérature à ce cas.

La modélisation locale linéaire est une approche statistique très importante dans la modélisation non paramétrique. Il s'agit d'une méthode d'estimation alternative à la méthode à noyau et qui a des avantages certains par rapport à cette dernière, notamment dans la partie biais de l'estimateur ou les problèmes aux bords. En effet, certains des problèmes que comporte la méthode du noyau, peuvent être diminués

lorsque la méthode par polynômes locaux est utilisée. Il est entre autres possible d'affirmer que la méthode par polynômes locaux est préférable.

L'approche locale linéaire fonctionnelle a fait l'objet d'une question ouverte dans la monographie de Ferraty et Vieu [31], et plusieurs auteurs ont essayé d'y apporter des réponses. Rappelons que l'approche locale linéaire est basée sur une approximation linéaire du modèle au voisinage de la variable observée. De plus, l'estimation non paramétrique dans l'analyse statistique, non paramétrique, des données fonctionnelles (NPFDA : Non Parametric Functional Data Analysis) repose sur une pondération des observations les plus proches de la variable observée. La première réponse a été donnée par Baïllo et Grané [2], qui ont donné la première version fonctionnelle de l'estimateur locale linéaire (L.L.E : Local Linear Estimator). Celle-ci est basée sur la structure hilbertienne de l'espace. Ils ont déterminé le biais et la variance de leur estimateur. Ensuite, Barrientos et al. [3] ont donné une version alternative, facile à utiliser même quand la covariable appartient à un espace semi-métrique. La version de ces derniers exploite bien la richesse de la structure topologique de l'espace semi-métrique et elle considère deux approches différentes pour quantifier la notion de voisinage. Comme propriétés asymptotiques de Barrientos et al. [3], la démonstration de la convergence presque-complète de l'estimateur construit et l'obtention de sa vitesse de convergence. Celle-ci se présente comme fonction du degré de régularité du modèle non paramétrique et de la fonction de concentration de la variable fonctionnelle. Tous les résultats précités ont été obtenus dans le cas où les observations sont indépendantes identiquement distribuées (i.i.d : independent and identically distributed). Une troisième version a été abordée par Berlinet et al. [6]. Ces derniers ont utilisé l'inverse locale de l'opérateur de covariance pour construire un L.L.E de la fonction de régression ainsi que sa dérivée. Leurs résultats asymptotiques portent sur la convergence en moyenne quadratique de l'estimateur construit. Chouaf et Laksaci [13], quant à eux, ont généralisé les résultats dus à Barrientos et al. [3] au cas où les observations sont spatio-fonctionnelles. Ils ont également obtenu la convergence presque-complète et sa vitesse. En ce qui concerne les paramètres statistiques conditionnels, Demongeot et al. [21] ont étudié l'L.L.E de la densité conditionnelle dans le cas où les observations

sont i.i.d. Ils ont montré la convergence presque-complète ponctuelle puis uniforme de l'estimateur construit en précisant, pour chaque mode, sa vitesse de convergence pour les deux types de convergence : la première est exprimée en fonction de l'hypothèse de concentration tandis que la deuxième est donnée en fonction de l'entropie de l'espace fonctionnel des observations. La convergence en moyenne quadratique du même estimateur a été obtenue par Rachdi et al. [62]. Dans ce dernier, les auteurs montrent la supériorité de l'L.L.E sur la méthode à noyau. Notons que tous ces résultats ont été obtenus dans le cas où les observations sont i.i.d et sous des conditions, considérées maintenant, classiques en NPFDA. Notons également que le cadre spatial a été abordé relativement à ce thème de la NPFDA. En effet, Laksaci et al. [53] ont donné la version spatiale de l'L.L.E de plusieurs modèles non paramétriques conditionnels, tels que la fonction de répartition conditionnelle, la densité conditionnelle, le mode conditionnel et la fonction de hasard conditionnelle. Ils ont montré également la convergence presque-complète de ces estimateurs. Messaci et al. [59] ont étudié l'L.L.E des quantiles conditionnels. Ils ont construit et étudié les propriétés asymptotiques d'un estimateur basé sur l'inverse de l'L.L.E de la fonction de répartition conditionnelle. Récemment, Demongeot et al [23] ont étudié l'L.L.E de l'opérateur de régression quand la variable explicative est dans un espace semi-métrique tandis que la variable réponse est hilbertienne, ils ont montré la convergence uniforme presque-complète et sa vitesse. Celle-ci est exprimée en fonction de la concentration de la variable explicative et de l'entropie.

Par ailleurs, il faut mentionner qu'il existe de nombreuses raisons pour les quelles il est difficile d'obtenir des données complètes dans les études impliquant des temps de survie. La censure et la troncature sont les deux causes de données incomplètes les plus répandues. La censure est un mécanisme qui empêche l'observation exacte du délai de survenue d'intérêt. On sait bien que ce délai appartient à un certain intervalle de temps. La troncature survient qu'on ne peut pas observer les individus de l'échantillon dont le délai de survenue appartient à un certain intervalle de temps, on observe donc un sous-échantillon. Dans ce cas les techniques classiques ne s'adaptent pas correctement aux données incomplètes.

1.2 Données Fonctionnelles : Etat de L'art

L'Analyse des Données Fonctionnelles (F.D.A : Functional Data Analysis) est une branche de la statistique qui est de plus en plus utilisée dans de nombreux domaines scientifiques appliqués tels que l'expérimentation biologique, la finance, la physique, etc. Une raison à cela est l'utilisation des nouvelles technologies de collecte de données qui augmentent le nombre d'observations dans un intervalle de temps. Valoriser et exploiter ces masses de données est un défi important pour produire de nouvelles connaissances. Il faut donc développer des méthodes et fournir des outils dédiés à un traitement systématique de ces données, avec une prise en compte adaptée de leur dimension temporelle.

En statistique, les analyses multivariées ont montré leur limite. En ce sens, certaines des raisons de cette inadéquation sont, par exemple, les suivantes : i) le nombre de pas d'observation p est plus grand que le nombre de réalisations n (cf. Hsing and Eubank ([40], p. 2)); ii) les réalisations ne sont pas observées sur les mêmes grilles de temps (cf. Sentürk and Müller [69]); iii) il y a de fortes corrélations temporelles au sein des réalisations (cf. Ferraty and Vieu ([31], p. 7)) et iv) la régularité et les dérivées des fonctions aléatoires observées jouent un rôle important dans l'étude des données (cf. Mas and Pumo [57]).

Une façon plus adaptée de traiter ce type de données est de les considérer comme des réalisations de processus stochastiques à temps continu (Bosq ([9], Ch 1)). Cela permet d'introduire les notions de données fonctionnelles et fonctions aléatoires. Les données fonctionnelles sont des échantillons de réalisations de fonctions aléatoires. Une fonction aléatoire représente une évolution, discrète ou à temps continu, d'une v.a. D'un point de vue mathématique, les fonctions aléatoires sont des fonctions mesurables définies sur un espace de probabilité avec des valeurs dans un espace de dimension infini (Ferraty and Vieu ([31], Ch 1)).

De plus, même si les données dont dispose le statisticien ne sont pas de nature fonctionnelle, celui-ci peut être amené à étudier des variables fonctionnelles construites à partir de son échantillon initial. Un exemple classique est celui où l'on observe

plusieurs échantillons de données réelles indépendantes et où l'on est ensuite amenés à comparer les densités de ces différents échantillons ou bien à considérer des modèles où elles interviennent (cf. Ramsay et Silverman, [64]). Dans le contexte particulier de l'étude des séries temporelles, l'approche introduite par Bosq [8] fait apparaître une suite de données fonctionnelles dépendantes qui modélisent la série chronologique observée. Cette approche consiste tout d'abord à considérer le processus non pas à travers sa forme discrétisée mais comme étant un processus à temps continu puis à le découper en un échantillon de courbes successives.

Rappelons que, le fléau de la dimension rend les vitesses de convergence très faibles. Une manière de tenter de remédier à cela est de chercher une topologie qui restitue de façon pertinente les proximités entre les données. Cela peut être fait, par exemple, à l'aide d'une semi-métrique de projection basée sur les composantes principales fonctionnelles, les décompositions selon une base de Fourier, d'ondelettes, de splines,... . Lorsque la variable explicative est à valeurs dans un espace de Hilbert séparable, Ferraty et Vieu ([32], Lemme 13-6) ont montré que l'on peut définir de manière générale une semi-métrique de projection qui permet de se ramener à des probabilités de petites boules de type fractal. On condense ainsi les données en réduisant leur dimension et on contourne ainsi le fléau de la dimension. En effet, on revient à des vitesses de convergence en puissance de n . Dans d'autres situations, on peut être confronté à des données très lisses. Dans ce cas, il peut être intéressant d'utiliser plutôt des semi-métriques basées sur les dérivées (cf. Ferraty et Vieu, [32]). Ces semi-métriques peuvent également être utiles lorsque les données présentent un shift vertical artificiel (i.e non informatif vis-à-vis des réponses). Elles ont alors pour effet d'éliminer ces décalages verticaux qui nuisent à la qualité de la prédiction. Enfin, on peut envisager d'autres types de phénomènes comme, à titre d'exemple, les décalages horizontaux (cf. Dabo-Niang et al., [16]).

D'une façon générale, l'analyse de tout type de données nécessite la définition de la notion de distance entre celles-ci. Il est bien connu que dans un espace vectoriel de dimension finie toutes les métriques sont équivalentes. Ceci n'est plus le cas quand l'espace d'observations est de dimension infinie. C'est pourquoi le choix de la métrique (et donc de la topologie associée) est un élément crucial pour l'étude des v.a's

fonctionnelles.

On peut donc affirmer, que le choix de la semi-métrie permet à la fois de prendre en compte des situations plus variées et de pouvoir contourner le fléau de la dimension. Ce choix ne doit cependant pas être pris à la légère mais, doit prendre en compte, non seulement la nature des données mais aussi la nature du problème étudié.

1.3 Brève Présentation des Résultats

Dans cette section, nous donnons une brève présentation des différents résultats obtenus dans ce projet de thèse.

1.3.1 Présentation de l'estimateur

Soient $(Y_i)_{i=1,\dots,n}$ n v.a.'s indépendantes et de même loi que $Y \in \mathbb{R}$, et soient $(X_i)_{i=1,\dots,n}$ n v.a.'s α -mélangentes et de même loi que $X \in \mathcal{F}$ où (\mathcal{F}, d) est un espace semi-métrique de dimension éventuellement infinie. Y_i représente notre $i^{\text{ème}}$ variable d'intérêt et X_i sa covariable associée. On introduit de même les v.a.'s de censure $(C_i)_{i=1,\dots,n}$ indépendantes et de même loi que $C \in \mathbb{R}$. Dans le modèle de censure on n'observe que $(X_i, T_i, \delta_i)_{i=1,\dots,n}$, où $T_i = \min(Y_i, C_i)$ et $\delta_i = \mathbb{I}_{(Y_i \leq C_i)}$. Pour tout $x \in \mathcal{F}$ on définit la fonction de repartition conditionnelle de Y sachant $X = x$, par

$$F(y|x) = P(Y \leq y|X = x), \quad \forall y \in \mathbb{R}.$$

On suppose que cette probabilité conditionnelle est absolument continue par rapport à la mesure de Lebesgue dont, on désigne par $\zeta(\cdot|x)$ la densité conditionnelle de Y sachant $X = x$ et par $\zeta^{(j)}(\cdot|x)$ la dérivée d'ordre j de cette densité conditionnelle. Nous supposons qu'il existe un ensemble compact $\Omega \subset \mathbb{R}$ tel que $\zeta(\cdot|x)$ a un mode unique $\theta(x) \subset \Omega$, où

$$\theta(x) = \arg \sup_{y \in \Omega} \zeta(y|x).$$

l'estimateur de la densité conditionnelle par la méthode local linéaire est défini par :

$$\widehat{\zeta}_n(y|x) = \frac{\sum_{i,j=1}^n \delta_i \overline{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)},$$

où

$\overline{G}_n(y)$ estimateur de Kaplan-Meier de G (fonction de répartition de C),

$$W_{ij}(x) = \beta(X_i, x) \left[\beta(X_i, x) - \beta(X_j, x) \right] K(h_K^{-1} \delta(x, X_i)) K(h_K^{-1} \delta(x, X_j)),$$

$\beta(\cdot, \cdot)$ et $\delta(\cdot, \cdot)$ sont des fonctions de localisation définies de \mathcal{F}^2 dans \mathbb{R} tel que $\forall \xi \in \mathcal{F}$, $\beta(\xi, \xi) = 0$, $d(\cdot, \cdot) = |\delta(\cdot, \cdot)|$, K et H sont des noyaux et h_H (resp h_K) est une suite de nombres réels positifs, avec la convention $0/0 = 0$. Un estimateur du mode conditionnel $\theta(x)$ est défini comme la v.a $\widehat{\theta}_n(x)$ maximisant l'estimateur conditionnelle $\widehat{\zeta}_n(y|x)$ de $\zeta(y|x)$, c'est-à-dire

$$\widehat{\theta}_n(x) = \arg \sup_{y \in \Omega} \widehat{\zeta}_n(y|x).$$

Par la suite on désignera par V_x un voisinage de x où $x \in \mathcal{F}$ et par $B(x, r) = \{x' \in \mathcal{F} / |\delta(x', x)| \leq r\}$ la boule de centre x et de rayon r .

1.3.2 Résultats sur la consistance du modèle

Sous des hypothèses techniques assez générales en NPFDA et des hypothèses sur le coefficient de mélange fort. Nous établissons la convergence presque sûre de l'estimateur de la densité conditionnelle et du mode conditionnelle. Des inégalités exponentielles de type *Fuk – Nagaev* sont utilisées pour la preuve.

Théorème : Sous certaines hypothèses. On a

$$\sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)| = O(h_K^{b_1} + h_H^{b_2}) + O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}}\right), p.s.$$

Théorème : Sous certaines hypothèses. On a

$$|\widehat{\theta}_n(x) - \theta(x)| = O(h_K^{b_1/2} + h_H^{b_2/2}) + O\left(\left(\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}\right)^{\frac{1}{4}}\right), p.s.$$

où $\phi_x(h) = \mathbb{P}(-h \leq \delta(X, x) \leq h)$, $\varphi_x(h) = \max_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h) \times B(x, h))$ et $\chi_x(h) = \max(\phi_x^2(h), \varphi_x(h))$.

Une discussion sur les hypothèses et les détails des démonstration sont exposés au Chapitre 3.

1.3.3 Résultats sur la normalité Asymptotique de l'estimateur

Le théorème suivant donne la normalité asymptotique de l'estimateur locale linéaire de la densité conditionnelle pour des données fonctionnelles censurées à droite.

Théorème : Sous certaines hypothèses. On a

$$(nh_H\phi_x(h_K))^{1/2} \left[\text{diag}(1, h_K) \begin{pmatrix} \widehat{\zeta}_n(y|x) - \zeta(y|x) \\ \widehat{b} - \Psi_0'(0) \end{pmatrix} - \frac{h_K^2}{2} \Psi_0''(0) \mathbf{S}^{-1} \mathbf{U} \right. \\ \left. - \frac{h_H^2}{2} \psi_2(x) \int t^2 H(t) dt \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \xrightarrow{D} N \left(0, \overline{G}^{-1}(y) \zeta(y|x) \int H^2(t) dt \mathbf{S}^{-1} \mathbf{V} \mathbf{S}^{-1} \right).$$

Corollaire : Sous certaines hypothèses. On a

$$(nh_H\phi_x(h_K))^{1/2} (\widehat{\zeta}_n(y|x) - \zeta(y|x) - \frac{h_K^2}{2} \Psi_0''(0) \frac{N(1,2)}{M_1} - \frac{h_H^2}{2} \psi_2(x) \int t^2 H(t) dt) \\ \xrightarrow{D} N(0, V_{HK}^x(y)),$$

où

$$V_{HK}^x(y) = \overline{G}^{-1}(y) \zeta(y|x) \frac{M_2}{M_1^2} \int H^2(t) dt.$$

Les termes Ψ_0 , $N(1,2)$, M_1 , M_2 , \mathbf{U} , \mathbf{S} et \mathbf{V} sont définies dans le Chapitre 4. De même, la démonstration et les détails des hypothèses imposées pour aboutir aux résultats cités.

1.4 Le Plan de la Thèse

Dans cette thèse, nous nous intéressons à l'estimation non paramétrique de la densité conditionnelle et les paramètres qui en découlent, comme le mode conditionnel, pour des variables aléatoires fonctionnelles censurées à droite. La thèse se présente en cinq chapitres, décrits successivement comme suit :

Les Chapitres un et deux sont introductifs. Dans le premier on présente un panorama général de l'analyse statistique des données fonctionnelles, l'état de l'art

des variables fonctionnelles, la densité conditionnelle, les différents types de dépendances (la dépendance forte : α - *mixing*) ainsi que le principe de la méthode d'estimation utilisé dans ce travail. Tandis que le deuxième chapitre propose une présentation des données incomplètes précisément les données censurées à droite ainsi que l'estimateur de Kaplan-Meier.

Dans le troisième chapitre, nous introduisons un estimateur local linéaire de la densité conditionnelle. Sous des hypothèses de régularité et des conditions de dépendance faible (mélange fort), nous établissons la convergence presque sûre et la vitesse de convergence lorsque la variable réponse est censurée et la variable explicative est fonctionnelle. Par la suite, nous utilisons les résultats obtenus pour déterminer les propriétés asymptotiques de l'estimateur locale linéaire du mode conditionnel.

Le quatrième chapitre constitue une suite du précédent où on développe une étude sur la normalité asymptotique de l'estimateur local linéaire de la densité conditionnelle sous les mêmes hypothèses.

Le dernier chapitre sera consacré à l'étude de simulation. En premier temps, nous vérifions l'influence de censure sur la performance de notre estimateur à travers des ensembles de données simulées. En second temps sur quelques données réelles applicable en médecine et comme exemple la transplantation rénale (Greffes de rein).

Finalement, la thèse s'achève par une conclusion générale ainsi que quelques perspectives.

Concepts de Base de l'Analyse de Survie

Contents

2.1	Durée de Vie	26
2.2	Données Incomplètes	28
2.2.1	Données Tronquées	28
2.2.2	Données Censurées	29
2.3	Les fonctions d'intérêt	31
2.3.1	L'estimateur de Kaplan-Meier	32

Dans ce chapitre, on présente quelques définitions qui nous sont utiles en analyse de survie.

2.1 Durée de Vie

L'analyse de survie est le nom d'un ensemble de techniques statistiques relatives à la modélisation des données de durée de vie. Ces méthodes sont utilisées pour décrire, quantifier et comprendre le comportement stochastique d'un événement, souvent appelé échec ou décès (même si l'événement peut être en réalité un "succès", comme la guérison après le traitement), survenant après une durée appelée durée de vie. Cet événement intervient au plus une seule fois pour chaque individu. Des

exemples classiques sont la panne de composants électroniques en fiabilité industrielle, la fin d'une grève ou d'une période de chômage en économie, la résolution d'une tâche spécifique en expérimentation psychologique ou, en médecine, la rechute ou la mort d'un patient. En sociologie, le choix est vaste avec les successions des événements de vie : mariage, naissance du premier enfant, divorce... .

La spécificité des durées de survie est de correspondre à des v.a.s positives et de comporter des observations dans la plupart des études prospectives, les individus sont suivis pendant une durée d'observation fixée à l'avance. Pour les sujets pour lesquels l'événement d'intérêt a lieu pendant la période d'observation, on dispose du délai exact d'apparition de cet événement d'intérêt, mesuré depuis une date initiale qu'il faut spécifier sans ambiguïté (date de randomisation dans les essais cliniques, par exemple). Cependant, à la fin de la période d'observation, certains sujets n'auront pas eu l'événement. On n'aura alors pour ces sujets qu'une information incomplète, à savoir que le délai d'apparition de l'événement est plus grand que la durée d'observation. De telles données sont dites censurées à droite et la durée d'observation constitue le délai de censure. Comme les sujets rentrent dans l'étude à des dates différentes, chaque sujet a un délai d'apparition et un délai de censure qui lui sont propres. Le délai observé est donc le minimum du délai d'apparition de l'événement et du délai de censure. Cela correspond à un mécanisme de censure à droite, qui est le cas le plus courant dans les études prospectives et qui est le cas que nous considérons dans toute la suite. Les sujets qui n'ont pas eu l'événement d'intérêt pendant la période d'observation sont dits "exclus vivants" à la fin de l'étude . Une autre cause de censure à droite, qu'on essaie de limiter au maximum dans les études, correspond aux sujets dits "perdus de vue", qui sont les individus qui ont quitté l'étude avant l'apparition de l'événement d'intérêt et dont on n'a plus de nouvelles à la date de fin d'étude. En présence de données censurées, le délai d'apparition de l'événement est inconnu et tout ce que nous savons, c'est que le délai de survie s'est produit avant, entre ou après certains instants précis.

2.2 Données Incomplètes

Une des caractéristiques des données de survie est l'existence d'observations incomplètes. En effet, les données sont souvent recueillies partiellement, notamment, à cause des processus de censure et de troncature. Les données censurées ou tronquées proviennent du fait qu'on n'a pas accès à toute l'information. Au lieu d'observer des réalisations i.i.d de durée Y , on observe la réalisation de la variable Y soumise à diverses perturbations indépendantes ou non de l'événement étudié. Les mécanismes de censure et de troncature peuvent survenir simultanément.

2.2.1 Données Tronquées

La troncature empêche l'observation de la variable Y entièrement (en général les valeurs extrêmes), et engendre une perte d'information (on n'étudie qu'un sous-échantillon). Il y a troncature si l'observation de la variable d'intérêt Y n'a lieu que conditionnellement à un autre événement. On dit qu'il y a :

1. Troncature à gauche lorsque Y n'est observable que si elle est supérieure à un seuil C positif fixé ou aléatoire. Il s'agit d'un modèle qui est tout d'abord apparu en astronomie, où des échantillons sont composés d'objets astraux d'une certaine zone. Les luminosités absolue et apparente d'un objet astral sont respectivement définies comme étant sa brillance observée à une distance fixe et de puis la Terre et l'on n'observe que les objets qui sont suffisamment brillants, c'est-à-dire ceux pour lesquels la luminosité $M \geq m$, m étant la variable de troncature. Dans ce cas, nous disposons de N objets dans l'échantillon, mais nous ne sommes capables d'observer que les n objets suffisamment brillants.

2. Troncature à droite lorsque Y n'est observable que si elle est inférieure à C .

3. Troncature par intervalle lorsque Y est tronquée à droite et à gauche. On rencontre ce type de troncature lors de l'étude des patients d'un registre : les patients diagnostiqués avant la mise en place du registre ou répertoriés après la consultation du registre ne seront pas inclus dans l'étude.

Lyndell-Bell [56] proposa une estimation non-paramétrique de la fonction de répartition de Y dans le cadre du modèle de troncature et les propriétés asymptotiques : la loi forte et la normalité asymptotique ont été étudiées par Woodroffe [75].

2.2.2 Données Censurées

L'analyse des données censurées nécessite une méthodologie adaptée permettant de prendre en compte l'information contenue dans le délai de censure. Les procédures usuelles qui tiennent compte de la censure supposent presque toujours l'indépendance des variables durées de survie et délais de censure. La censure est de plus supposée la plupart du temps non informative, dans le sens où sa distribution ne dépend pas des paramètres qui interviennent dans la distribution de la variable durée de survie. Dans toute la suite, ces hypothèses sont conservées. Ces hypothèses sont raisonnables si les données censurées sont dues à des sujets qui n'ont pas expérimenté l'événement à la fin de l'étude (sujets "exclus vivants"). Elles sont moins évidemment vérifiées quand les données censurées correspondent à des sujets qui ont quitté l'étude avant la fin ("perdus de vue"). On en distingue Les différents types de censure :

1 - Censure de type I : fixée : Au lieu d'observer les variables Y_1, Y_2, \dots, Y_n qui nous intéressent, on observe Y_i lorsqu'elle est inférieure à une durée fixe C , sinon on sait seulement que Y_i est supérieure à C . On observe donc une variable T_i telle que $T_i = \min(Y_i, C)$. Ce mécanisme de censure est fréquemment rencontré dans les applications industrielles.

Exemple 1. *En biologie, on peut tester l'efficacité d'une molécule sur un lot de souris (les souris vivantes au bout d'un temps μ sont sacrifiées)*

2 - Censure de type II : attente (jusqu'au r^{ime} décès : On observe les durées de vie de n patients jusqu'à ce que r d'entre eux soient décédés et on s'arrête à ce moment. Si on ordonne les Y_1, Y_2, \dots, Y_n , on obtient les statistiques d'ordre $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$. La date de la censure est alors $Y_{(r)}$ et on observe $T_{(1)} = Y_{(1)}, T_{(2)} = Y_{(2)}, \dots, T_{(r)} = Y_{(r)}, T_{(r+1)} = Y_{(r)}, \dots, T_{(n)} = Y_{(r)}$.

Exemple 2. *(Droesbeke et al. [26]) : Pour tester la fiabilité d'un système complexe,*

on met en état de fonctionnement n systèmes du même type, et on s'arrête lorsque la r^{ime} panne est observée.

3 - Censure de type III : aléatoire : On en distingue deux cas :

(i) - **Censure aléatoire à droite :** Dans ce modèle de censure, on observe le couple (T, δ) où T est la durée observée et δ est une variable binaire représentant la nature de cette durée qui prend la valeur 1 s'il s'agit d'une vraie durée de vie et 0 si c'est une censure. Autrement dit, à chaque individu i , $i = 1, 2, \dots, n$ est associé un couple de v.a.s durées (Y_i, C_i) , dont seule la plus petite est observée, Y_i v.a représente la durée de survie et C_i exprime l'instant de censure.

Exemple 3. *Lorsqu'une étude est menée afin d'analyser les temps de guérison d'un certain remède sur les patients atteints d'une maladie et que le patient meurt avant d'être guéri, le temps auquel l'individu décède devient alors une borne inférieure du temps de guérison, c'est-à-dire une donnée censurée à droite.*

(ii) - **Censure aléatoire à gauche :** On dit qu'il y a censure à gauche si au lieu d'observer Y_1, \dots, Y_n on observe (T_i, δ_i) où $T_i = \max(Y_i, C_i)$ et $\delta_i = \mathbb{I}_{\{Y_i \geq C_i\}}$ pour $i = 1, \dots, n$ et C_i est une censure aléatoire.

Exemple 4. *Une ethnologie étudie la durée d'apprentissage d'une tâche. Cette durée est une v.a Y et C est l'âge de l'enfant. Pour les enfants qui savent déjà accomplir la tâche, C censure Y à gauche car pour eux Y est inconnu mais inférieur à C .*

4 - Censure double ou mixte : Il y a censure mixte ou double dans un échantillon de données s'il y a à la fois censure à gauche et à droite dans cet échantillon. Les données sont censurées à la fois à droite et à gauche. Plusieurs modèles non paramétriques ont été présentés pour l'étude de la double censure. Par exemple, le modèle de Turnbull [73] est le plus utilisé, et plusieurs travaux sont basés sur ce modèle.

5 - Censure par intervalle : Une date est censurée par intervalle si au lieu d'observer avec certitude le temps de l'événement, la seule information disponible est qu'il a eu lieu entre deux dates connues. On retrouve ce modèle

en général dans des études de suivi médical où les patients sont contrôlés périodiquement.

Exemple 5. *Dans le cas d'un suivi de cohorte, les personnes sont souvent suivies par intermittence (pas en continu), on sait alors uniquement que l'événement s'est produit entre ces deux temps d'observations.*

2.3 Les fonctions d'intérêt

Outre la densité f , la fonction de répartition F et la fonction caractéristique, les fonctions les plus utilisées en analyse de la survie sont la fonction de survie S , la fonction de risque instantané h et la fonction de risque cumulé H .

Définition 2.3.1. *On appelle fonction de survie S_Y la probabilité que la durée de vie Y soit supérieure à y :*

$$\forall y \in \mathbb{R}; S_Y(y) = \mathbb{P}(Y > y) = 1 - F_Y(y).$$

Notons que si la loi de Y admet une densité f_Y par rapport à la mesure de Lebesgue,

$$\forall y \in \mathbb{R}; S_Y(y) = \int_y^{+\infty} f_Y(l)dl \text{ et } f_Y(y) = -S'_Y(y) \text{ p.p.}$$

Définition 2.3.2. *On appelle fonction de risque instantané h_Y la fonction définie pour y dans \mathbb{R}^+ par*

$$h_Y(y) = \begin{cases} \lim_{l \rightarrow 0} \frac{1}{l} \mathbb{P}(y < Y \leq y + l | Y > y) & \text{si } y \text{ est tel que } \mathbb{P}(Y > y) > 0 \\ +\infty & \text{sinon.} \end{cases}$$

La fonction de risque peut avoir des formes très différentes mais est nécessairement positive sur \mathbb{R} . Supposons maintenant que Y soit une variable continue, on observe alors que

$$\forall y \in \mathbb{R}^+, h_Y(y) = \frac{f_Y(y)}{S_Y(y)} = -\frac{\partial}{\partial y} \ln(S_Y(y)),$$

en posant $c/0 = +\infty$ pour tout $c > 0$. La définition de h_Y montre que pour l assez petit, $lh_Y(y)$ s'interprète comme la probabilité de survenue de l'événement d'intérêt dans l'intervalle $]y; y + l]$ sachant que cet événement ne s'est pas encore produit à l'instant y . Cette fonction traduit donc l'évolution dans le temps du risque de survenue de l'événement d'intérêt.

Définition 2.3.3. On appelle fonction de risque cumulé H_Y la fonction définie pour y dans \mathbb{R}^+ par

$$H_Y(y) = \int_0^y h_Y(s) ds = -\ln(S(y)),$$

qui vaut $+\infty$ quand $S(y) = 0$.

Des définitions précédentes on déduit que pour tout $y \in \mathbb{R}^+$, on a la relation

$$f_Y(y) = h_Y(y) \exp(-H_Y(y)).$$

En conclusion, les trois fonctions précédentes permettent de caractériser la loi de Y et les unes sont déductibles des autres. Cependant, c'est l'interprétation de la fonction de risque instantané qui permettra le plus souvent de guider le choix d'un modèle pour des données de durée de vie.

2.3.1 L'estimateur de Kaplan-Meier

Soit Y_1, \dots, Y_n un échantillon de v.a's indépendantes représentant les durées d'intérêt, de fonction de répartition F , de fonction de survie S et C_1, \dots, C_n un échantillon représentant les temps de censure, que l'on suppose indépendants des durées d'intérêt, de fonction de répartition G . Soit $(T_i, \delta_i)_{1 \leq i \leq n}$, l'échantillon réellement observé où $T_i = Y_i \wedge C_i$ et $\delta_i = \mathbb{I}_{Y_i \leq C_i}$ et soit $(T_{(i)}, \delta_{(i)})_{1 \leq i \leq n}$ sa statistique d'ordre croissant. L'estimateur le plus utilisé de S , qui généralise l'estimateur empirique de la survie $1 - F_n$ au cas censuré, lorsqu'il n'y a pas d'ex-aequo (c'est-à-dire des temps de décès identiques pour plusieurs sujets), est l'estimateur proposé par Kaplan et Meier [45] et défini par :

$$\widehat{S}_{KM}(y) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{\delta_{(i)}}{n - i + 1}\right)^{1_{\{T_{(i)} \leq y\}}} & \text{si } y < T_{(n)} \\ 0 & \text{sinon.} \end{cases}$$

L'estimateur de Kaplan-Meier \overline{G}_n de G s'obtient de la même manière par

$$\overline{G}_n(y) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\right)^{1_{\{T_{(i)} \leq y\}}} & \text{si } y < T_{(n)} \\ 0 & \text{sinon.} \end{cases}$$

L'estimateur de Kaplan-Meier converge presque sûrement et uniformément (Földes et al. [33]). Sous certaines conditions de régularité, il converge en loi vers un processus gaussien (voir Breslow et Crowley [10]). Les propriétés mathématiques de l'estimateur de Kaplan-Meier peuvent également être trouvées au chapitre 7 de Shorack et Wellner [70].

Functional data : Local linear estimation of a conditional density function under random censorship

This chapter was the subject of a publication in
Arabian Journal of Mathematics

<https://link.springer.com/article/10.1007/s40065-020-00282-1>

Contents

3.1 Introduction	36
3.2 Construction of the conditional density estimator	37
3.3 Assumptions and main results	40
3.3.1 Almost sure consistency of the conditional density estimator	40
3.3.2 Almost sure consistency of the conditional mode estimator .	43
3.4 Proofs	43

In this chapter, we study nonparametric L.L.E of the conditional density of a randomly censored scalar response variable given a functional random covariate. We establish under general conditions the pointwise almost sure convergence with rates of this estimator under α -mixing dependence. Finally, in order to show the applicability of our results, we conducted a computational study, firstly on simulated data and secondly on real data concerning Kidney transplant.

3.1 Introduction

Conditional density plays an important role; not only in exploring relationships between responses and covariates, but also in financial econometrics (see Ait-Sahalia [1]). A vast variety of papers use the estimators of conditional densities as building blocks. These papers include those of Robinson [65], Tjøstheim [72], among others. However, in all of these papers, the conditional density function is indirectly estimated. Hyndman et al [42] have studied the kernel estimator of conditional density estimator and its bias corrected version. There are many advantages of using local linear regression, such as the lack of boundary modifications, high minimax efficiency, ease of implementation. Then, Bashtannyk and Hyndman [4] have suggested several simple and useful rules for selecting bandwidths for conditional density estimation. Hall et al. [35] applied the cross-validation technique to estimate the conditional density. Fan and Yim [28] proposed a consistent data-driven bandwidth selection procedure for estimating the conditional density functions.

In the last decade, the kernel method has been largely used for nonparametric functional data study, in this context we refer to the monograph of Ferraty and Vieu [31]. Since then, many interesting publications have appeared. According to literature, results on the local linear modeling in the functional data setting are limited. Baillo and Grané [2] proposed a L.L.E of the regression operator when the explanatory variable takes values in a Hilbert space. Then, when the explanatory variable takes values in a semi-metric space. Demongeot et al. [20] presented L.L.E of the conditional density when the data are functional. Further, Messaci et al. [59] used the same approach to estimate the conditional quantile of a scalar response given a functional explanatory variable in the i.i.d case. In these papers, it is assumed that the observations are complete.

Censored data analysis is a major issue in survival studies. Censored data, truncated data, missing data, current status data, are among the complex data structures in which only partial information on the variable of interest is available (see Kaplan

and Meier [45]). In real data applications, censoring is a condition in which the value of a measurement or observation is only partially known. For this, we observe the censored lifetime (C_i) for $i = 1, \dots, n$ of the variable under study instead of the real lifetime (Y_i) for $i = 1, \dots, n$ (which has a continuous distribution function (d.f) $F(\cdot)$). We assume that $\{Y_i, i \geq 1\}$ is a stationary sequence of lifetimes and $\{C_i, i \geq 1\}$ is a sequence of i.i.d censoring random variable (r.v) with common unknown continuous d.f $G(\cdot)$, and we observe only the n pairs $\{(T_i, \delta_i), i = 1, \dots, n\}$, where $T_i = Y_i \wedge C_i$ and $\delta_i = \mathbb{I}_{\{Y_i \leq C_i\}}$ (with \wedge denoting minimum and \mathbb{I}_A denoting the indicator function on a set A).

The distinguishing characteristic of censoring has attracted the attention of many researchers. Beran [5] introduced a nonparametric estimate of the conditional survival function and showed some consistency results. Many other properties of the conditional distribution have been broadly studied in the literature (see Stute [71]). Furthermore, in the same context, many results on conditional quantile and conditional mode have been given (see Horrigue and Ould Saïd [39], Khardani and Thiam [47]). The present study extends the result of Demongeot et al. [20] to censored data under general conditions. We establish the almost sure consistencies with convergence rates of the conditional density estimator when the explanatory variable is of functional type.

3.2 Construction of the conditional density estimator

Consider n pairs of r.v's (X_i, Y_i) for $i = 1, \dots, n$ drawn from the pair (X, Y) with values in $\mathcal{F} \times \mathbb{R}$, where \mathcal{F} is a semi-metric space equipped with a semi-metric d . In this paper, we consider the problem of nonparametric estimation of the conditional Y given $X = x$ when the responses variable (Y_i) for $i = 1, \dots, n$ are right censored and when the observations (X_i, Y_i) for $i = 1, \dots, n$ are strongly mixing. Furthermore, we denote by (C_i) for $i = 1, \dots, n$ the censoring r.v's which are i.i.d with a common unknown continuous d.f G . Thus, we observe the triplets (X_i, T_i, δ_i) for $i = 1, \dots, n$, where $T_i = Y_i \wedge C_i$ and $\delta_i = \mathbb{I}_{\{Y_i \leq C_i\}}$ (with \wedge denoting minimum and \mathbb{I}_A denoting the indicator function on a set A). We suppose that (Y_i) for $i = 1, \dots, n$ and (C_i) for

$i = 1, \dots, n$ are independent which ensures the identifiability of the model. In the case of complete data, we adopt the fast functional locally modeling, introduced by Barrientos-Marin et al. [3] for regression analysis, that is, we estimate the conditional density $\zeta(\cdot|x)$ by \widehat{a} which is obtained by minimizing the following quantity

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n \left(h_H^{-1} H(h_H^{-1}(y - T_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1} \delta(x, X_i)),$$

where $\beta(\cdot, \cdot)$ is a known function from \mathcal{F}^2 into \mathbb{R} , such that $\forall \xi \in \mathcal{F}$, $\beta(\xi, \xi) = 0$, where K and H are kernels, $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers, and $\delta(\cdot, \cdot)$ is chosen as a function of $\mathcal{F} \times \mathcal{F}$, such that $d(\cdot, \cdot) = |\delta(\cdot, \cdot)|$. Here, we denote \widehat{a} by $\zeta_n(\cdot|x)$. Then, the expression of $\zeta_n(\cdot|x)$ is given as

$$\zeta_n(y|x) = \frac{\sum_{i,j=1}^n W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)},$$

where

$$W_{ij}(x) = \beta(X_i, x) \left[\beta(X_i, x) - \beta(X_j, x) \right] K(h_K^{-1} \delta(x, X_i)) K(h_K^{-1} \delta(x, X_j)),$$

with the convention $0/0 = 0$.

We assume that there exists a certain compact set $\Omega \subset \mathbb{R}$ such that $\zeta(y|x)$ has an unique mode $\theta(x)$ on Ω , where

$$\theta(x) = \underset{y \in \Omega}{\operatorname{argsup}} \zeta(y|x).$$

In the censored case, we adapt the idea of Carbonez et al. [12], Kohler et al. [50], and Khardani et al. [46] to the infinite dimension case by using a smooth distribution function $H(\cdot)$ instead of a step function. Then we get the "pseudo"-estimator of $\zeta(\cdot|x)$

$$\tilde{\zeta}_n(y|x) = \frac{\sum_{i,j=1}^n \delta_i \overline{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)} \doteq \frac{\tilde{f}_n^x(y)}{\gamma_n(x)},$$

where

$$\tilde{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \overline{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i)).$$

and

$$\gamma_n(x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]} \sum_{i \neq j} W_{ij}(x).$$

The d.f G , of the censoring r.v's, is estimated by the Kaplan-Meier [45] estimator defined as

$$\bar{G}_n(y) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\right)^{\mathbb{I}_{\{T_{(i)} \leq y\}}} & \text{if } y < T_{(n)} \\ 0 & \text{otherwise,} \end{cases}$$

where $T_{(1)} < T_{(2)} < \dots < T_{(n)}$ are the order statistics of T_i and $\delta_{(i)}$ is the concomitant of $T_{(i)}$, which is known to be uniformly convergent to \bar{G} .

Therefore a feasible estimator of $\zeta(\cdot|x)$ is given by

$$\widehat{\zeta}_n(y|x) = \frac{\widehat{f}_n^x(y)}{\gamma_n(x)}$$

where

$$\widehat{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i)).$$

Then, a natural estimator of $\theta(x)$ is defined by

$$\widehat{\theta}_n(x) = \arg \sup_{y \in \Omega} \widehat{\zeta}_n(y|x)$$

3.3 Assumptions and main results

Let $\mathcal{F}_i^k(Z)$ denote the σ -algebra generated by $\{Z_j, i \leq j \leq k\}$.

Définition 3.3.1. Let $\{Z_i, i = 1, 2, \dots\}$ be a strictly stationary sequence of r.v's. Given a positive integer n , set

$$\alpha(n) = \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_1^k(Z) \text{ and } B \in \mathcal{F}_{k+n}^\infty(Z), k \in \mathbb{N}^* \right\}.$$

The sequence is said to be α -mixing (strong mixing) if the mixing coefficient $\alpha(n) \rightarrow 0$ as $n \rightarrow \infty$.

This condition was introduced by Rosenblatt [66]. The strong-mixing condition is reasonably weak and has many practical applications, (see Doukhan [25], for more details).

3.3.1 Almost sure consistency of the conditional density estimator

Our first result concerns the almost sure convergence of the L.L.E of the conditional density. We introduce some conditions that are required to state our asymptotic result. Throughout the paper, x denotes a fixed point in \mathcal{F} , N_x denotes a fixed neighborhood of x . For any df L , let $\tau_L := \sup\{y : L(y) < 1\}$ be its support's right endpoint and assume that $\theta(x) \in \Omega \subset (-\infty, \tau]$, where $\tau < \tau_G \wedge \tau_F$, $B(x, r) = \{x' \in \mathcal{F} / |\delta(x', x)| \leq r\}$ and $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$.

Note that our nonparametric model is quite general in the sense that we just need the following assumptions

(H1) For any $r > 0$, $\phi_x(r) := \phi_x(-r, r) > 0$

(H2) (i) The conditional density ζ^x is such that : there exist $b_1 > 0, b_2 > 0$, $\forall (y_1, y_2) \in \Omega^2$ and $\forall (x_1, x_2) \in N_x \times N_x$

$$|\zeta(y_1|x_1) - \zeta(y_2|x_2)| \leq C_x \left(|\delta^{b_1}(x_1, x_2)| + |y_1 - y_2|^{b_2} \right),$$

where C_x is a positive constant depending on x .

(ii) $\zeta(\cdot|x)$ is twice differentiable, its second derivative $\zeta^{(2)}(\cdot|x)$ is continuous on a neighborhood of $\theta(x)$ and $\zeta^{(2)}(\theta(x)|x) < 0$.

(H3) The function $\beta(\cdot, \cdot)$ is such that :

$$\forall x' \in \mathcal{F}, C_1 |\delta(x, x')| \leq |\beta(x, x')| \leq C_2 |\delta(x, x')|, \text{ where } C_1 > 0, C_2 > 0.$$

(H4) The sequence $(X_i, Y_i)_{i \in \mathbb{N}}$ satisfies : $\exists a > 0, \exists c > 0, \forall n \in \mathbb{N} \alpha(n) \leq cn^{-a}$ and

$$\max_{i \neq j} \mathbb{P}\left((X_i, X_j) \in B(x, h) \times B(x, h)\right) = \varphi_x(h) > 0.$$

(H5) The conditional density of (Y_i, Y_j) given (X_i, X_j) exists and is bounded.

(H6) K is a positive, differentiable function with support $[-1, 1]$.

(H7) H is a positive, bounded, Lipschitzian continuous function, such that

$$\int |t|^{b_2} H(t) dt < \infty \text{ and } \int H^2(t) dt < \infty.$$

(H8) The bandwidth h_K satisfies : there exists an integer n_0 , such that

$$\forall n > n_0, -\frac{1}{\phi_x(h_K)} \int_{-1}^1 \phi_x(zh_K, h_K) \frac{d}{dz} (z^2 K(z)) dz > C_3 > 0,$$

and

$$h_K \int_{B(x, h_K)} \beta(u, x) dP(u) = o\left(\int_{B(x, h_K)} \beta^2(u, x) dP(u)\right),$$

where $dP(x)$ is the cumulative distribution.

(H9) $\lim_{n \rightarrow \infty} h_H = 0$ and $\exists \beta_1 > 0$ such that $\lim_{n \rightarrow \infty} n^{\beta_1} h_H = \infty$.

(H10) i) $\lim_{n \rightarrow \infty} h_K = 0$, $\lim_{n \rightarrow \infty} \frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)} = 0$,

ii) $C n^{\frac{(3-a)}{(a+1)} + \frac{3\beta_1+1}{a+1}} \log n [\log_2 n]^{6/(a+1)} \leq h_H \chi_x^{1/2}(h_K)$,

where $\chi_x(h) = \max(\phi_x^2(h), \varphi_x(h))$.

Remarks on the assumptions : Most of the assumptions are common in NPFDA context. More precisely, assumption (H1) is usually used in NPFDA and it is linked with the topological structure of the functional space, \mathcal{F} , of the explanatory variable X (see Ferraty and Vieu [31] for more discussions). Further, assumptions (H2) and (H3) are mild regularity assumptions on the conditional density function. Finally, conditions (H4), (H5), and (H10) are technical assumptions (see Ferraty and Vieu [31] for the constant local method case).

Theorem 3.3.1. *Under assumptions (H1)-(H10), we have*

$$\sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)| = O(h_K^{b_1} + h_H^{b_2}) + O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}}\right), \text{ a.s.}$$

Proof of Theorem 3.3.1 The proof is a direct consequence of the following decomposition $\forall y \in \Omega$,

$$\begin{aligned} \widehat{\zeta}_n(y|x) - \zeta(y|x) &= \frac{\widehat{f}_n^x(y) - \tilde{f}_n^x(y)}{\gamma_n(x)} + \frac{\tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)]}{\gamma_n(x)} \\ &+ \frac{\mathbb{E}[\tilde{f}_n^x(y)] - \zeta(y|x)}{\gamma_n(x)} + \zeta(y|x) \frac{(1 - \gamma_n(x))}{\gamma_n(x)}, \end{aligned} \quad (3.1)$$

where

$$\widehat{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - Y_i)),$$

$$\tilde{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - Y_i)),$$

and

$$\gamma_n(x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]} \sum_{i \neq j} W_{ij}(x).$$

and the following Lemmas 3.1-3.4. ■

Lemma 3.1. (cf. [20])

Under assumptions (H1), (H3), (H4),(H6),(H8); and (H10), we have

$$1 - \gamma_n(x) = O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n \phi_x^2(h_K)}}\right), \text{ a.co.}^1$$

and

$$\exists \delta > 0, \text{ such that } \sum_{i=1}^{\infty} \mathbb{P}(\gamma_n(x) < \delta) < \infty.$$

Lemma 3.2. Under assumptions (H1), (H2) and (H7), we obtain

$$\sup_{y \in \Omega} |\mathbb{E}[\tilde{f}_n^x(y)] - \zeta(y|x)| = O(h_K^{b_1} + h_H^{b_2}), \text{ a.s.}$$

Lemma 3.3. Under assumptions of Theorem 3.3.1, we get

$$\sup_{y \in \Omega} |\tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)]| = O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}}\right), \text{ a.co.}$$

Lemma 3.4. Under assumptions (H1), (H6), (H8),(H9) and (H10), we have

$$\sup_{y \in \Omega} |\widehat{f}_n^x(y) - \tilde{f}_n^x(y)| = O\left(\sqrt{\frac{\log(\log n)}{n}}\right), \text{ a.s.}$$

1. Let $(z_n)_{n \in \mathbb{N}}$ be a sequence of real r.v's, we say that z_n converges almost completely (a.co.) to zero if and only if $\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon) < \infty$. Moreover, let $(u_n)_{n \in \mathbb{N}^*}$ be a sequence of positive real numbers, we say that $z_n = O(u_n)$, a.s. if and only if $\exists \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon u_n) < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability (see Sarda and Vieu [68]).

3.3.2 Almost sure consistency of the conditional mode estimator

The convergence rate of the L.L.E of the conditional mode is a direct consequence of the previous result. Thus, in addition to the previous conditions, we assume that

(H11) There exists $\theta(x) \subset \Omega$ such that $\zeta(y|x) < \zeta(\theta(x)|x)$, for all $y \neq \theta(x), y \in \Omega$.

Then, the asymptotic behaviour of $\widehat{\theta}_n(x)$ is given in the following Theorem.

Theorem 3.3.2. *Assume that (H1)-(H11) hold, we have*

$$|\widehat{\theta}_n(x) - \theta(x)| = O\left(h_K^{b_1/2} + h_H^{b_2/2}\right) + O\left(\left(\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}\right)^{\frac{1}{4}}\right), \text{ a.s.}$$

3.4 Proofs

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. Moreover, we put, for any $x \in \mathcal{F}$, and for all $i = 1, \dots, n$,

$$K_i(x) = K(h^{-1}\delta(x, X_i)), \beta_i(x) = \beta(X_i, x) \text{ and } H_i(y) = H(h_H^{-1}(y - Y_i)).$$

Proposition 3.4.1. *(Ferraty and Vieu [31], page 237)*

Assume that $\{U_i, i \geq 1\}$ are identically distributed, with strong mixing coefficient $\alpha(n) = O(n^{-a}), a > 1$, such that $|U_1|$ is bounded. Then for each $r > 1$ and $\varepsilon > 0$,

$$P\left\{\left|\sum_{i=1}^n U_i\right| > \varepsilon\right\} \leq C\left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{-r/2} + Cnr^{-1}\left(\frac{2r}{\varepsilon}\right)^{a+1},$$

where

$$S_n^2 = \sum_{1 \leq i, j \leq n} |\text{cov}(U_i, U_j)|.$$

Proof of Lemma 3.2. The bias term is not affected by the dependence condition of (X_i, Y_i) . So, by the equiprobability of the couples (X_i, Y_i) , we have

$$\mathbb{E}[\tilde{f}_n^x(y)] - \zeta(y|x) = \frac{1}{h_H \mathbb{E}[W_{12}(x)]} \mathbb{E}\left[\delta_1 \overline{G}^{-1}(T_1) W_{12}(x) H_i(y)\right] - \zeta(y|x).$$

Using conditional expectation properties and the fact that

$$\mathbb{I}_{\{Y_1 \leq C_1\}} \varphi(T_1) = \mathbb{I}_{\{Y_1 \leq C_1\}} \varphi(Y_1), \quad (3.2)$$

for any measurable function φ , we have

$$\mathbb{E}[\tilde{f}_n^x(y)] - \zeta(y|x) = \frac{1}{h_H \mathbb{E}[W_{12}(x)]} \mathbb{E} \left[W_{12}(x) \mathbb{E}[H_i(y) \mathbb{I}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(T_1) | X_1] \right] - \zeta(y|x).$$

Since

$$\begin{aligned} \mathbb{E}[H_i(y) \mathbb{I}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(T_1) | X_1] &= \mathbb{E}[\mathbb{E}[H_i(y) \mathbb{I}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(Y_1) | Y_1] | X_1] \\ &= \mathbb{E}[H_i(y) \bar{G}^{-1}(Y_1) \mathbb{E}[\mathbb{I}_{\{Y_1 \leq C_1\}} | Y_1] | X_1] \\ &= \mathbb{E}[H_i(y) | X_1], \end{aligned}$$

and

$$\mathbb{E}[H_i(y) | X_1] = \int_{\mathbb{R}} H\left(\frac{y-z}{h_H}\right) \zeta(z|X) dz = h_H \int_{\mathbb{R}} H(t) \zeta(y - h_H t | X) dt,$$

then by assumptions (H2) and (H7) we get,

$$\begin{aligned} |\mathbb{E}[\tilde{f}_n^x(y)] - \zeta(y|x)| &= \left| \frac{1}{h_H \mathbb{E}[W_{12}(x)]} \mathbb{E} \left[W_{12}(x) h_H \int_{\mathbb{R}} H(t) \zeta(y - h_H t | X) dt \right] - \zeta(y|x) \right| \\ &= \left| \int_{\mathbb{R}} H(t) (\zeta(y - h_H t | X) - \zeta(y|x)) dt \right| \\ &\leq \int_{\mathbb{R}} H(t) |\zeta(y - h_H t | X) - \zeta(y|x)| dt \\ &\leq C \int_{\mathbb{R}} H(t) (h_K^{b_1} + |t|^{b_2} h_H^{b_2}) dt \\ &= O(h_K^{b_1}) + O(h_H^{b_2}), \end{aligned}$$

which proves Lemma 3.2. ■

Proof of Lemma 3.3.

Let Ω be a compact set. It can be covered by a finite number s_n of intervals of length l_n at some points $(z_k)_{k=1, \dots, s_n}$, that is,

$$\Omega \subset \bigcup_{k=1}^{s_n} (z_k - l_n, z_k + l_n),$$

with $l_n = n^{-\frac{3\gamma}{2} - \frac{1}{2}}$ and $s_n = O(l_n^{-1})$. Let

$$z_y = \arg \min_{z \in \{z_1, \dots, z_{s_n}\}} |y - z|, \quad (3.3)$$

and consider the following decomposition

$$\begin{aligned} \sup_{y \in \Omega} |\tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)]| &\leq \sup_{y \in \Omega} |\tilde{f}_n^x(y) - \tilde{f}_n^x(z_y)| + \sup_{y \in \Omega} |\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)]| \\ &\quad + \sup_{y \in \Omega} |\mathbb{E}[\tilde{f}_n^x(z_y)] - \mathbb{E}[\tilde{f}_n^x(y)]| \\ &:= \text{Im}_{1,n} + \text{Im}_{2,n} + \text{Im}_{3,n}. \end{aligned}$$

From (3.3) we have $|y - z_y| \leq l_n$, then under (H9) and Lemma 3.2, we get

$$\text{Im}_{1,n} = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \quad (3.4)$$

In the same way, we find

$$\text{Im}_{3,n} = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \quad (3.5)$$

Next, let us prove that

$$\text{Im}_{2,n} = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right).$$

For all $\eta > 0$, we have

$$\begin{aligned} &\mathbb{P} \left(\sup_{y \in \Omega} |\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ &= \mathbb{P} \left(\max_{z_y \in \{z_1, \dots, z_{s_n}\}} |\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ &\leq s_n \mathbb{P} \left(|\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \end{aligned}$$

Then, we have to show that

$$s_n \mathbb{P} \left(|\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) < \infty, \text{ for all } z_y \in \{z_1, \dots, z_{s_n}\}.$$

For this, we consider the following decomposition

$$\tilde{f}_n^x(z_y) = \underbrace{\frac{n^2 h_K^2 \phi_x^2(h_K)}{n(n-1)\mathbb{E}[W_{12}(x)]}}_{T_1} \left[\underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{\delta_j K_j(x) H_j(z_y)}{\overline{G}(T_j) h_H \phi_x(h_K)} \right)}_{T_2} \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{K_i(x) \beta_i^2(x)}{h_K^2 \phi_x(h_K)} \right)}_{T_3} \right. \\ \left. - \underbrace{\left(\frac{1}{n} \sum_{j=1}^n \frac{\delta_j K_j(x) \beta_j(x) H_j(z_y)}{\overline{G}(T_j) h_H h_K \phi_x(h_K)} \right)}_{T_4} \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \frac{K_i(x) \beta_i(x)}{h_K \phi_x(h_K)} \right)}_{T_5} \right],$$

which implies that,

$$\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] = T_1 \left((T_2 T_3 - \mathbb{E}[T_2 T_3]) - (T_4 T_5 - \mathbb{E}[T_4 T_5]) \right),$$

where,

$$\begin{aligned} T_2 T_3 - \mathbb{E}[T_2 T_3] &= (T_2 - \mathbb{E}[T_2])(T_3 - \mathbb{E}[T_3]) + (T_3 - \mathbb{E}[T_3])\mathbb{E}[T_2] \\ &\quad + (T_2 - \mathbb{E}[T_2])\mathbb{E}[T_3] + \mathbb{E}[T_2]\mathbb{E}[T_3] - \mathbb{E}[T_2 T_3], \end{aligned}$$

and

$$\begin{aligned} T_4 T_5 - \mathbb{E}[T_4 T_5] &= (T_4 - \mathbb{E}[T_4])(T_5 - \mathbb{E}[T_5]) + (T_5 - \mathbb{E}[T_5])\mathbb{E}[T_4] \\ &\quad + (T_4 - \mathbb{E}[T_4])\mathbb{E}[T_5] + \mathbb{E}[T_4]\mathbb{E}[T_5] - \mathbb{E}[T_4 T_5]. \end{aligned}$$

So, our claimed result is direct consequences of the following assertions

$$s_n \mathbb{P} \left\{ |T_i - \mathbb{E}[T_i]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right\} < \infty, \text{ for } i = 2, 3, 4, 5, \quad (3.6)$$

$$T_1 = O(1) \text{ and } \mathbb{E}[T_i] = O(1), \text{ for } i = 2, 3, 4, 5, \quad (3.7)$$

$$\text{Cov}(T_2, T_3) = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right), \quad (3.8)$$

and

$$\text{Cov}(T_4, T_5) = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \quad (3.9)$$

Concerning (3.6) : Observe that for $i = 3, 5$ has been already obtained in Lemma 3.1. Thus, we focus only on the case where $i = 2, 4$. For this, we use proposition 3.4.1. For $1 \leq i \leq n$ and $k = 0, 1$, let

$$\nabla_i^k = \frac{1}{h_K^k} \left(\frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) - \mathbb{E} \left[\frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) \right] \right). \quad (3.10)$$

Then, it can be seen that

$$T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}] = \frac{1}{nh_H \phi_x(h_K)} \sum_{i=1}^n \nabla_i^k, \quad \text{for } k = 0, 1.$$

Under assumptions (H1), (H3), (H5), and (H7), we have

$$\begin{aligned} \frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) &\leq \frac{C}{h_K^k} \frac{1}{\overline{G}(\tau)} K_i(x) |\delta(X_i, x)|^k \\ &\leq \frac{C}{h_K^k} \frac{1}{\overline{G}(\tau)} K(h_K^{-1} \delta(x, X_i)) |\delta(X_i, x)|^k \mathbb{I}_{]-1, 1[}(h_K^{-1} \delta(x, X_i)) \\ &\leq \frac{C}{\overline{G}(\tau)} K(h_K^{-1} \delta(x, X_i)) \mathbb{I}_{B(x, h_K)}(X_i) \\ &\leq C' \mathbb{I}_{B(x, h_K)}(X_i). \end{aligned} \quad (3.11)$$

Next, let us calculate

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\nabla_i^k, \nabla_j^k) = S_n^{2*} + n \text{Var}[\nabla_1^k],$$

where,

$$S_n^{2*} = \sum_{i=1}^n \sum_{i \neq j}^n \text{Cov}(\nabla_i^k, \nabla_j^k).$$

Using Equations (3.2) and (3.11), and the conditional expectation properties, we have

$$\begin{aligned} \text{Var}[\nabla_1^k] &= \text{Var} \left[\frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) \right] \\ &\leq C' \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i^2(z_y) | X_1]] + C \mathbb{E}^2[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i(z_y) | X_1]] \\ &= C' \phi_x(h_K) \mathbb{E}[H_i^2(z_y) | X_1] + C \phi_x^2(h_K) \mathbb{E}^2[H_i(z_y) | X_1] \end{aligned}$$

$$\leq C' \chi_x^{1|2}(h_K) \mathbb{E}[H_i^2(z_y)|X_1] + C \chi_x(h_K) \mathbb{E}^2[H_i(z_y)|X_1].$$

Finally, using the fact that ;

$$\mathbb{E}[H_i^2(z_y)|X_1] = O(h_H) \text{ and } \mathbb{E}[H_i(z_y)|X_1] = O(h_H),$$

we get

$$n \text{Var}[\nabla_1^k] = O(nh_H \chi_x^{(1/2)}(h_K)).$$

Then, from Equation (3.10), using again the conditional expectation, and from the fact that

$$\mathbb{E}[\delta_i \delta_j | Y_i, Y_j] = \bar{G}(T_i) \bar{G}(T_j),$$

we get under (H4) and (H5), for $i \neq j$

$$\begin{aligned} \text{Cov}(\nabla_i^k, \nabla_j^k) &= \text{Cov}\left(\frac{1}{h_K^k} \frac{\delta_i}{\bar{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x), \frac{1}{h_K^k} \frac{\delta_j}{\bar{G}(T_j)} K_j(x) H_j(z_y) \beta_j^k(x)\right) \\ &= \mathbb{E}\left[\frac{1}{h_K^{2k}} \frac{\delta_i \delta_j}{\bar{G}(T_i) \bar{G}(T_j)} K_i(x) K_j(x) H_i(z_y) H_j(z_y) \beta_i^k(x) \beta_j^k(x)\right] \\ &\quad - \mathbb{E}\left[\frac{1}{h_K^k} \frac{\delta_i}{\bar{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x)\right] \mathbb{E}\left[\frac{1}{h_K^k} \frac{\delta_j}{\bar{G}(T_j)} K_j(x) H_j(z_y) \beta_j^k(x)\right] \\ &\leq C [\mathbb{E}[\mathbb{I}_{B(x, h_K) \times B(x, h_K)}(X_i, X_j) \mathbb{E}[H_i(z_y) H_j(z_y) | X_i X_j]] \\ &\quad + \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i(z_y) | X_i]] \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_j) \mathbb{E}[H_j(z_y) | X_j]]] \\ &\leq Ch_H^2 (\varphi_x(h_K) + \phi_x^2(h_K)) \\ &\leq Ch_H^2 \chi_x(h_K). \end{aligned} \tag{3.12}$$

Now, following Masry [58], we define

$$S_1 = \{(i, j) \text{ such that } 1 \leq i - j \leq m_n\},$$

and

$$S_2 = \{(i, j) \text{ such that } m_n + 1 \leq i - j \leq n - 1\},$$

where $m_n \rightarrow \infty$, as $n \rightarrow \infty$. Thus,

$$s_n^{2*} \leq \sum_{S_1} \text{Cov}(\nabla_i^k, \nabla_j^k) + \sum_{S_2} \text{Cov}(\nabla_i^k, \nabla_j^k) := \gamma_{1,n} + \gamma_{2,n}.$$

We then get, from Equation (3.12)

$$\gamma_{1,n} \leq C n m_n h_H^2 \chi_x(h_K).$$

For $\gamma_{2,n}$, we use the Davydov-Rio's inequality for bounded mixing processes

$$\forall i \neq j \quad |\text{Cov}(\nabla_i^k, \nabla_j^k)| \leq C \alpha(|i - j|).$$

Therefore, using $\sum_{j \geq x+1} j^{-a} \leq \int_{u \geq x} u^{-a} = [(a-1)x^{a-1}]^{-1}$, and the first part of (H4), we get

$$\gamma_{2,n} \leq C n m_n^{-a+1}.$$

Thus,

$$s_n^{2*} \leq C(n m_n h_H^2 \chi_x(h_K) + n m_n^{1-a}).$$

By choosing $m_n = (h_H^2 \chi_x(h_K))^{-1/a}$, we obtain

$$s_n^{2*} = O(n h_H^2 \chi_x(h_K))^{(a-1)/a}.$$

Finally, as $a > 2$, then

$$S_n^2 = O(n h_H \chi_x^{1/2}(h_K)). \quad (3.13)$$

Taking $\varepsilon = \eta \frac{\sqrt{S_n^2 \log n}}{n h_H \phi_x(h_K)}$, making use of Proposition 3.4.1 and Equation (3.13), we get

$$\begin{aligned} \mathbb{P}\left\{\left|T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}]\right| > \varepsilon\right\} &= \mathbb{P}\left\{\left|\sum_{i=1}^n \nabla_i^k\right| > \varepsilon n h_H \phi_x(h_K)\right\} \\ &\leq C \left(1 + \frac{\varepsilon^2 n^2 h_H^2 (\phi_x(h_K))^2}{S_n^2 r}\right)^{-r/2} \\ &\quad + C n r^{-1} \left(\frac{2r}{\varepsilon n h_H \phi_x(h_K)}\right)^{a+1} \\ &\leq C \exp\left(-r/2 \log\left(1 + \frac{\eta^2 \log n}{r}\right)\right) \\ &\quad + C n r^{-1} \left(\frac{r}{\eta}\right)^{a+1} (n h_H \chi_x^{1/2}(h_K) \log n)^{-(a+1)/2} \\ &\leq C(\varphi_1 + \varphi_2). \end{aligned} \quad (3.14)$$

Next, using Equation (3.14) with $r = C \log n (\log_2 n)^{\frac{1}{a}}$ and Taylor series expansion of $\log(1+x)$, we obtain

$$\varphi_1 \leq C \exp(-C\eta^2 \log n) = Cn^{-C\eta^2}, \quad (3.15)$$

and

$$\varphi_2 \leq Cn^{(1-a)/2} \eta^{-(a+1)} (h_H \chi_x^{1/2}(h_K))^{-(a+1)/2} (\log n)^{(a-1)/2} \log_2 n. \quad (3.16)$$

Now, from Equations (3.15) and (3.16), using assumption (H10)(ii), we have $s_n \varphi_2 = O(n^{-1} \log^{-1} n \log_2^{-2} n)$ which is the general term of a convergent Bertrand series.

In the same way, we can choose η such that $s_n \varphi_1$ is the general term of a convergent series.

Finally, we find

$$s_n \mathbb{P} \left\{ \left| T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right\} < \infty.$$

Concerning (3.7) : from Demongeot et al. [20], we have

$$T_1 = O(1) \text{ and } \mathbb{E}[T_i] = O(1), \text{ for } i = 3, 5.$$

Now, we have to study the case where $i = 2, 4$.

Since the pairs $(X_i, Y_i), i = 1, \dots, n$ are identically distributed, we obtain

$$\mathbb{E}[T_2] = \frac{\mathbb{E}[\delta_1 \bar{G}^{-1}(T_1) K_1(x) H_1(z_y)]}{h_H \phi_x(h_K)},$$

and

$$\mathbb{E}[T_4] = \frac{\mathbb{E}[\delta_1 \bar{G}^{-1}(T_1) K_1(x) \beta_1(x) H_1(z_y)]}{h_K h_H \phi_x(h_K)}.$$

We have to evaluate

$$\mathbb{E} \left[\delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right], \text{ for } l = 0, 1.$$

Using the conditional expectation properties and Equation (3.2), we have for all $l = 0, 1$,

$$\mathbb{E} \left[\delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right] = O(h_H \mathbb{E} [K_i(x) \beta_i^l(x)]).$$

Using Lemma 3 given in Barrientos-Marin et al. [3], we obtain

$$\mathbb{E} \left[\delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right] = O(h_H h_K^l \phi_x(h_K)). \quad (3.17)$$

From Equation (3.17), we find

$$\mathbb{E}[T_i] = O(1), \text{ for } i = 2, 4.$$

Concerning (3.8) and (3.9) : following similar steps as in the proof of (3.13), we get

$$\begin{aligned} \text{Cov}(T_2, T_3) &= O \left(\frac{\chi_x^{(1/2)}(h_K)}{n \phi_x^2(h_K)} \right) = o \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ \text{and Cov}(T_4, T_5) &= O \left(\frac{\chi_x^{(1/2)}(h_K)}{n \phi_x^2(h_K)} \right) = o \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \end{aligned}$$

Hence,

$$\sup_{y \in \Omega} \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| = o_{a.co.} \left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \quad (3.18)$$

Finally, proof of Lemma 3.3 can be achieved by considering Equations (3.4), (3.5), and (3.18). ■

Proof of Lemma 3.4. Observe that

$$\begin{aligned} \left| \widehat{f}_n^x(y) - \tilde{f}_n^x(y) \right| &\leq \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \left| \delta_i W_{ij}(x) H_i(y) \left(\frac{1}{\bar{G}_n(T_i)} - \frac{1}{\bar{G}(T_i)} \right) \right| \\ &\leq \frac{\sup_{y \in \Omega} |\bar{G}_n(t) - \bar{G}(t)|}{\bar{G}_n(\tau)} \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}^{-1}(T_i) W_{ij}(x) H_i(y) \\ &= \frac{\sup_{y \in \Omega} |\bar{G}_n(t) - \bar{G}(t)|}{\bar{G}_n(\tau)} \tilde{f}_n^x(y) \end{aligned}$$

Since $\bar{G}(\tau) > 0$ in conjunction with the strong law of large numbers and the law of the iterated logarithm on the censoring law (see formula (4.28) in Deheuvels and Einmahl [19]), the result is an immediate consequence of Lemmas 3.2 and 3.3 ■

Proof of Theorem 3.3.2. One can easily show that

$$|\zeta(\widehat{\theta}_n(x)|x) - \zeta(\theta(x)|x)| \leq 2 \sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)| \quad (3.19)$$

Using Taylor's expansion $\zeta(\cdot|x)$, we get

$$\zeta(\widehat{\theta}_n(x)|x) - \zeta(\theta(x)|x) = \frac{1}{2}(\widehat{\theta}_n(x) - \theta(x))^2 \zeta^{(2)}(\theta^*(x)|x), \quad (3.20)$$

where $\theta^*(x)$ is between $\theta(x)$ and $\widehat{\theta}_n(x)$. Combining Equations (3.19) and (3.20), we obtain

$$|\widehat{\theta}_n(x) - \theta(x)| \leq 2 \sqrt{\frac{\sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)|}{\zeta^{(2)}(\theta^*(x)|x)}}.$$

The almost sure consistency of $\widehat{\theta}_n(x)$ follows, then, immediately from Theorem 3.3.1. ■

Local linear estimation of the conditional density for functional dependent and censored data : The Asymptotic Normality

This chapter has been accepted for publication in
South African Statistical Journal

Contents

4.1	Introduction	55
4.2	The model	57
4.3	Notations and hypotheses	58
4.4	Main results : Asymptotic normality	61
4.5	Appendix	62

This chapter present a study of the asymptotic behavior of the nonparametric L.L.E of the conditional density when the interest variable is subject to random right censoring of scalar response variable given a functional random covariate taking values in a semi-metric space. Under some regularity conditions, the joint asymptotic normality of the estimators of the conditional density is established.

4.1 Introduction

The conditional density function plays an important role in the nonparametric prediction. In addition, it provides the most informative summary of the relationships between a variable of interest Y and a covariate X . There is an extensive literature on the conditional density function estimation when the data are either independent or dependent and in finite or infinite dimensional spaces. For example, Hyndman et al. [42] studied kernel estimator of the conditional density and its bias-corrected version. Fan et al. [29] developed a direct estimation method via an innovative "double-kernel" local linear approach. Bashtannyk and Hyndman [4] and Hyndman and Yao [41] proposed several simple and useful rules for selecting bandwidths for the conditional density estimation. Hall et al. [35] applied the cross-validation technique to estimate the conditional density. Fan and Yim [28] proposed a consistent data-driven bandwidth selection procedure in estimating the conditional density functions. In these papers, it is assumed that the data are fully observed.

In the case of finite dimensional data, it is well known that the kernel method is inferior to the local linear fitting because of the limitations including large bias, non-adaptation as well as boundary effects. Recently, some results on the local linear modeling in the functional data setting have been done. Baillo and Grane [2] first proposed a L.L.E of the regression operator when the explanatory variable takes values in a Hilbert space. When the explanatory variable takes values in a semi-metric space, Barrientos-Marin et al. [3] proposed another alternative version of the L.L.E of the regression operator in the i.i.d setup, which was called as a locally modelled regression estimator. They found that the estimator made its computation easy and fast while keeping good performance. Then, this method has been employed to estimate the conditional density (see Demongeot et al. [21] and Rachdi et al. [62]), the conditional distribution (see Demongeot et al. [22]) and the conditional quantile (see Messaci et al. [59]) of a scalar response given a functional explanatory variable in the i.i.d setting.

In some fields as reliability or survival analysis, the r.v Y (which has a common unknown continuous d.f F) can be regarded as the lifetime of patients under study.

In reality it is not possible to observe the survival time of all patients, and often some of them are still alive at the end of the study, withdraw, or die from other causes than those addressed by the study. In those cases, we observe another r.v C called censoring. Then, assuming that $\{Y_i, i \geq 1\}$ is a stationary sequence which satisfy some kind of dependency and $\{C_i, i \geq 1\}$ is a sequence of i.i.d censoring r.v's with common unknown continuous d.f G and we observe only the n pairs (T_i, δ_i) for $i = 1, \dots, n$, where $T_i = \min(Y_i, C_i)$ and $\delta_i = \mathbb{I}_{(Y_i \leq C_i)}$. We suppose that (Y_i) and (C_i) for $i = 1, \dots, n$ are independent which ensures the identifiability of the model.

Moreover, in this case the distribution J of T_1 satisfies $1 - J = (1 - F)(1 - G)$.

Now let X be a r.v taking values in \mathcal{F} , where \mathcal{F} is a semi-metric space equipped with a semi-metric d . A semi metric space (\mathcal{F}, d) satisfies all the conditions of a metric space except it needs not satisfy that $d(x_1, x_2) = 0, x_1, x_2 \in \mathcal{F} \Rightarrow x_1 = x_2$. For example, set $\mathcal{F}_m = C^m[0, 1]$, i.e., the set of functions with continuous m^{th} derivatives on $[0, 1]$, $m \geq 0$; and the semi-metric $d_m(., .)$ is defined as

$$d_m(x_1, x_2) = \left\{ \int_0^1 [x_1^{(m)}(t) - x_2^{(m)}(t)]^2 dt \right\}^{1/2}, x_1, x_2 \in \mathcal{F}_m,$$

where $x_1^{(m)}(.)$ and $x_2^{(m)}(.)$ denote the m^{th} derivatives of $x_1(.)$ and $x_2(.)$, respectively. In the case that $m = 3$, it is easily seen that if $x_1(t) = t, x_2(t) = t^2, t \in [0, 1]$, then $d_3(x_1, x_2) = 0$. Then, (\mathcal{F}_3, d_3) is a semi-metric space .

The α -mixing (*strong mixing*) condition is the weakest among mixing conditions known in the literature and it has an important role in a number of applications with survival data [see Kang and Koehler [44] or Cai et al. [11]]. We begin by recalling the definition of the strong mixing property. For this we introduce the following notations. Let $\mathcal{F}_i^k(V)$ denote the σ - algebra generated by $\{V_j, i \leq j \leq k\}$.

In this paper we are interested to establish the asymptotic normality of the L.L.E of the conditional density when the response variable is subject to random censoring and the observations are an α -mixing processes. In the case of finite dimensional data, Fan et al.[29] established the joint asymptotic normality of conditional density and its derivative under stationary ρ -mixing processes, Liang and Baek [55] also get similar results under left-truncated and α -mixing data. Recently, Xiong et al.[76] established the asymptotic normality of the L.L.E of the conditional density

for functional time series data. This work will extend their results to α -mixing data for functional censored data.

4.2 The model

We assume that there exists a regular version of the conditional probability of Y given X , which is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} and has two-times continuously differentiable density, denoted by $\zeta(y|x)$. When the data are complete, Local polynomial smoothing is based on the assumption that functional parameter is smooth enough to be locally well approximated by a polynomial. In functional statistics, there are several ways for extending the local linear ideas (cf. Barrientos et al.[3], Baïllo. and Grané [2]). Here we adopt the fast functional locally modeling, that is, we estimate the conditional density $\zeta(y|x)$ by \widehat{a} where the couple $(\widehat{a}, \widehat{b})$ is obtained by minimizing the following quantity :

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n \left(h_H^{-1} H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1} \delta(x, X_i)),$$

where $\beta(.,.)$ and $\delta(.,.)$ are locating functions defined from \mathcal{F}^2 to \mathbb{R} , such that $\forall x \in \mathcal{F}$, $\beta(x, x) = 0$ and $d(.,.) = |\delta(.,.)|$, with K and H are kernels and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is chosen as a sequence of positive real numbers.

In the censored case, we adapt the idea of Carbonez et al. [12], Kohler et al. [50], and Khardani et al. [46] to the infinite dimension case by using a smooth d.f $H(.)$ instead of a step function, In practice $\overline{G}(.):= 1 - G(.)$ is unknown. Then, we replace by its Kaplan-Meier [45] estimate $\overline{G}_n(.)$ given by

$$\overline{G}_n(y) := 1 - G_n(.) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1} \right)^{\mathbb{1}_{\{T_{(i)} \leq y\}}} & \text{if } y < T_{(n)} \\ 0 & \text{otherwise,} \end{cases}$$

where $T_{(1)} < T_{(2)} < \dots < T_{(n)}$ are the order statistics of T_i and $\delta_{(i)}$ is the concomitant of $T_{(i)}$, which is known to be uniformly convergent to \overline{G} . Then the estimator of $\zeta(y|x)$ is defined as \widehat{a} where the couple $(\widehat{a}, \widehat{b})$ minimizes the following quantity :

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n \left(h_H^{-1} \delta_i \overline{G}_n^{-1}(T_i) H(h_H^{-1}(y - T_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1} \delta(x, X_i)). \quad (4.1)$$

Here we denote \widehat{a} by $\widehat{\zeta}_n(y|x)$. In addition, the estimator \widehat{b} may similarly be called as the estimator of the *derivative* of $\zeta(y|x)$.

In what follows, we put, for any $x \in \mathcal{F}$, and for all $i = 1, \dots, n$:

$$K_i = K(h_K^{-1} \delta(x, X_i)), \quad \beta_i = \beta(X_i, x) \text{ and } H_i = H(h_H^{-1}(y - T_i)).$$

Let

$$\mathbf{X} = \begin{pmatrix} 1 & \beta_1 \\ \vdots & \vdots \\ 1 & \beta_n \end{pmatrix}, \quad \mathbf{H}^* = \begin{pmatrix} h_H^{-1} \delta_1 \overline{G}_n^{-1}(T_1) H_1 \\ \vdots \\ h_H^{-1} \delta_n \overline{G}_n^{-1}(T_n) H_n \end{pmatrix}, \quad \mathbf{W} = \text{diag}(K(h_K^{-1} \delta(x, X_i))).$$

Then, from (4.1) a simple algebra shows that

$$(\widehat{a}, \widehat{b})^t = (\mathbf{X}^t \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{W} \mathbf{H}^*.$$

Denote by

$$s_{nj} = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n \left(\frac{\beta_i}{h_K}\right)^j K_i, \quad t_{nj} = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n \left(\frac{\beta_i}{h_K}\right)^j K_i h_H^{-1} \delta_i \overline{G}_n^{-1}(T_i) H_i,$$

$$\mathbf{S}_n = \begin{pmatrix} s_{n0} & s_{n1} \\ s_{n1} & s_{n2} \end{pmatrix} \text{ and } \mathbf{t}_n = \begin{pmatrix} t_{n0} \\ t_{n1} \end{pmatrix}.$$

Then

$$(\widehat{\zeta}_n(y|x), \widehat{b})^t = \text{diag}(1, h_K^{-1}) \mathbf{S}_n^{-1} \mathbf{t}_n.$$

Clearly, by a simple algebra, we get explicitly the following definition of $\widehat{\zeta}_n(y|x)$:

$$\widehat{\zeta}_n(y|x) = \frac{\sum_{i,j=1}^n \delta_i \overline{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)},$$

where

$$W_{ij}(x) = \beta(X_i, x) (\beta(X_i, x) - \beta(X_j, x)) K(h_K^{-1} \delta(x, X_i)) K(h_K^{-1} \delta(x, X_j)),$$

with the convention $0/0 = 0$.

4.3 Notations and hypotheses

In what follows, for any d.f L , let $\tau_L := \sup\{t : L(t) < 1\}$ be its support's right endpoint. Observe that $\tau_J = \min(\tau_F, \tau_G)$ and consider $\tau < \tau_J$. Let Ω be a compact subset of $(-\infty, \tau]$.

In the sequel, let C, C_1 , and C_2 denote generic finite positive constants, whose values are unimportant and may change from line to line. Set

$$\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(x, X) \leq r_1),$$

$$\psi_l(\cdot) = \frac{\partial^l \zeta(y|\cdot)}{\partial y^l}$$

and

$$\Psi_l(s) = \mathbb{E}[\psi_l(X) - \psi_l(x) | \delta(x, X) = s], \text{ for some } l \in \{0, 2\}.$$

We will assume the following hypotheses :

(H1) For any $r > 0$, $\phi_x(r) := \phi_x(-r, r) > 0$ and there exists a function $\chi_x(\cdot)$ such that :

$$\forall t \in [-1, 1], \lim_{h \rightarrow 0} \frac{\phi_x(-h, th)}{\phi_x(h)} = \chi_x(t).$$

(H2) For any $l \in \{0, 2\}$, the quantities $\Psi_l'(0)$ and $\Psi_l''(0)$ exist, where Ψ_l' and Ψ_l'' denotes the first and the second derivatives of Ψ_l , respectively.

(H3) $(Y_i, X_i)_{i \geq 1}$ is a sequence of stationary α -mixing r.v's with coefficient $\alpha(n)$.

(H4) The mixing coefficient $\alpha(n)$ satisfies :

(i)- $\alpha(n) = O(n^{-\lambda})$ for some $\lambda > 3$.

(ii)- There exist positive integers $q := q_n$ such that $q = o((n h_H \phi_x(h_K))^{1/2})$ and

$$\lim_{n \rightarrow \infty} (n (h_H \phi_x(h_K))^{-1})^{1/2} \alpha(q) = 0.$$

(H5) The locating operator $\beta(\cdot, \cdot)$ satisfies the following three conditions :

$$\left\{ \begin{array}{l} (i) \forall z \in \mathcal{F}, C_1 |\delta(x, z)| \leq |\beta(x, z)| \leq C_2 |\delta(x, z)|, \text{ where } 0 < C_1 < C_2; \\ (ii) \sup_{u \in B(x, r)} |\beta(u, x) - \delta(x, u)| = o(r) \\ \text{and} \\ (iii) h_K \int_{B(x, h_K)} \beta(u, x) dP_X(u) = o\left(\int_{B(x, h_K)} \beta^2(u, x) dP_X(u) \right), \end{array} \right.$$

where $B(x, r) = \{z \in \mathcal{F} / |\delta(x, z)| \leq r\}$ is a ball centered at x with radius r and $P_X(u)$ is the probability distribution of X .

(H6) (i)- K is a positive, differentiable function supported within $[-1, 1]$. Its derivative K' satisfies $K'(t) < 0$, for $-1 \leq t < 1$, and $K(1) > 0$.

(ii)- The r.v $\delta(X, x)$ is measurable with respect to the σ -field generated by the r.v $\beta(X, x)$.

(H7) H is a positive function, integrable, bounded, symmetric and such that :

$$\int H(t)dt = 1 \text{ and } \int t^2 H(t)dt < \infty.$$

(H8) The bandwidths h_K and h_H satisfy :

$$\lim_{n \rightarrow \infty} h_K = 0, \lim_{n \rightarrow \infty} h_H = 0 \text{ and } \lim_{n \rightarrow \infty} n h_H \phi_x(h_K) = \infty.$$

(H9) (i)- $\sup_{i \neq j} \mathbb{P}[(X_i, X_j) \in B(x, h_K) \times B(x, h_K)] \leq f(x)g(h_K)$ as $h_K \rightarrow 0$, where $g(h_K) \rightarrow 0$ as $h_K \rightarrow 0$ and $f(x)$ is a nonnegative functional in $x \in \mathcal{F}$. We assume that the ratio $g(h_K)/\phi_x^2(h_K)$ is bounded.

(ii)- For all $j > 1$, the joint conditional density $\zeta(\cdot, \cdot | \cdot, \cdot)$ of (Y_1, Y_j) given (X_1, X_j) exists on $\mathbb{R} \times \mathbb{R} \times \mathcal{F} \times \mathcal{F}$ and satisfies $\zeta(y_1, y_j | x_1, x_j) \leq C$ for $(y_1, y_j, x_1, x_j) \in \mathbb{R} \times \mathbb{R} \times B(x; r_0) \times B(x; r_0)$, where $r_0 > 0$.

(iii)- For all $j > 1$, the conditional density $\zeta(\cdot | \cdot, \cdot)$ of Y_1 given (X_1, X_j) exists on $\mathbb{R} \times \mathcal{F} \times \mathcal{F}$ and satisfies $\zeta(y_1 | x_1, x_j) \leq C$ for $(y_1, x_1, x_j) \in \mathbb{R} \times B(x; r_0) \times B(x; r_0)$.

(iv)- For all $j > 1$, the conditional density $\zeta(\cdot | \cdot, \cdot)$ of Y_j given (X_1, X_j) exists on $\mathbb{R} \times \mathcal{F} \times \mathcal{F}$ and satisfies $\zeta(y_j | x_1, x_j) \leq C$ for $(y_j, x_1, x_j) \in \mathbb{R} \times B(x; r_0) \times B(x; r_0)$.

(H10) $\lim_{n \rightarrow \infty} n h_H^5 \phi_x(h_K) = 0$ and $\lim_{n \rightarrow \infty} n h_H h_K^4 \phi_x(h_K) = 0$.

The assumption (H1) characterizes the concentration property of the probability measure of the functional variable X , while the assumption (H2) is a regularity condition which characterizes the functional space of our model. Assumptions (H3) and (H4-(i)) specify the model and the rate of mixing coefficient. Conditions in (H4) allow us to employ Bernstein's big-block and small-block technique to prove asymptotic normality for an α -mixing sequence. Assumption (H5-(i)) is necessary to control the shape of the local functional object β ; assumption (H5-(ii)) is unrestrictive and assumption (H5-(iii)) is a pivotal hypothesis on the local performance of the operator β . Assumption (H6-(i)), (H7) and (H8) are used commonly in the literature. The assumption (H9) is mainly technical, which is employed to simplify the calculations of covariances in the proof. Finally, the assumption (H10) is used to remove the bias term.

4.4 Main results : Asymptotic normality

To give the main result, we list some notations. In the sequel, we denote :

$$\mathbf{M}_a = K^a(1) - \int_{-1}^1 (K^a(u))' \chi_x(u) du$$

where $a > 0$,

$$\mathbf{N}(a, b) = K^a(1) - \int_{-1}^1 (u^b K^a(u))' \chi_x(u) du$$

for all $a > 0$ and $b > 1$;

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\mathbf{N}(1,2)}{\mathbf{M}_1} \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \frac{\mathbf{M}_2}{\mathbf{M}_1^2} & 0 \\ 0 & \frac{\mathbf{N}(2,2)}{\mathbf{M}_1^2} \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} \frac{\mathbf{N}(1,2)}{\mathbf{M}_1} \\ \frac{\mathbf{N}(1,3)}{\mathbf{M}_1} \end{pmatrix}.$$

Theorem 4.4.1. *Suppose that assumptions (H1)-(H9) hold. Then*

$$(nh_H \phi_x(h_K))^{1/2} \left[\text{diag}(1, h_K) \begin{pmatrix} \widehat{\zeta}_n(y|x) - \zeta(y|x) \\ \widehat{b} - \Psi_0'(0) \end{pmatrix} - \frac{h_K^2}{2} \Psi_0''(0) \mathbf{S}^{-1} \mathbf{U} \right. \\ \left. - \frac{h_H^2}{2} \psi_2(x) \int t^2 H(t) dt \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \xrightarrow{D} N \left(0, \overline{G}^{-1}(y) \zeta(y|x) \int H^2(t) dt \mathbf{S}^{-1} \mathbf{V} \mathbf{S}^{-1} \right). \quad (4.2)$$

Corollary 4.1. *Suppose that assumptions (H1)-(H9) hold. Then*

$$(nh_H \phi_x(h_K))^{1/2} (\widehat{\zeta}_n(y|x) - \zeta(y|x) - \frac{h_K^2}{2} \Psi_0''(0) \frac{\mathbf{N}(1,2)}{\mathbf{M}_1} - \frac{h_H^2}{2} \psi_2(x) \int t^2 H(t) dt) \\ \xrightarrow{D} N(0, V_{HK}^x(y)), \quad (4.3)$$

where

$$V_{HK}^x(y) = \overline{G}^{-1}(y) \zeta(y|x) \frac{\mathbf{M}_2}{\mathbf{M}_1^2} \int H^2(t) dt.$$

Now to construct confidence intervals for $\zeta(y|x)$ we need to remove the bias term and obtain a plug-in estimator of

$$\overline{G}^{-1}(y) \frac{\zeta(y|x)}{h_H} \frac{\mathbf{M}_2}{\mathbf{M}_1^2 n \phi_x(h_K)} \int H^2(t) dt, \quad (4.4)$$

Corollary 4.2. *Suppose that assumptions (H1)-(H10) hold. Then*

$$\left(\frac{nh_H \phi_x(h_K)}{V_{HK}^x(y)} \right)^{1/2} (\widehat{\zeta}_n(y|x) - \zeta(y|x)) \xrightarrow{D} N(0, 1).$$

On the other hand : by assumptions (H1),(H5) and (H6)(i) we know that : $\mathbf{M}_2/\mathbf{M}_1^2 n\phi_x(h_K)$ can be estimated by $\mathbb{E}(K_1^2)/\mathbb{E}^2(K_1)$, and by applying the kernel estimator of $\zeta(y|x)$ and the Kaplan-Meier estimator of $\overline{G}^{-1}(y)$ given above, the quantity (4.4) can be estimated by :

$$\overline{G}_n^{-1}(y) \frac{\widehat{\zeta}_n(y|x)}{h_H} \frac{\mathbb{E}(K_1^2)}{\mathbb{E}^2(K_1)} \int H^2(t)dt \doteq \hat{\sigma}^2(y/x).$$

Then, we approximate $(1 - \gamma)$ confidence interval of $\zeta(y|x)$ by :

$$\left[\widehat{\zeta}_n(y|x) - u_{1-\gamma/2} \hat{\sigma}(y/x); \widehat{\zeta}_n(y|x) + u_{1-\gamma/2} \hat{\sigma}(y/x) \right],$$

where $u_{1-\gamma/2}$ denotes $1 - \gamma/2$ quantile of standard Normal $N(0,1)$.

4.5 Appendix

To proof the main result we need the following lemmas.

Lemma 4.1. (see Rachdi et al. [62])

Suppose that assumptions (H1), (H5) and (H6)(i) hold. Then :

- (a) $\mathbb{E}[K_1^a] = \mathbf{M}_a \phi_x(h_k) + o(\phi_x(h_k))$, for $a > 0$;
- (b) $\mathbb{E}[K_1^a \beta_1] = o(h_k \phi_x(h_k))$, for all $a > 0$;
- (c) $\mathbb{E}[K_1^a \beta_1^b] = N(a, b) h_k^b \phi_x(h_k) + o(h_k^b \phi_x(h_k))$, for all $a > 0, b > 1$.

Lemma 4.2. (see Xiong et al. [76])

Suppose that assumptions (H1)-(H9) hold, then :

$$s_{n0} \xrightarrow{P} 1, s_{n1} \xrightarrow{P} 0, s_{n2} \xrightarrow{P} \frac{\mathbf{N}(1, 2)}{\mathbf{M}_1}, s_{n3} \xrightarrow{P} \frac{\mathbf{N}(1, 3)}{\mathbf{M}_1}.$$

Lemma 4.3. (see Volkonskii [74])

Let V_1, \dots, V_m be α -mixing r.v's measurable with respect to the σ -algebra

$\mathcal{F}_{i_1}^{j_1}, \dots, \mathcal{F}_{i_m}^{j_m}$, respectively, with $1 \leq i_1 < j_1 < \dots < j_m \leq n$, $1 \leq w \leq i_{l+1} - j_l$ and $|V_j| \leq 1$ for $l, j = 1, 2, \dots, m$. Then

$$\left| \mathbb{E}\left[\prod_{j=1}^m V_j\right] - \prod_{j=1}^m \mathbb{E}[V_j] \right| \leq 16(m-1)\alpha_w,$$

where $\mathcal{F}_a^b = \sigma\{V_i, a \leq i \leq b\}$ and α_w is the mixing coefficient.

Lemma 4.4. (see Davydov [17], Corollary, p 692)

Suppose that X and Y are r.v's, such that $\mathbb{E}|X|^p < \infty$, $\mathbb{E}|Y|^q < \infty$, where $p, q > 1$, $p^{-1} + q^{-1} < 1$, then

$$|\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \leq 8\mathbb{E}^{p-1}|X|^p\mathbb{E}^{q-1}|Y|^q \left\{ \sup_{A \in \sigma(X), B \in \sigma(Y)} |p(A \cap B) - p(A)p(B)| \right\}^{1-p^{-1}-q^{-1}}.$$

Lemma 4.5. Suppose that assumptions (H2), (H5)-(H7) and (H9) hold, then :

$$(a) h_H^{-2} \mathbb{E}[K_1 K_j \delta_1 \bar{G}^{-1}(T_1) H_1 \delta_j \bar{G}^{-1}(T_j) H_j] = O(g(h_K));$$

$$h_H^{-1} \mathbb{E}[K_1 K_j \delta_1 \bar{G}^{-1}(T_1) H_1] = O(g(h_K));$$

$$h_H^{-1} \mathbb{E}[K_1 K_j \delta_j \bar{G}^{-1}(T_j) H_j] = O(g(h_K)) \mathbb{E}[K_1 K_j] = O(g(h_K)), \text{ for all } j > 1$$

$$(b) h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \delta_1 \bar{G}^{-c}(T_1) H_1^c] = \bar{G}^{1-c}(y) \int H^c(t) dt \left\{ \psi_0(x) \mathbb{E}[K_1^a \beta_1^b] + \Psi_0'(0) \mathbb{E}[K_1^a \beta_1^{b+1}] \right\} \\ + \frac{h_H^2}{2} \bar{G}^{1-c}(y) \int t^2 H^c(t) dt \left\{ \psi_2(x) \mathbb{E}[K_1^a \beta_1^b] \right. \\ \left. + \Psi_2'(0) \mathbb{E}[K_1^a \beta_1^{b+1}] + o(\mathbb{E}[K_1^a \beta_1^{b+1}]) + o(h_H^2 \mathbb{E}[K_1^a \beta_1^b]) \right\};$$

for $a > 0$, $b = 0$ and $c = 0$, or $b > 1$ and $c > 1$.

Proof of Theorem 4.4.1. let

$$T_{nj} = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n \left(\frac{\beta_i}{h_K} \right)^j K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i - \mathbb{E}[\zeta(y|X_i, \beta_i)]),$$

$$t_{nj}^* = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n \left(\frac{\beta_i}{h_K} \right)^j K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i - \mathbb{E}[\zeta(y|X_i, \beta_i)]),$$

$$\mathbf{T}_n = (T_{n0}, T_{n1})^\tau \text{ and } \mathbf{t}_n^* = (t_{n0}^*, t_{n1}^*)^\tau.$$

In view of (H2), when $|\beta_i| \leq Ch_K$; by the Taylor expansion, we get

$$\Psi_0(\beta_i) = \Psi_0'(0)\beta_i + \frac{1}{2}\Psi_0''(0)\beta_i^2 + o(\beta_i^2).$$

Then

$$\begin{aligned} \mathbb{E}[\zeta(y|X_i, \beta_i)] &= \zeta(y|x) + \mathbb{E}[\zeta(y|X_i) - \zeta(y|x)|\beta_i] \\ &= \zeta(y|x) + \Psi_0(\beta_i) \\ &= \zeta(y|x) + \Psi_0'(0)\beta_i + \frac{1}{2}\Psi_0''(0)\beta_i^2 + o(\beta_i^2). \end{aligned} \tag{4.5}$$

Therefore

$$(\mathbb{E}[\zeta(y|X_1, \beta_1)], \dots, \mathbb{E}[\zeta(y|X_n, \beta_n)])^\tau = \mathbf{X} \begin{pmatrix} \zeta(y|x) \\ \Psi_0'(0) \end{pmatrix} + \frac{1}{2} \Psi_0''(0) (\beta_1^2, \dots, \beta_n^2)^\tau + (o(\beta_1^2), \dots, o(\beta_n^2))^\tau.$$

Then

$$\begin{aligned} \mathbf{S}_n^{-1} \mathbf{t}_n^* &= \text{diag}(1, h_K) \begin{pmatrix} \widehat{\zeta}_n(y|x) - \zeta(y|x) \\ \widehat{\Psi}_0'(0) - \Psi_0'(0) \end{pmatrix} - \frac{h_K^2}{2} \Psi_0''(0) \mathbf{S}_n^{-1} \begin{pmatrix} s_{n2} \\ s_{n3} \end{pmatrix} \\ &\quad - o(h_K^2) \mathbf{S}_n^{-1} \begin{pmatrix} s_{n2} \\ s_{n3} \end{pmatrix} - \mathbf{S}_n^{-1} (\mathbf{T}_n - \mathbf{t}_n^*). \end{aligned}$$

Lemma 4.2 implies that

$$\mathbf{S}_n \xrightarrow{P} \mathbf{S},$$

and

$$\mathbf{S}_n^{-1} \begin{pmatrix} s_{n2} \\ s_{n3} \end{pmatrix} \xrightarrow{P} \mathbf{S}^{-1} \mathbf{U}.$$

The rest proofs are divided into the following two steps.

step1 : We verify that

$$\mathbf{S}_n^{-1} (\mathbf{T}_n - \mathbf{t}_n^*) = O_p\left(\sqrt{\frac{\log(\log n)}{n}}\right).$$

Note that, for $0 \leq j \leq 1$, we have

$$\begin{aligned} |T_{nj} - t_{nj}^*| &= \left| \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n \left(\frac{\beta_i}{h_K}\right)^j K_i h_H^{-1} \delta_i H_i [\overline{G}_n^{-1}(T_i) - \overline{G}^{-1}(T_i)] \right| \\ &\leq \sup_{t \in \Omega} |\overline{G}_n^{-1}(t) - \overline{G}^{-1}(t)| \frac{1}{nh_K^j \mathbb{E}[K_1]} \sum_{i=1}^n |\beta_i^j K_i h_H^{-1} H_i|. \end{aligned}$$

Then from $\frac{1}{nh_K^j \mathbb{E}[K_1]} \sum_{i=1}^n |\beta_i^j K_i h_H^{-1} H_i| = O_p(1)$, lemma 4.2 and the law of the iterated logarithm on the censoring law (see formula (4.28) in Deheuvels and Einmahl [19]), one obtains the result.

step2 : We prove that

$$(nh_H \phi_x(h_K))^{1/2} \left\{ t_n^* - \frac{h_H^2}{2} \psi_2(x) \int t^2 H(t) dt \begin{pmatrix} 1 \\ 0 \end{pmatrix} + o(h_H^2) \right\} \xrightarrow{D} N(0, \overline{G}^{-1}(y) \zeta(y|x) \int H^2 dt \mathbf{V}) \quad (4.6)$$

For any given vector of real numbers $\mathbf{a} = (a_0, a_1)^\tau \neq 0$, set

$$U_i = \frac{(h_H \phi_x(h_K))^{1/2}}{\mathbb{E}[K_1]} (a_0 + a_1 \frac{\beta_i}{h_K}) K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i - \mathbb{E}[\zeta(y|X_i, \beta_i)]), 1 \leq i \leq n$$

then

$$(nh_H \phi_x(h_K))^{1/2} \mathbf{a}^\tau t_n^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n (U_i - \mathbb{E}[U_i]) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E}[U_i]. \quad (4.7)$$

From the Cramer-Wold theorem and Equation (4.7), Equation (4.6) will hold if we can prove the following two claims :

$$\textbf{Claim 1} : \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E}[U_i] = (nh_H \phi_x(h_K))^{1/2} \left\{ \frac{a_0 h_H^2 \psi_2(x)}{2} \int t^2 H(t) dt + o(h_H^2) \right\};$$

$$\textbf{Claim 2} : \frac{1}{\sqrt{n}} \sum_{i=1}^n (U_i - \mathbb{E}[U_i]) \xrightarrow{D} N(0, \Delta^2(y|x)), \Delta^2(y|x) := \bar{G}^{-1}(y) \zeta(y|x) \int H^2(t) dt \mathbf{a}^\tau \mathbb{V} \mathbf{a}.$$

Claim 1 :

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E}[U_i] &= \sqrt{n} \mathbb{E}[U_1] \\ &= \frac{a_0 (nh_H \phi_x(h_K))^{1/2}}{\mathbb{E}[K_1]} \{h_H^{-1} \mathbb{E}[K_1 \delta_1 \bar{G}^{-1}(T_1) H_1] - \mathbb{E}[K_1 \psi_0(X_1)]\} \\ &\quad + \frac{a_1 (nh_H \phi_x(h_K))^{1/2}}{h_K \mathbb{E}[K_1]} \{h_H^{-1} \mathbb{E}[K_1 \beta_1 \delta_1 \bar{G}^{-1}(T_1) H_1] - \mathbb{E}[K_1 \beta_1 \psi_0(X_1)]\}, \end{aligned}$$

For $a = c = 1$, we obtain from Lemma 4.5 that :

$$\begin{aligned} &h_H^{-1} \mathbb{E}[K_1 \beta_1^b \delta_1 \bar{G}^{-1}(T_1) H_1] - \mathbb{E}[K_1 \beta_1^b \psi_0(X_1)] \\ &= \frac{h_H^2}{2} \int t^2 H(t) dt \{ \psi_2(x) \mathbb{E}[K_1 \beta_1^b] + \Psi_2'(0) \mathbb{E}[K_1 \beta_1^{b+1}] + o(\mathbb{E}[K_1 \beta_1^{b+1}]) \} + o(h_H^2 \mathbb{E}[K_1 \beta_1^b]). \end{aligned}$$

Then

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E}[U_i] &= \frac{a_0 (nh_H \phi_x(h_K))^{1/2}}{\mathbb{E}[K_1]} \mathbb{E}[K_1] \left\{ \frac{h_H^2}{2} \int t^2 H(t) dt \psi_2(x) + o(h_H^2) \right\} \\ &\quad + \frac{a_1 (nh_H \phi_x(h_K))^{1/2}}{h_K \mathbb{E}[K_1]} \left\{ \frac{h_H^2}{2} \int t^2 H(t) dt \{ \mathbb{E}[K_1 \beta_1] \psi_2(x) + \mathbb{E}[K_1 \beta_1^2] \Psi_2'(0) \} \right. \\ &\quad \left. + o(\mathbb{E}[K_1 \beta_1^2]) \right\} + o(h_H^2 \mathbb{E}[K_1 \beta_1]), \end{aligned}$$

Claim 1 follows from Lemma 4.1.

Claim 2 : The proof is similar to the proof of Theorem 4.1 in Xiong et al. [76].

Assumption (H4)-(ii) implies that there is a sequence of positive integers $\delta_n \rightarrow \infty$,

such that $\delta_n q_n = o((nh_H \phi_x(h_K))^{1/2})$ and $\delta_n (n(h_H \phi_x(h_K))^{-1})^{1/2} \alpha(q_n) \rightarrow 0$. Let $\omega = \lfloor \frac{n}{r+q} \rfloor$ and $r = \lfloor (nh_H \phi_x(h_K))^{1/2} / \delta_n \rfloor$. Then

$$q/r \rightarrow 0, \omega \alpha(q) \rightarrow 0, \omega q/n \rightarrow 0, r/n \rightarrow 0, r/(nh_H \phi_x(h_K))^{1/2} \rightarrow 0 \quad (4.8)$$

Next we will employ Bernstein's big-block and small-block procedure. Partition the set $\{1, 2, \dots, n\}$ into $2\omega + 1$ subsets with large block of size r and small block of size q . Let

$$W_i = U_i - \mathbb{E}[U_i], 1 \leq i \leq n, \zeta_{mn} = \sum_{i=k_m}^{k_m+r-1} W_i, \zeta'_{mn} = \sum_{i=l_m}^{l_m+q-1} W_i, \zeta''_{mn} = \sum_{i=\omega(r+q)+1}^n W_i,$$

where $k_m = (m-1)(r+q) + 1, l_m = (m-1)(r+q) + r + 1, m = 1, \dots, \omega$. Then

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n (U_i - \mathbb{E}[U_i]) &= n^{-1/2} \sum_{i=1}^n W_i \\ &= n^{-1/2} \left\{ \sum_{m=1}^{\omega} \zeta_{mn} + \sum_{m=1}^{\omega} \zeta'_{mn} + \zeta''_{\omega n} \right\} \\ &:= n^{-1/2} \{ \mathbb{I}m_{1n} + \mathbb{I}m_{2n} + \mathbb{I}m_{3n} \}. \end{aligned}$$

Then it suffices to show that

$$n^{-1} \mathbb{E}(\mathbb{I}m_{2n})^2 \rightarrow 0, n^{-1} \mathbb{E}(\mathbb{I}m_{3n})^2 \rightarrow 0, Var[n^{-1/2} \mathbb{I}m_{1n}] \rightarrow \Delta^2(y|x), \quad (4.9)$$

$$\left| \mathbb{E}[\exp(it \sum_{m=1}^{\omega} n^{-1/2} \zeta_{mn})] - \prod_{m=1}^{\omega} \mathbb{E}[\exp(it n^{-1/2} \zeta_{mn})] \right| \rightarrow 0, \quad (4.10)$$

$$A_n(\varepsilon) = \frac{1}{n} \sum_{m=1}^{\omega} \mathbb{E}[\zeta_{mn}^2 \mathbb{I}(|\zeta_{mn}| > \varepsilon \Delta(y|x) \sqrt{n})] \rightarrow 0, \forall \varepsilon > 0. \quad (4.11)$$

We first prove (4.9).

$$\begin{aligned} n^{-1} \mathbb{E}(\mathbb{I}m_{2n})^2 &= \frac{1}{n} \sum_{m=1}^{\omega} \sum_{i=l_m}^{l_m+q-1} \mathbb{E}[W_i^2] + \frac{2}{n} \sum_{m=1}^{\omega} \sum_{l_m \leq i < j \leq l_m+q-1} Cov(W_i, W_j) \\ &+ \frac{2}{n} \sum_{1 \leq i < j \leq \omega} Cov(\zeta'_{in}, \zeta'_{jn}) \\ &:= J_{1n} + J_{2n} + J_{3n}. \end{aligned}$$

Calculate $\mathbb{E}[W_1^2] = \mathbb{E}[U_1^2] - \mathbb{E}^2[U_1]$.

$$\begin{aligned}
\mathbb{E}[U_1^2] &= \frac{h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \mathbb{E}[(a_0 + a_1 \frac{\beta_1}{h_K})^2 K_1^2 (h_H^{-1} \delta_1 \bar{G}^{-1}(T_1) H_1 - \mathbb{E}[\zeta(y|X_1, \beta_1)])^2] \\
&= \frac{a_0^2 h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \delta_1 \bar{G}^{-2}(T_1) H_1^2] - 2h_H^{-1} \mathbb{E}[K_1^2 \delta_1 \bar{G}^{-1}(T_1) H_1 \mathbb{E}[\zeta(y|X_1, \beta_1)]] \\
&\quad + \mathbb{E}[K_1^2 \mathbb{E}^2[\zeta(y|X_1, \beta_1)]]\} \\
&\quad + \frac{a_1^2 h_H \phi_x(h_K)}{h_K^2 \mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \beta_1^2 \delta_1 \bar{G}^{-2}(T_1) H_1^2] - 2h_H^{-1} \mathbb{E}[K_1^2 \beta_1^2 \delta_1 \bar{G}^{-1}(T_1) H_1 \mathbb{E}[\zeta(y|X_1, \beta_1)]] \\
&\quad + \mathbb{E}[K_1^2 \beta_1^2 \mathbb{E}^2[\zeta(y|X_1, \beta_1)]]\} \\
&\quad + \frac{2a_0 a_1 h_H \phi_x(h_K)}{h_K \mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \beta_1 \delta_1 \bar{G}^{-2}(T_1) H_1^2] - 2h_H^{-1} \mathbb{E}[K_1^2 \beta_1 \delta_1 \bar{G}^{-1}(T_1) H_1 \mathbb{E}[\zeta(y|X_1, \beta_1)]] \\
&\quad + \mathbb{E}[K_1^2 \beta_1 \mathbb{E}^2[\zeta(y|X_1, \beta_1)]]\}.
\end{aligned}$$

Equation(4.5), assumptions (H6)(i), (H5) and (H2) imply that

$$K_1^2 \mathbb{E}[\zeta(y|X_1, \beta_1)] = K_1^2 O(1) \text{ and } K_1^2 \mathbb{E}^2[\zeta(y|X_1, \beta_1)] = K_1^2 O(1).$$

Then, we get

$$\begin{aligned}
\mathbb{E}[U_1^2] &= \frac{a_0^2 h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \delta_1 \bar{G}^{-2}(T_1) H_1^2] \\
&\quad - 2h_H^{-1} \mathbb{E}[K_1^2 \delta_1 \bar{G}^{-1}(T_1) H_1] O(1) + \mathbb{E}[K_1^2] O(1)\} \\
&\quad + \frac{a_1^2 h_H \phi_x(h_K)}{h_K^2 \mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \beta_1^2 \delta_1 \bar{G}^{-2}(T_1) H_1^2] \\
&\quad - 2h_H^{-1} \mathbb{E}[K_1^2 \beta_1^2 \delta_1 \bar{G}^{-1}(T_1) H_1] O(1) + \mathbb{E}[K_1^2 \beta_1^2] O(1)\} \\
&\quad + \frac{2a_0 a_1 h_H \phi_x(h_K)}{h_K \mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1^2 \beta_1 \delta_1 \bar{G}^{-2}(T_1) H_1^2] \\
&\quad - 2h_H^{-1} \mathbb{E}[K_1^2 \beta_1 \delta_1 \bar{G}^{-1}(T_1) H_1] O(1) \\
&\quad + \mathbb{E}[K_1^2 \beta_1] O(1)\},
\end{aligned}$$

which along with Lemma 4.1 and Lemma 4.5(b) implies

$$\begin{aligned}
\mathbb{E}[U_1^2] &\rightarrow a_0^2 \frac{\mathbf{M}_2}{\mathbf{M}_1^2} \bar{G}^{-1}(y) \int H^2(t) dt \psi_0(x) + a_1^2 \frac{\mathbf{N}(2, 2)}{\mathbf{M}_1^2} \bar{G}^{-1}(y) \int H^2(t) dt \psi_0(x) \\
&= a_0^2 \frac{\mathbf{M}_2}{\mathbf{M}_1^2} \bar{G}^{-1}(y) \int H^2(t) dt \zeta(y|x) + a_1^2 \frac{\mathbf{N}(2, 2)}{\mathbf{M}_1^2} \bar{G}^{-1}(y) \int H^2(t) dt \zeta(y|x) \\
&= \bar{G}^{-1}(y) \zeta(y|x) \int H^2(t) dt a^\tau \mathbb{V}a. \tag{4.12}
\end{aligned}$$

From **Claim 1** we have $\mathbb{E}[U_1] \rightarrow 0$: Then

$$\mathbb{E}[W_1^2] = \mathbb{E}[U_1^2] - \mathbb{E}^2[U_1] \rightarrow \Delta^2(y|x), \quad (4.13)$$

which yields that $J_{1n} = O(\omega q/n) \rightarrow 0$ by Equation (4.8). From the definition of $\mathbb{I}m_{2n}$, we know

$$|J_{2n}| \leq \frac{2}{n} \sum_{1 \leq i < j \leq n} |Cov(W_i, W_j)| = \frac{2}{n} \sum_{1 \leq i < j \leq n} |Cov(U_i, U_j)|,$$

$$|J_{3n}| \leq \frac{2}{n} \sum_{1 \leq i < j \leq n} |Cov(W_i, W_j)| = \frac{2}{n} \sum_{1 \leq i < j \leq n} |Cov(U_i, U_j)|.$$

Therefore, to prove $J_{2n} = o(1)$ and $J_{3n} = o(1)$, we need only to prove that

$$\frac{1}{n} \sum_{1 \leq i < j \leq n} |Cov(U_i, U_j)| \rightarrow 0. \quad (4.14)$$

Take $c_n = [h_H \phi_x(h_K)]^{-(1-\frac{1}{\lambda})/\eta}$, for some $1 - 1/\lambda < \eta < \lambda - 2$.

Then we set

$$G_1 = \{(1, j) : j \in \{1, \dots, n\}, 1 \leq j-1 \leq c_n\}$$

$$G_2 = \{(1, j) : j \in \{1, \dots, n\}, c_n + 1 \leq j-1 \leq n-1\}$$

According to the above splitting, we get

$$\frac{1}{n} \sum_{1 \leq i < j \leq n} |Cov(U_i, U_j)| = \sum_{j \in G_1} \left(1 - \frac{j-1}{n}\right) |Cov(U_1, U_j)| + \sum_{j \in G_2} \left(1 - \frac{j-1}{n}\right) |Cov(U_1, U_j)|. \quad (4.15)$$

Note that

$$|Cov(U_i, U_j)| \leq |\mathbb{E}[U_1 U_j]| + \mathbb{E}^2[U_1]. \quad (4.16)$$

Equation (4.5), assumptions (H6)(i), (H5) and (H2) imply that

$$K_1 \mathbb{E}[\zeta(y|X_1, \beta_1)] = K_1 O(1) \text{ and } \frac{\beta_1}{h_K} K_1 = K_1 O(1). \quad (4.17)$$

Then

$$\begin{aligned}
|\mathbb{E}[U_1 U_j]| &\leq \mathbb{E}|U_1 U_j| \\
&= \frac{h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \mathbb{E} \left| (a_0 + a_1 \frac{\beta_1}{h_K}) K_1 (h_H^{-1} \delta_1 \bar{G}^{-1}(T_1) H_1 - \mathbb{E}[\zeta(y|X_1, \beta_1)]) \right. \\
&\quad \left. (a_0 + a_1 \frac{\beta_j}{h_K}) K_j (h_H^{-1} \delta_j \bar{G}^{-1}(T_j) H_j - \mathbb{E}[\zeta(y|X_j, \beta_j)]) \right| \\
&\leq \frac{h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \mathbb{E} [O(1) K_1 K_j (h_H^{-1} \delta_1 \bar{G}^{-1}(T_1) H_1 + O(1)) (h_H^{-1} \delta_j \bar{G}^{-1}(T_j) H_j + O(1))] \\
&= O(1) \frac{h_H \phi_x(h_K)}{\mathbb{E}^2[K_1]} \{h_H^{-2} \mathbb{E}[K_1 K_j \delta_1 \bar{G}^{-1}(T_1) H_1 \delta_j \bar{G}^{-1}(T_j) H_j] \\
&\quad + h_H^{-1} \mathbb{E}[K_1 K_j \delta_1 \bar{G}^{-1}(T_1) H_1] + h_H^{-1} \mathbb{E}[K_1 K_j \delta_j \bar{G}^{-1}(T_j) H_j] + \mathbb{E}[K_1 K_j]\},
\end{aligned}$$

which with Lemma 4.5(a) and assumption (H9)(i) implies that

$$|\mathbb{E}[U_1 U_j]| = O(h_H \phi_x(h_K)). \quad (4.18)$$

In addition, **Claim 1** yields that

$$\mathbb{E}^2[U_1] = O(h_H^5 \phi_x(h_K)). \quad (4.19)$$

Then, Equations (4.16), (4.18) and (4.19) yield that

$$\text{Cov}(U_1, U_j) = O(h_H \phi_x(h_K)). \quad (4.20)$$

Equation (4.20) implies that

$$\begin{aligned}
\sum_{j \in G_1} (1 - \frac{j-1}{n}) |\text{Cov}(U_1, U_j)| &= O(1) \sum_{j=1}^{c_n} h_H \phi_x(h_K) = O(c_n h_H \phi_x(h_K)) \\
&= O(1) [h_H \phi_x(h_K)]^{1-(1-\frac{1}{\lambda})/\eta} \rightarrow 0.
\end{aligned} \quad (4.21)$$

On the other hand, it follows from Lemma 4.4 that

$$|\text{Cov}(U_1, U_j)| \leq 8 [\mathbb{E}|U_1|^{2\lambda}]^{1/\lambda} [\alpha(j-1)]^{1-\frac{1}{\lambda}}. \quad (4.22)$$

The C^r -inequality and Equation (4.17) imply that

$$\begin{aligned}
\mathbb{E}[|U_1|]^{2\lambda} &= \frac{(h_H \phi_x(h_K))^\lambda}{\mathbb{E}^{2\lambda}[K_1]} \mathbb{E} \left| (a_0 + a_1 \frac{\beta_1}{h_K}) K_1 (h_H^{-1} \delta_1 \bar{G}^{-1}(T_1) H_1 - \mathbb{E}[\zeta(y|X_1, \beta_1)]) \right|^{2\lambda} \\
&= \frac{(h_H \phi_x(h_K))^\lambda}{\mathbb{E}^{2\lambda}[K_1]} \mathbb{E} \left| O(1) K_1 (h_H^{-1} \delta_1 \bar{G}^{-1}(T_1) H_1 + O(1)) \right|^{2\lambda} \\
&= O(1) \frac{(h_H \phi_x(h_K))^\lambda}{\mathbb{E}^{2\lambda}[K_1]} \{ \mathbb{E}[h_H^{-2\lambda} K_1^{2\lambda} \delta_1 \bar{G}^{-2\lambda}(T_1) H_1^{2\lambda}] + \mathbb{E}[K_1^{2\lambda}] \} \\
&= O(1) \frac{(h_H \phi_x(h_K))^\lambda}{\mathbb{E}^{2\lambda}[K_1]} \{ h_H^{1-2\lambda} \mathbb{E}[h_H^{-1} K_1^{2\lambda} \delta_1 \bar{G}^{-2\lambda}(T_1) H_1^{2\lambda}] + \mathbb{E}[K_1^{2\lambda}] \},
\end{aligned}$$

again by Lemma 4.1 and Lemma 4.5(b),

$$\mathbb{E}|U_1|^{2\lambda} = O((h_H \phi_x(h_K))^{1-\lambda}). \quad (4.23)$$

It follows from Equations (4.22) and (4.23) that

$$\begin{aligned} & \sum_{j \in G_2} \left(1 - \frac{j-1}{n}\right) |\text{Cov}(U_1, U_j)| \\ &= O(1) \sum_{j \in G_2} [\alpha(j-1)]^{1-\frac{1}{\lambda}} (h_H \phi_x(h_K))^{-(1-\frac{1}{\lambda})} \\ &\leq O(1) \sum_{m \geq C_n+1} m^{-(\lambda-1)} (h_H \phi_x(h_K))^{-(1-\frac{1}{\lambda})} \\ &\leq O(1) C_n^{-(\lambda-2)} (h_H \phi_x(h_K))^{-(1-\frac{1}{\lambda})} \\ &\leq O(1) (h_H \phi_x(h_K))^{-(1-\frac{1}{\lambda})} [h_H \phi_x(h_K)]^{\frac{(1-\frac{1}{\lambda})(\lambda-2)}{\eta}} \rightarrow 0. \end{aligned} \quad (4.24)$$

Equations (4.15), (4.21) and (4.24) imply Equation (4.14). Then, we have $n^{-1}E(\text{Im}_{2n})^2 \rightarrow 0$.

As to $n^{-1}E(\text{Im}_{3n})^2$, by Equation (4.8),(4.12) and (4.14),

$$\begin{aligned} n^{-1}\mathbb{E}(\text{Im}_{3n})^2 &\leq \frac{1}{n} \sum_{i=\omega(r+q)+1}^n \mathbb{E}[W_i^2] + \frac{2}{n} \sum_{1 \leq i < j \leq n} |\text{Cov}(W_i, W_j)| \\ &\leq \frac{1}{n} \sum_{i=\omega(r+q)+1}^n \mathbb{E}[U_i^2] + \frac{2}{n} \sum_{1 \leq i < j \leq n} |\text{Cov}(U_i, U_j)| \\ &\leq C \cdot \frac{n - \omega(r+q)}{n} + \frac{2}{n} \sum_{1 \leq i < j \leq n} |\text{Cov}(U_i, U_j)| \rightarrow 0. \end{aligned}$$

Besides, since $\omega r/n \rightarrow 1$, it follows from Equations (4.13) and (4.14) that

$$\begin{aligned} \text{Var}[n^{-1/2} \text{Im}_{1n}] &= \frac{1}{n} \mathbb{E}(\text{Im}_{1n})^2 \\ &= \frac{1}{n} \sum_{m=1}^{\omega} \sum_{i=k_m}^{k_m+r-1} \mathbb{E}[W_i^2] + \frac{2}{n} \sum_{m=1}^{\omega} \sum_{k_m \leq i < j \leq k_m+r-1} |\text{Cov}(W_i, W_j)| \\ &+ \frac{2}{n} \sum_{1 \leq i < j \leq n} |\text{Cov}(\zeta_{in}, \zeta_{jn})| \\ &= \frac{\omega r}{n} \mathbb{E}[W_i^2] + O\left(\frac{1}{n} \sum_{1 \leq i < j \leq n} |\text{Cov}(U_i, U_j)|\right) \rightarrow \Delta^2(y|x). \end{aligned}$$

So Equation (4.9) is proved.

As to Equation (4.10), according to Equation (4.8) and the Lemma 4.3, we have

$$\left| \mathbb{E}[\exp(it \sum_{m=1}^{\omega} n^{-1/2} \zeta_{mn})] - \prod_{m=1}^{\omega} \mathbb{E}[\exp(itn^{-1/2} \zeta_{mn})] \right| \leq 16(\omega - 1)\alpha(q + 1) \leq 16\omega\alpha(q) \rightarrow 0.$$

Finally, we establish Equation (4.11). Thinking about $\max_{1 \leq m \leq \omega} |\zeta_{mn}|$, from equation (4.17), assumptions (H6)-(i), (H7) and Lemma 4.1(a), we have

$$\begin{aligned} & \frac{(h_H \phi_x(h_K))^{1/2}}{r} \max_{1 \leq m \leq \omega} |\zeta_{mn}| \\ \leq & \frac{(h_H \phi_x(h_K))^{1/2}}{r} \max_{1 \leq m \leq \omega} \sum_{i=k_m}^{k_m+r-1} |W_i| \\ \leq & \frac{(h_H \phi_x(h_K))^{1/2}}{r} \max_{1 \leq m \leq \omega} \sum_{i=k_m}^{k_m+r-1} \left| \frac{(h_H \phi_x(h_K))^{1/2}}{\mathbb{E}[K_1]} \right| \\ \times & \left[|(a_0 + a_1 \frac{\beta_i}{h_K}) K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i - \mathbb{E}[\zeta(y|X_i, \beta_i)])| \right. \\ + & \left. |\mathbb{E}[(a_0 + a_1 \frac{\beta_i}{h_K}) K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i - \mathbb{E}[\zeta(y|X_i, \beta_i)])]| \right] \\ = & \frac{(h_H \phi_x(h_K))^{1/2}}{r} \max_{1 \leq m \leq \omega} \sum_{i=k_m}^{k_m+r-1} \frac{(h_H \phi_x(h_K))^{1/2}}{\mathbb{E}[K_1]} [O(1) K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i + O(1)) \\ + & O(1) \mathbb{E}[K_i (h_H^{-1} \delta_i \bar{G}^{-1}(T_i) H_i + O(1))] \\ = & \frac{h_H \phi_x(h_K)}{r} \max_{1 \leq m \leq \omega} \sum_{i=k_m}^{k_m+r-1} O(h_H^{-1}) \\ = & O(1). \end{aligned}$$

Therefore, we can get

$$\max_{1 \leq m \leq \omega} |\zeta_{mn}| = O\left(r / \sqrt{h_H \phi_x(h_K)}\right),$$

which leads that for large n ,

$$\mathbb{I}(|\zeta_{mn}| > \varepsilon \Delta(y|x) \sqrt{n}) = 0,$$

by the fact that $r / (h_H \phi_x(h_K))^{1/2} \rightarrow 0$ in Equation (4.8). Therefore $A_n(\varepsilon) \rightarrow 0$. \blacksquare

Proof of Lemma 4.5

(a) Under assumptions (H2),(H5)-(H7) and (H9) and by a simple calculations we can show the results.

(b) Take conditional expectation,

$$h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \delta_1 \overline{G}^{-c}(T_1) H_1^c] = h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \mathbb{E}[\delta_1 \overline{G}^{-c}(T_1) H_1^c | X_1]],$$

according to $\mathbb{E}[\delta_1 | Y_1] = \overline{G}(Y_1)$, we find

$$\begin{aligned} \mathbb{E}[\delta_1 \overline{G}^{-c}(T_1) H_1^c | X_1] &= \mathbb{E}[\mathbb{E}[\delta_1 \overline{G}^{-c}(T_1) H_1^c | Y_1] | X_1] \\ &= \mathbb{E}[\overline{G}^{-c}(Y_1) H_1^c \mathbb{E}[\delta_1 | Y_1] | X_1] \\ &= \mathbb{E}[\overline{G}^{1-c}(Y_1) H_1^c | X_1], \end{aligned}$$

which yield that

$$\begin{aligned} h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \delta_1 \overline{G}^{-c}(T_1) H_1^c] &= h_H^{-1} \mathbb{E}[K_1^a \beta_1^b] \mathbb{E}[\overline{G}^{1-c}(Y_1) H_1^c | X_1] \\ &= h_H^{-1} \mathbb{E}[K_1^a \beta_1^b] \int \overline{G}^{1-c}(s) H^c\left(\frac{y-s}{h_H}\right) \zeta(s | X_1) ds \\ &= \mathbb{E}[K_1^a \beta_1^b] \int \overline{G}^{1-c}(y - h_H t) H^c(t) \zeta(y - h_H t | X_1) dt. \end{aligned} \quad (4.25)$$

Applying the Taylor expansion of order 2 to $\zeta(\cdot | X_1)$ in y , we obtain

$$\begin{aligned} \zeta(y - h_H t | X_1) &= \zeta(y | X_1) - h_H t \frac{\partial \zeta(y | X_1)}{\partial y} + \frac{h_H^2 t^2}{2} \frac{\partial^2 \zeta(y | X_1)}{\partial^2 y} + o(h_H^2 t^2) \\ &= \psi_0(X_1) - h_H t \psi_1(X_1) + \frac{h_H^2 t^2}{2} \psi_2(X_1) + o(h_H^2 t^2), \end{aligned}$$

which combined Equation (4.25), assumption(H7) and from the fact that $G(\cdot)$ is continuous, we obtain

$$\begin{aligned} h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \delta_1 \overline{G}^{-c}(T_1) H_1^c] &= \overline{G}^{1-c}(y) \int H^c(t) dt \mathbb{E}[K_1^a \beta_1^b \psi_0(X_1)] \\ &\quad + \frac{h_H^2}{2} \overline{G}^{1-c}(y) \int t^2 H^c(t) dt \mathbb{E}[K_1^a \beta_1^b \psi_2(X_1)] + o(h_H^2 \mathbb{E}[K_1^a \beta_1^b]). \end{aligned} \quad (4.26)$$

Now, following Rachdi et al. [62], we show that

$$\begin{aligned} \mathbb{E}[K_1^a \beta_1^b \psi_l(X_1)] &= \psi_l(x) \mathbb{E}[K_1^a \beta_1^b] + \mathbb{E}[K_1^a \beta_1^b (\psi_l(X_1) - \psi_l(x))] \\ &= \psi_l(x) \mathbb{E}[K_1^a \beta_1^b] + \mathbb{E}[K_1^a \beta_1^b \mathbb{E}[\psi_l(X_1) - \psi_l(x) | \beta_1]] \\ &= \psi_l(x) \mathbb{E}[K_1^a \beta_1^b] + \mathbb{E}[K_1^a \beta_1^b \Psi_l(\beta_1)] \end{aligned}$$

and since $\Psi_l(0) = 0$ for $l \in \{0, 2\}$ and assumption (H2), by the Taylor expansion of order 1, we obtain

$$\mathbb{E}[K_1^a \beta_1^b \Psi_l(\beta_1)] = \Psi_l'(0) \mathbb{E}[K_1^a \beta_1^{b+1}] + o(\mathbb{E}[K_1^a \beta_1^{b+1}]),$$

which together with equation (4.26) implies

$$\begin{aligned}
h_H^{-1} \mathbb{E}[K_1^a \beta_1^b \delta_1 \overline{G}^{-c}(T_1) H_1^c] &= \overline{G}^{1-c}(y) \int H^c(t) dt \{ \psi_0(x) \mathbb{E}[K_1^a \beta_1^b] + \Psi_0'(0) \mathbb{E}[K_1^a \beta_1^{b+1}] \} \\
&+ \frac{h_H^2}{2} \overline{G}^{1-c}(y) \int t^2 H^c(t) dt \{ \psi_2(x) \mathbb{E}[K_1^a \beta_1^b] + \Psi_2'(0) \mathbb{E}[K_1^a \beta_1^{b+1}] \} \\
&+ o(\mathbb{E}[K_1^a \beta_1^{b+1}]) + o(h_H^2 \mathbb{E}[K_1^a \beta_1^b]).
\end{aligned}$$

■

Computational studies

Contents

5.1 On simulated data	75
5.2 Real data application	79

In this chapter, a simulation study is carried out to investigate the finite sample performance of the L.L.E $\widehat{\zeta}_n(y|x)$ of the conditional density function under right-censored and functional dependent data.

5.1 On simulated data

As common to all, the applicability of asymptotic normality result requires a practical estimation of the asymptotic bias and variance. For this, we neglect the bias term and we use a plug-in approach to construct an estimator of the asymptotic variance of the conditional density function given by :

$$\overline{G}^{-1}(y) \frac{\zeta(y|x)}{h_H} \frac{M_2}{M_1^2 n \phi_x(h_K)} \int H^2(t) dt, \tag{5.1}$$

To test the effectiveness of the asymptotic normality result and to attain this purpose, let us consider the following regression model where the response is a scalar : $Y_i = r(X_i) + \epsilon_i$, $i = 1, \dots, n$, where ϵ_i is the error generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}(\epsilon_{i-1} + \eta_i), \quad i = 1, \dots, n,$$

with $\{\eta_i\}_i$ is a sequence of i.i.d r.v's normally distributed with a variance equal to 0.1. The explanatory variables are constructed by :

$$X(t) = A(2 - \cos(\pi t W)) + (1 - A) \cos(\pi t W), \quad t \in [0, 1],$$

where W is generated from a gaussian distribution $\mathcal{N}(0, 1)$ and A is a r.v Bernoulli distributed with parameter $p = 0.5$. X_i 's are generated from 300 curves and are plotted in Figure 5.1.

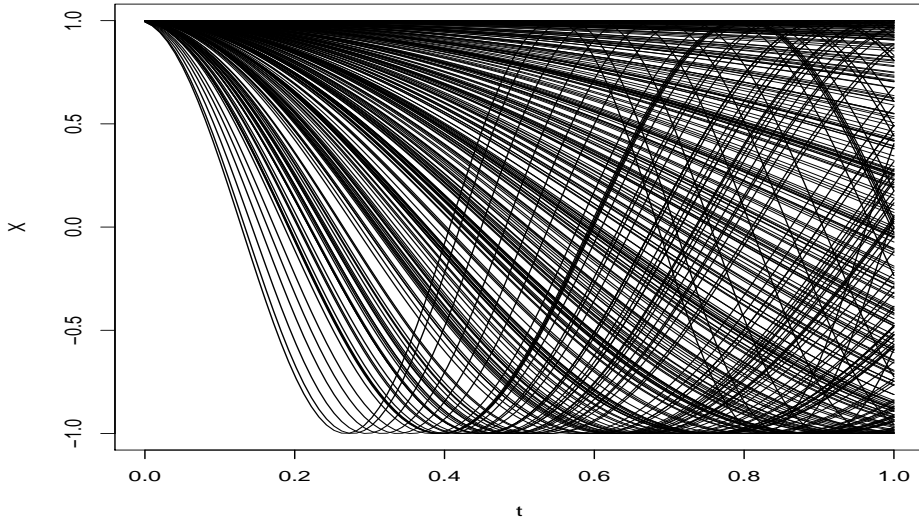


FIGURE 5.1 – A set of 300 simulated curves.

On the other hand, n i.i.d r.v's $(C_i)_{i=1, \dots, n}$ are simulated through the exponential distribution $\mathcal{E}(\lambda)$ and for $i = 1, \dots, n = 300$, the scalar response Y_i is computed by considering the following operator :

$$r(X) = 4 \exp \left\{ \frac{1}{2 + \int_0^{\pi/2} |X_i(t)|^2} \right\}$$

Given $X = x$, we can easily see that Y is as a gaussian distribution $\mathcal{N}(r(x), 0.2)$. Then, we can get the corresponding conditional density, which is explicitly defined by

$$\zeta(y|x) = \frac{1}{\sqrt{2\pi \times 0.3}} \exp \left\{ -\frac{1}{2 \times 0.3} (y - r(x))^2 \right\}$$

Therefore, the conditional mode, the conditional mean $r(x)$, and the conditional

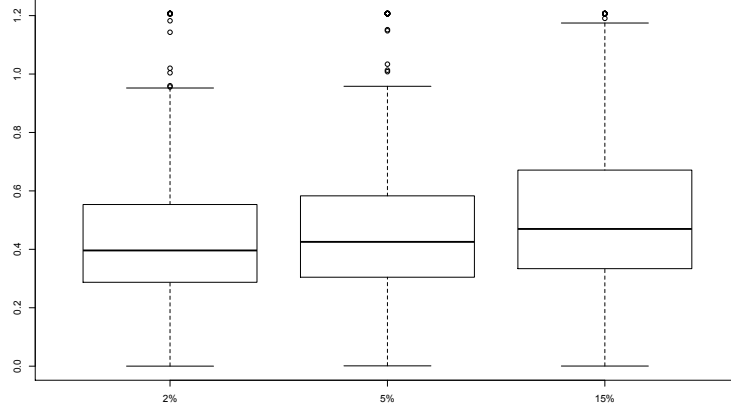


FIGURE 5.2 – Distribution of the MSE obtained for different CR for $n = 300$.

median functions will coincide and will be equal to $r(x)$, for any fixed x . Our purpose, now, consists in evaluating the accuracy of the conditional mode function estimator based on randomly censored data. The computation of this estimator is based on the observed data $(X_i, T_i, \delta_i)_{i=1, \dots, n}$, where $T_i = \min(Y_i, C_i)$ and $\delta_i = \mathbb{I}_{\{Y_i \leq C_i\}}$. In this simulation study, we present only results of the case where $i = 2$ and $q = 1$. For this, we take $K_0(s) = 3(1-s^2)\mathbb{I}_{[-1,1]}$, $K(1) > 0$, and $K_1(s) = 3(1-s^2)\mathbb{I}_{[-1,1]}$. Elsewhere, as it is well-known in F.D.A, the choice of the metric and the smoothing parameters have crucial roles in the computational issues. To optimize these choices on this illustration, we use the local cross-validation procedure method in the aim of choosing smoothing parameters h_K and h_H (see Laksaci et al. [53]).

Another important point for insuring good behavior of the considered methods is to use locating functions δ and/or β that are well adapted to the kind of data that we have to deal with. Here, it is clear that the shape of the curves (cf. Figure 5.1) allows us to use the locating functions σ and β defined by the derivatives of the curves. More precisely, we take

$$\delta(x, x') = \left(\int_0^1 (x^{(i)}(t) - x'^{(i)}(t))^2 dt \right)^{1/2} \quad \text{and} \quad \beta(x, x') = \int_0^1 \alpha(t)(x_i(t) - x'_i(t))^2 dt,$$

where $x^{(i)}$ denotes the i th derivative of the curve x and α is the eigenfunction of the empirical covariance operator $\frac{1}{n} \sum_{j=1}^n (X_j^{(i)} - \bar{X}^{(i)})(X_j^{(i)} - \bar{X}^{(i)})$ associated with the

TABLE 5.1 – Numerical summary of the distribution of the MSE , for $N=400$, obtained for $n = 50, 100, 200$ and 300 .

	$CR = 2\%$	$CR = 7\%$	$CR = 16\%$	$CR = 49\%$
$n = 50$	0.392	0.472	0.673	2.431
$n = 100$	0.301	0.332	0.472	1.964
$n = 200$	0.220	0.292	0.346	1.632
$n = 300$	0.102	0.150	0.215	1.214

q -greatest eigenvalues. The performance of the conditional mode estimator $\widehat{\theta}_n(x)$ is evaluated on $N = 400$ replications using different sample sizes $n = 50, 100, 200$, and 300 . The mean square error (MSE) is considered here, such that, for a fixed x , $MSE = \frac{1}{400} \sum_{m=1}^{400} (\widehat{\theta}_{n,m}(x) - r(x))^2$. Figure 5.2 displays the distribution of the obtained MSE given by the N replications. It can be observed that the proposed estimator performs well, especially when the sample size increases. This conclusion is confirmed by Table 5.1 which provides a numerical summary of the distribution of the MSE , with different censored rates (CR). In the second part of the simulation studies, we are interested in the evaluation of the prediction accuracy of the conditional median with different CR . A sample $(X_i, Y_i)_{i=1, \dots, 550}$ of size $n = 550$ generated from the model described above, is considered for this purpose. We split this sample into two parts : a learning subsample $\{(X_i, Y_i); i = 1, \dots, 500\}$ which is used to calculate the predictor (the conditional mode in this case) and a testing subsample $\{(X_i, Y_i); i = 501, \dots, 550\}$ used to evaluate the performance of the predictor. The prediction accuracy is measured, for different values of CR , by using the Mean Absolute Error (MAE), defined as $MAE = \frac{1}{50} \sum_{i=501}^{550} |Y_i - \widehat{\theta}_n(X_i)|$, as well as the MSE such that $MSE = \frac{1}{50} \sum_{i=501}^{550} (Y_i - \widehat{\theta}_n(X_i))^2$. We can see that the prediction accuracy of the conditional mode decreases as the CR increases. For censoring distributions $\mathcal{E}(2) - 1$, ($CR = 7\%$, $MAE = 0.204$, $MSE = 0.483$), $\mathcal{E}(2)$, ($CR = 16\%$, $MAE = 0.363$, $MSE = 0.62$), and $\mathcal{E}(2) + 5$, ($CR = 49\%$, $MAE = 1.052$, $MSE = 2.291$).

5.2 Real data application

A useful tool in survival analysis is the hazard rate, which reflects the instantaneous probability that a duration will end within the next time instant. Among the most used examples in survival analysis : survival times of patients, Stanford Heart Transplant, durations between subsequent transactions in a financial security,... For application on real data, we apply the local linear method via the Kidney transplant data (see Klein and Moeschberger [49] and <https://www.agence-biomedecine.fr>). The bandwidth selection is given by plug-in rules and more advanced selection methods like cross-validation are likely to further improve the performance of the local linear hazard rate estimator. To use the Kidney transplant data, we need to describe three fundamental parameters that are (1)– the survival times in days of patients following kidney transplant as the response, (2) race (black / white) and (3) age in years for each patient as the covariate. We propose a new method based on the functional local linear approach. More precisely, we use the conditional hazard rate. The methodology of this study is given by the following description : In the first, we consider the subsamples of white males and white females. At the second step we take 432 patients as the first group and has a censoring rate of 83%. The second group involves 280 patients with a 86% censoring rate. The survival times of white patients vary between 1 day and 9.4 years. The average age of the male patients is slightly less than 44 years, whereas the female patients are almost 41 years old on average. We find only small differences between the various plug-in bandwidths for the conditional hazard rate, where we have taken age as the conditioning variable.

In conclusion, we find only small differences between the various plug-in bandwidths for the conditional hazard rate, where we have taken age as the conditioning variable (see Fig. 5.3). Figure (A) and Figure (D), based on the normal reference rule, depict the conditional hazard rate as a function of time and consider several values of age. Figure. (A) and Figure (B) zoom in on the period shortly after the transplantation, whereas Figure (C) and Figure (D) display the hazard rate over a longer time.

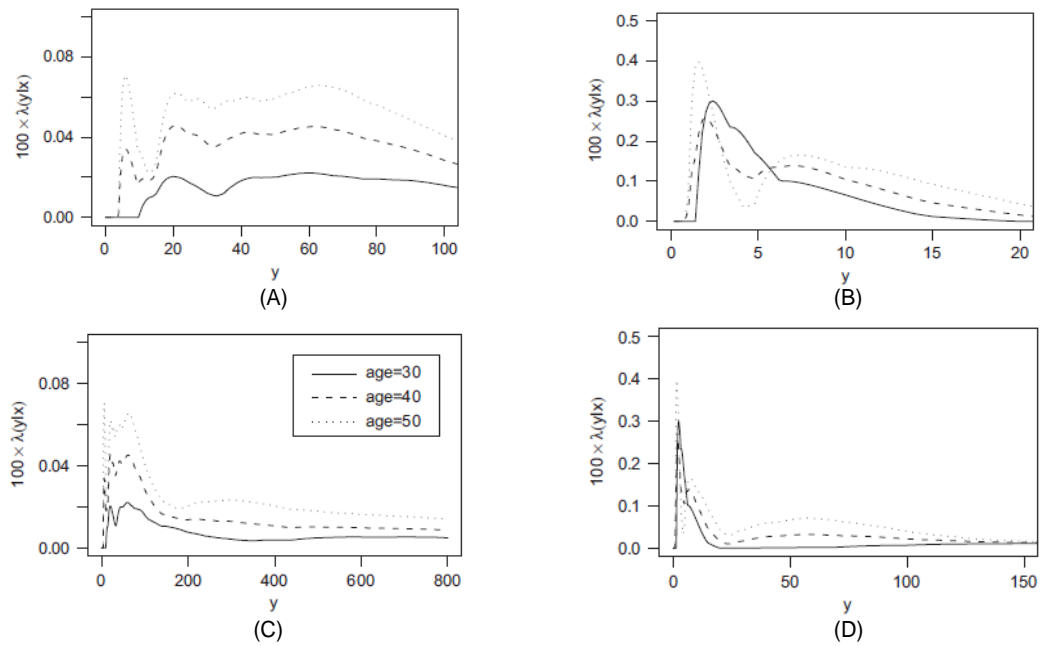


FIGURE 5.3 – Estimates of the conditional hazard rate after the transplantation

We see that the hazard rate is strongly non-monotonic for both male and female patients. As expected, the hazard rate is higher for older patients. Interestingly, the differences in the hazard rates of younger and older women diminish after about 150 days. For men the differences between the various age groups persist longer. Also, the risk of dying shortly after the transplantation is higher for men than for women. On the other hand, for female patients the risk of dying in a later stage is lower than for males.

Conclusion et Perspectives

Conclusion

Le but de ce travail était de présenter une méthode d'estimation pour la densité conditionnelle alternative à la méthode à noyau afin de palier aux problèmes du bord créés par cette dernière, lorsque la variable d'intérêt est soumise à une censure aléatoire à droite, le cas de censure le plus fréquent.

La réalité pratique de l'étude des durées de vie, nous a amené à considérer le cas où les données présentent une forme de dépendance. Nous avons fait le choix de l' α -*m*lange qui est une forme de dépendance raisonnablement faible qui implique toutes les autres formes de mélange et qui est surtout très répandue, par exemple, les processus AR et ARMA sont α -*m*langeants.

Dans un premier temps, nous avons construit un nouvel estimateur pour la densité conditionnel. L'estimateur proposé est une généralisation au cas de censure de l'estimateur par local linéaire introduit par madani et al [20]. Comme résultats asymptotiques nous avons établi la convergence presque sûre sur un compact, en donnant sa vitesses. Nous illustrons les résultats obtenus par des simulations sur des données simulées puis réelles.

Nous nous sommes intéressé, dans un deuxième temps, à l'étude de la normalité asymptotique de l'estimateure définit dans le premier cas avec les même hypothèses. Nous donnons quelques applications de ces résultats notamment à la prévision et à la construction des intervalles de confiance.

Perspectives

Au cours de ce travail de recherche nous avons vu que de nombreux problèmes restent encore à étudier. C'est pour ceci que nous allons nous atteler, dans ce paragraphe, à lister quelques questions ouvertes à traiter dans le futur (à court terme ou à long terme) en vue d'améliorer et d'étendre nos résultats.

L'L.L.E, par la méthode des k plus proches voisins, peut être considérée comme une méthode d'estimation alternative qui reste encore à développer pour nos modèles.

Il serait intéressant d'établir un résultat de type Berry-Esseen pour la densité conditionnelle dans un modèle de données incomplètes et associées.

la généralisation de nos résultats sur la méthode par polynômes locaux d'ordre supérieur strictement à 1.

on peut également envisager d'étendre nos résultats sur la densité conditionnel pour des données tronqué.

D'autres questions ouvertes peuvent être abordées suite à notre travail, comme par exemple, changer la méthode d'estimation et faire une étude comparative avec les estimateurs qui existent dan la littérature.

Bibliographie

- [1] Y. Ait-Sahalia. Transition densities for interest rate and other non-linear diffusions. *J. Finance*. **54**, 1361–1395, (1999).
- [2] A. Baillo and A. Grané. Local linear regression for functional predictor and scalar response. *Journal of Multivariate Analysis*. **100**, 102–111, (2009).
- [3] J. Barrientos-Marin, F. Ferraty and P. Vieu. Locally Modelled Regression and Functional Data. *J. of Nonparametric Statistics*. **22**, (3), 617–632, (2009).
- [4] D.M. Bashtannyk and R. J. Hyndman. Bandwidth selection for kernel conditional density estimation. *Comp. Statist. Data Anal*. **36**, 279–298, (2001).
- [5] R. Beran. Nonparametric regression with randomly censored survival data, Technical Report, Department of Statistics, University of California, Berkeley, CA, (1981).
- [6] A. Berlnet, and R. Servien. Necessary and sufficient condition for the existence of a limit distribution of the nearest-neighbour density estimator. *J. Nonparametr. Stat.*. **23**, (3), 633-643, (2011).
- [7] J.R. Blum, , D.L. Hanson, , L.H. Koopmans. On the strong law of large numbers for a class of stochastic processes. *Z. Wahrsch. verw. Gebiete*. **2**, 1–11, (1963).
- [8] D. Bosq. Modelization, nonparametric estimation and prediction for continuous time processes. In Nonparametric functional estimation and related topics (Spetses, 1990). *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* **19**, Kluwer Acad. Publ., Dordrecht.. 509– 529, (1991).

- [9] D. Bosq. Linear Processes in Function Spaces : Theory and Applications Lecture Notes in Statistics. *Springer-Verlag, New York*. 149, (2000).
- [10] N. Breslow and J. Crowley. A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.*. 2, (3), 437–453, (1974).
- [11] T. Cai, L.J. Wei and M. Wilcox. Semiparametric regression analysis for clustered survival data. *Biometrika*. 87, 867–878, (2000).
- [12] A. Carbonez, L. Györfy and E.C. VanderMeulin. Partitioning estimates of a regression function under random censoring. *Stat.Decisions*. 13, 21–37, (1995).
- [13] A. Chouaf, and A. Laksaci. On the functional local linear estimate for spatial regression. *Stat. Risk Model.*. 29, (3), 189-214, (2012).
- [14] S. Dabo-Niang, et A. Laksaci, . Propriétés asymptotiques d’un estimateur à noyau du mode conditionnel pour variable explicative fonctionnelle. *Ann. I.S.U.P.*. 51, 27–42, (2007).
- [15] S. Dabo-Niang, , F. Ferraty, et P. Vieu. Estimation du mode dans un espace vectoriel semi-normé. (French) [Mode estimation in a semi-normed vector space]. *C. R. Math. Acad. Sci. Paris*. 339, (9), 659–662, (2004a).
- [16] S. Dabo-Niang, F. Ferraty et P. Vieu. Mode estimation for functional random variable and its application for curves classification. *Far East J. Theor. Stat.*. 18, (1), 93–119, (2006).
- [17] Yu. A. Davydov. On convergence of distributions generated by stationary stochastic processes. *Theory Probab. Appl.* 13, (4), 691-696, (1968).
- [18] J. Dedecker, P. Doukhan, G. Lang, J.R. Leön, S. Louhichi and C. Priour. Weak Dependence : Models, Theory, and Applications. Lecture Notes in Statistics 190. *NewYork :Springer-Verlag*. (2007).
- [19] P. Deheuvels and H. Einmahl. Functional limit laws for the increments of Kaplan-Meier product limit processes and applications. *Ann. probab.* 28, 1301–1335, (2000).
- [20] J. Demongeot, A. Laksaci, F. Madani and M. Rachdi. A fast functional locally modelled of the conditional density and mode for functional time series *Recent Advances in Functional Data Analysis and Related Topics*, (2011) 19 pages.

- [21] J. Demongeot, A. Laksaci, F. Madani, M. Rachdi, Functional data : local linear estimation of the conditional density and its application. *Statistics.* **47**, (1), 26–44, (2013).
- [22] J. Demongeot, A. Laksaci, M. Rachdi, A. Abdali, , S. Rahmani, On the Local Linear Modelization of the Conditional Distribution for Functional Data. *Sankhya A.* **76**, (2), 328–355, (2014).
- [23] J. Demongeot, A. Laksaci, A. Naceri and M. Rachdi. Local linear regression modelization when all variables are curves. *Statistics and Probability Letters.* **121**, 37–44, (2017).
- [24] J. C. Deville. Méthodes statistiques et numériques de l’analyse harmonique. , , . *Ann. Insee.* **15**, 3–101, (1974).
- [25] P. Doukhan. Mixing properties and Examples. Lecture Notes in Statistics. *New York : Springer Verlag.* (1994).
- [26] J.J. Dreesbeke, B. Fichet and P. Tassi. Analyse statistique des durées de vie. *Economica.* (1989).
- [27] Sam. Efromovich. Conditional density estimation in a regression setting. *Ann. Statist.* **35**,(6), 2504–2535, (2007).
- [28] J. Fan and T.H. Yim. A cross validation method for estimating conditional densities. *J. Biometrika.* **91**, 819–834, (2004).
- [29] J.Fan, Q.Yao and H.Tong. Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *J. Biometrika.* **83**, 189–206, (1996).
- [30] F. Ferraty and P. Vieu. Nonparametric functional data analysis. Theory and Practice. *Springer Series in Statistics. New York.* (2006).
- [31] F. Ferraty, and P. Vieu. Nonparametric functional data analysis. *Theory and Practice. Springer Verlag.* **21**, 97–130, (2006).
- [32] F. Ferraty and P. Vieu. Nonparametric modelling for functional data. *Springer-Verlag, New York.* (2006a).
- [33] A. Földes, L. Rejto et B. B. Winter. Strong consistency properties of nonparametric estimators for randomly censored : I : the product-limit estimator. *Periodica Mathematica Hungarica.* **11**, (3), 233–250, (1980).

- [34] H. Gebelein. Das statistische problem der korrelation als variations und eigenwert problem und sein zusammenhang mit der ausgleichsrechnung. *Z. Angew. Math. Mech.* **21**, 364–379, (1941).
- [35] P. Hall, J. Racine and Q. Li. Cross-validation and the estimation of conditional probability densities. *J. Amer. Statist. Association.* **99**, 1015–1026, (2004).
- [36] P. Hall, C. C. Heyde . Martingale limit theory and its application. *New York : Academic Press.* (1980).
- [37] H.O. Hirschfeld. A connection between correlation and contingency. *Math. Proceed. of the Cambridge Philosophical Soc.,* **31**, 520–524, (1935).
- [38] I. Holmstrom. On method for parametric representation of the state of the atmosphere. *Tellus.* **15**, 127–149, (1961).
- [39] W. Horrigue and E. Ould-Saïd. Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Oper. Stoch. Equ.* **19**, 131–156, (2011).
- [40] T. Hsing and R. Eubank. Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators. *Wiley Series in Probability and Statistics. John Wiley Sons, Ltd., Chichester.* (2015).
- [41] R.J. Hyndman and Q. Yao. Nonparametric estimation and symmetry tests for conditional density functions. *J. Nonparametric Statist.* **14**, 259–278, (2002).
- [42] R.J. Hyndman, D.M. Bashtannyk and G.K. Grunwald. Estimating and visualizing conditional densities. *J. Comp. Graph. Statist.* **5**, 315–336, (1996).
- [43] I.A. Ibragimov. Some limit theorems from stationary processes. *Theory Proc. Appl.* **7**, 349–382, (1962).
- [44] SS. Kang and KJ. Koehler. Modification of the Greenwood formula for correlated response times. *Biometrics.* **53**, 885–899, (1997).
- [45] E.L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *JASA.* **53**, 457–481, (1958).
- [46] S. Khardani, M. Lemdani and E. Ould-Saïd. Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. *J.Kor.Stat.Soc.* **39**, 455–469, (2010).

- [47] S. Khardani and B. Thiam. Strong consistency result of a nonparametric conditional mode estimator under random censorship for functional regressors. *Communications in statistics -theory and methods* . **45**, 1863–1875, (2016).
- [48] M. Kirkpatrick, and N. Heckman. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* **27**, (4), 429–450, (1989).
- [49] J.P. Klein and M.L. Moeschberger. Survival Analysis : Techniques for Censored and Truncated Data. *Springer, Berlin*. (2004).
- [50] M. Kohler, K. Mathe, and M. Pinter. Prediction from randomly right censored data. *J.Mult.Anal.* **80**, 73–100, (2002).
- [51] A.N. Kolmogorov. Über die analytischen methoden inder wahrschein lichkeitsrechnung. *Math. Ann.* **104**, 415–458, (1931).
- [52] A. Laksaci. Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle. *C. R. Math. Acad. Sci. Paris.* **345**, 171–175, (2007).
- [53] A. Laksaci, F. Madani and M. Rachdi. Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Communications in Statistics-Theory and Methods*. **42**, (19), 3544–3570, (2013).
- [54] A. Laksaci, et A. Yousfate. Estimation fonctionnelle de la densité de l'opérateur de transition d'un processus de Markov à temps discret. *C. R. Acad. Sci., Paris*. **334**, (11), 1035–1038, (2002).
- [55] H. Y. Liang, and J. I. Baek. Asymptotic normality of conditional density estimation with left-truncated and dependent data. *Statistical Papers*. **57**, (1), 1–20, (2016).
- [56] D. Lynden-Bell. A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Not. R. Astronomical Soc.* **155**, (1), 95–118, (1971).
- [57] A. Mas and B. Pumo. Functional linear regression with derivatives. *Journal of Nonparametric Statistics*. **21**, (1), 19-40, (2009).

- [58] E. Masry. Recursive probability density estimation for weakly dependent process. *IEEE Trans. Information Theory* 32. 254–267, (1986).
- [59] F. Messaci, N. Nemouchi, I. Ouassou and M. Rachdi. Local polynomial modelling of the conditional quantile for functional data. *Statistical Methods and Applications*. **24**, (4), 597–622, (2015).
- [60] E. A. Nadaraya. On estimating regression. *Theory of Probability and its Applications* . **10**, 186-190, (1964).
- [61] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* . **33**, 1065-1076, (1962).
- [62] M. Rachdi, A. Laksaci, J. Demongeot, A. Abdali, F. Madani. Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics and Data Analysis*. **73**, 53–68, (2014).
- [63] J. Ramsay, and B. Silverman. Functional Data Analysis (Second Edition). *Springer-Verlag, New York*. (2005).
- [64] J. Ramsay and B. Silverman. Applied functional data analysis : Methods and case studies. *Springer-Verlag, New York*.. (2002).
- [65] P. M. Robinson. Consistent Nonparametric Entropy-Based Testing. *Rev. Econ. Studies*. **58**, 437–453, (1991).
- [66] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Mathematics*. **42**, 43–47, (1956).
- [67] G. Roussas. On some properties of nonparametric estimates of probability density functions. *Bull. Soc. Math. Grèce(N.S.)*. **9**, (1), 29–43, (1968).
- [68] P. Sarda and P. Vieu. Kernel Regression. *Wiley, New York*. 43–70, (2000).
- [69] D. Sentürk and H.-G. Müller. Functional varying coefficient models for longitudinal data. *Journal of the American Statistical Association*. **105**, (491), 1256-1264, (2010).
- [70] G.R. Shorack and J.A. Wellner. Empirical Processes with Applications to Statistics. *John Wiley Sons*.. (1986).

- [71] W. Stute. Distributional convergence under random censorship when covariates are present *Scand. J. Statist.* **23**, 461–471, (1996).
- [72] D. Tjøstheim. Non-linear time series. *J. Statist.* **21**, 97–130, (1994).
- [73] B.W. Turnbull. Nonparametric estimation of a survivorship function with doubly censored data. *J. Amer. Statist. Assoc.* **69**, 169–173, (1974).
- [74] V. A. Volkonskii, Y. A. Rozanov. Some limit theorems for random functions. *Theory of Probability and its Applications.* **4**, (2), 178–197, (1959).
- [75] M. Woodroofe. Estimating a distribution function with truncated data. *Ann.Statist.* 163–177, (1985).
- [76] X. Xiong, P. Zhou, and C. Ailian. Asymptotic normality of the local linear estimation of the conditional density for functional time series data. *Communications in Statistics - Theory and Methods.* **47**, 3418–3440, (2017).
- [77] E. Youndjé. Estimation non paramétrique de la densité conditionnelle par la méthode du noyau. *PhD Thesis from the Rouen University (in French)*.. (1993).