Mohamed Khider University (Algeria)

and

Littoral Côte d'Opale University (France)

A Thesis Presented for the Degree of

DOCTOR OF SCIENCES

In the Field of **Statistics**

By

Djabrane YAHIA

Title

Conditional Quantile for Truncated Dependent data

Examination Committee Members

Brahim MEZERDI	Professor	Univ. Biskra (Algeria)	Chairman
Abdelhakim NECIR	Professor	Univ. Biskra (Algeria)	Co-Director of Thesis
Elias OULD SAID	Professor	Univ. Littoral. (France)	Co-Director of Thesis
Ali GANNOUN	Professor	Univ. Montpellier 2 (France)	Examinator
Zaher MOHDEB	Professor	Univ. Constantine (Algeria)	Examinator
Mustapha RACHDI	Professor	Univ. Grenoble 2 (France)	Examinator
Abdelkader TATACHAK	MC (A)	Univ. USTHB (Algeria)	Invited

19 September 2010

To my

Dear Wife

&

My Son

Mohamed-Abdelaziz

Remerciements

Je souhaite exprimer ici ma reconnaissance envers mes deux Directeurs de thèse, le Professeur Abdelhakim Necir de l'Université de Biskra et le Professeur Elias Ould Saïd de l'Université du Littoral Côte d'Opale (France), pour leurs conseils bénéfiques et leurs apports précieux tout au long de la réalisation de ma thèse. Ils n'ont ménagé aucun effort pour m'accompagner dans ma recherche et m'ont permis, par leurs connaissances et ouvertures envers mes idées, d'améliorer mes aptitudes dans ce domaine. Merci de votre disponibilité, de votre grande patience et de vos conseils fort judicieux...

Je remercie infiniment le Professeur Brahim Mezerdi de l'Université de Biskra de me faire l'honneur de présider le Jury de ma thèse.

Je tiens également à exprimer ma gratitude envers les Professeurs Ali Gannoun et Zaher Mohdeb d'avoir accepté de rapporter cette thèse malgré leur occupation diverses. Le Professeur Mustapha Rachdi a bien accepté de participer au Jury de cette thèse, qu'il trouve içi toute ma gratitude. Messieur Abdelkader Tatachak a également bien accepté de participer à la soutenance de cette thèse, je le remercie énormément.

Mes remerciements sont adressés également à tous les membres de Département de Mathématiques et ceux du Laboratoire de Mathématiques Appliquées et en particulier à mon collègue et ami Djamel Meraghni, son aide, ses conseils et son soutien m'ont été précieux.

Je ne pourrais terminer cette dédicace sans remercier chaleureusement mes parents, ma femme et toute ma famille, pour leur confiance sans faille en mes capacités, leur appui et leur amour sans limite. Sans eux, je ne me serais jamais rendu aussi loin.

Enfin, je tiens à rendre hommage à l'un des piliers de notre département de Mathématiques, mon enseignant Prof. Seid Bahlali, qui nous a quitté le 31 janvier 2010 suite à un malaise cardiaque, en plein exercice de ses fonctions à l'intérieur de l'université Mohamed Khider de Biskra : Nous n'oublierons jamais sa joie de vivre et les longues heures passées à ses côté. Que Dieu puisse l'accueillir en son vaste paradis.

Cotutelle

Cette thèse à été préparée sous la convention de Cotutelle entre :

l'Université Mohamed Khider - Biskra (Algérie)

et l'Université du Littoral Côte d'Opale (France).

Accord des Commissions des Thèses

Cette thèse est soutenue avec l'accord de la commission des thèses de la région Nord Pas de Calais (France) et l'accord du Comité Scientifique du Département de Mathématiques de l'Université Mohamed Khider Biskra (Algérie).

Abstract

In this thesis we study some asymptotic properties of the kernel conditional quantile estimator when the interest variable is subject to random left truncation. The uniform strong convergence rate of the estimator is obtained. In addition, it is shown that, under regularity conditions and suitably normalized, the kernel estimate of the conditional quantile is asymptotically normally distributed.

Our interest in conditional quantile estimation is motivated by it's robusteness, the constructing of the confidence bands and the forecasting from time series data. Our results are obtained in a more general setting (strong mixing) which includes time series modelling as a special case.

Keywords: Asymptotic normality; Conditional quantile; Kernel estimate; Strong mixing; Strong uniform consistency; Truncated data.

Résumé

Dans cette thèse nous étudions certaines propriétés asymptotiques de l'estimateur à noyau du quantile conditionnel lorsque la variable d'intérêt est soumise à une troncature aléatoire à gauche. La convergence uniforme presque sûre avec vitesse de l'estimateur est obtenue. En outre, il est démontré que, sous des conditions de régularité, l'estimateur à noyau du quantile conditionnel convenablement normalisé est asymptotiquement normal.

L'intérêt principal dans l'étude de l'estimation des quantiles conditionnels est sa robustesse, la construction des intervalles de confiance et la prévision à partir des données de séries chronologiques. Nos résultats sont obtenus dans un cadre général (mélangeance forte), qui inclut des modèles populaires de séries financières et économétriques comme cas particulier.

Mots-clés: Convergence uniforme forte; Données tronquées; Estimation à noyau; Normalité asymptotique; Mélangeance forte; Quantiles conditionnels; Vitesse de convergence.

Achieved Works

1. A strong uniform convergence rate of a kernel conditional quantile estimator under random left-truncation and dependent data.
   Electronic Journal of Statistics, Vol. 3 (2009) 426–445.
   (with E. Ould Saïd and A. Necir, 2009)

2. Asymptotic normality of a kernel conditional quantile estimator under strong mixing hypothesis and left-truncation.
   (To appear in *Communications in Statistics, Theory & Method*).
   (with E. Ould Saïd, 2010)

Avant-propos

Dans l'analyse de survie on doit souvent modéliser le lien entre la fonction de survie et un certain nombre de facteurs (covariables ou variables explicatives). Dans les études de durée de vie il se peut que pour certains individus on n'arrive pas à observer l'événement d'intérêt. Par exemple, au cours d'une étude sur les fumeurs, il est intéressant de savoir comment le temps de survie (la variable d'intérêt) est influencé par l'âge auquel la personne a commencé à fumer. Les personnes sont suivies pendant une certaine période de temps. Un fumeur qui décède avant le début de l'étude est systématiquement exclu de l'échantillon et donne lieu à ce qu'on appelle une observation tronquée à gauche. En revanche, un fumeur non décédé avant la fin de l'étude donne lieu à ce qu'on appelle une observation censurée à droite.

Une durée de vie est une variable aléatoire (va) souvent positive, précisément c'est le temps nécessaire de passer d'un état A à un état B. Il n'est pas rare, donc, que les données à traiter ne soient pas complètes, dans ce cas les techniques classiques ne s'adaptent pas correctement aux données incomplètes. La littérature est beaucoup plus riche en ce qui concerne la censure que la troncature qui est plus récente. Dans cette thèse, nous nous intéressons particulièrement, à la troncature gauche qui est le cadre dans le quel nous avons apporté de nouveaux résultats.

Le modèle de trancature est apparu tout d'abord en astronomie, mais il est observé dans plusieurs domaines comme la médecine, l'épidémiologie, la biométrie et l'économie. La recherche d'objets cachés qui devront être assez grand pour être détectés, comme les réserves de pétrole est un champ d'application pour les données tronquées. De plus, les enquêtes de suivi médical où la troncature gauche peut apparaître si le temps d'origine de la durée de vie précède le temps d'origine de l'étude.

En cas de troncature gauche, nous ne sommes capable d'observer que les durées de vie Y pour les quelles $Y \geq T$, ici T est la variable de troncature. Dans ce cas, nous disposons d'un échantillon de taille n, dont la variable d'intérêt Y est observable, cet échantillon est extrait d'un échantillon de plus grand taille N inconnue. Les résultats statistiques doivent être donnés en considérant la population dont est extrait le $N-$échantillon et non le $n-$échantillon. Il n'est pas possible d'avoir un échantillon représentant toute la population considérée, car lorsque $Y < T$ rien ne peut être observé. Ceci implique qu'il y a plusieurs mesures de probabilités et nécessite beaucoup de précaution pour énoncer les résultats asymptotiques.

Lorsque l'on travaille avec des données tronquées, la proportion de la population avec laquelle nous pouvons disposer d'une observation joue un rôle important en estimation sous troncature. Cette probabilité notée $\mu = P(Y \geq T)$ pourrait être estimée par la quantité $\frac{n}{N}$ mais malheureusement cet estimateur ne peut pas être calculé, car N est inconnue et la taille n de l'échantillon observé est elle-même une variable aléatoire de loi binomiale $\mathcal{B}(N, \mu)$. En utilisant des estimateurs produit-limites de Lynden-Bell (1971), He et Yang (1998) donnent un estimateur calculable de μ ainsi que des résultats de convergence asymptotique dont la normalité. Nous rappelons dans le deuxième chapitre de cette thèse, les principaux résultats concernant l'estimation sous la troncature gauche.

Bien que notre intérêt dans l'estimation non paramétrique soit motivé par la construction des intervalles de confiance à partir des données de séries chronologiques, nous présentons nos résultats dans un cadre plus général (mélange fort) qui inclut des modèles des séries chronologiques comme cas particulier. Dans le premier chapitre, nous rappelons les concepts de base sur les mélanges avec certaines propriétés liant les différents coefficients de mélanges. Il existe plusieurs types de mélanges qui sont définis à partir de coefficients, notés, selon les cas, α, β, ρ, ψ et ϕ. Parmi toutes ces formes de mélanges le $\alpha-$mélange est le plus faible. Tout résultat énoncé pour des données $\alpha-$mélangeantes sera valable pour des données soumises à une autre forme de mélange, car toutes suites de va's β, ρ, ψ ou $\phi-$mélangeante sera donc forcément $\alpha-$mélangeante.

Gorodetskii (1977) et Withers (1981) dérivent les conditions pour lesquelles un processus linéaire est mélangeant. En fait, sous des hypothèses classiques le modèle linéaire autorégressif et généralement les modèles bilinéaires de séries chronologiques sont fortement mélangeant avec des coefficients de mélange à décroissance exponentielle. Auestad et Tjϕstheim (1990) donnent des discussions éclairantes sur le rôle des-mélange pour l'identification du modèle dans l'analyse de séries temporelles non linéaires. En outre, Masry et Tjϕstheim (1995-97) ont montré que sous certaines conditions douces, les deux processus autorégressif conditionnellement hétéroscédastique $ARCH$ et les processus autorégressifs non linéaires additifs avec variables exogènes, qui sont particulièrement populaires dans la finance, sont stationnaires et mélangeant.

Récemment, des nouveaux développements ont eu lieu dans la théorie de statistique non paramétrique. Des résultats asymptotiques ont été obtenus pour certains estimateurs et prédicteurs pour des données incomplètes (sous troncature ou censure). Rappelons les travaux de Ould Saïd et Lemdani (2006) pour la fonction de régression sous troncature, Ould Saïd et Sadki (2008) concernant les quantiles conditionnels dans un modèle de censure à droite, Ould Saïd et Tatachak (2009) pour le mode conditionnel sous troncature à gauche et finalement Lemdani et al. (2009) ont étudié la fonction des quantiles conditionnels pour des données tronquées mais dans le cas de va's indépendantes et identiquement distribués (i.i.d.).

Les médianes et quantiles conditionnels sont fréquemment utilisés dans l'analyse des données de séries chronologiques avec des queues lourdes pour leurs propriétés de robustesse. Il est bien connu, que la moyenne est sensible aux valeurs aberrantes (voir Hampel et al. 1986), il peut être judicieux d'utiliser la médiane, qui est un cas particulier du quantile, plutôt que la moyenne pour prévoir l'avenir puisque la médiane est très robuste contre les valeurs aberrantes, en particulier la fonction médiane conditionnelle pour distribution asymétrique.

Dans cette thèse nous étudions les propriétés asymptotiques de l'estimateur à noyau du quantile conditionnel lorsque la variable d'intérêt est soumise à la troncature gauche. Notre intérêt pour l'estimation des quantiles conditionnels est motivé par la construction des intervalles de confiance et de la prévision à partir des données de séries chronologiques. Nos résultats sont dérivés dans un cadre plus général, de stationnarité et de forte dépendance (i.e., $\alpha-$mixing). Ce type de dépendance modélise beaucoup de processus en particulier les modèles $ARMA$ ou $ARCH$ souvent rencontrés en finance et économétrie.

Le chapitre deux, est consacré aux rappels des résultats existants sur l'estimation non paramétrique dans le cas du modèle tronqué aléatoirement à gauche. Plus précisément ces résultats concernent les propriétés de convergence des estimateurs de la probabilité d'observer la variable d'intérêt Y (tronquée par la variable T) ainsi que les fonctions de répartition correspondantes notée F et G respectivement. l'estimateur à noyau à été introduit par Ould Saïd et Lemdani (2006, *Ann. Instit. Statist. Math.*) qui est rappelé ici. Il est bien connu que les quantiles et les quantiles conditionnels sont une alternative à la prévision par la moyenne et moyenne conditionnelle respectivement. L'estimateur à noyau du quantile conditionnel en présence de troncature aléatoire à gauche, a été introduit par Lemdani, Ould Saïd et Poulin (2009, *J. of Multivariate Analysis*) dont les propriétés asymptotiques de convergence et de normalité ont été établies dans le cas i.i.d.

Le troisième chapitre est consacré à l'extension des résultats de convergence uniforme presque sûre. Ici le triplet de données (Y, T, X) est supposé satisfaire une condition de mélange fort, pour relaxer la condition i.i.d, supposée dans l'article cité précédemment. La condition de mélange fort est satisfaite, en général, par des processus du type $ARMA$ ou $ARCH$ ainsi que leur extension $GARCH(1,1)$. Ici, nous précisons que nous ne pouvons pas supposer que les données d'origine (qui est le $N-$échantillon) satisfait une certaine forme de dépendance. En effet, nous ne savons pas si les données observées sont $\alpha-$mélangeante ou non. Et s'ils le sont, nous ne connaissons pas le coefficient. Par conséquent, nous supposons que les données observées satisfont une sorte de condition de mélange. Sous certaines hypothèses sur le noyau, la fenêtre et la régularité de la fonction des quantiles on montre la convergence uniforme presque sûre avec vitesse de convergence de l'estimateur à noyau du quantile conditionnel en présence de troncature aléatoire à gauche dans le cas de mélange. Ce travail a fait l'objet d'une publication dans la revue *Electronic Journal of Statistics*, 2009, Vol.3, 426–445.

Le quatrième chapitre de cette thèse, traite de la normalité asymptotique de l'estimateur à noyau pour le même modèle. Il est montré que cet estimateur convenablement normalisé, converge en loi vers une variable aléatoire normale centrée réduite, où la variance asymptotique est explicitement donnée. De même pour la normalité, nos hypothèses permettent d'obtenir les mêmes vitesses que le cas i.i.d. Des applications aux prévisions et aux intervalles de confiances sont également établis. Ce travail a fait l'objet d'une publication, qui est sous presse, dans la revue *Communications in Statistics, Theory & Method*.

Contents

Introduction

Recently new developments have taken place in the theory of nonparametric statistics. Asymptotic results have been obtained and special behavior of estimators and predictors for incomplete data (under truncation and/or censoring) has been pointed out, mentioning in fully nonparametric estimation the percursor work of Ould Saïd and Lemdani [55] who established the asymptotic properties of the regression function kernel estimator under pure truncation, Ould Saïd and Sadki [56] which they study the conditional quantile for right censorship model, Ould Saïd and Tatachak [57] concerning the kernel conditional mode function from randomly left-truncated model and Lemdani *et al.* [45, 46] where they study the quantile and conditional quantile functions under left-truncation but for independent and identically distributed random variables.

However, the independence assumption for the observations is not always adequate in applications, especially for sequentially collected economic data, which often exhibit evident dependence. Our focus in the present thesis is to study the strong uniform convergence and the asymptotic normality of the kernel conditional quantile estimator used by Lemdani *et al.* [46] for the left truncation model when the data exhibit some kind of dependence. Although our interest in conditional quantile estimation is motivated by the forecasting from time series data, our results are derived where the observations exhib some kind of dependence (it is assumed that the lifetime observations with multivariate covariates from a stationary strong mixing process).

Conditional medians and quantiles are frequently used in analyzing time series data with heavy tails for their robustness properties. It is well known from the robustness that the mean is sensible to outliers (see Hampel *et al.* [32]); it may be sensible to use the median, which is a particular case of the quantile, rather than the mean to forecast future since the median is highly resistant against outliers, especially the conditional median function for asymmetric distribution, which can provide a useful alternative to the ordinary regression based on the mean.

The nonparametric estimation of conditional quantile has first been considered in the case of complete data (no truncation). Roussas [69] showed the convergence and asymptotic normality of kernel estimates of conditional quantile under Markov assumptions. For independent and identically distributed random variables, Stone [73] proved the weak consistency of kernel estimates. The uniform consistency was studied by Schlee [72] for strong mixing case. The asymptotic normality in the iid case has been established by Samanta [71]. Many other authors considered this problem; without pretending to the exhaustiveness, we quote Battacharya and Gangopadhay [2]. Jones and Hall [39], Mehra *et al.* [52], Chaudhuri [13], Fan *et al.* [19], Welsh [84] and Xiang [88]. Hounda[13] dealt with the strong mixing case and proved the uniform convergence and asymptotic normality of an estimate of the conditional quantile by considering the particular case of stationary strong mixing process. Furthermore, Qui and Wu [63] obtained the asymptotic normality of an estimator for a conditional quantile using the empirical likelihood method and a linear fitting when some auxiliary information is available. Finally Gannoun *et al.* [26] gave a smooth nonparametric conditional median predictor, based on double kernel methods and established its asymptotic normality and proposed an extension to the conditional quantile.

In censoring case, Beran [3] introduced a nonparametric estimate of the conditional survival function and prove some consistency results which were later exposed and extended by Dabrowska [15] in the iid case, and Lecoutre and Ould Saïd [43] studied the consistency in the strong mixing case. Dabrowska [15] established a Bahadur representation of kernel quantile estimator and Xiang [87] obtained the deficiency of the sample quantile estimator with respect to a kernel estimator using coverage probability. Other large samples properties of the conditional distribution have been studied extensively in the literature (see, e.g., Stute [75] and Van Keilegomand and Veraverbeke [80, 81, 82]). In the recent paper of Ould Saïd [54] (see also Kohler *et al.* [40] and Carbonnez *et al.* [12]), who established a strong uniform convergence rate of a kernel conditional quantile estimator under iid censorship model.

In the random left-truncation model, Gürler *et al.* [30] established a Bahadur-type representation for the quantile function and asymptotic normality. Its extension to time series analysis have been obtained by Lemdani *et al.* [45]. A nonparametric regression function estimator with randomly truncated data is considered in [28], [42], [35] and [55]. In the same way, Lemdani *et al.* [46] introduce a kernel conditional quantile estimator and prove its almost sure (a.s.) consistency and asymptotic normality in the independent and identically distributed case.

Although our interest in nonparametric estimation is motivated by the constructing of the confidence intervals from time series data, we introduce our results in a more general setting (strong mixing) which includes time series modeling as a special case. Among various mixing conditions used in literature, $\alpha-$mixing is reasonably weak, and is known to be fulfilled for many stochastic processes including many time series models. Gorodetskii [29] and Withers [85] derived the conditions under which a linear process is $\alpha-$mixing. In fact, under very mild assumptions linear autoregressive and mor generally bilinear time series models are strongly mixing with mixing coefficients decaying exponentially. Auestad and Tjϕstheim [1] provided illuminating discussions on the role of $\alpha-$mixing for model identification in nonlinear time series analysis. Further, Masry and Tjϕstheim [50, 50] showed that under some mild conditions, both autoregressive conditional heteroscedastic process and nonlinear additive autoregressive processes with exogenous variables, which are particularly popular in finance, are stationary and $\alpha-$mixing.

After having recalled the main basic concepts on truncated and mixing data in the first chapter, we give in a second chapter, some important and useful results existing in the literature for the random left truncation model. In the third chapter, under strong mixing hypotheses, the strong uniform convergence with rates of the kernel conditional quantile and that of the conditional distribution function is established under random left truncation and dependent data. In the fourth chapter of this thesis, we give the asymptotic normality of the kernel conditional quantile estimator still for the left-truncated and dependent data.

Chapter 1

Basic concepts

1.1 Incomplete data

One hears by lifetime, the random variable, often positive. Indeed this variable is observed in several domains as the astronomy, medicine, epidemiology, biometry, and the economy... A lifetime is therefore, in a general case, the time that it is necessary to pass from a state A to a state B. It is not rare that data to treat are not complete, in this case a classical techniques don't adjust correctly to the incomplete data. Since our work carries on the incomplete data, and in order to give back easy the reading of this thesis, we give some definitions and examples of the incomplete data.

1.1.1 Censoring

Definition 1.1.1 *Censoring is when an observation is incomplete due to some random case. The cause of the censoring must be independent of the event of interest if we are to use standard methods of analysis.*

Example 1.1.2 *Lung cancer patients are recruited to a study to test the effect of a drug on their survival from lung cancer.*
a) *takes part in the study until her death at time T_a. Her survival time is uncensored.*
b) *takes part in the study until time T_b. He then leaves the study. His survival time is censored. We know it is at least T_b but we don't know it precisely.*
c) *takes part in the study until time T_c. She then is hit by a car and dies. Her*

survival time with regard to the event of interest, namely death through lung cancer, is also censored. We know is it at least T_c.

Commonest form of censoring is **Right censoring**. Subjects followed until some time, at which the event has yet to occur, but then talks no further part in the study. This may be because:

- the subject dies from another cause, independently of the cause of interest,
- the study ends while the subject survives, or
- the subject is lost to the study, by dropping out, moving to a different area, etc.

If our data contain only uncensored and right-censored data, we can represent all individuals by the triple (i, t_i, δ_i) :

- i indexes subjects,
- t_i is the time at which the death or censoring event occurs to individual i, and
- δ_i is an indicator: $\delta_i = 1$ if i is uncensored and $\delta_i = 0$ if censored.

Remark 1.1.3 *Left censoring is much rare, in this case, event of interest already occurred at the observation time, but it is not known exactly when. Examples of left censoring include: infection with a sexually transmitted disease such as HIV/AIDS and time at which teenagers begin to drink alcohol.*

Definition 1.1.4 *Interval censoring: exact time event occurs is not known precisely, but an interval bounding this time is known. Examples of interval censoring include: infection with HIV/AIDS with regular testing and failure of a machine during the Chinese new Year.*

1.1.2 Truncation

The censored data are not the unique type of incomplete data. The other classic case is the one of the so-called truncated data, that modeling the lifetime by a variable Y that must be big enough to be observed. Must of it in fact to be bigger than the truncation variable T. Therefore contrarily to the censored data, variables are not still observed being given that if $Y < T$, nor Y nor truncation T can not be observed. It is a model that first appeared in astronomy where is composed of astral objects. The truncated data are frequently used on the lifetime study. At the end of 1980, some statistical studies were undertaken on the time of incubation of the virus of the AIDS, that is the time during which a person is seropositive without to develop the illness as much.

Definition 1.1.5 *Truncation is a variant of censoring but different which occurs when the incomplete nature of the observation is due to a systematic selection process inherent to the study design.*

Randomly truncated data frequently arise in medical studies, other application areas include economics, insurance and astronomy... In a broad sense, random truncation corresponds to biased sampling, where only partial or incomplete data are available about the variable of interest. One has two type of truncation, as follows:

i) Right truncation: only individuals with event time less than some threshold are included in the study. As example, if you ask a group of smoking school pupils at what age they started smoking, you necessarily have truncated data, as individuals who start smoking after leaving school are not included in the study.

ii) Left truncation: due to structure of the study design, we can only observe those individuals whose event time is greater than some truncation threshold. As example, imagine you wish to study how long people who have been hospitalized for a heart attack survive taking some treatment at home. The start time is taken to be the time of the heart attack. Only those individuals who survive their stay in hospital are able to be included in the study.

1.2 Mixing conditions

For many phenomena of the real world, observations in the past and present may have considerable influence on observations in the near future, but rather weak influence on observations in the far future. Random sequences that satisfy strong mixing conditions are used to model such phenomena.

In the reality, the treated data present a certain form of dependence or mixing, and they exists several form of mixing that expresses themselves according to coefficients, noted: α, β, ρ, ψ and ϕ. Among those, the alpha-mixing is weakest and is therefore least restraining. Thus, all results statement for alpha mixing data will be valid for the submissive data to another type of mixing.

1.2.1 Definitions and properties

There is large literature on basic properties of strong mixing conditions. For the approximation of mixing sequences by martingale differences, see the book by Hall and Heyde [31]. For the direct approximation of mixing random variables by independent ones, see [53], [62] and [66, Chapter 5], for mixing proprieties of linear processes, see [18] and [68]. For a recent developpement see Bradley [8].

Mixing conditions, as introduced by Rosenblatt [89] are weak dependence conditions in terms of the $\sigma-$algebras generated by a random sequence. In order to define such conditions we first introduce the conditions relative to sub $\sigma-$algebras $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ on an abstract probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{L}^2(\mathcal{A})$ denote the space of square integrable and $\mathcal{A}-$measurable random variables. Define the following measures of dependence :

$$\alpha(\mathcal{A}, \mathcal{B}) := \sup |P(A \cap B) - P(A)P(B)|, \quad A \in \mathcal{A},\ B \in \mathcal{B},$$
$$\phi(\mathcal{A}, \mathcal{B}) := \sup |P(B|A) - P(B)|, \quad A \in \mathcal{A},\ B \in \mathcal{B} \text{ and } P(A) > 0$$
$$\psi(\mathcal{A}, \mathcal{B}) := \sup \left| \frac{P(A \cap B)}{P(A)P(B)} - 1 \right|, \quad A \in \mathcal{A},\ B \in \mathcal{B},\ P(A) > 0 \text{ and } P(B) > 0$$
$$\rho(\mathcal{A}, \mathcal{B}) := \sup |corr(f, g)|, \quad f \in \mathcal{L}^2(\mathcal{A}),\ g \in \mathcal{L}^2(\mathcal{B})$$
$$\beta(\mathcal{A}, \mathcal{B}) := \sup \tfrac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|,$$

where the supremum is taken over all pairs of finite partitions $\{A_1, ..., A_I\}$ and $\{B_1, ..., B_J\}$ of Ω such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j.

The following inequality give the ranges of possible values of those measures of dependence:

$$\begin{aligned}
&0 \leq \alpha(\mathcal{A},\mathcal{B}) \leq \tfrac{1}{4}, \quad 0 \leq \phi(\mathcal{A},\mathcal{B}) \leq 1,\\
&0 \leq \psi(\mathcal{A},\mathcal{B}) \leq \infty, \quad 0 \leq \rho(\mathcal{A},\mathcal{B}) \leq 1 \text{ and}\\
&0 \leq \beta(\mathcal{A},\mathcal{B}) \leq 1.
\end{aligned}$$

Each of the following equalities is equivalent to the condition that $\mathcal{A}$ and $\mathcal{B}$ are independent:

$$\begin{aligned}
&\alpha(\mathcal{A},\mathcal{B}) = 0, \quad \phi(\mathcal{A},\mathcal{B}) = 0,\\
&\psi(\mathcal{A},\mathcal{B}) = 0, \quad \rho(\mathcal{A},\mathcal{B}) = 0 \text{ and}\\
&\beta(\mathcal{A},\mathcal{B}) = 0.
\end{aligned}$$

Finally, the measures of dependence satisfy the following inequalities:

$$\begin{cases}
2\alpha(\mathcal{A},\mathcal{B}) \leq \beta(\mathcal{A},\mathcal{B}) \leq \phi(\mathcal{A},\mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A},\mathcal{B}),\\
4\alpha(\mathcal{A},\mathcal{B}) \leq \rho(\mathcal{A},\mathcal{B}) \leq \psi(\mathcal{A},\mathcal{B}),\\
\rho(\mathcal{A},\mathcal{B}) \leq 2\left[\phi(\mathcal{A},\mathcal{B})\right]^{1/2}\left[\phi(\mathcal{B},\mathcal{A})\right]^{1/2} \leq 2\left[\phi(\mathcal{A},\mathcal{B})\right]^{1/2}.
\end{cases} \tag{1.1}$$

The first and second inequalities are elementary, the third inequality was shown by Peligrad[59] with an extension of the arguments used by Cogburn [14] and Ibrahgimov [38] to show the inequality $\rho(\mathcal{A},\mathcal{B}) \leq 2\left[\phi(\mathcal{A},\mathcal{B})\right]^{1/2}$ (see also Doob [17, p222, lemma 7.1]).

1.2.2 Strong mixing conditions

Suppose $X := (X_k; k \in \mathbb{Z})$ is a (not necessarily stationary) sequence of random variables. For $-\infty \leq I \leq J \leq \infty$, define the $\sigma-$field

$$\mathcal{F}_I^J := \sigma\left(X_k;\quad I \leq k \leq J, \quad (k \in \mathbb{Z})\right).$$

Here and below, the notation $\sigma(\cdots)$ means the $\sigma-$field $\subset \mathcal{F}$ generated by $(\cdots)$. For each $n \geq 1$, define the following dependence coefficients :

$$\begin{aligned}
&\alpha(n) := \sup_{j\in\mathbb{Z}} \alpha\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right), \quad \phi(n) := \sup_{j\in\mathbb{Z}} \phi\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right),\\
&\psi(n) := \sup_{j\in\mathbb{Z}} \psi\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right), \quad \rho(n) := \sup_{j\in\mathbb{Z}} \rho\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right) \text{ and}\\
&\beta(n) := \sup_{j\in\mathbb{Z}} \beta\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right).
\end{aligned}$$

The random sequence X is said to be :

- $\alpha-$mixing (or strong mixing) if $\alpha(n) \to 0$ as $n \to \infty$,

- $\phi-$mixing if $\phi(n) \to 0$ as $n \to \infty$,
- $\psi-$mixing if $\psi(n) \to 0$ as $n \to \infty$,
- $\rho-$mixing if $\rho(n) \to 0$ as $n \to \infty$ and
- $\beta-$mixing (or absolutely regular) if $\beta(n) \to 0$ as $n \to \infty$.

The strong mixing condition was introduced by Rosenblatt [69]. The $\phi-$mixing condition was introduced by Ibragimov [37], and was also studied by Cogburn [14]. The $\psi-$mixing condition had its origin in a paper by Blum *et al.* [5] studying a different condition based on the same measure of dependence, and it took its present form in the paper of Philipp [61]. The $\rho-$mixing condition was introduced by Kolmogorov and Rozanov [41]. The absolute regularity condition was introduced by Volkonskii and Rozanov [78, 79]. In the special case where the sequence X is strictly stationary, one has simply

$$\alpha(n) := \sup \alpha\left(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_{n}^{\infty}\right),$$

and the same holds for the other dependence coefficients.

Remark 1.2.1 *It needs to be kept in mind that two barely different phrases are used with quite different meanings: The phrase "strong mixing condition" (singular), or simply "strong mixing" refers to $\alpha-$mixing ($\alpha(n) \to 0$) as above. In contrast, the phrase "strong mixing conditions" (plural) refers to all mixing conditions that are at least as strong as (i.e. that imply) $\alpha-$mixing. The latter phrase "strong mixing conditions" is intended to distinguish from a broad class of "mixing conditions" from ergodic theory that are weaker than $\alpha-$mixing (See e.g. Petersen [60]).*

Finally, from (1.1), The following relations hold:

$$\phi - mixing \Longrightarrow \left\{ \begin{array}{c} \rho - mixing \\ \beta - mixing \end{array} \right\} \Longrightarrow \alpha - mixing$$

and no reverse implication holds in general.

Remark 1.2.2 *For more details on the mixing conditions, one can consult for example: Doukhan [18], Bosq [7], Rio [66], Bradley [8] and Dedecker et al. [16].*

Chapter 2

Estimation under random left-truncation model

In this chapter, we present some important and useful results existing in the literature for the random left truncation model :

2.1 Random left-truncation model

Let $(\mathcal{Y}_j, \mathcal{T}_j)$, $1 \leq j \leq N$, be a sequence of iid random vectors such that $(\mathcal{Y}_j)$ is independent of $(\mathcal{T}_j)$. Let F and G denote the respective common distribution functions of the $\mathcal{Y}_j$ values and $\mathcal{T}_j$ values. In the *random left truncation model (RLT)*, the *rv of interest* $\mathcal{Y}$ is interfered by *the truncation rv* $\mathcal{T}$, in such a way that both $\mathcal{Y}_j$ and $\mathcal{T}_j$ are observable when $\mathcal{Y}_j \geq \mathcal{T}_j$.

If there were no truncation, we could think of the observations as $(\mathcal{Y}_j, \mathcal{T}_j)\,;\, 1 \leq j \leq N$, where the sample size N is deterministic, but unknown. Under RLT, however, some of these vectors would be missing and for notational convenience, we shall denote $(Y_i, T_i)\,;\, 1 \leq i \leq n$, $(n \leq N)$ the observed subsequence subject to $Y_i \geq T_i$ from the $N-$sample. As a consequence of truncation, the size of actually observed sample, n, is a binomial rv with parameters N and $\mu := \mathbb{P}(\mathcal{Y} \geq \mathcal{T})$. It is clear that, the parameter μ represent the probability that we observe the rv of interest $\mathcal{Y}$, however, if $\mu = 0$ no data can be observed. Therefore, we suppose throughout the thesis that $\mu > 0$. By the strong law of large numbers we have, as $N \to \infty$

$$\hat{\mu}_n := \frac{n}{N} \to \mu, \quad \mathbb{P}-a.s. \tag{2.1}$$

Conditionally on the value of n, these observed random vectors are still iid. Under RLT sampling scheme, the conditional joint distribution (Stute [74]) of (Y,T) becomes

$$\begin{aligned} J^*(y,t) &= \mathbf{P}\left(Y \le y, T \le t\right) = \mathbb{P}\left(Y \le y, T \le t | Y \ge T\right) \\ &= \mu^{-1}\int_{-\infty}^{y} G(t \wedge u) dF(u) \end{aligned}$$

where $t \wedge u := \min(t,u)$. The marginal distribution are defined by

$$F^*(y) \ := \ J^*(y,\infty) = \mu^{-1}\int_{-\infty}^{y} G(u)dF(u)$$

and

$$\begin{aligned} G^*(t) \ := \ J^*(\infty,t) &= \mu^{-1}\int_{-\infty}^{\infty} G(t \wedge u) dF(u) \\ &= \mu^{-1}\int_{-\infty}^{t} \left(1 - F(u)\right) dG\left(u\right), \end{aligned}$$

which are estimated by

$$F_n^*(y) = n^{-1}\sum_{i=1}^{n}\mathbf{1}_{\{Y_i \le y\}} \quad \text{and} \quad G_n^*(t) = n^{-1}\sum_{i=1}^{n}\mathbf{1}_{\{T_i \le t\}}$$

respectively, where $\mathbf{1}_A$ denotes the indicator function of the set A. Let $C\left(\cdot\right)$ be a function defined by

$$\begin{aligned} C\left(y\right) &= \mathbb{P}\left(T \le y \le Y | Y \ge T\right) \ := \ G^*(y) - F^*(y) \\ &= \mu^{-1}G\left(y\right)\left[1 - F\left(y\right)\right], \end{aligned}$$

with empirical estimator

$$C_n(y) = G_n^*(y) - F_n^*(y-) = n^{-1}\sum\nolimits_{i=1}^{n}\mathbf{1}_{\{T_i \le y \le Y_i\}}.$$

The nonparametric maximum likelihood estimators of F and G are the product-limit estimators (Lynden-Bell [48]) given by

$$F_n(y) = 1 - \prod_{i/Y_i \le y}\left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)}\right] \quad \text{and} \quad G_n(y) = \prod_{i/T_i > y}\left[\frac{nC_n(T_i) - 1}{nC_n(T_i)}\right]. \qquad (2.2)$$

2.2 Estimation of the truncation probability

For any df L, denotes the left and right endpoint of its support by $a_L := \inf\{x : L(x) > 0\}$ and $b_L := \sup\{x : L(x) < 1\}$,respectively. Consequently, μ is identifiable only if $a_G \leq a_F$ and $b_G \leq b_F$. Note that the estimator $\hat{\mu}_n$ defined in equation (2.1) cannot be calculated (since N is unknown). Another estimator, namely

$$\mu_n = \frac{G_n(y)\left[(1 - F_n(y-))\right]}{C_n(y)}, \tag{2.3}$$

is used, where $F_n(y-)$ denotes the left-limite of $F_n(\cdot)$ at y.

Lemma 2.2.1 (Liang *et al.* [42]) *Let $\{Y_i,\ i \geq 1\}$ be a stationary $\alpha-$mixing sequence of rv's with mixing coefficient $\alpha(n) = O(n^{-\nu})$ for some $\nu > 3$. Then*

$$\sup_y |C_n(y) - C(y)| = O\left((\log\log n/n)^{1/2}\right), \qquad a.s. \tag{2.4}$$

$$\sup_y |F_n(y) - F(y)| = O\left((\log\log n/n)^{1/2}\right), \qquad a.s. \tag{2.5}$$

$$\sup_y |G_n(y) - G(y)| = O\left((\log\log n/n)^{1/2}\right), \qquad a.s. \tag{2.6}$$

$$\sup_y |\mu_n - \mu| = O\left((\log\log n/n)^{1/2}\right), \qquad a.s. \tag{2.7}$$

Remark 2.2.2 *Under iid setting, Woodroofe [86, Theorem 2] established the uniform consistency results of F_n and G_n :*

$$\sup_y \ |F_n(y) - F_0(y)| \overset{\mathbf{P}-a.s.}{\to} 0, \quad and \quad \sup_y \ |G_n(y) - G_0(y)| \overset{\mathbf{P}-a.s.}{\to} 0, \tag{2.8}$$

where F_0 denotes the conditional distribution of Y given $Y \geq a_G$ and G_0 is the conditional distribution of T given $T \leq b_F$. Therefore, F is identifiable ($F = F_0$) only when $a_G \leq a_F$, whereas G is identifiable ($G = G_0$) only when $b_G \leq b_F$. As pointed out by Ould Saïd and Lemdani [55], these are necessary but not sufficient identifiability conditions. He and Yang [34] proved that μ_n does not depend on y and its value can then be obtained for any y such that $C_n(y) \neq 0$. Furthermore, they showed -in the iid case- (see their Corollary 2.5) its $\mathbb{P}-a.s.$ consistency.

Remark 2.2.3 *Under $\alpha-$mixing structure, Sun and Zhou [76] expressed the product limit estimator F_n as an average of a sequence of bounded rv's plus a remainder term of order $O\left(n^{-1/2}\log^{-\varsigma} n\right)$ for some $\varsigma > 0$, and obtained similar results as those obtained for the Kaplan–Meier estimator for censored dependent data (see Cai [11]).*

The proof of Lemma 2.2.1 is based on the next result :

Lemma 2.2.4 (Cai and Roussas [9]) *Let* $\{\xi_n,\ n \geq 1\}$ *be a stationary* $\alpha-$*mixing sequence of rv's with df* $\mathcal{F}$ *and mixing coefficient* $\alpha(n) = O(n^{-\nu})$ *for some* $\nu > 3$, *and let* $\mathcal{F}_n$ *be the empirical df based on the segment* $\xi_1, \cdots, \xi_n$. *Then*

$$\limsup_{n\to\infty} \left\{ \left(\frac{n}{2\log\log n} \right)^{1/2} \sup_{x\in\mathbb{R}} |\mathcal{F}_n(x) - \mathcal{F}(x)| \right\} = 1, \qquad a.s.$$

Set

$$\Lambda(y) = \int_0^y \frac{dF^*(u)}{C(u)}, \quad \Lambda_n(y) = \int_0^y \frac{dF_n^*(u)}{C_n(u)}, \quad \bar{\Lambda}(y) = \int_y^\infty \frac{dG^*(u)}{C(u)}, \quad \bar{\Lambda}_n(y) = \int_y^\infty \frac{dG_n^*(u)}{C_n(u)}.$$

Obviously

$$\Lambda_n(y) = \sum_{i=1}^n \frac{\mathbf{1}_{\{Y_i \leq y\}}}{nC_n(Y_i)}, \qquad \bar{\Lambda}_n(y) = \sum_{i=1}^n \frac{\mathbf{1}_{\{T_i > y\}}}{nC_n(T_i)}.$$

Proof of Lemma 2.2.1. By applying Lemma 2.2.4 we have

$$\sup_y |F_n^*(y) - F^*(y)| = O(\theta_n) \quad a.s. \quad \text{and} \quad \sup_y |G_n^*(y) - G^*(y)| = O(\theta_n) \quad a.s. \tag{2.9}$$

where $\theta_n = (\log\log(n)/n)^{1/2}$.

We first verify (2.4). Since

$$C(y) = G^*(y) - F^*(y) \qquad \text{and} \qquad C_n(y) = G_n^*(y) - F_n^*(y-),$$

by (2.9) we get (2.4).

We now consider (2.5). We define

$$\bar{F}_n(y) = 1 - \prod_{i/Y_i \leq y} \left[1 - \frac{1}{nC_n(Y_i) + 1} \right].$$

By using inequality $|e^{-x} - e^{-y}| \leq |x-y|$ for $x, y \geq 0$ and expanding $\ln\left(1-\bar{F}_n\right)$, we get

$$\begin{aligned}
\left|1-\bar{F}_n - e^{-\Lambda_n(y)}\right| &\leq \left|\ln\left(1-\bar{F}_n\right) + \Lambda_n(y)\right| \\
&= \left|\sum_{i/Y_i\leq y} \frac{1}{nC_n(Y_i)} - \sum_{i/Y_i\leq y}\sum_{k\geq 0} \frac{1}{k\left[nC_n(Y_i)+1\right]^k}\right| \\
&\leq \frac{3}{2}\sum_{i/Y_i\leq y} \frac{1}{nC_n(Y_i)\left[nC_n(Y_i)+1\right]} \\
&\leq \frac{3}{2n}\int_0^y \frac{dF_n^*(u)}{C_n^2(u)}.
\end{aligned} \tag{2.10}$$

Since

$$\left|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i\right| \leq \sum_{i=1}^n |a_i - b_i| \qquad \text{for} \quad |a_i|, \ |b_i| \leq 1,$$

we have

$$\begin{aligned}
\left|F_n(y) - \bar{F}_n(y)\right| &= \left|(1-F_n(y)) - \left(1-\bar{F}_n(y)\right)\right| \\
&\leq \frac{1}{n}\int_0^y \frac{dF_n^*(u)}{C_n^2(u)}.
\end{aligned} \tag{2.11}$$

Noticing that by (2.10) and (2.11)

$$1 - F_n(y) - e^{-\Lambda_n(y)} = O\left(n^{-1}\right)\int_0^y \frac{dF_n^*(u)}{C_n^2(u)} \tag{2.12}$$

and $1 - F(y) = e^{-\Lambda(y)}$, we have
$|F_n(y) - F(y) - [1-F(y)]\left[\Lambda_n(y) - \Lambda(y)\right]|$

$$\begin{aligned}
&= \left|\left[e^{-\Lambda(y)} - e^{-\Lambda_n(y)}\right] - e^{-\Lambda(y)}\left[\Lambda_n(y) - \Lambda(y)\right] - 1 - F_n(y) - e^{-\Lambda_n(y)}\right| \\
&\leq e^{-\xi_{2n}(y)}\left[\Lambda_n(y) - \Lambda(y)\right]^2 + \left|1 - F_n(y) - e^{-\Lambda_n(y)}\right|,
\end{aligned} \tag{2.13}$$

where $\xi_{2n}(y)$ is between $\xi_{1n}(y)$ and $\Lambda_n(y)$, and $\xi_{1n}(y)$ is between $\Lambda_n(y)$ and $\Lambda(y)$.

Note that

$$\begin{aligned}
\Lambda_n(y) - \Lambda(y) &= \int_0^y \frac{dF_n^*(u)}{C_n(u)} - \int_0^y \frac{dF^*(u)}{C(u)} \\
&= \int_0^y \left(\frac{1}{C(u)} - \frac{1}{C_n(u)} \right) dF_n^*(u) + \int_0^y \frac{d\left(F_n^*(u) - F^*(u)\right)}{C(u)} \\
&= \int_0^y \frac{C(u) - C_n(u)}{C(u)\, C_n(u)} dF_n^*(u) + \frac{F_n^*(y) - F^*(y)}{C(y)} \\
&+ \int_0^y \frac{F_n^*(u) - F^*(u)}{C^2(u)} dC(u),
\end{aligned}$$

which, together with (2.4) and (2.9), implies

$$\sup_y |\Lambda_n(y) - \Lambda(y)| = O(\theta_n) \qquad a.s. \tag{2.14}$$

Therefore, from (2.12)–(2.14) we conclude

$$\sup_y |F_n(y) - F(y)| = O(\theta_n), \qquad a.s., \tag{2.15}$$

proving (2.5).

Next we prove (2.6), denote

$$\bar{G}_n(y) = 1 - \prod_{i/T_i > y} \left[1 - \frac{1}{nC_n(T_i) + 1} \right].$$

Similarly to the proof of (2.5) we have

$$\begin{aligned}
\left| G_n(y) - e^{-\bar{\Lambda}_n(y)} \right| &\leq \left| \bar{G}_n(y) - e^{-\bar{\Lambda}_n(y)} \right| + \left| G_n(y) - \bar{G}_n(y) \right| \\
&\leq \left| \ln\left(\bar{G}_n(y)\right) + \bar{\Lambda}_n(y) \right| + \left| G_n(y) - \bar{G}_n(y) \right| \\
&= \frac{3}{2} \sum_{i/T_i > y} \frac{1}{nC_n(T_i)\left[nC_n(T_i) + 1\right]} + \sum_{i/T_i > y} \frac{1}{nC_n(T_i)\left[nC_n(T_i) + 1\right]} \\
&\leq O\left(n^{-1}\right) \int_y^\infty \frac{dG_n^*(u)}{C_n^2(u)}.
\end{aligned} \tag{2.16}$$

Note that

$$\bar{\Lambda}_n(y) = \int_y^\infty \frac{dG^*(u)}{C(u)} = \int_y^\infty \frac{dG(u)}{C(u)} = -\ln(G(y)),$$

which implies $G(u) = e^{-\bar{\Lambda}(y)}$. Hence we have
$\left|(G_n(y) - G(y)) + G(y)\left(\bar{\Lambda}_n(y) - \bar{\Lambda}(y)\right)\right|$

$$\begin{aligned}
&\le \left|G_n(y) - e^{-\bar{\Lambda}_n(y)}\right| + \left|\left(e^{-\bar{\Lambda}(y)} - e^{-\bar{\Lambda}_n(y)}\right) - e^{-\bar{\Lambda}(y)}\left(\bar{\Lambda}_n(y) - \bar{\Lambda}(y)\right)\right| \\
&= \left|G_n(y) - e^{-\bar{\Lambda}_n(y)}\right| + e^{-\xi_{4n}(y)}\left|\bar{\Lambda}_n(y) - \bar{\Lambda}(y)\right|\left|\xi_{3n}(y) - \bar{\Lambda}(y)\right|, \qquad (2.17)
\end{aligned}$$

where $\xi_{4n}(y)$ is between $\xi_{3n}(y)$ and $\bar{\Lambda}(y)$, and $\xi_{3n}(y)$ is between $\bar{\Lambda}_n(y)$ and $\bar{\Lambda}(y)$.
We observe that

$$\begin{aligned}
\left|\bar{\Lambda}_n(y) - \bar{\Lambda}(y)\right| &= \int_y^\infty \frac{dG_n^*(u)}{C_n(u)} - \int_y^\infty \frac{dG^*(u)}{C(u)} \\
&= \int_y^\infty \frac{C(u) - C_n(u)}{C(u)C_n(u)} dG_n^*(u) + \frac{G_n^*(y) - G^*(y)}{C(y)} \\
&+ \int_y^\infty \frac{G_n^*(u) - G^*(u)}{C^2(u)} dC(u). \qquad (2.18)
\end{aligned}$$

Therefore, from (2.4), (2.9) and (2.16)–(2.18) we conclude (2.6), that is,

$$\sup_y |G_n(y) - G(y)| = O(\theta_n), \qquad a.s.$$

Finally we prove (2.7). We observe that

$$\begin{aligned}
\mu_n - \mu &= \frac{G_n(y)[1 - F_n(y-)]}{C_n(y)} - \frac{G(y)[1 - F(y-)]}{C(y)} \\
&= \frac{1}{C_n(y)C(y)}\{C(y)[1 - F_n(y-)][G_n(y) - G(y)] \\
&+ C(y)G(y)[F(y) - F_n(y-)] - [C_n(y) - C(y)]G(y)[1 - F(y-)]\}.
\end{aligned}$$

Hence, from the continuity of F and equations (2.4), (2.5) and (2.6) we obtain (2.7). □

2.3 Estimation of the covariate's density

Now, in addition to the considered previously variables $\mathcal{Y}$ and $\mathcal{T}$, we consider a random vector $\mathcal{X} \in \mathbb{R}^d$ of covariates, assumed to be absolutely continuous with distribution function $V(\cdot)$ and continuous density $v(\cdot)$. We could think of the observations as $(\mathcal{X}_j, \mathcal{Y}_j, \mathcal{T}_j)$; $1 \leq j \leq N$, where the sample size N is deterministic, but unknown. Under RLT, however, some of these vectors would be missing and for notational convenience, we shall denote $(\mathbf{X}_i, Y_i, T_i)$; $1 \leq i \leq n$, $(n \leq N)$ the observed subsequence subject to $Y_i \geq T_i$ from the $N-$sample. From now on, T is assumed to be independent of $(\mathbf{X}, Y)$ and $(\mathbf{X} \leq \mathbf{x})$ stands for $(X_1 \leq x_1, \cdots, X_d \leq x_d)$.

We build here estimators of $V(\cdot)$ and $v(\cdot)$. Firstly, the naturel kernel estimator of the covarite's denstiy $v(\cdot)$ is given by

$$v_N(\mathbf{x}) := \frac{1}{Nh_N^d} \sum_{j=1}^{N} K_d\left(\frac{\mathbf{x} - \mathcal{X}_j}{h_N}\right), \tag{2.19}$$

where $K_d : \mathbb{R}^d \to \mathbb{R}$ is a fixed kernel with $\int_{\mathbb{R}^d} K_d = 1$ and $(h_N)_{N \geq 1}$ a nonnegative bandwidth sequence tending to zero as N grows to infinity. Note that we can no longer use the kernel estimator $v_N(\cdot)$, since only $(n \leq N)$ observations are made. On the other hand,

$$v_n^*(\mathbf{x}) := \frac{1}{Nh_n^d} \sum_{i=1}^{n} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) \tag{2.20}$$

is an estimator of the conditional density $v^*(\cdot)$ (subject to $Y \geq T$). To overcome this difficulty, we first consider the following trivariate conditional joint distribution $\mathcal{H}^*$ of $(\mathbf{X}, Y, T)$:

$$\begin{aligned}
\mathcal{H}^* &= \mathbf{P}(\mathbf{X} \leq \mathbf{x}, Y \leq y, T \leq t) \\
&= \mathbb{P}(\mathbf{X} \leq \mathbf{x}, Y \leq y, T \leq t | Y \geq T) \\
&= \mu^{-1} \int_{u \leq \mathbf{x}} \int_{a_G \leq v \leq y} G(t \wedge u) \mathbf{F}(d\mathbf{u}, dv),
\end{aligned}$$

where $\mathbf{F}(\cdot, \cdot)$ is the joint distribution function of $(\mathbf{X}, Y,)$. Taking $t = +\infty$, the observed pair then has the following distribution $\mathbf{F}^*(\cdot, \cdot)$:

$$\begin{aligned}
\mathbf{F}^*(\mathbf{x}, y) &= \mathcal{H}^*(\mathbf{x}, y, +\infty) \\
&= \mu^{-1} \int_{u \leq \mathbf{x}} \int_{a_G \leq v \leq y} G(u) \mathbf{F}(d\mathbf{u}, dv).
\end{aligned} \tag{2.21}$$

By differentiating (2.21), we get

$$\mathbf{F}(d\mathbf{x},dy)=\frac{1}{\mu^{-1}G\,(y)}\mathbf{F}^*(d\mathbf{x},dy),\qquad \text{for } y\geq a_G. \tag{2.22}$$

Hence

$$\mathbf{f}(\mathbf{x},y)=\frac{1}{\mu^{-1}G\,(y)}\mathbf{f}^*(\mathbf{x},y). \tag{2.23}$$

Integrating (2.22) over y, we obtain the df of $\mathbf{X}$:

$$V(\mathbf{x})=\mu^{-1}\int_{u\leq\mathbf{x}}\int_{a_G\leq y}\frac{1}{G\,(y)}\mathbf{F}^*(d\mathbf{u},dy).$$

A natural estimator of $V(\mathbf{x})$ is then given by

$$V_n(\mathbf{x})=\frac{\mu_n}{n}\sum_{i=1}^{n}\frac{1}{G_n\,(Y_i)}\mathbf{1}_{(X_i\leq x)}. \tag{2.24}$$

Note that in (2.24) and in the sequel, the sum is taken only for i such that $G\,(Y_i)\neq 0$. Finally (2.24) yields the density estimator of $\mathbf{X}$ as

$$\begin{aligned} v_n(\mathbf{x}) &= \frac{1}{h_n^d}\int_{\mathbb{R}^d}K_d\left(\frac{\mathbf{x}-\mathbf{u}}{h_n}\right)V_n(d\mathbf{u}) \\ &= \frac{\mu_n}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right), \end{aligned} \tag{2.25}$$

where $(h_n)_{n\geq 1}$ a positive bandwidth sequence tending to zero as n grows to infinity.

Adopting the same methodology, and while observing (2.23), we get an estimator of $\mathbf{F}(\mathbf{x},y)$ as follows

$$\mathbf{F}_n(\mathbf{x},y)=\frac{\mu_n}{n}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}\mathbf{1}_{(X_i\leq x,Y_i\leq y)}.$$

According to (2.23), we define the kernel estimate of joint probability density function $\mathbf{f}(\mathbf{x},y)$ as follows

$$\begin{aligned} \mathbf{f}_n(\mathbf{x},y) &= \frac{1}{h_n^d\ell_n}\int_{\mathbb{R}^d\times\mathbb{R}}K_d\left(\frac{\mathbf{x}-\mathbf{u}}{h_n}\right)K_0\left(\frac{y-v}{\ell_n}\right)\mathbf{F}_n(d\mathbf{u},dv) \\ &= \frac{\mu_n}{nh_n^d\ell_n}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)K_0\left(\frac{y-Y_i}{\ell_n}\right), \end{aligned} \tag{2.26}$$

where $K_0:\mathbb{R}\to\mathbb{R}$ is a fixed kernel with $\int_{\mathbb{R}}K_0=1$ and $(\ell_n)_{n\geq 1}$ is defined as $(h_n)_{n\geq 1}$ above.

Chapter 3

A strong uniform convergence rate of a kernel conditional quantile estimator under random left-truncation and dependent data

Elias Ould Saïd, Djabrane Yahia, Abdelhakim Necir

Abstract: In this chapter[1] we study some asymptotic properties of the kernel conditional quantile estimator with randomly left-truncated data which exhibit some kind of dependence. We extend the result obtained by Lemdani, Ould Saïd and Poulin [46, J.MA. 2009] in the iid case. The uniform strong convergence rate of the estimator under strong mixing hypothesis is obtained.

AMS 2000 subject classifications: Primary 62G05, 62G20.
Keywords and phrases : Kernel estimator, Quantile function, Rate of convergence, Strong mixing, Strong uniform consistency, Truncated data.

[1]This chapter corresponds to the paper appeared in *Electronic Jornal of Statistics*, Vol. 3 (2009) 426–445.

3.1 Introduction

Let $\mathcal{Y}$ and $\mathcal{T}$ be two real random variables (rv) with unknown cumulative distribution functions (df) F and G respectively, both assumed to be continuous. Let $\mathcal{X}$ be a real-valued random covariable with df V and continuous density v. Under random left-truncation (RLT), the rv of interest $\mathcal{Y}$ is interfered by the truncation rv $\mathcal{T}$, in such a way that $\mathcal{Y}$ and $\mathcal{T}$ are observed only if $\mathcal{Y} \geq \mathcal{T}$. Such data occur in astronomy and economics (see Woodroofe [86], Feigelson and Babu [20], Wang *et al.* [83], Tsai *et al.* [77]) and also in epidemiology and biometry (see, e.g., He and Yang [33]).

If there were no truncation, we could think of the observations as $(\mathcal{X}_j, \mathcal{Y}_j, \mathcal{T}_j)$; $1 \leq j \leq N$, where the sample size N is deterministic, but unknown. Under RLT, however, some of these vectors would be missing and for notational convenience, we shall denote $(\mathbf{X}_i, Y_i, T_i)$; $1 \leq i \leq n$, $(n \leq N)$ the observed subsequence subject to $Y_i \geq T_i$ from the $N-$sample.

As a consequence of truncation, the size of actually observed sample, n, is a binomial rv with parameters N and
$\mu := \mathbb{P}(\mathcal{Y} \geq \mathcal{T}) > 0$. By the strong law of large numbers we have, as $N \to \infty$

$$\hat{\mu}_n := \frac{n}{N} \to \mu, \quad \mathbb{P}-a.s. \tag{3.1}$$

Now we consider the joint df $\mathbf{F}(\cdot,\cdot)$ of the random vector $(\mathcal{X}, \mathcal{Y})$ related to the $N-$sample and suppose it is of class $\mathcal{C}^1$. The conditional df of $\mathcal{Y}$ given $\mathcal{X} = \mathbf{x}$, that is $\mathbf{F}(y|\mathbf{x}) = \mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y} \leq y\}} | \mathcal{X} = \mathbf{x}\right]$ which may be rewritten into

$$\mathbf{F}(\cdot|x) = \frac{\mathbf{F}_1(\mathbf{x}, \cdot)}{v(\mathbf{x})} \tag{3.2}$$

where $\mathbf{F}_1(x, \cdot)$ is the first derivative of $F(\mathbf{x}, \cdot)$ with respect to $\mathbf{x}$. For all fixed $p \in (0, 1)$, the p^{th} conditional quantile of $\mathbf{F}$ given $\mathcal{X} = \mathbf{x}$ is defined by

$$q_p(\mathbf{x}) := \inf \{y \in \mathbb{R} : \mathbf{F}(y|\mathbf{x}) \geq p\} .$$

It is well known that the quantile function can give a good description of the data (see, Chaudhuri *et al.* [13]), such as robustness to heavy-tailed error distributions and outliers, especially the conditional median function $q_{1/2}(\mathbf{x})$ for asymmetric distribution, which can provide a useful alternative to the ordinary regression based on the mean. The nonparametric estimation of conditional quantile has first been considered in the case of complete data (no truncation). Roussas [69] showed the convergence and asymptotic normality of kernel estimates of conditional quantile under Markov assumptions. For independent and identically distributed (iid) rv's, Stone [73] proved the weak consistency of kernel estimates. The uniform consistency was studied by Schlee [72] and Gannoun [23]. The asymptotic normality has been established by Samanta [71]. Mehra *et al.* [52] proposed and discussed certain smooth variants (based both on single as well as double kernel weights) of the standard conditional quantile estimator, proved the asymptotic normality and found an almost sure (a.s.) convergence rate, whereas Xiang [87] gave the asymptotic normality and a law of the iterated logarithm for a new kernel estimator. In the dependent case, the convergence of nonparametric estimation of quantile was proved by Gannoun [24] and Boente and Fraiman [6].

In the RLT model, Gürler *et al.* [30] gave a Bahadur-type representation for the quantile function and asymptotic normality. Its extension to time series analysis was obtained by Lemdani *et al.* [45].

The aim of this paper is to establish a strong uniform convergence rate for the kernel conditional quantile estimator with randomly left-truncated data under $\alpha-$mixing conditions whose definition is given below. Hence, we extend the obtained result by Lemdani *et al.* [46] in the iid case.

First, let $\mathcal{F}_i^k(Z)$ denotes the σ-field of events generated by $\{Z_j,\ i \leq j \leq k\}$. For easy reference, let us recall the following definition.

Definition 3.1.1 *Let $\{Z_i,\ i \geq 1\}$ denotes a sequence of random variables. Given a positive integer n, set:*

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

The sequence is said to be $\alpha-$mixing (strongly mixing) if the mixing coefficient $\alpha(n) \to 0$.

Among various mixing conditions used in the literature, $\alpha-$mixing is reasonably weak and has many practical applications (see, e.g. Doukhan [18] or Cai ([10, 11] for more details). In particular, Masry and Tjϕstheim [50] proved that, both ARCH processes and nonlinear additive AR models with exogenous variables, which are particularly popular in finance and econometrics, are stationary and $\alpha-$mixing.

The rest of this chapter is organized as follows. In Section 2, we recall a definition of the kernel conditional quantile estimator with randomly left-truncated data. Assumptions and main results are given in Section 3. Section 4 is devoted to application to prediction. Finally, the proofs of the main results are postponed to Section 5 with some auxiliary results and their proofs.

3.2 Definition of the estimator

In the sequel, the letters C and C' are used indiscriminately as generic constants. Note also that, N is unknown and n is known (although random), our results will not be stated with respect to the probability measure $\mathbb{P}$ (related to the $N-$sample) but will involve the conditional probability $\mathbf{P}$ (related to the $n-$sample). Also $\mathbb{E}$ and $\mathbf{E}$ will denote the expectation operators related to $\mathbb{P}$ and $\mathbf{P}$, respectively. Finally, we denote by a superscript $(^*)$ any df that is associated to the observed sample.

The estimation of conditional df is based on the choice of weights. For the complete data, the well-known Nadaraya-Watson weights are given by

$$W_{i,N}(\mathbf{x}) = \frac{K_d\left\{(x-\mathcal{X}_i)/h_N\right\}}{\sum_{i=1}^N K_d\left\{(x-\mathcal{X}_i)/h_N\right\}} = \frac{\left(Nh_N^d\right)^{-1} K_d\left\{(x-\mathcal{X}_i)/h_N\right\}}{v_N(x)} \tag{3.3}$$

that are measurable functions of x depending on $\mathcal{X}_1, ..., \mathcal{X}_N$, with the convention $0/0 = 0$. The kernel K_d is a measurable function on $\mathbb{R}^d$ and (h_N) a nonnegative sequence which tends to zero as N tends to infinity. The regression estimator based on the N-sample is then given by

$$r_N(x) = \frac{\left(Nh_N^d\right)^{-1} \sum_{i=1}^n Y_i K\left\{(x-\mathcal{X}_i)/h_N\right\}}{v_N(x)} \tag{3.4}$$

where v_N is a well known kernel estimator of v based on the $N-$sample. As N is unknown, then $v_N(\cdot)$ cannot be calculated and therefore $r_N(\cdot)$. On the other hand, based on the $n-$sample, the kernel estimator

$$v_n^*(x) = \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{x-\mathbf{X}_i}{h_n}\right) \tag{3.5}$$

is an estimator of the conditional density $v^*(x)$ (given $\mathcal{Y} \geq \mathcal{T}$), see Ould Saïd and Lemdani [55].

Under RLT sampling scheme, the conditional joint distribution (Stute, [74]) of (Y, T) becomes

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

where $t \wedge u := \min(t, u)$. The marginal distribution and their empirical versions are defined by

$$F^*(y) = \mu^{-1} \int_{-\infty}^{y} G(u) dF(u), \qquad F_n^*(y) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{Y_i \leq y\}},$$

$$G^*(t) = \mu^{-1} \int_{-\infty}^{\infty} G(t \wedge u) dF(u) \qquad \text{and} \qquad G_n^*(t) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq t\}},$$

where $\mathbf{1}_A$ denote the indicator function of the set A.

In the sequel we use the following consistent estimator

$$\mu_n = \frac{G_n(y)\left[(1 - F_n(y-))\right]}{C_n\left(y\right)}, \tag{3.6}$$

for any y such that $C_n\left(y\right) \neq 0$, where $F_n(y-)$ denotes the left-limite of $F_n(\cdot)$ at y. Here F_n and G_n are the product-limit estimators (Lynden-Bell, [48]) for F and G, respectively i.e.,

$$F_n(y) = 1 - \prod_{i/Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)}\right], \qquad G_n(y) = \prod_{i/T_i > y} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)}\right]$$

and $C_n(y) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq y \leq Y_i\}}$ is the empirical estimator of

$$C\left(y\right) = \mathbb{P}\left(T \leq y \leq Y | Y \geq T\right).$$

He and Yang [34] proved that μ_n does not depend on y and its value can then be obtained for any y such that $C_n\left(y\right) \neq 0$. Furthermore, they showed in the iid case (see their Corollary 2.5) its $\mathbb{P}-a.s.$ consistency.

Suppose that one observes the n triplets $(\mathbf{X}_i, Y_i, T_i)$ among the N ones and for any df L, denote the left and right endpoint of its support by $a_L := \inf\left\{x : L(x) > 0\right\}$

and $b_L := \sup\{x : L(x) < 1\}$, respectively. Then under the current model, as discussed by Woodroofe [86], F and G can be estimated completely only if

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

In order to estimate the marginal density v we have to take into account the truncation and the estimator

$$v_n(x) = \frac{\mu_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{x - \mathbf{X}_i}{h_n}\right) \tag{3.7}$$

is considered in Ould Saïd-Lemdani [55]. Note that in this formula and the forthcoming, the sum is taken only for i such that $G_n(Y_i) \neq 0$.

Then, adapting Ould Saïd-Lemdani's weights, we get the following estimator of the conditional df of $\mathcal{Y}$ given $\mathcal{X} = x$

$$\begin{aligned}
\mathbf{F}_n(y|x) &= \mu_n \sum_{i=1}^{n} \widetilde{W}_{i,n}(x) G_n^{-1}(Y_i) H\left(\frac{y - Y_i}{h_n}\right) \\
&= \frac{\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{x - \mathbf{X}_i}{h_n}\right) H\left(\frac{y - Y_i}{h_n}\right)}{\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{x - \mathbf{X}_i}{h_n}\right)} \\
&=: \frac{\mathbf{F}_{1,n}(x, y)}{v_n(x)},
\end{aligned} \tag{3.8}$$

where H is a df defined on $\mathbb{R}$ and

$$\mathbf{F}_{1,n}(x, y) = \frac{\mu_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{x - \mathbf{X}_i}{h_n}\right) H\left(\frac{y - Y_i}{h_n}\right) \tag{3.9}$$

is an estimator of $\mathbf{F}_1(x, y)$. As the latter is continuous, it is clear that it is better to define a smooth estimator by using a continuous function $H(\cdot)$ instead of a step function $I_{\{\cdot\}}$. We point out here that the estimators (3.8) and (3.9) have been already defined in Lemdani *et al.* [46].

Then a natural estimator of the p^{th} conditional quantile $q_p(x)$ is given by

$$q_{p,n}(x) := \inf\{y \in \mathbb{R} : \mathbf{F}_n(y|x) \geq p\}, \tag{3.10}$$

which satisfies $\mathbf{F}_n(q_{p,n}(x)|x) = p$.

3.3 Assumptions and main results

In what follows, we focus our attention on the case of univariate covaraible (i.e., $d = 1$) and denote X for $\mathbf{X}$ and K for K_1. Assume that $0 = a_G < a_F$ and $b_G \leq b_F$. We consider two real numbers a and b such that $a_F < a < b < b_F$. Let Ω be a compact subset of $\Omega_0 = \{x \in \mathbb{R} | v(x) > 0\}$ and $\gamma := \inf_{x\in\Omega} v(x) > 0$.
We introduce some assumptions, gathered below for easy reference needed to state our results.

$(K1)$ K is a positive-valued, bounded probability density, Hölder continuous with exponent $\beta > 0$ and satisfying

$$|u| K(u) \to 0 \quad as \quad \|u\| \to +\infty.$$

$(K2)$ H is a df with $\mathcal{C}^1-$probability density $H^{(1)}$ which is positive, bounded and has compact support. It is also Hölderian with exponent β.

$(K3)$ $i)$ $H^{(1)}$ and K are second-order kernels,
$ii)$ $\int K^2(r)dr < \infty$.

$(M1)$ $\{(X_i, Y_i), \ i \geq 1\}$ is a sequence of stationary α-mixing random variables with coefficient $\alpha(n)$.

$(M2)$ $\{T_i; \ i \geq 1\}$ is a sequence of iid truncating variables independent of $\{(X_i, Y_i), \ i \geq 1\}$ with common continuous df G.

$(M3)$ There exists $\nu > 5 + 1/\beta$ for some $\beta > 1/7$ such that $\forall n$, $\alpha(n) = O(n^{-\nu})$.

$(D1)$ The conditional density $v^*(\cdot)$ is twice continuously differentiable.

$(D2)$ The joint conditional density $v^*(\cdot,\cdot)$ of (X_i, X_j) exists and satisfies

$$\sup_{r,s} |v^*(r,s) - v^*(r)v^*(s)| \leq C < \infty,$$

for some constant C not depending on (i, j).

$(D3)$ The joint conditional density of (X_i, Y_i, X_j, Y_j), denoted by $f^*(\cdot,\cdot,\cdot,\cdot)$, exists and satisfies for any constant C,

$$\sup_{r,s,t,u} |f^*(r,s,t,u) - f^*(r,s)f^*(t,u)| \leq C < \infty.$$

$(D4)$ The joint density $f(\cdot,\cdot)$ is bounded and twice continuously differentiable.

$(D5)$ The marginal density $v(\cdot)$ is locally Lipschitz continuous over Ω_0.

The bandwidth $h_n =: h$ satisfies:

$(H1)$

$$h \downarrow 0, \qquad \frac{\log n}{nh} \to 0 \quad \textit{and} \quad h = o\left(1/\log n\right), \qquad \text{as } n \to \infty,$$

$(H2)$

$$Cn^{\frac{(3-\nu)\beta}{\beta(\nu+1)+4\beta+1}+\eta} < h < C'n^{\frac{1}{1-\nu}},$$

where η satisfies

$$\frac{2}{\beta(\nu+1)+4\beta+1} < \eta < \frac{(\nu-3)\beta}{\beta(\nu+1)+4\beta+1} + \frac{1}{1-\nu}$$

and β and ν are as in $(M3)$.

Remark 3.3.1 *Assumptions* (K) *are quite usual in kernel estimation. Conditions* $(D1)$, $(D4)$ *and* $(D5)$ *are needed in the study of the bias term.* $(D2)$ *and* $(D3)$ *are needed for covariance calculus and take similar forms to those used under mixing. Hypothesis* $(H2)$ *is used in Ould Saïd and Tatachak [57] and is needed to establish Lemma 3.5.1 and Lemma 3.5.4. Assumptions* (M) *concern the mixing processes structure which are standard in such situation. The choice of* β *seems to be surprising, but it is only technical choice which permit us to make one of the variance term to be negligible.*

Remark 3.3.2 *Here we point out that we can not suppose that the original data (that is the* N*-sample) satisfies some kind of dependency. Indeed, we do not know if the observed data are* α*-mixing or are not. And if they are, we do not know the coefficient. Therefore, we suppose that the observed data satisfy some kind of mixing condition.*

Remark 3.3.3 *As we are interested in the number* n *of observations (*N *is unknown), we give asymptotics as* $n \to \infty$ *unless otherwise specified. Since* $n \leq N$, *this implies* $N \to \infty$ *and these results also hold under* $\mathbb{P}-a.s.$ $\quad N \to \infty$.

Our first result, stated in Proposition 3.3.4, is the uniform almost sure convergence with rate of the conditional df estimator defined in (3.8).

Proposition 3.3.4 *Under assumptions* (K), (M), (D) *and* (H), *we have,*

$$\sup_{x\in\Omega}\sup_{a\leq y\leq b}|\mathbf{F}_n(y|x)-\mathbf{F}(y|x)|=O\left(\max\left\{\sqrt{\frac{\log n}{nh}},h^2\right\}\right),\quad \mathbf{P}-a.s.\quad n\to\infty.$$

The second result deals with the strong uniform convergence with rate of the kernel conditional quantile estimator $q_{p,n}(.)$ which is given in the following theorem.

Theorem 3.3.5 *Under the assumptions of Proposition 3.3.4 and for each fixed* $p\in(0,1)$ *if the function* q_p *satisfies for given* $\varepsilon>0$ *there exists* $\beta>0$ *such that*

$$\forall\eta_p:\Omega\to\mathbb{R},\ \sup_{x\in\Omega}\left|q_p(x)-\eta_p(x)\right|\geq\varepsilon\Rightarrow\sup_{x\in\Omega}\left|\mathbf{F}\left(q_p(x)\right)-\mathbf{F}\left(\eta_p(x)\right)\right|\geq\beta, \tag{3.11}$$

we have

$$\lim_{n\to\infty}\sup_{x\in\Omega}|q_{p,n}(x)-q_p(x)|=0,\quad \mathbf{P}-a.s.$$

Furthermore, we have,

$$\sup_{x\in\Omega}|q_{p,n}(x)-q_p(x)|=O\left(\max\left\{\sqrt{\frac{\log n}{nh}},h^2\right\}\right),\qquad \mathbf{P}-a.s.\quad as\quad n\to\infty.$$

3.4 Applications to prediction

It is well known, from the robustness theory that the median is more robust than the mean, therefore the conditional median, $\mu(x)=q_{1/2}(x)$, is a good alternative to the conditional mean as a predictor for a variable Y given $X=x$. Note that the estimation of $\mu(x)$ is given by $\mu_n(x)=q_{\frac{1}{2},n}(x)$. Using this considerations and section 2, we want to predict the non observed rv Y_{n+1} (which corresponds to some modality of our problem), from available data $X_1,\ldots,X_n$. Given a new value X_{n+1}, we can predict the corresponding response Y_{n+1} by

$$\widehat{Y}_{n+1}=\mu_n(X_{n+1})=q_{1/2,n}(X_{n+1}).$$

Nevertheless to say, that the theoretical predictor is given by

$$\mu(X_{n+1})=q_{1/2}(X_{n+1}).$$

Applying the above Theorem, we have the following corollary:

Corollary 3.4.1 *Under the assumptions of Theorem 3.3.5, we have*

$$|q_{p,n}(x)-q_p(x)|=0,\qquad \mathbf{P}-a.s.\quad as\quad n\to\infty.$$

3.5 Proofs

We need some auxiliary results and notations to prove our results. The first lemma gives the uniform convergence with rate of the estimator $v_n^*(x)$ defined in (3.5).

Lemma 3.5.1 *Under* $(K1)$, $(K3)$, (M), $(D1)$, $(D2)$ *and* (H) *we have for*

$$\sup_{x\in\Omega}|v_n^*(x)-v^*(x)| = O\left(\max\left\{\sqrt{\frac{\log n}{nh}}, h^2\right\}\right), \qquad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. We have

$$\begin{aligned}\sup_{x\in\Omega}|v_n^*(x)-v^*(x)| &\le \sup_{x\in\Omega}|v_n^*(x)-\mathbf{E}\left[v_n^*(x)\right]| + \sup_{x\in\Omega}|\mathbf{E}\left[v_n^*(x)\right]-v^*(x)| \\ &=: \mathcal{I}_{1n}+\mathcal{I}_{2n}. \end{aligned} \tag{3.12}$$

We begin by study the variance term $\mathcal{I}_{1n}$. The idea consists in using an exponential inequality taking into account the $\alpha-$mixing structure. The compact set Ω can be covered by a finite number l_n of intervals of length $\omega_n = (n^{-1}h^{1+2\beta})^{\frac{1}{2\beta}}$, where β is the Hölder exponent. Let $I_k := I(x_k, \omega_n)$; $k = 1, ..., l_n$, denote each interval centered at some points x_k. Since Ω is bounded, there exists a constant C such that $\omega_n l_n \le C$. For any x in Ω, there exists I_k which contains x such that $|x-x_k| \le \omega_n$. We start by writing

$$\triangle_i(x) := \frac{1}{nh}\left\{K\left(\frac{x-X_i}{h}\right) - \mathbf{E}\left[K\left(\frac{x-X_1}{h}\right)\right]\right\}.$$

Clearly, we have

$$\begin{aligned}\sum_{i=1}^{n}\triangle_i(x) &= \{(v_n^*(x)-v_n^*(x_k)) - (\mathbf{E}\left[v_n^*(x)\right]-\mathbf{E}\left[v_n^*(x_k)\right])\} \\ &\quad + (v_n^*(x_k)-\mathbf{E}\left[v_n^*(x_k)\right]) \\ &=: \sum_{i=1}^{n}\widetilde{\triangle}_i(x) + \sum_{i=1}^{n}\triangle_i(x_k).\end{aligned}$$

Hence

$$\begin{aligned}\sup_{x\in\Omega}\left|\sum_{i=1}^{n}\triangle_i(x)\right| &\le \max_{1\le k\le l_n}\sup_{x\in I_k}\left|\sum_{i=1}^{n}\widetilde{\triangle}_i(x)\right| + \max_{1\le k\le l_n}\left|\sum_{i=1}^{n}\triangle_i(x_k)\right| \\ &=: S_{1n}+S_{2n}.\end{aligned} \tag{3.13}$$

First, we have under assumption $(K1)$,

$$\begin{aligned}\sup_{x\in I_k}\left|\sum_{i=1}^{n}\widetilde{\triangle}_i(x)\right| &\leq \frac{1}{nh}\sum_{i=1}^{n}\left|K\left(\frac{x-X_i}{h}\right)-K\left(\frac{x_k-X_i}{h}\right)\right| \\ &+\frac{1}{h}\mathbf{E}\left[\left|K\left(\frac{x-X_1}{h}\right)-K\left(\frac{x_k-X_1}{h}\right)\right|\right] \\ &\leq \frac{2\sup_{x\in I_k}|x-x_k|^{\beta}}{h^{1+\beta}} \\ &\leq C\omega_n^{\beta}h^{-1-\beta}=O\left((nh)^{-1/2}\right).\end{aligned}$$

Hence, by $(H1)$ and for n large enough, we get $S_1=o_{\mathbf{P}}(1)$.
We now turn to the term S_{2n} in (3.13). Under $(K1)$, the rv's $U_i = nh\triangle_i(x_k)$ are centered and bounded. The use of the well known Fuk-Nagaev's inequality (see Rio [66, formula 6.19b, page 87]) slightly modified in Ferraty and Vieu [22, proposition A.11-ii), page 237], allows one to get, for all $\varepsilon>0$ and $r>1$

$$\begin{aligned}\mathbf{P}\left\{\max_{1\leq k\leq l_n}\left|\sum_{i=1}^{n}\triangle_i(x_k)\right|>\varepsilon\right\} &\leq \sum_{k=1}^{l_n}\mathbf{P}\left\{\left|\sum_{i=1}^{n}\triangle_i(x_k)\right|>\varepsilon\right\} \\ &\leq C\omega_n^{-1}\left\{\frac{n}{r}\left(\frac{r}{\varepsilon nh}\right)^{\nu+1}+\left(1+\frac{\varepsilon^2n^2h^2}{rs_n^2}\right)^{-\frac{r}{2}}\right\} \\ &=: \mathcal{I}_{11n}+\mathcal{I}_{12n} \qquad (3.14)\end{aligned}$$

where

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

Putting

$$r=(\log n)^{1+\delta},\ \text{where } \delta>0, \quad \text{and } \varepsilon=\varepsilon_0\sqrt{\frac{\log n}{nh}},\ \text{for some } \varepsilon_0>0. \qquad (3.15)$$

We have

$$\begin{aligned}Q_{1n}&=C(n^{-1}h^{1+2\beta})^{\frac{-1}{2\beta}}\frac{n}{(\log n)^{1+\delta}}\left(\frac{(\log n)^{1+\delta}}{\varepsilon_0\sqrt{nh\log n}}\right)^{\nu+1} \\ &=n^{1-\frac{\nu+1}{2}+\frac{1}{2\beta}}h^{-\left(\frac{1}{2\beta}+1+\frac{\nu+1}{2}\right)}(\log n)^{\upsilon(1+\delta)-\frac{\upsilon+1}{2}}\varepsilon_0^{-(\nu+1)}.\end{aligned}$$

Note that under $(M3)$, it is easy to see the following modified assumption $(H'2)$ of $(H2)$ hold,

$$Cn^{\frac{(3-\nu)\beta}{\beta(\nu+1)+2\beta+1}+\eta}<h<C'n^{\frac{1}{1-\nu}}, \qquad (3.16)$$

where η satisfies

$$\frac{1}{\beta(\nu+1)+2\beta+1} < \eta < \frac{(\nu-3)\beta}{\beta(\nu+1)+2\beta+1} + \frac{1}{1-\nu} \tag{3.17}$$

β and ν are as in $(M3)$.
Then, from the left-hand side of (3.16)

$$\mathcal{I}_{11n} \leq C'(\log n)^{\nu(1+\delta)-\frac{v+1}{2}} n^{-1-\frac{\eta}{2\beta}(\beta(\nu+1)+2\beta+1-\frac{1}{\eta})}.$$

Hence, for any η as in (3.17), $\mathcal{I}_{11n}$ is bounded by the term of a finite-sum series. Before we focus on $\mathcal{I}_{11n}$, we have to study the asymptotic behavior of

$$\begin{aligned} s_n^2 &= \sum_{i=1}^{n} Var(U_i) + \sum_{i\neq j} |Cov(U_i, U_j)| \\ &=: s_n^{var} + s_n^{cov}. \end{aligned}$$

First, by $(K1;1)$, $(D1)$ and a change of variable, we obtain

$$\begin{aligned} s_n^{var} &= nVar(U_1) \\ &= n\left\{\mathbf{E}K^2\left(\frac{x_k - X_1}{h}\right) - \mathbf{E}^2\left[K\left(\frac{x_k - X_1}{h}\right)\right]\right\} \\ &= O\left(nh\right). \end{aligned} \tag{3.18}$$

On the other hand, a change of variable, $(K1)$, $(M1)$ and $(D2)$ lead to

$$\begin{aligned} |Cov(U_i, U_j)| &= |\mathbf{E}\left[U_i U_j\right]| \\ &\leq \iint K\left(\frac{x_k - r}{h}\right) K\left(\frac{x_k - s}{h}\right) |v^*(r,s) - v^*(r)v^*(s)|\, drds \\ &= O\left(h^2\right). \end{aligned} \tag{3.19}$$

Note also that, these covariances can be controlled by means of the usual Davydov covariance inequality for mixing processes (see Rio [66, formula 1.12a, page 10]; or Bosq [7, formula 1.11, page 22]). We have

$$\forall i \neq j, \quad |Cov(U_i, U_j)| \leq C\alpha\left(|i-j|\right). \tag{3.20}$$

To evaluate s_n^{cov}, we use the technique developed by Masry [49]. Taking $\varphi_n = \left\lceil (n^{-1}h)^{-1/\nu} \right\rceil$ (where $\lceil . \rceil$ denotes the smallest integer greater than the argument), we can write

$$s_n^{cov} = \sum_{0<|i-j|\leq\varphi_n} |Cov(U_i, U_j)| + \sum_{|i-j|>\varphi_n} |Cov(U_i, U_j)|. \tag{3.21}$$

First, applying the upper bound (3.19) to the first covariance term in (3.21), we get

$$\sum_{0<|i-j|\leq\varphi_n} |Cov(U_i, U_j)| \leq Cnh^2\varphi_n. \tag{3.22}$$

For the second term, thanks to (3.20) we get

$$\begin{aligned}\sum_{|i-j|>\varphi_n} |Cov(U_i, U_j)| &\leq C \sum_{|i-j|>\varphi_n} \alpha\left(|i-j|\right) \\ &\leq Cn^2\alpha\left(\varphi_n\right). \end{aligned} \tag{3.23}$$

According to the right-hand side of $(H'2)$, using $(M3)$, (3.22) and (3.23), we get

$$s_n^{cov} = O(nh). \tag{3.24}$$

Finally, (3.18) and (3.24) lead directly to $s_n^2 = O\left(nh\right).$
This is enough to study the quantity $\mathcal{I}_{12n}$, since for ε and r as in (3.15) and Taylor expansion of $\log(1+x)$ allows us to write that

$$\begin{aligned}\mathcal{I}_{12n} &= C\omega_n^{-1}\exp\left[-\frac{r}{2}\log\left(1+\frac{\varepsilon_0^2 nh\log n}{rs_n^2}\right)\right] \\ &\leq Cn^{\frac{1}{2\beta}-C'\varepsilon_0^2}h^{-(1+\frac{1}{2\beta})} \\ &= Cn^{\frac{1}{2\beta}-C'\varepsilon_0^2}h^{-\frac{1}{2\beta}(\beta(\nu+1)+2\beta+1)}h^{\frac{v+1}{2}}.\end{aligned}$$

By using $(H'2)$ and $(M3)$, the later can be made as a general term of a convergent series. Hence $\sum_{n\geq1}\left(\mathcal{I}_{11n}+\mathcal{I}_{12n}\right)<\infty$, and therefore by Borel-Cantelli's Lemma, we have

$$\mathcal{I}_{1n} = O\left(\sqrt{\frac{\log n}{nh}}\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty$$

On the other hand, the bias term $\mathcal{I}_{2n}$ does not depend on the mixing structure. We prove its convergence by using a change of variable and a Taylor expansion (see Lemdani *et al.* [46, Lemma 6.1]). We get, under $(K3)$ and $(D1)$

$$\mathcal{I}_{2n} = O\left(h^2\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty$$

Hence, replacing $\mathcal{I}_{1n}$ and $\mathcal{I}_{2n}$ in (3.12), we get the result. □
The following Lemma is Lemma 5.2 in Ould Saïd and Tatachak [57], in which they state a rate of convergence for μ_n under α-mixing hypothesis, which is interesting in itself, similar to that established in the iid case by He and Yang [34].

Lemma 3.5.2 *Under assumptions* (M)*, we have*

$$\sup_{x\in\Omega}|\mu_n-\mu| = O\left(\sqrt{\frac{\log\log n}{n}}\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. See Lemma 5.2 in Ould Saïd and Tatachak [57]. □

Adapting (3.9), define

$$\tilde{F}_{1,n}(x,y) := \frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{x-X_i}{h}\right)H\left(\frac{y-Y_i}{h}\right). \tag{3.25}$$

Lemma 3.5.3 *Under the assumptions of Lemma 3.5.1 and* $(K2)$*, we have*

$$\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{F}_{1,n}(x,y)-\tilde{F}_{1,n}(x,y)\right| = O\left(\sqrt{\frac{\log\log n}{n}}\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. Under $(K2)$, the df H is bounded by 1. Hence

$$\left|\mathbf{F}_{1,n}(x,y)-\tilde{F}_{1,n}(x,y)\right| \le \left\{\frac{|\mu_n-\mu|}{G_n(a_F)}+\frac{\mu\sup_{y\ge a_F}|G_n(y)-G(y)|}{G_n(a_F)G(a_F)}\right\}|v_n^*(x)|.$$

From Lemma *3.5.2*, we have

$$|\mu_n-\mu| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \mathbf{P}-a.s. \quad n\to\infty.$$

Moreover, $G_n(a_F) \overset{\mathbf{P}-a.s.}{\to} G(a_F) > 0$. In the same way and using Lemma 3.4 in Liang *et al.* [42] (see Lemma 2.2.1, Chapter 2) we get

$$\sup_{y\ge a_F}|G_n(y)-G(y)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Combining these last results with Lemma *3.5.1*, we achieve the proof. □

Lemma 3.5.4 *Under assumptions* (K)*,* (M)*,* $(D3)$*,* $(D4)$*, and* (H)*, we have,*

$$\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\tilde{F}_{1,n}(x,y)-\mathbf{E}\left[\tilde{F}_{1,n}(x,y)\right]\right| = O\left(\sqrt{\frac{\log n}{nh}}\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. The proof is analogous to Lemma *3.5.1*, we give only the leading lines. As Ω and $[a,\ b]$ are compact sets, then they can be covered by finite numbers l_n and d_n of intervals $I_1,...,I_{l_n}$ and $J_1,...,J_{d_n}$ of length ω_n as in Lemma *3.5.1* and

$\lambda_n = (n^{-1}h^{2\beta})^{\frac{1}{2\beta}}$ and centers $x_1, ..., x_{l_n}$ and $y_1, ..., y_{d_n}$ respectively. Since Ω and $[a,\ b]$ are bounded, there exist two constant C_1 and C_2 such that $l_n\omega_n \leq C_1$ and $d_n\lambda_n \leq C_2$. Hence for any $(x,y) \in \Omega \times [a,b]$ there exists x_k and y_j such that $||x - x_k|| \leq \omega_n$ and $|y - y_j| \leq \lambda_n$. Thus we have the following decomposition

$$\begin{aligned}
&\sup_{x\in\Omega}\sup_{y\in[a,\ b]}\left|\tilde{F}_{1,n}(x,y) - \mathbb{E}\left[\tilde{F}_{1,n}(x,y)\right]\right| \\
&\leq \max_{1\leq k\leq l_n}\sup_{x\in I_k}\sup_{y}\left|\tilde{F}_{1,n}(x,y) - \tilde{F}_{1,n}(x_k,y)\right| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}\sup_{y\in J_j}\left|\tilde{F}_{1,n}(x_k,y) - \tilde{F}_{1,n}(x_k,y_j)\right| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}\left|\tilde{F}_{1,n}(x_k,y_j) - \mathbb{E}\left[\tilde{F}_{1,n}(x_k,y_j)\right]\right| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}\sup_{y\in J_k}\left|\mathbb{E}\left[\tilde{F}_{1,n}(x_k,y_j)\right] - \mathbb{E}\left[\tilde{F}_{1,n}(x_k,y)\right]\right| \\
&+ \max_{1\leq k\leq l_n}\sup_{x\in I_k}\sup_{y}\left|\mathbb{E}\left[\tilde{F}_{1,n}(x_k,y)\right] - \mathbb{E}\left[\tilde{F}_{1,n}(x,y)\right]\right| \\
&=: \mathcal{J}_{1n} + \mathcal{J}_{2n} + \mathcal{J}_{3n} + \mathcal{J}_{4n} + \mathcal{J}_{5n}.
\end{aligned}$$

Concerning $\mathcal{J}_{1n}$ and $\mathcal{J}_{5n}$, assumptions $(K1)$, $(K2)$ yield

$$\begin{aligned}
\sup_{x\in I_k}\sup_{y}\left|\tilde{F}_{1,n}(x,y) - \tilde{F}_{1,n}(x_k,y)\right| &\leq \frac{C\mu\omega_n^\beta}{G(a_F)h^{1+\beta}}\sup_{y}\left|H\left(\frac{y-Y_i}{h}\right)\right| \\
&= O\left((nh)^{-1/2}\right).
\end{aligned}$$

Hence, by $(H1)$ we get

$$\sqrt{\frac{nh}{\log n}}\sup_{x\in\Omega}\sup_{y\in[a,\ b]}\left|\tilde{F}_{1,n}(x,y) - \tilde{F}_{1,n}(x_k,y)\right| = o(1). \tag{3.26}$$

Similarly, we obtain for $\mathcal{J}_{2n}$ and $\mathcal{J}_{4n}$

$$\begin{aligned}
\sup_{y\in J_j}\left|\tilde{F}_{1,n}(x_k,y) - \tilde{F}_{1,n}(x_{k,y_j})\right| &\leq \frac{C\mu\lambda_n^\beta}{G(a_F)h^{1+\beta}}\left|K\left(\frac{x-X_i}{h}\right)\right| \\
&= O\left((nh^2)^{-1/2}\right).
\end{aligned}$$

Again, by $(H1)$ we get

$$\sqrt{\frac{nh}{\log n}}\sup_{x\in\Omega}\sup_{y\in[a,\ b]}\left|\tilde{F}_{1,n}(x_k,y) - \tilde{F}_{1,n}(x_k,y_j)\right| = o(1). \tag{3.27}$$

As to $\mathcal{J}_3$, for all $\varepsilon > 0$ we have

$$\mathbb{P}\left\{\max_{1\le k\le l_n}\max_{1\le j\le d_n}\left|\tilde{F}_{1,n}(x_k,y_j)-\mathbb{E}\left[\tilde{F}_{1,n}(x_k,y_j)\right]\right|>\varepsilon\right\}$$
$$\le l_n\, d_n\mathbb{P}\left\{\left|\tilde{F}_{1,n}(x_k,y_j)-\mathbb{E}\left[\tilde{F}_{1,n}(x_k,y_j)\right]\right|>\varepsilon\right\}. \tag{3.28}$$

Set, for any $i \ge 1$

$$\Psi_i(x_k,y_j)=\frac{\mu}{nh}\left\{\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_j-Y_i}{h}\right)\right.$$
$$\left.-\ \mathbb{E}\left[\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_j-Y_i}{h}\right)\right]\right\}.$$

Under $(K1)$ and $(K2)$, the rv's $V_i = nh\Psi_i(x_k,y_j)$ are centered and bounded by $\frac{2\mu M_0 M_1}{G(a_F)} =: C < \infty$. Then, applying again Fuck-Nagaev' inequality, we obtain for all $\epsilon > 0$ and $r > 1$,

$$\mathbb{P}\Big\{\max_{1\le k\le l_n}\max_{1\le j\le d_n}\left|\sum_{i=1}^{n}\Psi_i(x_k,y_j)\right|>\epsilon\Big\}$$
$$=\mathbb{P}\Big\{\max_{1\le k\le l_n}\max_{1\le j\le d_n}\big|\sum_{i=1}^{n}V_i\big|>nh\epsilon\Big\}$$
$$\le C_1C_2\left(\omega_n\lambda_n\right)^{-1}\left\{\frac{n}{r}\left(\frac{2r}{\epsilon nh}\right)^{1+\nu}+\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{-\frac{r}{2}}\right\}$$
$$\le Cn^{\frac{1}{\beta}}h^{-\left(\frac{1}{2\beta}+2\right)}\frac{n}{r}\left(\frac{r}{\epsilon nh}\right)^{1+\nu}+Cn^{\frac{1}{\beta}}h^{-\left(\frac{1}{2\beta}+2\right)}\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{-\frac{r}{2}}$$
$$=:\mathcal{J}_{31n}+\mathcal{J}_{32n}, \tag{3.29}$$

where

$$s_n^2=\sum_{1\le i\le n}\ \sum_{1\le j\le n}\left|Cov(U_i,U_j)\right|.$$

By taking ε and r as in (3.15), we get

$$\mathcal{J}_{31n}=C\epsilon_0^{-(1+\nu)}n^{1+\frac{1}{\beta}-\frac{1+\nu}{2}}(\log n)^{v(1+\delta)-\frac{\nu+1}{2}}h^{-\frac{1}{2\beta}(1+4\beta+\beta(1+\nu))}.$$

Then, Using $(H1)$ and $(H2)$ we get

$$\mathcal{J}_{31n}\le C\epsilon_0^{-(1+\nu)}n^{-1-\frac{\eta}{2\beta}\left(1+4\beta+\beta(1+\nu)-\frac{2}{\eta}\right)}(\log n)^{v(1+\delta)-\frac{\nu+1}{2}}.$$

Hence, the condition upon β and for any η as in $(H2)$, $\mathcal{J}_{31n}$ is the general term of a finite-sum series.

Let us now examine the term $\mathcal{J}_{32n}$. First, we have to calculate

$$s_n^2 = nVar(V_1) + \sum\nolimits_{i \neq l} \left| cov(V_i, V_l) \right|.$$

We have

$$\begin{aligned} Var(V_1) &= \mathbb{E}\left[\frac{\mu^2}{G^2(Y_1)} K^2\left(\frac{x_k - X_1}{h}\right) H^2\left(\frac{y_j - Y_1}{h}\right)\right] \\ &- \mathbb{E}^2\left[\frac{\mu}{G(Y_1)} K\left(\frac{x_k - X_1}{h}\right) H\left(\frac{y_j - Y_1}{h}\right)\right] \\ &=: \mathcal{V}_1 - \mathcal{V}_2. \end{aligned}$$

Remark that

$$\begin{aligned} \mathbb{E}\left[\frac{\mu^2}{G^2(Y_1)} H^2\left(\frac{y_j - Y_1}{h}\right) \Big| X_1\right] &= \int \frac{\mu^2}{G^2(y_1)} H^2\left(\frac{y_j - y_1}{h}\right) f^*(y_1 \big| X_1) dy_1 \\ &= \int \frac{\mu}{G(y_1)} H^2\left(\frac{y_j - y_1}{h}\right) f(y_1 \big| X_1) dy_1 \\ &= \mathbb{E}\left[\frac{\mu}{G(Y_1)} H^2\left(\frac{y_j - Y_1}{h}\right)\right]. \end{aligned}$$

Then

$$\begin{aligned} \mathcal{V}_1 &= \mathbb{E}\left[\frac{\mu}{G(Y_1)} K^2\left(\frac{x_k - X_1}{h}\right) H^2\left(\frac{y_j - Y_1}{h}\right)\right] \\ &\leq \frac{\mu}{G(a_F)} \mathbb{E}\left[K^2\left(\frac{x_k - X_1}{h}\right)\right] \\ &\leq \frac{\mu h}{G(a_F)} \int K^2(r) v^*(x_k - rh) dz. \end{aligned}$$

Under $(K3; ii)$ and $(D1)$, we have $\mathcal{V}_1 = O(h)$. An analogous development gives that $\mathcal{V}_2 = O(h^2)$, which implies $Var(\Delta_1) = O(h)$.

On the one hand, $(M1)$, $(K1)$ and $(K2)$ lead to

$$\begin{aligned}\Big|cov\left(V_i,V_j\right)\Big| &= \Big|\int\int\int\int\int \frac{\mu^2}{G(r)G(t)}K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_j-r}{h}\right)\\
&\times K\left(\frac{x_k-s}{h}\right)H\left(\frac{y_j-t}{h}\right)f^*_{1,i,j+1,j+1}(u,r,s,t)dudrdsdt\\
&-\int\int\frac{\mu}{G(r)}K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_j-r}{h}\right)f^*(u,r)dudr\\
&\times\int\int\frac{\mu}{G(t)}K\left(\frac{x_k-s}{h}\right)H\left(\frac{y_j-t}{h}\right)f^*(s,t)dsdt\Big|\\
&\le\frac{\mu^2}{G^2(a_F)}\int\int\int\int\Big|K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_j-r}{h}\right)K\left(\frac{x_k-s}{h}\right)\\
&\times H\left(\frac{y_j-t}{h}\right)\left(f^*_{1,i,j+1,j+1}(u,r,s,t)dudrdsdt-f^*(u,r)dudrf^*(s,t)dsdt\right)\Big|.\end{aligned}$$

Using assumption $(D3)$ and by change of variable, it follows that

$$\Big|cov\left(V_i,V_j\right)\Big| = O\left(h^4\right). \tag{3.30}$$

On the other hand, from a result in Bosq [7, p. 22], we have

$$\Big|cov\left(V_i,V_j\right)\Big| = O\left(\alpha(|i-j|)\right). \tag{3.31}$$

Then to evaluate $\sum_{i\neq l}\big|cov(V_i,V_j)\big|$, the idea is to introduce a sequence of integers φ_n the same as in Lemma 3.5.1, and using (3.30) for the nearest and (3.31) for the farest integer i and j. Then we get

$$\begin{aligned}\sum_{i\neq j}\big|cov(V_i,V_j)\big| &= \sum\sum_{0<|i-j|\le\varphi_n}\big|cov(V_i,V_j)\big|\\
&+\sum\sum_{|i-j|>\varphi_n}\big|cov(V_i,V_j)\big|\\
&\le\sum\sum_{0<|i-j|\le\varphi_n}h^4+\sum\sum_{|i-j|>\varphi_n}\alpha(|i-j|)\\
&\le Cn\varphi_n h^4+Cn^2\alpha\left(\varphi_n\right).\end{aligned}$$

The right-hand side of $(H2)$ and $(M3)$, one has

$$\sum_{i\neq j}\big|cov(V_i,V_j)\big| = O\left(nh\right).$$

So $s_n^2 = O\left(nh\right)$.
Consequently, by taking ϵ and r as in (3.15) and using Taylor expansion of

$\log(1+x)$, the term $\mathcal{J}_{32n}$ becomes

$$\begin{aligned}\mathcal{J}_{32n} &\leq Cn^{\frac{1}{\beta}}h^{-(2+\frac{1}{2\beta})}\exp\left\{-\frac{1}{2}\epsilon_0^2\log n\right\}\\ &= Cn^{\frac{1}{\beta}-C\epsilon_0^2}h^{-(2+\frac{1}{2\beta})}.\end{aligned}$$

By using $(H2)$ and $(M3)$, the later can be made as a general term of summable series. Thus $\sum_{n\geq 1}(\mathcal{J}_{31n}+\mathcal{J}_{32n}) < \infty$. Then by Borel-Cantelli's Lemma, the first term of (3.29) goes to zero a.s. and for n large enough, we have $\mathcal{J}_{3n} = O\left(\sqrt{\dfrac{\log n}{nh}}\right)$, this completes the proof of the Lemma. □

Lemma 3.5.5 *Under assumptions* $(K3)$ *and* $(D4)$*, we have*

$$\sup_{x\in\Omega}\sup_{a\leq y\leq b}\left|\mathbf{E}\left[\tilde{F}_{1,n}(x,y)\right]-\mathbf{F}_1(x,y)\right| = O\left(h^2\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. The bias terms do not depend on the mixing structure. The proof of Lemma 3.5.5 is similar to that of Lemma 6.2 in Lemdani *et al.* [46], hence its proof is omitted. □

The next Lemma gives the uniform convergence with rate of the estimator $v_n(x)$ defined in (3.7).

Lemma 3.5.6 *Under the assumptions of Lemma 3.5.1 and condition* $(D5)$*, we have*

$$\sup_{x\in\Omega}|v_n(x)-v(x)| = O\left(\max\left\{\sqrt{\frac{\log n}{nh}},h^2\right\}\right), \quad \mathbf{P}-a.s. \quad as \quad n\to\infty.$$

Proof. Adapting (3.7), define

$$\tilde{v}_n(x) = \frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{x-X_i}{h}\right). \tag{3.32}$$

We have

$$\begin{aligned}\sup_{x\in\Omega}|v_n(x)-v(x)| &\leq \sup_{x\in\Omega}|v_n(x)-\tilde{v}_n(x)|\\ &+\sup_{x\in\Omega}|\tilde{v}_n(x)-\mathbf{E}\left[\tilde{v}_n(x)\right]|\\ &+\sup_{x\in\Omega}|\mathbf{E}\left[\tilde{v}_n(x)\right]-v(x)|\\ &=: \mathcal{L}_{1n}+\mathcal{L}_{2n}+\mathcal{L}_{3n}.\end{aligned}$$

For the first term, using analogous framework as in Lemma *3.5.3*, we get

$$\mathcal{L}_{1n} = O\left(\sqrt{\frac{\log\log n}{n}}\right), \quad \mathbf{P}-a.s. \quad as\ n \to \infty. \tag{3.33}$$

In addition, by using the same approach as for $\mathcal{I}_{1n}$ in the proof of Lemma *3.5.1*, we can show that, for n large enough

$$\mathcal{L}_{2n} = O\left(\sqrt{\frac{\log n}{nh}}\right), \quad \mathbf{P}-a.s. \quad as\ n \to \infty. \tag{3.34}$$

Finally, a change of variable and a Taylor expansion, we get, under $(K3)$ and $(D5)$

$$\begin{aligned}
\mathbf{E}\left[\tilde{v}_n(x)\right] - v(x) &= \mathbf{E}\left[\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_1)}K\left(\frac{x-X_1}{h}\right)\right] - v(x)\\
&= \frac{1}{h}\int K\left(\frac{x-u}{h}\right)v(u)du - v(x)\\
&= \frac{h^2}{2}\int r^2 K\left(r\right)v''(\tilde{x})dr
\end{aligned}$$

with $\tilde{x} \in [x - rh, x]$, which yields that

$$\mathcal{L}_{3n} = O\left(h^2\right), \quad \mathbf{P}-a.s. \quad as\ n \to \infty. \tag{3.35}$$

Combining (3.33), (3.34) and (3.35) permit to conclude the proof. □

Proof of Proposition 3.3.4 In view of (3.8), we have the following classical decomposition

$$\begin{aligned}
&\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{F}_n(y|x) - \mathbf{F}(y|x)\right|\\
&\qquad \le \frac{1}{\beta - \sup_{x\in\Omega}|v_n(x) - v(x)|}\left\{\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{F}_{1,n}(x,y) - \mathbf{F}_1(x,y)\right|\right.\\
&\qquad\qquad \left. + \gamma^{-1}\sup_{x\in\Omega}\sup_{a\le y\le b}|F(y|x)|\sup_{x\in\Omega}|v_n(x) - v(x)|\right\}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{F}_{1,n}(x,y) - \mathbf{F}_1(x,y)\right| \le &\sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{F}_{1,n}(x,y) - \tilde{F}_{1,n}(x,y)\right|\\
&+ \sup_{x\in\Omega}\sup_{a\le y\le b}\left|\tilde{F}_{1,n}(x,y) - \mathbf{E}\left[\tilde{F}_{1,n}(x,y)\right]\right|\\
&+ \sup_{x\in\Omega}\sup_{a\le y\le b}\left|\mathbf{E}\left[\tilde{F}_{1,n}(x,y)\right] - \mathbf{F}_1(x,y)\right|.
\end{aligned}$$

In conjunction with Lemmas 3.5.1–3.5.6, we conclude the proof. $\square$

We now embark on the proof of Theorem 3.3.5.

Proof of Theorem *3.3.5* Let $x \in \Omega$. As $\mathbf{F}_n(\cdot|x)$ and $\mathbf{F}(\cdot|x)$ are continuous, we have

$\mathbf{F}(q_p(x)|x) = \mathbf{F}_n(q_{p,n}(x)|x) = p$. Then

$$\begin{aligned} |\mathbf{F}(q_{p,n}(x)|x) - \mathbf{F}(q_p(x)|x)| &\leq |\mathbf{F}(q_{p,n}(x)|x) - \mathbf{F}_n(q_{p,n}(x)|x)| \\ &\quad + |\mathbf{F}_n(q_{p,n}(x)|x) - \mathbf{F}(q_p(x)|x)| \qquad (3.36) \\ &\leq |\mathbf{F}(q_{p,n}(x)|x) - \mathbf{F}_n(q_{p,n}(x)|x)| \\ &\leq \sup_{a \leq y \leq b} |\mathbf{F}_n(y|x) - \mathbf{F}(y|x)| \,. \qquad (3.37) \end{aligned}$$

The consistency of $q_{p,n}(x)$ follows then immediately from Proposition 3.3.4 in conjunction with the inequality

$$\sum_n \left\{ \sup_{x \in \Omega} |q_{p,n}(x) - q_p(x)| \geq \varepsilon \right\} \leq \sum_n \left\{ \sup_{x \in \Omega} \sup_{a \leq y \leq b} |\mathbf{F}_n(y|x) - \mathbf{F}(y|x)| \geq \beta \right\}.$$

For the second part, a Taylor expansion of $F(\cdot|\cdot)$ in neighborhood of q_p, implies that

$$\mathbf{F}(q_{p,n}(x)|x) - \mathbf{F}(q_p(x)|x) = (q_{p,n}(x) - q_p(x))\, f\left(\tilde{q}_p(x)|x\right) \qquad (3.38)$$

where $\tilde{q}_p$ is between q_p and $q_{p,n}$ and $f(\cdot|x)$ is the conditional density of Y given $X = x$. Then, from the behavior of $\mathbf{F}(q_{p,n}(x)|x) - \mathbf{F}(q_p(x)|x)$ as n goes to infinity, it is easy to obtain asymptotic results for the sequence $(q_{p,n}(x) - q_p(x))$. By (3.38) we have

$$\sup_{x \in \Omega} |q_{p,n}(x) - q_p(x)| \left| f\left(\tilde{q}_p(x)|x\right) \right| \leq \sup_{x \in \Omega} \sup_{a \leq y \leq b} |\mathbf{F}_n(y|x) - \mathbf{F}(y|x)| \,.$$

The result follows from $(D4)$ and the Proposition 3.3.4. Here we point out that, if $f\left(\tilde{q}_p(x)|x\right) = 0$, for some $x \in \Omega$, we should increase the order of Taylor expansion to obtain the consistency of $q_{p,n}(x)$ (with an adapted rate). $\square$

Chapter 4

Asymptotic normality of a kernel conditional quantile estimator under strong mixing hypothesis and left-truncation

Elias Ould Saïd, Djabrane Yahia

Abstract: In this chapter, we consider the estimation of the conditional quantile when the interest variable is subject to left truncation. It is shown that, under regularity conditions, kernel estimate of the conditional quantile is asymptotically normally distributed, when the data exhibit some kind of dependence. We use asymptotic normality to construct confidence bands for predictors based on the kernel estimate of the conditional median.

Subject classifications: Primary 62G20; Secondary 62G08, 62E20
Keywords: Asymptotic normality; Conditional quantile; Kernel estimate; Strong mixing; Truncated data.

---

[2]This chapter corresponds to the paper which is to appear in *Communications in Statistics, Theory & Method* (2010).

4.1 Introduction

Let $\mathcal{Y}$ and $\mathcal{T}$ be two real random variables (rv) with unknown cumulative distribution functions (df) F and G respectively, both assumed to be continuous. Let $\mathcal{X}-3.4mm\mathcal{X}$ be a random vector of covariates taking its values in $\mathbb{R}^d$ with df V and continuous density v. When no truncation is present, we could think of the observations as $(\boldsymbol{\mathcal{X}}_j, \mathcal{Y}_j, \mathcal{T}_j)\,;\ 1 \le j \le N$, where the sample size N is deterministic, but unknown. Under random left-truncation (RLT), the rv of interest $\mathcal{Y}$ is interfered by the truncation rv $\mathcal{T}$, in such a way that $\mathcal{Y}$ and $\mathcal{T}$ are observed only if $\mathcal{Y} \ge \mathcal{T}$. Therefore, for notational convenience, we shall denote $(\mathbf{X}_i, Y_i, T_i)\,;\ 1 \le i \le n$, $(n \le N)$ the observed subsequence that is $Y_i \ge T_i$ from the $N-$sample. As a consequence of truncation, the size of the actually observed sample, n, is a binomial rv with parameters N and $\mu := \mathbb{P}(\mathcal{Y} \ge \mathcal{T}) > 0$. Such data occur in astronomy and economics (see Woodroofe [86], Feigelson and Babu [20] and also in epidemiology and biometry (see, e.g., He and Yang [33]).

Consider the joint df $\mathbf{F}(\cdot,\cdot)$ of the random vector $(\boldsymbol{\mathcal{X}},\, \mathcal{Y})$ related to the $N-$sample and suppose it is of class $\mathcal{C}^1$. The conditional df of $\mathcal{Y}$ given $\boldsymbol{\mathcal{X}} = \mathbf{x} =: (x_1, ..., x_d)^t$, that is $\mathbf{F}(y|\mathbf{x}) = \mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y} \le y\}}|\boldsymbol{\mathcal{X}} = \mathbf{x}\right]$ which may be rewritten into

$$\mathbf{F}(\cdot|\mathbf{x}) = \frac{\mathbf{F}_1(\mathbf{x},\cdot)}{v(\mathbf{x})}, \quad \mathbf{F}_1(\mathbf{x},\cdot) = \frac{\partial \mathbf{F}(\mathbf{x},\cdot)}{\partial \mathbf{x}} := \frac{\partial^d \mathbf{F}(\mathbf{x},\cdot)}{\partial x_1 ... \partial x_d}. \tag{4.1}$$

For all fixed $p \in (0,\, 1)$, the p^{th} conditional quantile of $\mathbf{F}$ given $\boldsymbol{\mathcal{X}} = \mathbf{x}$ is defined by

$$q_p(\mathbf{x}) := \inf\left\{y \in \mathbb{R} : \mathbf{F}(y|\mathbf{x}) \ge p\right\}.$$

It is well known that the conditional quantiles can give a good description of the data (see, Chaudhuri *et al.* [13]), such as robustness to heavy-tailed error distributions and outliers, especially the conditional median function $q_{1/2}(\mathbf{x})$ for asymmetric distributions, which can provide a useful alternative to the ordinary regression based on the mean.

In the RLT model, Gürler *et al.* [30] establish a Bahadur-type representation for the quantile function and asymptotic normality. Its extension to time series analysis have been obtained by Lemdani *et al.* [45]. Ould Saïd and Lemdani [55] study a nonparametric regression function estimator with RLT data. In the same way, Lemdani *et al.* [46] introduce a kernel conditional quantile estimator and prove its almost sure (a.s.) consistency and asymptotic normality in the iid case.

Under strong mixing hypotheses, the strong uniform convergence with rates of the kernel conditional quantile and that of the conditional df is established by Ould Saïd *et al.* [58]. In this chapter, our purpose is to study the asymptotic normality of the kernel conditional quantile estimator with RLT data. Although our interest in conditional quantile estimation is motivated by the forecasting from time series data, our results are derived where the observations exhib some kind of dependence.

First, let $\mathcal{F}_i^k(Z)$ denotes the σ-field of events generated by $\{Z_j,\ i \leq j \leq k\}$. For easy reference, let us recall the following definition.

Definition 4.1.1 *Let $\{Z_i,\ i \geq 1\}$ denotes a sequence of random variables. Given a positive integer n, set:*

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

The sequence is said to be $\alpha-$mixing (strongly mixing) if the mixing coefficient $\alpha(n) \to 0$.

Strong-mixing condition is reasonably weak and has many practical applications (see, e.g. Doukhan[18], Cai [10, 11] and Dedecker *et al.*[16] for more details). In particular, Masry and Tójstheim [50] proved that, both ARCH processes and nonlinear additive autoregressive models with exogenous variables, which are particularly popular in finance and econometrics, are stationary and $\alpha-$mixing.

The rest of the chapter is organized as follows. In Section 2, we recall a definition of the kernel conditional quantile estimator in the RLT model, the assumptions and our main results. In Section 3, we derive from our results the asymptotic normality of a predictor and propose a confidence bands for the conditional quantile function. Finally, the proofs of the main results are postponed to Section 4 with some auxiliary results and their proofs.

4.2 The model, the assumptions and the main results

In the sequel, the notation $(\mathbf{X} \leq \mathbf{x})$ stands for $(X_1 \leq x_1, ..., X_d \leq x_d)$. Note also that, since N is unknown and n is known (although random), our results will not be stated with respect to the probability measure $\mathbb{P}$ (related to the $N-$sample) but will involve the conditional probability $\mathbf{P}$ (related to the $n-$sample). Also $\mathbb{E}$ and $\mathbf{E}$ will denote the expectation operators related to $\mathbb{P}$ and $\mathbf{P}$, respectively. Finally, we denote by a superscript (*) any df that is associated to the observed sample.

Suppose that the n triplets $(\mathbf{X}_i, Y_i, T_i)$ are observed among the N ones. For any df L, denote the left and right endpoints of its support by $a_L := \inf \{u : \ L(u) > 0\}$ and $b_L := \sup \{u : \ L(u) < 1\}$, respectively. Then under the current model, as discussed by Woodroofe [86], F and G can be estimated completely only if

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

Under RLT sampling scheme, the conditional joint distribution (Stute [74]) of (Y, T) becomes

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

where $t \wedge u := \min(t, u)$. The marginal distributions and their empirical versions are defined by

$$F^*(y) = \mu^{-1} \int_{-\infty}^{y} G(u) dF(u), \qquad F_n^*(y) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{Y_i \leq y\}},$$

$$G^*(t) = \mu^{-1} \int_{-\infty}^{\infty} G(t \wedge u) dF(u) \qquad \text{and} \qquad G_n^*(t) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq t\}},$$

where $\mathbf{1}_A$ denotes the indicator function of the set A.
In the sequel we use the following consistent estimator

$$\mu_n = \frac{G_n(y)\left[(1 - F_n(y-))\right]}{C_n\left(y\right)}, \tag{4.2}$$

for any y such that $C_n(y) \neq 0$, where $F_n(y-)$ denotes the left-limite of F_n at y. Here F_n and G_n are the product-limit estimators (Lynden-Bell [48]) for F and G, respectively i.e.,

$$F_n(y) = 1 - \prod_{i/Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)}\right], \quad G_n(y) = \prod_{i/T_i > y} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)}\right],$$

where $C_n(y) = n^{-1}\sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq y \leq Y_i\}}$ is the empirical estimator of

$$C(y) = \mathbb{P}(T \leq y \leq Y | Y \geq T).$$

Remark 4.2.1 *In the iid case, He and Yang [34] proved that μ_n does not depend on y and its value can then be obtained for any y such that $C_n(y) \neq 0$. Furthermore, they showed (see their Corollary 2.5) that*

$$\mu_n \overset{P-a.s.}{\rightarrow} \mu, \quad \text{as } n \rightarrow \infty.$$

The estimation of conditional df is based on the choice of weights. For complete data, the well-known Nadaraya-Watson weights are given by

$$W_{i,N}(\mathbf{x}) = \frac{K_d\{(\mathbf{x} - \mathcal{X}_i)/h_N\}}{\sum_{i=1}^{N} K_d\{(\mathbf{x} - \mathcal{X}_i)/h_N\}} =: \frac{\left(Nh_N^d\right)^{-1} K_d\{(\mathbf{x} - \mathcal{X}_i)/h_N\}}{v_N(\mathbf{x})}, \tag{4.3}$$

that are measurable functions of $\mathbf{x}$ depending on $\mathcal{X}_1, \cdots, \mathcal{X}_N$, with the convention $0/0 = 0$. The kernel K_d is a measurable function on $\mathbb{R}^d$ and (h_N) a nonnegative sequence which tends to zero as N tends to infinity. The corresponding estimator $v_N(\cdot)$ of $v(\cdot)$ is based on the $N-$sample and cannot therefore be calculated. On the other hand

$$v_n^*(\mathbf{x}) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) \tag{4.4}$$

is an estimator of the conditional density $v^*(\mathbf{x})$ (given $\mathcal{Y} \geq \mathcal{T}$).

In order to estimate the marginal density $v(\cdot)$ we have to take into account the truncation and the estimator

$$v_n(\mathbf{x}) = \frac{\mu_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) \tag{4.5}$$

is considered in Ould Saïd and Lemdani [55]. Note that in this formula and the forthcoming, the sum is taken only for i such that $G_n(Y_i) \neq 0$.

Then, adapting Ould Saïd-Lemdani's weights, we get the following estimator of the conditional df of $\mathcal{Y}$ given $\boldsymbol{\mathcal{X}} = \mathbf{x}$

$$\mathbf{F}_n(y|\mathbf{x}) = \frac{\displaystyle\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) H\left(\frac{y-Y_i}{h_n}\right)}{\displaystyle\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)}$$
$$=: \frac{\mathbf{F}_{1,n}(\mathbf{x}, y)}{v_n(\mathbf{x})}, \tag{4.6}$$

where H is a distribution function defined on $\mathbb{R}$, and

$$\mathbf{F}_{1,n}(\mathbf{x}, y) = \frac{\mu_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) H\left(\frac{y-Y_i}{h_n}\right) \tag{4.7}$$

is an estimator of $\mathbf{F}_1(\mathbf{x}, y)$.

We point out here that the estimator (4.6) and (4.7) have been already defined in Lemdani *et al.* [46] and used in Ould Saïd *et al.* [58]. Then a natural estimator of the p^{th} conditional quantile $q_p(\mathbf{x})$, is given by

$$q_{p,n}(\mathbf{x}) := \inf\{y \in \mathbb{R} : \mathbf{F}_n(y|\mathbf{x}) \geq p\} \tag{4.8}$$

which satisfies $\mathbf{F}_n(q_{p,n}(\mathbf{x})|\mathbf{x}) = p$.

Finally, considering the density $H^{(1)}$ and (4.6), we easily get (see Lemdani *et al.* [46, Remark 4.1]) an estimator of the conditional density of $\mathcal{Y}$ given $\boldsymbol{\mathcal{X}} = \mathbf{x}$ (defined by $f\,(y|\cdot) = \dfrac{\partial \mathbf{F}(y|.)}{\partial y}$)

$$f_n(y|\mathbf{x}) = \frac{f_{1,n}(\mathbf{x}, y)}{v_n(\mathbf{x})}, \tag{4.9}$$

where

$$f_{1,n}(\mathbf{x}, y) = \frac{\mu_n}{nh_n^{d+1}} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) H^{(1)}\left(\frac{y-Y_i}{h_n}\right) \tag{4.10}$$

is an estimator of $f\,(\mathbf{x}, y) = \dfrac{\partial \mathbf{F}_1(\mathbf{x}, y)}{\partial y}$, and $H^{(1)}$ is the derivative of H.

Throughout this paper, C denotes a positive constant which might take different values at different place. Assume that $0 = a_G < a_F$ and $b_G \leq b_F$. We consider

two real numbers a and b such that $a_F < a < b < b_F$. Let Ω be a compact subset of $\Omega_0 = \left\{ \mathbf{x} \in \mathbb{R}^d \,\middle\|\, v(\mathbf{x}) > 0 \right\}$ and $\gamma := \inf_{\mathbf{x} \in \Omega} v(\mathbf{x}) > 0$.
We introduce our assumptions, gathered below for easy reference.

$(K1)$ K_d is a bounded probability density, Hölder continuous with exponent $\beta > 0$ and satisfying

$$\|u\|^d K_d(u) \to 0 \quad as \quad \|u\| \to \infty.$$

$(K2)$ H is a df with $\mathcal{C}^1-$probability density $H^{(1)}$ and compact support.

$(K3)$ $H^{(1)}$ and K_d are second-order kernels.

$(M1)$ The observed sequence $\{(X_i, Y_i),\ i \geq 1\}$ is of stationary α-mixing random variables with coefficient $\alpha(n)$.

$(M2)$ $\{T_i;\ i \geq 1\}$ is a sequence of iid truncating variables independent of $\{(X_i, Y_i),\ i \geq 1\}$ with common continuous df G.

$(M3)$ There exists $\nu > 5 + 1/\beta$ for some $\beta > 1/7$ such that $\forall n$, $\alpha(n) = O(n^{-\nu})$.

$(D1)$ The conditional density $v^*(\cdot)$ is twice continuously differentiable.

$(D2)$ The marginal density $v(\cdot)$ is locally Lipschitz continuous over Ω_0.

$(D3)$ The joint density $f(\cdot, \cdot)$ is bounded and twice continuously differentiable.

$(D4)$ For all $j > 1$, the joint conditional density $f^*_{1,j+1}(\cdot, \cdot)$ of (X_1, X_{j+1}) exists and satisfies

$$\sup_{r,s} \left| f^*_{1,j+1}(r, s) - v^*(r)v^*(s) \right| \leq C < \infty,$$

for some constant C not depending on (i, j). The joint conditional density of (X_1, X_{j+1}, Y_{j+1}) and that of $(X_1, Y_1, X_{j+1}, Y_{j+1})$ are denoted by $f^*_{1,j+1,j+1}(\cdot, \cdot, \cdot)$ and $f^*_{1,1,j+1,j+1}(\cdot, \cdot, \cdot, \cdot)$ respectively.

$(H1)$ The bandwidth h_n satisfies:

(a) $nh_n^2/\log n \to \infty$ and $h_n = o(1/\log n)$, as $n \to \infty$,

(b) $h_n^{d+1} < Cn^{\frac{1}{1-\nu}}$ and (**c**) $h_n > Cn^{\frac{1}{2}\frac{(3-\nu)\beta}{\beta(\nu+1)+2\beta+1}+\eta}$ where η satisfies

$$\frac{1}{\beta(\nu+1)+2\beta+1} < \eta < \frac{(\nu-3)\beta}{\beta(\nu+1)+2\beta+1} + \frac{1}{1-\nu},$$

β and ν are the same as in $(M3)$.

$(H2)$ There exists a sequence $(m_n)_{n\geq 1}$, $1 \leq m_n \leq n$, such that as $n \to \infty$

(**a**) $m_n \to \infty, \quad m_n h \to 0$

(**b**) $\left(1/h^{\delta}\right) \sum_{l=m_n}^{\infty} (\alpha(l))^{\delta} \to 0, \quad$ with $\delta \in (0,1)$.

$(H3)$ Let (M_n) and (N_n) be subsequences of (n) tending to infinity such that:

(**a**) $M_n + N_n \leq n, \quad \frac{r_n M_n}{n} \to 1$ and $\frac{r_n N_n}{n} \to 0$,

(**b**) $M_n (nh_n)^{-1/2} \to 0$,

(**c**) $r_n \alpha(N_n) \to 0, \quad$ as $\quad n \to \infty$

where (r_n) be the largest positive integer for which $r_n (M_n + N_n) \leq n$.

Remark 4.2.2 (*Comments on the assumptions*) *Assumptions (K) are quite usual in kernel estimation. Assumptions $(D1)-(D3)$ are needed in the study of the bias term. $(D4)$ is needed for covariance calculus and takes similar forms to those used in complete data under dependence. Note also that, it is satisfied in the iid case. Assumptions (M) is related to mixing coefficient. Assumptions $(H1: a-b)$ are used in Ould Saïd et al. [58] to prove the uniform a.s. convergence of $\mathbf{F}_n(y|\mathbf{x}) - \mathbf{F}(y|\mathbf{x})$ and is needed here to prove the uniform a.s. convergence of $f_n(y|\mathbf{x}) - f(y|\mathbf{x})$, which is used in the proof of the asymptotic normality. Assumptions $(H2)$ and $(H3)$ deal with real sequences. They are used in Louani and Ould Saïd [47] and take part in establishing our results.*

Remark 4.2.3 *We point out here, that if we suppose that the original observations (related to $N-$sample) are dependent, we do not know, which dependence are the observed data. Then, we suppose in $(M1)$ that the observed sequence (related to $n-$sample) is alpha-mixing.*

Remark 4.2.4 *Assumption $(K2)$ implies that the kernel $H^{(1)}$ is bounded by a constant $M_0 > 0$. In the same way, under $(K1)$, we put $M_1 = \|K\|_{\infty}$.*

Remark 4.2.5 *As we are interested in the number n of observations (N is unknown), we give asymptotics as $n \to \infty$ unless otherwise specified. Since $n \leq N$ this implies $N \to \infty$ and these results also hold under $\mathbb{P} - a.s.$ as $N \to \infty$.*

Our first result, stated in Proposition 4.2.6, is the uniform a.s. convergence of $f_n(y|\mathbf{x}) - f(y|\mathbf{x})$ with rate of the conditional density estimator defined in (4.9).

Proposition 4.2.6 *Under assumptions (K), (M), (D) and $(H1)$, we have,*

$$\sup_{x\in\Omega}\sup_{a\leq y\leq b}|f_n(y|\mathbf{x}) - f(y|\mathbf{x})| = O\left(\max\left\{\sqrt{\frac{\log n}{nh_n^{d+1}}}, h_n^2\right\}\right), \quad \mathbf{P} - a.s. \quad n \to \infty.$$

The next result states the pointwise asymptotic normality of the conditional df estimator defined in (4.6). Let

$$\Sigma(\mathbf{x}, y) = \begin{pmatrix} \Sigma_0(\mathbf{x}, y) & \Sigma_1(\mathbf{x}, y) \\ \Sigma_1(\mathbf{x}, y) & \Sigma_2(\mathbf{x}) \end{pmatrix}$$

where

$$\Sigma_k(\mathbf{x}, y) = \int_{-\infty}^{y} H^{2-k}\left(\frac{y - \mathbf{s}}{h}\right)\frac{f(\mathbf{x}, s)}{G(s)}d\mathbf{s} \quad \text{for } k = 0, 1 \quad \text{and} \quad \Sigma_2(\mathbf{x}) = \int \frac{f(\mathbf{x}, s)}{G(s)}ds.$$

The second result deals with the strong uniform convergence with rate of the kernel conditional quantile estimator $q_{p,n}(.)$ which is given in the following theorem.

Proposition 4.2.7 *Under the assumptions (K), (M), (D) and (H), we have*

$$\sqrt{nh_n}\left(\mathbf{F}_n(y|\mathbf{x}) - \mathbf{F}(y|\mathbf{x})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2(\mathbf{x}, y)\right), \quad as \quad n \to \infty$$

where $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution and $\mathcal{N}(0, \sigma^2(\mathbf{x}, y))$ is the Gaussian distribution with zero mean and variance given by

$$\sigma^2(\mathbf{x}, y) = \kappa\frac{\Sigma_0(\mathbf{x}, y)\, v^2(\mathbf{x}) + \Sigma_2(\mathbf{x})\,\mathbf{F}_1^2(\mathbf{x}, y) - 2\Sigma_1(\mathbf{x}, y)\,\mathbf{F}_1(\mathbf{x}, y)v(\mathbf{x})}{\mu v^4(\mathbf{x})}$$

with $\kappa = \int K(r)\, dr$.

Our main result is given in the following Theorem.

Theorem 4.2.8 *Under the assumptions of Proposition 4.2.7, we have, for each* $p \in (0,1)$ *and for any* $\mathbf{x} \in \Omega_0$ *such that* $f_n(q_p(\mathbf{x})|\mathbf{x}) \neq 0$

$$\left(\frac{nh_n}{\sigma_q^2(\mathbf{x}, q_p(\mathbf{x}))}\right)^{1/2} (q_{p,n}(\mathbf{x}) - q_p(\mathbf{x})) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ 1), \quad as \quad n \to \infty$$

where

$$\sigma_q^2(\mathbf{x}, q_p(\mathbf{x})) = \frac{\sigma^2(\mathbf{x}, q_p(\mathbf{x}))}{f^2(q_p(\mathbf{x})|\mathbf{x})}.$$

Remark 4.2.9 *Note on the one hand that* $\Sigma_2(\mathbf{x}) \geq v(x)$ *and on the other hand, by Cauchy-Schwartz inequality, we have* $\Sigma_1^2(\mathbf{x}, y) < \Sigma_0(\mathbf{x}, y)\Sigma_2(\mathbf{x})$*. Therefore,* $\Sigma(\mathbf{x}, y)$ *is positive definite as soon as* $v(\mathbf{x}) > 0$*.*

4.3 Application to prediction

In this section we recall some situations and conditions that some usual processes satisfy the strong mixing conditions:
1) Gaussian process :
Let $X = (X_n,\ n \in \mathbb{N})$ be a stationary Gaussian process where X has a spectral density f of the form

$$f(e^{it}) = |p(e^{it})|^2 exp[u(e^{it}) + \bar{v}(e^{it})], \quad t \in [-\pi,\ \pi],$$

where $p(\cdot)$ is a polynomial, u and v are continuous real functions on the unit circle in the complex plane, and $\bar{v}$ is the conjugate of v. Then the process X_n is strong mixing.

2) Countable-state Markov chains :
Let $X = (X_n,\ n \in \mathbb{N})$ be a strictly stationary Markov chain, irreductible and aperiodic whose state space is an at most countable set. Then the process X_n is strong mixing.
An example of Markov chain with space state $\{0,\ 1\}$ which is strong mixing with exponential coefficient is given in Bradley [8, vol 1, pp 215-216].

3) Linear process (see Withers [85])
Let $Z_i,\ i \in \mathbb{N}$ be independent r.v. on $\mathbb{R}$ with characteristic functions ϕ_i such that

$$\max_i \int |\phi(u)|du < \infty \qquad \text{and} \qquad \gamma := \max_i \mathbb{E}|Z_i|^\delta < \infty \text{ for some } \delta > 0.$$

Let (g_i) be a sequence of complex numbers and for all t put $X_{nt} = \sum_{i=0}^n g_i Z_{t-i}$. Suppose that

$$K := \sup_{k\ s\ m\geq 1} \sup_{\nu} \max_t \left| \frac{\partial}{\partial \nu_t} \mathbb{P}\left(W + \nu \in \bigcup_{i=1}^s D_i \right) \right| < \infty,$$

where $D_i = X_(a_{it}, b_{it}),\quad \nu = (\nu_k, \cdots, \nu_{k+m-1}),\quad W = (W_k, \cdots, W_{k+m-1})$ and $W_t = X_{t-1,\,t}$, then the process X_t is strong mixing with coefficient $\alpha(k)$ is such that

$$\alpha(k) \leq 2(4K+\gamma)\alpha_0(k) \quad \text{where} \quad \alpha_0(k) = \sum_{i=k}^{\infty} |g_i|^\delta.$$

Remark 4.3.1 *Some particular case of linear processes can be given:*
3.i) Under the same conditions as before, if $g_k = O\left(k^{-v}\right)$ where

$$v > 1 + \delta^{-1} + \max(1, \delta^{-1}),$$

then the process X_t is strong mixing with $\alpha(k) = O\left(k^{-\varepsilon}\right)$ where $\varepsilon = (v\delta - \max(1,\delta))\left(1+\delta\right)^{-1} - 1 > 0$.
3.ii) Under the same conditions as before, if $g_k = O\left(e^{-vk}\right)$ where $v > 0$, then the process X_t is strong mixing with $\alpha(k) = O\left(e^{-v\lambda k}\right)$ where the $\lambda = \delta\left(1+\delta\right)^{-1}$.

A main application is given by the general $ARMA(p,q)$ process :
Let $\rho_j,\ j = 1, \cdots, p$ be the coefficients of autoregressive part and suppose $r = \max_{j=1,\ldots,p} \left|\rho_j\right| < 1$. Then, the $ARMA(p,q)$ process is equivalent to $X_t = \sum_{j=0}^{\infty} g_j Z_{t-j}$ and under the same condition as above, the process X_t is strong mixing with $\alpha(k) = O\left(r_0^{\lambda k}\right)$ for $r_0 > r$ and $\lambda = \delta\left(1+\delta\right)^{-1}$.

In what follow we apply Theorem 4.2.8 to the problem of prediction. It is well known, from the robustness theory that conditional median estimators are more robust than those of the classical conditional mean. Therefore the conditional median, $\mu\left(\mathbf{x}\right) = q_{1/2}(\mathbf{x})$, is a good alternative to the conditional mean as a predictor for a variable Y given $\mathbf{X} = \mathbf{x}$, specially in the case when the conditional density is asymmetric or has heavy tails. Let $(U_i)_{i\in\mathbb{N}}$ be a real-valued stationary and strong mixing process. The prediction aims at evaluating U_{m+1} given $U_1, ..., U_m$. To this end, set $\mathbf{X}_i = (U_i, ..., U_{i+d-1})$ and $Y_i = U_{i+d}$, $i = 1, ..., n$, where $n = m - d + 1$. The predictor estimator of U_{m+1} is defined by

$$\widehat{U}_{m+1} = \mu_n(\mathbf{X}_n) = q_{1/2,n}(\mathbf{X}_n),$$

where $q_{1/2,n}(\mathbf{X}_n)$ is given by (4.8).
The following Corollary is a consequence of Theorem 4.2.8.

Corollary 4.3.2 *Under the assumptions of Theorem 4.2.8, we have*

$$\left(\frac{nh_n}{\sigma_q^2\left(\mathbf{X}_n, q_{1/2}\left(\mathbf{X}_n\right)\right)}\right)^{1/2}\left(q_{1/2,n}(\mathbf{X}_n) - q_{1/2}(\mathbf{X}_n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ 1), \quad as \quad n \to \infty.$$

A plug-in-type estimate $\widehat{\sigma_q}^2\left(x, q_{1/2,n}(\mathbf{x})\right)$ for the asymptotic variance $\sigma_q^2\left(x, q_{1/2}(\mathbf{x})\right)$ can easily be obtained by using (4.4), (4.7), (4.10) and the estimators

$$\widehat{\Sigma}_1(\mathbf{x}, y) = \frac{\mu_n}{nh_n}\sum_{i=1}^{n}\frac{\mathbf{1}_{\{Y_i \le y\}}}{G_n^2(Y_i)}K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right) \quad \text{and} \quad \widehat{\Sigma}_2(\mathbf{x}) = \frac{\mu_n}{nh_n}\sum_{i=1}^{n}\frac{1}{G_n^2(Y_i)}K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)$$

of $\Sigma_1(\mathbf{x}, y)$ and $\Sigma_2(\mathbf{x})$ respectively. Then we get from Corollary 4.3.2 :

Corollary 4.3.3 *Under the assumptions of Theorem 4.2.8, we have*

$$\left(\frac{nh_n}{\widehat{\sigma_q}^2\left(x, q_{1/2,n}(\mathbf{x})\right)}\right)^{1/2}\left(q_{1/2,n}(\mathbf{x}) - q_{1/2}(\mathbf{x})\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ 1), \quad as \quad n \to \infty.$$

From Corollary 4.3.3, we get for each fixed $\zeta \in (0, 1)$, the following approximate $(1-\zeta)\,\%$ confidence interval for $q_{1/2}(\mathbf{x})$

$$\left[q_{1/2,n}(\mathbf{x}) - \frac{t_{1-\zeta/2}\widehat{\sigma}_q\left(x, q_{1/2,n}(\mathbf{x})\right)}{\sqrt{nh_n}},\ q_{1/2,n}(\mathbf{x}) + \frac{t_{1-\zeta/2}\widehat{\sigma}_q\left(\mathbf{x}, q_{1/2,n}(\mathbf{x})\right)}{\sqrt{nh_n}}\right]$$

where $t_{1-\zeta/2}$ denotes the $(1-\zeta/2)$ quantile of the standard normal distribution.

4.4 Proofs

We need some auxiliary results and notations to prove our results. Firstly, adapting (4.10), define

$$\tilde{f}_{1,n}(\mathbf{x}, y) = \frac{\mu}{nh_n^{d+1}} \sum_{i=1}^{n} \frac{1}{G(Y_i)} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right) H^{(1)}\left(\frac{y - Y_i}{h_n}\right). \tag{4.11}$$

We have

Lemma 4.4.1 *Under Assumptions,* (K) , (M) , $(D1)$ *and* $(H1 : a)$ *we have*

$$\sup_{x\in\Omega} \sup_{a\leq y\leq b} \left| f_{1,n}(\mathbf{x}, y) - \tilde{f}_{1,n}(\mathbf{x}, y)\right| = O\left(\sqrt{\frac{\log\log n}{nh_n^2}}\right), \quad \mathbf{P} - a.s. \quad as \quad n \to \infty.$$

Proof. Firstly, using $(K2)$ we can show that

$$\left| f_{1,n}(\mathbf{x}, y) - \tilde{f}_{1,n}(\mathbf{x}, y)\right| \leq \left\{ \frac{|\mu_n - \mu|}{G_n(a_F)} + \frac{\mu \sup_{y\geq a_F} |G_n(y) - G(y)|}{G_n(a_F)G(a_F)} \right\} \frac{M}{h_n} |v_n^*(\mathbf{x})| \,.$$

Recall that by (4.4) and $(D1)$, $v_n^\star$ is bounded. Furthermore from Lemma 5.2 in Ould Saïd and Tatachak [57], we have

$$|\mu_n - \mu| = O\left(\sqrt{\frac{\log\log n}{n}}\right), \qquad \mathbf{P} - a.s.$$

Moreover, $G_n(a_F) \overset{\mathbf{P}-a.s.}{\rightarrow} G(a_F) > 0$. In the same way, and using Lemma 3.4 in Liang *et al.* [42] (see Lemma 2.2.1, Chapter 2) we get

$$\sup_{y\geq a_F} |G_n(y) - G(y)| = O\left(\sqrt{\frac{\log\log n}{n}}\right), \qquad \mathbf{P} - a.s.$$

An immediate consequence of Lemma 4.1 in Ould Saïd *et al.* [58, see Lemma 3.5.1, chapter 3] gives the result. □

Lemma 4.4.2 *Under Assumptions,* (K) , (M) , $(D1)$, $(D4)$ *and* $(H1)$ *we have*

$$\sup_{x\in\Omega} \sup_{a\leq y\leq b} \left| \tilde{f}_{1,n}(\mathbf{x}, y) - \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}, y)\right]\right| = O\left(\sqrt{\frac{\log n}{nh_n^{d+1}}}\right), \quad \mathbf{P}{-}a.s. \quad as \quad n \to \infty.$$

Proof. The proof of this lemma makes use of the covering technique and the Fuk-Nagaev's inequality for strong mixing data (see Rio[66, formula 6.19b, page 87]). The compact set Ω can be covered by a finite number l_n of balls of radius $\omega_n = (n^{-1}h_n^{d(1+2\beta)})^{\frac{1}{2\beta}}$, where β is the Hölder exponent. Let $B_k := B(\mathbf{x}_k, \omega_n)$; $k = 1, ..., l_n$, denote each ball centered at some points $\mathbf{x}_k$. Since Ω is bounded, there exists a constant C such that $\omega_n l_n \leq C$. For any $\mathbf{x}$ in Ω, there exists B_k which contains $\mathbf{x}$ is that $|\mathbf{x} - \mathbf{x}_k| \leq \omega_n$.

Set, for any $i \geq 1$ and any $(\mathbf{x}, y) \in \mathbb{R}^d \times \mathbb{R}$

$$\begin{aligned} Z_i(\mathbf{x}, y) \ := \ & \frac{\mu}{nh_n^{d+1}} \left\{ \frac{1}{G(Y_i)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n} \right) H^{(1)} \left(\frac{y - Y_i}{h_n} \right) \right. \\ & \left. - \mathbf{E} \left[\frac{1}{G(Y_i)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_1}{h_n} \right) H^{(1)} \left(\frac{y - Y_i}{h_n} \right) \right] \right\}. \end{aligned}$$

Clearly, we have

$$\begin{aligned} \sum_{i=1}^{n} Z_i(\mathbf{x}, y) = & \left\{ \tilde{f}_{1,n}(\mathbf{x}, y) - \tilde{f}_{1,n}(\mathbf{x}_k, y) - \left(\mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}, y) \right] - \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}_k, y) \right] \right) \right\} \\ & + \left(\tilde{f}_{1,n}(\mathbf{x}_k, y) - \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}_k, y) \right] \right) \\ =: & \sum_{i=1}^{n} \tilde{Z}_i(\mathbf{x}, y) + \sum_{i=1}^{n} Z_i(\mathbf{x}_k, y), \end{aligned}$$

one then has

$$\sup_{\mathbf{x} \in \Omega} \left| \sum_{i=1}^{n} Z_i(\mathbf{x}, y) \right| \leq \max_{1 \leq k \leq q_n} \underbrace{\sup_{\mathbf{x} \in B_k} \left| \sum_{i=1}^{n} \tilde{Z}_i(\mathbf{x}, y) \right|}_{A_n} + \max_{1 \leq k \leq q_n} \underbrace{\left| \sum_{i=1}^{n} Z_i(\mathbf{x}_k, y) \right|}_{B_n}.$$

First, we have

$$\begin{aligned} A_n \leq & \frac{1}{nh_n^{d+1}} \sum_{i=1}^{n} \left| \frac{\mu}{G(Y_i)} H^{(1)} \left(\frac{y - Y_i}{h_n} \right) \right| \left| K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n} \right) - K_d \left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_n} \right) \right| \\ & + \frac{1}{h_n^{d+1}} \mathbf{E} \left[\left| \frac{\mu}{G(Y_i)} H^{(1)} \left(\frac{y - Y_i}{h_n} \right) \right| \left| K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n} \right) - K_d \left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_n} \right) \right| \right] \\ =: & \ A_{1n} + A_{2n}. \end{aligned}$$

Assumptions $(K1)$ and $(K2)$, yield

$$\begin{aligned} A_{1n} \leq & \frac{\mu M_0 \left| x - \mathbf{x}_k \right|^{\beta}}{G(a_F) h_n^{d(1+\beta)+1}} \leq C \omega_n^{\beta} h_n^{-d(1+\beta)-1} \\ = & \ O\left(\left(nh_n^{d+2} \right)^{-1/2} \right). \end{aligned}$$

Similar argument as above, lead to the same bound for A_{2n}. Hence, by $(H1:a)$ and for all n large enough, we get $A_n = o_{\mathbf{P}}(1)$.
Now, we focus on B_n, under $(K1)$ and $(K2)$, the rv's $W_i := nh_n^{d+1}Z_i(\mathbf{x}_k, y)$ are centered and may be bounded by

$$\frac{2\mu M_0 M_d}{G(a_F)} \leq C < \infty.$$

The use of the well known Fuk-Nagaev's inequality slightly modified in Ferraty and Vieu [22, proposition A.11-ii), page 237], allows one to get, for all $\varepsilon > 0$ and $\eta > 1$

$$\begin{aligned} \mathbf{P}\left\{\max_{1\leq k\leq q_n}\left|\sum_{i=1}^{n} Z_i(\mathbf{x}_k, y)\right| > \varepsilon\right\} &\leq \sum_{i=1}^{q_n}\mathbf{P}\left\{\left|\sum_{i=1}^{n} Z_i(\mathbf{x}_k, y)\right| > \varepsilon\right\} \\ &\leq C\omega_n^{-1}\left\{\frac{n}{\eta}\left(\frac{\eta}{\varepsilon n h_n^{d+1}}\right)^{\nu+1} + \left(1 + \frac{\varepsilon^2 n^2 h_n^{2(d+1)}}{\eta s_n}\right)^{-\frac{r}{2}}\right\} \\ &=: B_{1n} + B_{2n} \qquad (4.12) \end{aligned}$$

where

$$s_n = \sum_{1\leq i\leq n}\sum_{1\leq j\leq n}|Cov(W_i, W_j)|.$$

By taking

$$\eta = (\log n)^{1+\delta}, \text{ where } \delta > 0, \quad \text{and} \quad \varepsilon = \varepsilon_0\sqrt{\frac{\log n}{nh_n^{d+1}}} \quad \text{for some } \varepsilon_0 > 0. \qquad (4.13)$$

We get

$$B_{1n} = Cn^{1-\frac{\nu+1}{2}+\frac{1}{2\beta}}(\log n)^{\nu(1+\delta)-\frac{v+1}{2}}\left(\varepsilon_0^{\nu+1}h_n^{d\left(\frac{1}{2\beta}+1\right)}h_n^{(d+1)\frac{v+1}{2}}\right)^{-1}.$$

Then, using $(H1:c)$ we get

$$B_{1n} \leq C'(\log n)^{\nu(1+\delta)-\frac{v+1}{2}}n^{-1-\frac{\eta}{2\beta}(\beta(\nu+1)+2\beta+1-\frac{1}{\eta})},$$

Hence, for any η as in $(H1:c)$, B_{1n} is bounded by the term of a finite-sum series.

Let's now examine the term B_{2n}. Firstly, we have to get the asymptotic behavior of

$$\begin{aligned} s_n &= \sum_{i=1}^{n} Var(W_i) + \sum_{1\leq i<j\leq n}|Cov(W_i, W_j)| \\ &=: s_n^{var} + s_n^{cov}. \end{aligned}$$

Assumption $(K2)$ implies that, the kernel $H^{(1)}$ is bounded by a constant $M_0 > 0$. Hence, under $(K1)$, $(D1)$ and a change of variable we have

$$\begin{aligned}
s_n^{var} &= n\mathbf{E}\left[\frac{\mu^2}{G^2(Y_1)}K_d^2\left(\frac{\mathbf{x}_k - X_1}{h_n}\right)H^{(1)2}\left(\frac{y-Y_1}{h_n}\right)\right] \\
&- n\mathbf{E}^2\left[\frac{\mu}{G(Y_i)}K_d\left(\frac{\mathbf{x}_k - X_1}{h_n}\right)H^{(1)}\left(\frac{y-Y_1}{h_n}\right)\right] \\
&\leq \frac{\mu^2 M_0^2}{G^2(a_F)}\left(nh_n^{d+1} + nh_n^{2(d+1)}\right) \\
&= O\left(nh_n^{d+1}\right). \qquad (4.14)
\end{aligned}$$

On the other hand, $(M1)$, $(K1)$ and $(K2)$ lead to

$$\begin{aligned}
|Cov(W_i, W_j)| &= \Bigg|\int\int\int\int \frac{\mu}{G(r)}K_d\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_n}\right)H^{(1)}\left(\frac{y-r}{h_n}\right)\frac{\mu}{G(t)}K_d\left(\frac{\mathbf{x}_k - \mathbf{s}}{h_n}\right) \\
&\times H^{(1)}\left(\frac{y-t}{h_n}\right)f^*_{1,1,j+1,j+1}(u,r,s,t)d\mathbf{u}drd\mathbf{s}dt \\
&- \int\int \frac{\mu}{G(r)}K_d\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_n}\right)H^{(1)}\left(\frac{y-r}{h_n}\right)f^*(u,r)d\mathbf{u}dr \\
&\times \int\int \frac{\mu}{G(t)}K_d\left(\frac{\mathbf{x}_k - \mathbf{s}}{h_n}\right)H^{(1)}\left(\frac{y-t}{h_n}\right)f^*(s,t)d\mathbf{s}dt\Bigg| \\
&\leq \frac{\mu^2}{G^2(a_F)}\Bigg|\int\int\int\int K_d\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_n}\right)H^{(1)}\left(\frac{y-r}{h_n}\right)K_d\left(\frac{\mathbf{x}_k - \mathbf{s}}{h_n}\right) \\
&\times H^{(1)}\left(\frac{y-t}{h_n}\right)\left(f^*_{1,1,j+1,j+1}(u,r,s,t) - f^*(u,r)f^*(s,t)\right)d\mathbf{u}drd\mathbf{s}dt\Bigg|.
\end{aligned}$$

Using Assumption $(D4)$ and a change of variable, it follows that

$$|Cov(W_i, W_j)| = O\left(h_n^{2(d+1)}\right). \qquad (4.15)$$

Note also that, these covariances can be controlled by means of the usual Davydov covariance inequality for mixing processes (see Rio [66, formula 1.12a, page 10]; or Bosq [7, formula 1.11, page 22]). We have

$$\forall i \neq j, \quad |Cov(U_i, U_j)| \leq C\alpha\left(|i-j|\right). \qquad (4.16)$$

To evaluate s_n^{cov}, we use the technique developed by Masry [49]. We deduce easily that for $\varphi_n > 0$

$$\begin{aligned}
s_n^{cov} &= \sum_{0<|i-j|\leq\varphi_n}|Cov(W_{ik}, W_{jk})| + \sum_{|i-j|>\varphi_n}|Cov(W_{ik}, W_{jk})| \\
&= O(nh_n^{2(d+1)}\varphi_n) + O(n^2\alpha\left(\varphi_n\right)).
\end{aligned}$$

It suffice to take $\varphi_n = \left\lceil \left(n^{-1}h_n^{d+1}\right)^{-1/\nu} \right\rceil$ (where $\lceil . \rceil$ denotes the smallest integer greater than the argument), and use $(H1 : b)$, to get

$$s_n^{cov} = O(nh_n^{d+1}). \tag{4.17}$$

Then, (4.14) and (4.17) lead directly to

$$s_n = O\left(nh_n^{d+1}\right).$$

Consequently, by taking r and ε as in (4.13) and using Taylor expansion of $\log(1+x)$, the term B_{2n} becomes

$$B_{2n} \leq Cn^{\frac{1}{2\beta} - C'\varepsilon_0^2} h_n^{-d(1+\frac{1}{2\beta})},$$

which by an appropriate choice of ε_0 can be made $O\left(n^{-3/2}\right)$, which in tern is the general term of summable series. Thus $\sum_{n\geq 1}(B_{1n} + B_{2n}) < \infty$, and Borel-Cantelli Lemma allow us to conclude. □

Lemma 4.4.3 *Under Assumptions, (K) and $(D3)$ we have, for n large enough*

$$\sup_{x\in\Omega} \sup_{a\leq y\leq b} \left| \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}, y)\right] - f_1(\mathbf{x}, y) \right| = O\left(h_n^2\right), \quad \mathbf{P} - a.s.$$

Proof. The bias terms do not depend on the mixing structure. The proof is analogous to that of Lemma 4.8 in Lemdani *et al.* [46], therefore, it is omitted. □

Proof of Proposition 4.2.6 In view of (4.9), we have the following classical decomposition

$$\begin{aligned}
&\sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} |f_n(y|\mathbf{x}) - f(y|\mathbf{x})| \\
&\qquad \leq \frac{1}{\gamma - \sup\limits_{\mathbf{x}\in\Omega}|v_n(\mathbf{x}) - v_n(\mathbf{x})|} \left\{ \sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} |f_{1,n}(\mathbf{x}, y) - f_1(\mathbf{x}, y)| \right. \\
&\qquad\qquad \left. + \gamma^{-1} \sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} |f(y|\mathbf{x})| \sup_{\mathbf{x}\in\Omega} |v_n(\mathbf{x}) - v_n(\mathbf{x})| \right\}.
\end{aligned}$$

In addition, we have

$$\begin{aligned}
\sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} |f_{1,n}(\mathbf{x}, y) - f_1(\mathbf{x}, y)| &\leq \sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} \left| f_{1,n}(\mathbf{x}, y) - \tilde{f}_{1,n}(\mathbf{x}, y) \right| \\
&\quad + \sup_{\mathbf{x}\in\Omega} \sup_{a\leq y\leq b} \left| \tilde{f}_{1,n}(\mathbf{x}, y) - \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}, y)\right] \right| \\
&\quad + \sup \sup_{a\leq y\leq b} \left| \mathbf{E}\left[\tilde{f}_{1,n}(\mathbf{x}, y)\right] - f_1(\mathbf{x}, y) \right|.
\end{aligned}$$

The result follows straightforwardly from Lemma 4.4.1-4.4.3. □

In order to prove Theorem 4.2.8 we will prove the asymptotic normality of the estimator of the conditional df. Using (4.6), we write

$$\mathbf{F}_n(y|\mathbf{x}) = \frac{\mu_n^{-1}\mathbf{F}_{1,n}(\mathbf{x},y)}{\mu_n^{-1}v_n(\mathbf{x})}, \tag{4.18}$$

where, from (4.7) and (4.5),

$$\frac{\mathbf{F}_{1,n}(\mathbf{x},y)}{\mu_n} = \frac{1}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)H\left(\frac{y-Y_i}{h_n}\right)$$

and

$$\frac{v_n(\mathbf{x})}{\mu_n} = \frac{1}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right).$$

Ould Saïd and Lemdani [55] give three-terms decomposition of the differences $\mu_n^{-1}v_n(\mathbf{x}) - \mu^{-1}v(\mathbf{x})$;

$$\begin{aligned}\frac{v_n(\mathbf{x})}{\mu_n} - \frac{v(\mathbf{x})}{\mu} &= \frac{v_n(\mathbf{x})}{\mu_n} - \frac{\tilde{v}(\mathbf{x})}{\mu} + \frac{\tilde{v}_n(\mathbf{x})}{\mu} - \mathbf{E}\left[\frac{\tilde{v}_n(\mathbf{x})}{\mu}\right] \\ &\quad + \mathbf{E}\left[\frac{\tilde{v}_n(\mathbf{x})}{\mu}\right] - \frac{v(\mathbf{x})}{\mu} &(4.19)\\ &=: \Gamma_{n1}(\mathbf{x}) + \Gamma_{n2}(\mathbf{x}) + \Gamma_{n3}(x). &(4.20)\end{aligned}$$

Similarly, Lemdani *et al. [46]* give three-terms decomposition of the differences $\mu_n^{-1}\mathbf{F}_{1,n}(\mathbf{x},y) - \mu^{-1}\mathbf{F}(\mathbf{x},y)$;

$$\begin{aligned}\frac{F_{1,n}(\mathbf{x},y)}{\mu_n} - \frac{F(\mathbf{x},y)}{\mu} &= \frac{F_{1,n}(\mathbf{x},y)}{\mu_n} - \frac{\tilde{F}_{1,n}(\mathbf{x},y)}{\mu} + \frac{\tilde{F}_{1,n}(\mathbf{x},y)}{\mu} - \mathbf{E}\left[\frac{\tilde{F}_{1,n}(\mathbf{x},y)}{\mu}\right] \\ &\quad + \mathbf{E}\left[\frac{\tilde{F}_{1,n}(\mathbf{x},y)}{\mu}\right] - \frac{F(\mathbf{x},y)}{\mu} \\ &=: \Lambda_{n1}(\mathbf{x},y) + \Lambda_{n2}(\mathbf{x},y) + \Lambda_{n3}(\mathbf{x},y). &(4.21)\end{aligned}$$

We first consider the negligible terms in (4.20) and (4.21).

Lemma 4.4.4 *Under Assumptions of Lemma 4.4.1 and for any* $\mathbf{x}$, y *both* $\sqrt{nh_n^d}\Gamma_{n1}(\mathbf{x})$ *and* $\sqrt{nh_n^d}\Lambda_{n1}(\mathbf{x},y)$ *are* $o_{\mathbf{p}}(1)$ *as* $n\to\infty$.

Proof. Using Lemma 3.4 in Liang *et al.* [42] (see Lemma 2.2.1, Chapter 2) and Lemma 4.1 in Ould Saïd *et al.* [58], we get

$$\begin{aligned}\sqrt{nh_n^d}\Gamma_{n1}(\mathbf{x}) &\leq \sqrt{nh_n^d}\frac{\sup_{y\geq a_F}|G_n(y)-G(y)|}{G_n(a_F)G(a_F)}v_n^*(\mathbf{x}) \\ &= O_{\mathbf{P}}\left(\sqrt{h_n^d}\right).\end{aligned}$$

In the same way,

$$\sqrt{nh_n^d}\Lambda_{n1}(\mathbf{x},y) = O_{\mathbf{P}}\left(\sqrt{h_n^d}\right),$$

by using $(K2)$. □

Lemma 4.4.5 *Under Assumptions* (K), $(D3)$, $(D4)$, *and* $(H1:a)$, *for any* $\mathbf{x}$, y *both* $\sqrt{nh_n^d}\Gamma_{n3}(\mathbf{x})$ *and* $\sqrt{nh_n^d}\Lambda_{n3}(\mathbf{x},y)$ *are* $o_{\mathbf{P}}(1)$ *as* $n\to\infty$.

Proof. We have

$$\sqrt{nh_n^d}\Gamma_{n3}(\mathbf{x}) = \frac{1}{\mu}\left\{\mathbf{E}\left[\tilde{v}_n(\mathbf{x})\right]-v(\mathbf{x})\right\}.$$

Using this, the result is a direct consequence of Lemma 4.4 in Lemdani *et al.* [46] . Likewise, we can show $\sqrt{nh_n^d}\Lambda_{n3}(\mathbf{x},y) = o_{\mathbf{P}}(1)$. □

Now we consider the dominant terms $\Gamma_{n2}(\mathbf{x})$ and $\Lambda_{n2}(\mathbf{x},y)$ and prove that

$$\sqrt{nh_n^d}\left(\Lambda_{n2}(\mathbf{x},y)\ ,\ \Gamma_{n2}(\mathbf{x})\right)^T \xrightarrow{\mathcal{D}} \mathcal{N}\left(0\ ,\ \sigma^2(\mathbf{x},y)\right)$$

where the variance $\sigma^2(\mathbf{x},y)$ will be explicitly given later on.
We follow the same lines as Louani and Ould Saïd [47] for the kernel conditional mode estimator or as Berlinet *et al.* [4] for the conditional quantile estimator in the case of complete data (no truncation). Let $c=(c_1,c_2)^T$ be a pair of real numbers satisfying $c_1^2+c_2^2\neq 0$. Put

$$\sqrt{nh_n^d}\left(c_1\Gamma_{n2}(\mathbf{x})+c_2\Lambda_{n2}(\mathbf{x},y)\right) =: \left(nh_n^d\right)^{-1/2}\sum_{i=1}^n\Delta_i(\mathbf{x},y)$$

where

$$\begin{aligned}\Delta_i(\mathbf{x},y) &= \frac{c_1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)-c_1\mathbf{E}\left[\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)\right] \\ &+\frac{c_2}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)H\left(\frac{y-Y_i}{h_n}\right) \\ &-c_2\mathbf{E}\left[\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)H\left(\frac{y-Y_i}{h_n}\right)\right] \\ &=: c_1\Delta_{1,i}-c_1\mathbf{E}\left[\Delta_{1,i}\right]+c_2\Delta_{2,i}-c_2\mathbf{E}\left[\Delta_{2,i}\right].\end{aligned}$$

Lemma 4.4.6 *Let* $(\mathbf{x}, y) \in \mathbb{R}^d \times \mathbb{R}$, *under assumptions* $(K1)$, $(K2)$, $(D3)$, $(D4)$, *and* $(H1:a)$, *we have as* $n \to \infty$,

$$\frac{1}{h_n^d}\mathbf{E}\left[\Delta_1^2(\mathbf{x}, y)\right] \to \mu^{-1}\kappa c^T \Sigma(\mathbf{x}, y)\, c, \tag{4.22}$$

where κ *and* $\Sigma(\mathbf{x}, y)$ *are described as in Proposition 4.2.7,*

$$|Cov\left[\Delta_{1,1}, \Delta_{1,j+1}\right]| = O\left(h_n^{2d}\right), \tag{4.23}$$

$$|Cov\left[\Delta_{1,1}, \Delta_{2,j+1}\right]| = O\left(h_n^{2d}\right) \tag{4.24}$$

and

$$|Cov\left[\Delta_{2,1}, \Delta_{2,j+1}\right]| = O\left(h_n^{2d}\right). \tag{4.25}$$

Proof. Firstly, remark that

$$\frac{1}{h_n^d}\mathbf{E}\left[\Delta_1^2(\mathbf{x}, y)\right] = \frac{c_1^2}{h_n^d}Var\left[\Delta_{1,1}\right] + \frac{c_2^2}{h_n^d}Var\left[\Delta_{2,1}\right] + \frac{2c_1c_2}{h_n^d}Cov\left[\Delta_{1,i}, \Delta_{2,j}\right].$$

The first term of the right hand side of this latter equation is given in Ould Saïd and Lemdani [55, Lemma 6.9] . On the other hand, by Lemma 4.13 in Lemdani *et al.* [46] we get the second and last term. Particularly,

$$\frac{1}{h_n^d}Var\left[\Delta_{1,1}\right] = \frac{\kappa}{\mu}\Sigma_2(\mathbf{x}, y) + o(1),$$

$$\frac{1}{h_n^d}Var\left[\Delta_{1,1}\right] = \frac{\kappa}{\mu}\Sigma_0(\mathbf{x}, y) + o(1)$$

and

$$\frac{1}{h_n^d}Cov\left[\Delta_{1,i}, \Delta_{2,j}\right] = \frac{\kappa}{\mu}\Sigma_1(\mathbf{x}, y) + o(1).$$

To prove (4.23), we make use $(D4)$ and a change of variable, we obtain

$$\begin{aligned}
|Cov\left[\Delta_{1,1}, \Delta_{1,j+1}\right]| &\leq \frac{h_n^{2d}}{G_n^2(a_F)}\int\int K_d(\mathbf{u})\, K_d(\mathbf{r}) \\
&\quad \times \left|f_{1,j+1}^*(\mathbf{x} - h_n\mathbf{u}, \mathbf{x} - h_n\mathbf{r}) - v^*(\mathbf{x} - h_n\mathbf{u})\, v^*(\mathbf{x} - h_n\mathbf{r})\right| d\mathbf{u}d\mathbf{r} \\
&= O\left(h_n^{2d}\right).
\end{aligned}$$

In the same way, we get : $|Cov\left[\Delta_{1,1}, \Delta_{2,j+1}\right]| =$

$$\begin{aligned}
&\left|Cov\left[\frac{1}{G_n(Y_1)}K_d\left(\frac{\mathbf{x} - \mathbf{X}_1}{h_n}\right), \frac{1}{G_n(Y_{j+1})}K_d\left(\frac{\mathbf{x} - \mathbf{X}_{j+1}}{h_n}\right) H\left(\frac{y - Y_{j+1}}{h_n}\right)\right]\right| \\
&\leq \frac{1}{G_n^2(a_F)}\left|\int\int\int K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_n}\right) K_d\left(\frac{\mathbf{x} - \mathbf{r}}{h_n}\right) H\left(\frac{y - s}{h_n}\right) f_{1,j+1,j+1}^*(\mathbf{u}, \mathbf{r}, s)\, d\mathbf{u}d\mathbf{r}ds\right. \\
&\left. - \int K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_n}\right) v^*(\mathbf{u})\, d\mathbf{u}\int\int K_d\left(\frac{\mathbf{x} - \mathbf{r}}{h_n}\right) H\left(\frac{y - s}{h_n}\right) f_{1,j+1}^*(\mathbf{r}, s)\, d\mathbf{r}ds\right|.
\end{aligned}$$

Under $(K2)$, the kernel H is bounded by 1. Hence integrates over s, using $(D4)$ and a change of variable lead to get (4.24). Likewise, we obtain (4.25). □

Lemma 4.4.7 *Let $(\mathbf{x}, y) \in \mathbb{R}^d \times \mathbb{R}$, under the assumptions of Lemma 4.4.6 and condition $(H2)$, we have*

$$\frac{1}{nh_n^d} \sum_{1 \le i < j \le n}^{n} |\mathbf{E}[\Delta_i(\mathbf{x}, y) \;\; \Delta_j(\mathbf{x}, y)]| \to 0, \qquad as \;\; n \to \infty. \tag{4.26}$$

Proof. The proof is based on decomposition (4.27) hereafter,

$$\begin{aligned}
\frac{1}{nh_n^d} \sum_{1 \le i < j \le n}^{n} |\mathbf{E}[\Delta_i(\mathbf{x}, y) \;\; \Delta_j(\mathbf{x}, y)]| &= \frac{c_1^2}{nh_n^d} \sum_{1 \le i < j \le n}^{n} Cov[\Delta_{1,i}, \Delta_{1,j}] \\
&+ \frac{c_2^2}{nh_n^d} \sum_{1 \le i < j \le n}^{n} Cov[\Delta_{2,i}, \Delta_{2,j}] \\
&+ \frac{c_1 c_2}{nh_n^d} \sum_{1 \le i < j \le n}^{n} Cov[\Delta_{1,i}, \Delta_{2,j}] \\
&+ \frac{c_1 c_2}{nh_n^d} \sum_{1 \le i < j \le n}^{n} Cov[\Delta_{2,i}, \Delta_{1,j}] \\
&=: \mathcal{A}_{1n} + \mathcal{A}_{2n} + \mathcal{A}_{3n} + \mathcal{A}_{4n}.
\end{aligned} \tag{4.27}$$

We will prove that each term of the right hand side of (4.27) tends to 0 as n tends to infinity. In the sequel, we use technique developed by Masry [49] and used in Louani and Ould Saïd [47]. Define the sets S_1 and S_2 as follows

$$\begin{aligned}
S_1 &= \{(i, j) : i, j \in \{1, 2, ..., n\}, \; 1 \le j - i \le m_n\}, \\
S_2 &= \{(i, j) : i, j \in \{1, 2, ..., n\}, \; m_n + 1 \le j - i \le n - 1\}
\end{aligned}$$

where m_n is as in $(H2)$ and observe that $S_1 \cup S_2 = \{(i, j) : 1 \le i < j \le n\}$.
If $m_n \ge n - 1$, then $S_2 = \varnothing$. Therefore, using (4.23)

$$\begin{aligned}
\mathcal{A}_{1n} &= \frac{c_1^2}{nh_n^d} \sum_{j=2}^{n} \sum_{i=1}^{j-1} Cov\left[\frac{1}{G_n(Y_i)} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_n}\right), \frac{1}{G_n(Y_j)} K_d\left(\frac{\mathbf{x} - \mathbf{X}_j}{h_n}\right)\right] \\
&\le C m_n h_n^d.
\end{aligned}$$

Suppose now that $m_n \le n - 2$. It follows that

$$\mathcal{A}_{1n} = \frac{c_1^2}{nh_n^d} \sum_{(i,j) \in S_1} Cov[\Delta_{1,i}, \Delta_{1,j}] + \frac{c_1^2}{nh_n^d} \sum_{(i,j) \in S_2} Cov[\Delta_{1,i}, \Delta_{1,j}]. \tag{4.28}$$

By (4.23), we have

$$\begin{aligned}
&\frac{c_1^2}{nh_n^d}\sum_{(i,j)\in S_1} Cov\left[\Delta_{1,i},\Delta_{1,j}\right]\\
&=\frac{c_1^2}{nh_n^d}\sum_{i=1}^{n-m_n}\sum_{j=i+1}^{m_n+1} Cov\left[\frac{1}{G_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right),\frac{1}{G_n(Y_j)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_j}{h_n}\right)\right]\\
&\le C m_n h_n^d.
\end{aligned}$$

To bound the sum over S_2 in (4.28), we use moment inequality due to Rio [65]. Let p,q,r be integer numbers greater than 1 such that $\frac{1}{p}+\frac{1}{q}=1-\frac{1}{r}$. We have

$$\frac{c_1^2}{nh_n^d}\sum_{(i,j)\in S_1} Cov\left[\Delta_{1,i},\Delta_{1,j}\right]=\frac{c_1^2}{nh_n^d}\sum_{(i,j)\in S_1} 2^{1+\frac{1}{r}}\left(\alpha\left(i-j\right)\right)^{\frac{1}{r}}\left(\mathbf{E}\left|\Delta_{1,1}\right|^p\right)^{\frac{1}{p}}\left(\mathbf{E}\left|\Delta_{1,1}\right|^q\right)^{\frac{1}{q}}.$$

Moreover, under $(K1)$, we get for n large enough,

$$\begin{aligned}
\mathbf{E}\left|\Delta_{1,1}\right|^p &= \mathbf{E}\left|\frac{1}{G_n(Y_1)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_1}{h_n}\right)\right|^p\\
&\le \frac{1}{G_n^p(a_F)}\left|\int K_d^p\left(\frac{\mathbf{x}-\mathbf{u}}{h_n}\right)v^*(\mathbf{u})d\mathbf{u}\right|\\
&\le \frac{h_n^d}{\mu^p}\int\left|K_d\left(\mathbf{r}\right)\right|^p v(\mathbf{x}-h_n\mathbf{r})d\mathbf{r}\\
&\le C(\mathbf{x})h_n^d. \qquad (4.29)
\end{aligned}$$

where $C(\mathbf{x})$ is a constant possibly depending on x. Hence, using (4.29), it follows that

$$\begin{aligned}
\frac{c_1^2}{nh_n^d}\sum_{(i,j)\in S_1} Cov\left[\Delta_{1,i},\Delta_{1,j}\right] &\le \frac{c_1^2}{nh_n^d}\sum_{k=m_n+1}^{n-1}\sum_{i=1}^{n-k}\left(2^{1+\frac{1}{r}}\left(\alpha\left(k\right)\right)^{\frac{1}{r}}\left(C(\mathbf{x})h_n^d\right)^{1-\frac{1}{r}}\right)\\
&\le C(\mathbf{x})2^{1+\frac{1}{r}}h_n^{-\frac{d}{r}}\sum_{k=m_n+1}^{n-1}\left(\alpha\left(k\right)\right)^{\frac{1}{r}}.
\end{aligned}$$

Therefore, by assumption $(H2)$ we should have $\mathcal{A}_{1n}\to 0, \quad \text{as } n\to\infty$.
Finally, by using the same argument as for $\mathcal{A}_{1n}$. We can get that, each terms of the right hand side of (4.27) goes to 0 as n goes to infinity. The details are omitted. □

Lemma 4.4.8 *Let $(\mathbf{x},y)\in\mathbb{R}^d\times\mathbb{R}$, under the assumptions of Lemma 4.4.7, we have*

$$nh_n^d Var\left(c_1\Gamma_{n2}\left(\mathbf{x}\right)+c_2\Lambda_{n2}\left(\mathbf{x},y\right)\right)=\mu^{-1}\kappa c^T\Sigma\left(\mathbf{x},y\right)c, \qquad \textit{as } \; n\to\infty. \qquad (4.30)$$

Proof. Remark that

$$
\begin{aligned}
nh_n^d Var\left(c_1\Gamma_{n2}(\mathbf{x}) + c_2\Lambda_{n2}(\mathbf{x},y)\right) &= \frac{1}{nh_n^d} Var\left(\sum_{i=1}^{n} \Delta_i(\mathbf{x},y)\right) \\
&= \frac{c_1^2}{h_n^d} Var\left(\Delta_{1,1}\right) + \frac{2c_1^2}{nh_n^d} \sum_{1\le i<j\le n} Cov\left[\Delta_{1,i},\Delta_{1,j}\right] \\
&+ \frac{c_2^2}{h_n^d} Var\left(\Delta_{2,1}\right) + \frac{2c_2^2}{nh_n^d} \sum_{1\le i<j\le n} Cov\left[\Delta_{2,i},\Delta_{2,j}\right] \\
&+ \frac{2c_1c_2}{nh_n^d} \sum_{1\le i<j\le n} Cov\left[\Delta_{1,i},\Delta_{2,j}\right] \\
&+ \frac{2c_1c_2}{nh_n^d} \sum_{1\le i<j\le n} Cov\left[\Delta_{1,j},\Delta_{2,i}\right] \\
&+ \frac{2c_1c_2}{nh_n^d} Cov\left[\Delta_{1,1},\Delta_{2,1}\right].
\end{aligned}
$$

The result follows directly from Lemma 4.4.6 and 4.4.7. □

In order to establish the asymptotic normality for sums of dependent rv's, we use the Doob's small-block and large-block technique (see, Doob [17, pp. 228–232]) according to which the sum $\sum_{i=1}^{n} \Delta_i(\mathbf{x},y)$ is split up as follows. Partition $\{1,..,n\}$ into $2r_n+1$ subsets with large-block of size M_n and small-block of size N_n, where (M_n), (N_n) and (r_n) are three sequences of integer numbers described in Assumption $(H3)$ and set

$$
\sum_{i=1}^{n} \Delta_i(\mathbf{x},y) = S_n(\mathbf{x},y) + T_{1,n}(\mathbf{x},y) + T_{2,n}(\mathbf{x},y) \tag{4.31}
$$

where

$$
S_n(\mathbf{x},y) = \sum_{j=1}^{r_n} L_j(\mathbf{x},y),
$$

$$
T_{1,n}(\mathbf{x},y) = \sum_{j=1}^{r_n} L_j'(\mathbf{x},y) \quad \text{and} \quad T_{2,n}(\mathbf{x},y) = \sum_{j=(M_n+N_n)r_n+1}^{n} \Delta_j(\mathbf{x},y)
$$

with

$$
L_j(\mathbf{x},y) = \sum_{i=j(M_n+N_n)+1}^{j(M_n+N_n)+M_n} \Delta_i(\mathbf{x},y), \qquad 0 \le j \le r_n - 1
$$

and

$$
L_j'(\mathbf{x},y) = \sum_{i=j(M_n+N_n)+M_n+1}^{(j+1)(M_n+N_n)} \Delta_i(\mathbf{x},y), \qquad 0 \le j \le r_n - 1.
$$

We first show that $\left(nh_n^d\right)^{-1/2}\left(T_{1,n}\left(\mathbf{x},y\right)+T_{2,n}\left(\mathbf{x},y\right)\right)\xrightarrow{\mathbf{P}}0$, as $n\to\infty$, next we state that $\left(nh_n^d\right)^{-1/2}S_n\left(\mathbf{x},y\right)$ converge in distribution to Gaussian variable with 0 mean and a variance given explicitly.

Lemma 4.4.9 *Under the assumptions of Lemma 4.4.7 and assumption* $(H3:a)$*, we have*

$$\left(nh_n^d\right)^{-1/2}\left(T_{1,n}\left(\mathbf{x},y\right)+T_{2,n}\left(\mathbf{x},y\right)\right)\xrightarrow{\mathbf{P}}0,\qquad as\ \ n\to\infty. \tag{4.32}$$

Proof. By Markov inequality, it suffices to show that

$$\frac{1}{nh_n^d}\mathbf{E}\left[T_{1,n}^2\left(\mathbf{x},y\right)+T_{2,n}^2\left(\mathbf{x},y\right)\right]\to 0,\quad \text{as } n\to\infty.$$

First, observe that

$$\begin{aligned}
\frac{1}{nh_n^d}\mathbf{E}\left[T_{1,n}^2\left(\mathbf{x},y\right)\right] &= \frac{1}{nh_n^d}\sum_{j=1}^{r_n}\mathbf{E}\left[\left(L_j'\left(\mathbf{x},y\right)\right)^2\right]+\frac{2}{nh_n^d}\sum_{1\le i<j\le r_n}\mathbf{E}\left[L_i'\left(\mathbf{x},y\right)L_j'\left(\mathbf{x},y\right)\right]\\
&= \frac{r_nM_n}{nh_n^d}\mathbf{E}\left[\Delta_1^2\left(\mathbf{x},y\right)\right]+\frac{2}{nh_n^d}\sum_{j=1}^{r_n}\sum_{1\le i<j\le M_n}\mathbf{E}\left[\Delta_i\left(\mathbf{x},y\right)\Delta_j\left(\mathbf{x},y\right)\right]\\
&+\frac{2}{nh_n^d}\sum_{1\le k<l\le r_n}\mathbf{E}\left[\left(\sum_{i=k(M_n+N_n)+1}^{(k+1)(M_n+N_n)}\Delta_i\left(\mathbf{x},y\right)\right)\left(\sum_{j=l(M_n+N_n)+1}^{(l+1)(M_n+N_n)}\Delta_j\left(\mathbf{x},y\right)\right)\right]\\
&\le \frac{r_nM_n}{nh_n^d}\mathbf{E}\left[\Delta_1^2\left(\mathbf{x},y\right)\right]+\frac{2r_n}{nh_n^d}\sum_{1\le i<j\le n}\mathbf{E}\left|\Delta_i\left(\mathbf{x},y\right)\Delta_j\left(\mathbf{x},y\right)\right|\\
&+\frac{2}{nh_n^d}\sum_{1\le i<j\le n}\mathbf{E}\left|\Delta_i\left(\mathbf{x},y\right)\Delta_j\left(\mathbf{x},y\right)\right|.
\end{aligned}\tag{4.33}$$

Using $(H3:a)$ and (4.22), the first term in the right hand side of (4.33) converge to zero as $n\to\infty$. By Lemma 4.4.7, the second and last terms in (4.33) converge to zero as $n \to \infty$. Likewise, we have

$$\begin{aligned}
\frac{1}{nh_n^d}\mathbf{E}\left[T_{2,n}^2(\mathbf{x},y)\right] &= \frac{1}{nh_n^d}\mathbf{E}\left[\left(\sum_{i=(M_n+N_n)r_n+1}^{n}\Delta_i(\mathbf{x},y)\right)^2\right] \\
&= \frac{1}{nh_n^d}\left(n-(M_n+N_n)r_n\right)\mathbf{E}\left[\Delta_1^2(\mathbf{x},y)\right] \\
&+ \frac{2}{nh_n^d}\sum_{1\le i<j\le n-(M_n+N_n)r_n}\mathbf{E}\left[\Delta_i(\mathbf{x},y)\Delta_j(\mathbf{x},y)\right] \\
&\le \left(1-\frac{r_nM_n}{n}+\frac{r_nN_n}{n}\right)\frac{1}{h_n^d}\mathbf{E}\left[\Delta_1^2(\mathbf{x},y)\right] \\
&+ \frac{2}{nh_n^d}\sum_{1\le i<j\le n}\mathbf{E}\left|\Delta_i(\mathbf{x},y)\Delta_j(\mathbf{x},y)\right|.
\end{aligned}$$

Therefore, (4.22), Assumption $(H3:a)$ and Lemma 4.4.7 give the result. □

Lemma 4.4.10 *Under the assumptions of Lemma 4.4.9, we have*

$$\frac{1}{nh_n^d}\sum_{j=1}^{r_n}Var\left(L_j(\mathbf{x},y)\right) \to \mu^{-1}\kappa c^T\Sigma(\mathbf{x},y)\,c, \qquad as\ \ n\to\infty. \tag{4.34}$$

Proof. From (4.31), we have

$$\left(nh_n^d\right)^{-1/2}S_n(\mathbf{x},y) = \left(nh_n^d\right)^{-1/2}\sum_{i=1}^{n}\Delta_i(\mathbf{x},y) - \left(nh_n^d\right)^{-1/2}\left(T_{1,n}(\mathbf{x},y)+T_{2,n}(\mathbf{x},y)\right).$$

Using (4.22) and Lemma 4.4.9, we get

$$\frac{1}{nh_n^d}Var\left(S_n(\mathbf{x},y)\right) \to \mu^{-1}\kappa c^T\Sigma(\mathbf{x},y)\,c, \quad \text{as } n\to\infty. \tag{4.35}$$

On the other hand,

$$\frac{1}{nh_n^d}Var\left(S_n(\mathbf{x},y)\right) = \frac{1}{nh_n^d}\sum_{j=1}^{r_n}Var\left(L_j(\mathbf{x},y)\right) + \frac{2}{nh_n^d}\sum_{1\le i<j\le r_n}\mathbf{E}\left[L_i(\mathbf{x},y)L_j(\mathbf{x},y)\right]$$

where

$$\begin{aligned}
\frac{1}{nh_n^d}\sum_{1\le i<j\le r_n}\mathbf{E}\left[L_i(\mathbf{x},y)L_j(\mathbf{x},y)\right] &= \frac{1}{nh_n^d}\sum_{1\le i<j\le r_n}\mathbf{E}\left[\left(\sum_{k=i(M_n+N_n)+1}^{i(M_n+N_n)+M_n}\Delta_k(\mathbf{x},y)\right)\right. \\
&\times\left.\left(\sum_{l=j(M_n+N_n)+1}^{j(M_n+N_n)+M_n}\Delta_l(\mathbf{x},y)\right)\right] \\
&\le \frac{1}{nh_n^d}\sum_{1\le i<j\le n}\mathbf{E}\left|\Delta_i(\mathbf{x},y)\Delta_j(\mathbf{x},y)\right|.
\end{aligned}$$

Lemma 4.4.7 entails the desired result. □

Lemma 4.4.11 *Under the assumptions of Lemma 4.4.7and condition* $(H3)$, *we have*

$$\left(nh_n^d\right)^{-1/2} S_n(\mathbf{x},y) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0\ ,\ \mu^{-1}\kappa c^T\Sigma(\mathbf{x},y)\,c\right), \qquad as\ \ n\to\infty. \tag{4.36}$$

Proof. The proof is based on the approximation of $\left(nh_n^d\right)^{-1/2} S_n(\mathbf{x},y)$ by a sum of independent rv's. Let $Z_{n1}(\mathbf{x},y),...,Z_{nr_n}(\mathbf{x},y)$ be a sequence of iid rv's of the same distribution as $\left(nh_n^d\right)^{-1/2} L_1(\mathbf{x},y)$. Denote by $\phi_n(t)$ the characteristic function (cf) of $L_1(\mathbf{x},y)$. It follows that the cf of $\sum_{j=1}^{r_n} Z_{nj}(\mathbf{x},y)$ is $\phi_n^{r_n}\left(t\left(nh_n^d\right)^{-1/2}\right)$. Using inequality due to Roussas and Ioannides [69, Theorem 7.2], we have

$$\left|\mathbf{E}\left[\prod_{j=1}^{r_n}\exp\left\{it\left(nh_n^d\right)^{-1/2}L_i(\mathbf{x},y)\right\}\right]-\prod_{j=1}^{r_n}\mathbf{E}\left[\exp\left\{it\left(nh_n^d\right)^{-1/2}L_i(\mathbf{x},y)\right\}\right]\right|$$
$$\leq 16\,(r_n-1)\,\alpha(N_n),$$

where r_n and q_n are defined in $(H3)$. So, by $(H3:c)$

$$\left|\mathbf{E}\left[\prod_{j=1}^{r_n}\exp\left\{it\left(nh_n^d\right)^{-1/2}L_i(\mathbf{x},y)\right\}\right]-\phi_n^{r_n}\left(t\left(nh_n^d\right)^{-1/2}\right)\right|\to 0, \quad \text{as } n\to\infty$$

and the rv's $\sum_{j=1}^{r_n} Z_{nj}(\mathbf{x},y)$ and $\left(nh_n^d\right)^{-1/2} S_n(\mathbf{x},y)$ have the same asymptotic distribution. Hence, it suffices to show that $\phi_n^{r_n}\left(t\left(nh_n^d\right)^{-1/2}\right)$ converge to the cf of $\mathcal{N}\left(0\ ,\ \mu^{-1}\kappa c^T\Sigma(\mathbf{x},y)\,c\right)$ as $n\to\infty$. To this end, set

$$s_n^2(\mathbf{x},y)=\sum_{j=1}^{r_n} Var\left(Z_{nj}(\mathbf{x},y)\right)=r_n Var\left(Z_{n1}(\mathbf{x},y)\right)=\frac{r_n}{nh_n^d}Var\left(L_1(\mathbf{x},y)\right)$$

and

$$\tilde{Z}_{nj}(\mathbf{x},y)=\frac{Z_{nj}(\mathbf{x},y)}{s_n(\mathbf{x},y)}, \quad j=1,...,r_n.$$

Clearly, $\mathbf{E}\left[\tilde{Z}_{nj}(\mathbf{x},y)\right]=0$ and $\sum_{j=1}^{r_n} Var\left(\tilde{Z}_{nj}(\mathbf{x},y)\right)=1$.
In order to state the asymptotic normality, we have to show that the Lindberg condition is satisfied for the sequence $\left(\tilde{Z}_{nj}(\mathbf{x},y)\right)$, that is,

$$\forall\varepsilon>0, \qquad \varphi_n(\varepsilon):=\sum_{j=1}^{r_n}\int_{\{|z|\geq\varepsilon\}} z^2 dF_{nj}(z)\to 0, \quad \text{as } n\to\infty$$

where F_{nj} is the df of $\tilde{Z}_{nj}(\mathbf{x},y)$, noted F_{n1} because it is the same for all $j = 1,...,r_n$. Firstly, we have

$$\begin{aligned}\varphi_n(\varepsilon) &= r_n \int_{\{|z|\geq\varepsilon\}} z^2 dF_{n1}(z) \\ &= r_n \mathbf{E}\left[\tilde{Z}_{n1}^2(\mathbf{x},y) \times \mathbf{1}_{\left\{\left|\tilde{Z}_{n1}(\mathbf{x},y)\right|\geq\varepsilon\right\}}\right].\end{aligned}$$

On the other hand, we have

$$\Delta_i(\mathbf{x},y) \leq \frac{2c_1 M_d}{G(a_F)} + \frac{2c_2 M_d M_0}{G(a_F)} \leq C < \infty$$

Therefore, by Markov inequality, we get

$$\begin{aligned}\varphi_n(\varepsilon) &= \frac{r_n}{nh_n^d s_n^2(\mathbf{x},y)} \mathbf{E}\left[L_1^2(\mathbf{x},y) \times \mathbf{1}_{\left\{\left|\tilde{Z}_{n1}(\mathbf{x},y)\right|\geq\varepsilon\left(nh_n^d\right)^{1/2} s_n(\mathbf{x},y)\right\}}\right] \\ &\leq \frac{r_n^2 M_n^2 C^2}{nh_n^d s_n^2(\mathbf{x},y)} \mathbf{P}\left\{|L_1(\mathbf{x},y)| \geq \varepsilon\left(nh_n^d\right)^{1/2} s_n(\mathbf{x},y)\right\} \\ &\leq \frac{C^2}{\varepsilon^2} \frac{nh_n^d}{r_n Var(L_1(\mathbf{x},y))} \frac{M_n^2}{nh_n^d}. \end{aligned} \tag{4.37}$$

By $(H: b, c)$, the last term in (4.37) tends to zero as $n \to \infty$ and using Lemma 4.4.9 give the result. □

Lemma 4.4.12 *Under the assumptions of Lemma 4.4.11, we have*

$$\sqrt{nh_n^d}\left(c_1\Gamma_{n2}(\mathbf{x}) + c_2\Lambda_{n2}(\mathbf{x},y)\right)^T \xrightarrow{\mathcal{D}} \mathcal{N}\left(0\ ,\ \mu^{-1}\kappa\Sigma(\mathbf{x},y)\right)$$

Proof. It suffices to use the Cramér-Wold device and Lemmas 4.4.6-4.4.11 to get the result. □

Proof of Proposition 4.2.7 Consider the mapping Θ from $\mathbb{R}$ to $\mathbb{R}$ defined by $\Theta(\mathbf{x},y) = \mathbf{x}/y$ for $y \neq 0$. Since $\mathbf{F}_n(y|\mathbf{x})$ and $\mathbf{F}(y|\mathbf{x})$ are the respective images of $\mu_n^{-1}(\mathbf{F}_n(y|\mathbf{x}), v_n(\mathbf{x}))$ and $\mu^{-1}(\mathbf{F}(y|\mathbf{x}), v(\mathbf{x}))$ by Θ. We deduce from Lemmas 4.4.6-4.4.12 and from Mann-Wold's Theorem (see Rao [64, p 321]) that $\sqrt{nh_n^d}(\mathbf{F}_n(y|\mathbf{x}) - \mathbf{F}(y|\mathbf{x}))$ converge in distribution to $\mathcal{N}\left(0,\ \mu^{-1}\kappa\nabla\Theta^T\Sigma(\mathbf{x},y)\nabla\Theta\right)$, where the gradient $\nabla\Theta$ is evaluated at $\mu^{-1}(\mathbf{F}(y|\mathbf{x}), v(\mathbf{x}))$. Simple algebra gives then the variance $\sigma^2(\mathbf{x},y)$. □

We embark now on the proof of Theorem 4.2.8.

Proof of Theorem 4.2.8. We make use of the property

$$\mathbf{F}(q_p(\mathbf{x})|\mathbf{x}) = p = \mathbf{F}_n(q_{p,n}(\mathbf{x})|\mathbf{x}).$$

A Taylor expansion of $\mathbf{F}_n(.|.)$ in neighborhood of q_p, implies that

$$q_{p,n}(\mathbf{x}) - q_p(\mathbf{x}) = \frac{\mathbf{F}(q_p(\mathbf{x})|\mathbf{x}) - \mathbf{F}_n(q_{p,n}(\mathbf{x})|\mathbf{x})}{f_n(\tilde{q}_{p,n}(\mathbf{x})|\mathbf{x})}$$

where $\tilde{q}_{p,n}$ is between q_p and $q_{p,n}$. The continuity of $f(.|\mathbf{x})$, the almost sure convergence of $q_{p,n}(\mathbf{x})$ to $q_p(\mathbf{x})$ (see [58, Theorem 3.1]) and Proposition 4.2.6 imply the convergence in probability of the above denominator to $f(q_p(\mathbf{x})|\mathbf{x})$. Proposition 4.2.7 is used to finish the proof. $\square$

Remark 4.4.13 *If the condition $f_n(\tilde{q}_{p,n}(\mathbf{x})|\mathbf{x}) \neq 0$ was not satisfied, we should have increase the order of Taylor expansion and to modify the proofs of the Theorem accordingly. Furthermore, we point out that our assumptions contain those in Ould Saïd et al.[58] which permit us to get the convergence of the conditional quantile estimator.*

Conclusion

In this thesis, we establish the uniform almost sure convergence and asymptotic normality of the estimator based on conditional quantiles for truncated and dependent data. Conditional medians and quantiles are frequently used in analyzing time series data with heavy tails for their robustness properties. Although our interest in conditional quantile estimation is motivated by the constructing of the confidence intervals and the forecasting from time series data.

To study a statistical model more practical for several applications, we are interested in the context of the left-truncated data. We therefore sought to relax this assumption by considering a form of dependency. We made the choice of alpha mixing, this type of dependency modeling many processes in particular are strongly mixing.

Our results are derived in a more general setting (strong mixing) which includes time series modeling as a special case. It is assumed that the lifetime observations with multivariate covariates from a stationary strong mixing process.

The progress of the quality of results is linked to that of probabilistic tools in particular that of exponential inequalities (i.e., Fuk-Nagaev inequality). The choice of this type of inequality and its use in the case of dependent data is justified by the fact that it can be adapted better and poses fewer technical problems than Bosq [7] or that of Bernstein type.

Appendix A. Cramér-Wold device

The characteristic function $t \to \mathbf{E}\left[\exp\left(it^T X\right)\right]$ of a vector X is determined by the set of all characteristic function $u \to \mathbf{E}\left[\exp\left(iu\left(t^T X\right)\right)\right]$ of all linear combinations $t^T X$ of the components of X. Therefore the continuity theorem implies that the weak convergence of vectors is equivalent to weak convergence of linear combinations.

$X_n \rightsquigarrow X$ if and only if $t^T X_n \rightsquigarrow t^T X$ for all $t \in \mathbb{R}^k$. This is known as the Cramér-Wold device. It allows to reduce all higher dimensional weak convergence problems to the one-dimensional case.
Example (Multivariate central limit theorem) *Let Y, Y_1, Y_2 be iid random vectors in IR^k with mean vector $\mu = E[Y]$ and covariance matrix* $\Sigma = E\left[(Y-\mu)(Y-\mu)^T\right]$. *Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Y_i - \mu) = \sqrt{n}\left(\bar{Y}_n - \mu\right) \xrightarrow{\mathcal{D}} \mathcal{N}_k(0, \Sigma), \quad \text{as } n \to \infty.$$

By the Cramér-Wold device the problem can be reduced to finding the limit distribution of the sequences of real-variables

$$t^T\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Y_i - \mu)\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(t^T Y_i - t^T\mu\right).$$

Since the random variable $t^T Y_1 - t^T\mu$, $t^T Y_2 - t^T\mu, \cdots$ are iid with zero mean and variance $t^T \Sigma t$, this sequence is asymptotically $\mathcal{N}_1\left(0, t^T\Sigma t\right)$.

Appendix B. Stochastic o and O symbols

It is convenient to have short expressions for terms that converge in probability to zero or are uniformly tight. The notation $o_p(1)$ (*'small "oh-P-one"'*) is short for a sequence of random vectors that converges to zero in probability. The expression $O_p(1)$ (*"big "oh-P-one"'*) denotes a sequence that is bounded in probability. More generally, for a given sequence of random variables R_n

$$X_n = o_p(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \text{and } Y_n \xrightarrow{P} 0,$$

$$X_n = O_p(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \text{and } Y_n = O_p(1).$$

This expresses that the sequence X_n converges in probability to zero or bounded in probability at *"rate"* R_n. For deterministic sequences X_n and R_n the stochastic oh-symbols reduce to usual o and O symbols, which will be applied without comment. For instance,

$$\begin{array}{ll} o_p(1) + o_p(1) = o_p(1), & o_p(1) + O_p(1) = O_p(1), \\ o_p(1) O_p(1) = o_p(1), & (1 + o_p(1))^{-1} = O_p(1), \\ o_p(R_n) = R_n o_p(1), & O_p(R_n) = R_n O_p(1), \\ o_p(O_p(1)) = o_p(1). & \end{array}$$

To see the validity of these "rules" it suffices to restate them in terms of explicitly named vectors, where each $o_p(1)$ and $O_p(1)$ should be replaced by a different sequence vectors that converge to zero or is bounded in probability. In this manner the first rule says; if $X_n \xrightarrow{P} 0$ and $Y_n \xrightarrow{P} 0$, then $X_n + Y_n \xrightarrow{P} 0$, this is an example of the continuous mapping theorem. The third rule is short for, if X_n is bounded in probability and $Y_n \xrightarrow{P} 0$, then $X_n Y_n \xrightarrow{P} 0$.

Appendix C. Notations and abbreviations

$a.s.$	almost sure convergence
rv	random variable
cf	characteristic function
iid	independent and identically distributed
cdf	cumulative distribution functions
RLT	random left-truncation
$\mathcal{Y}$	random variable of interest
$\mathcal{T}$	truncation random variable
$\mathcal{X}$	random covariable
$(\mathcal{X}, \mathcal{Y}, \mathcal{T})$	complete sample
N	sample size
(X, Y, T)	observed subsequence subject to $Y \geq T$
n	size of observed sample
μ	truncation probability
$\mathbb{P}$	probability measure related to the $N - sample$
$\mathbb{E}$	expectation operators related to $\mathbb{P}$
$\mathbf{P}$	the conditional probability related to the $n - sample$
$\mathbf{E}$	expectation operators related to $\mathbf{P}$
$\alpha\,(n)$	$\alpha -$ mixing coefficient
$q_p(x)$	p^{th} conditional quantile

Bibliography

[1] Auestad, B., Tjϕstheim, D. (1990) Identification of nonlinear time series: First order characterization and order determination, *Biometrika* **77**, 669–687.

[2] Battacharya, P.K., Gangopadhay, A. (1990) Asymptotic normality of convergent estimates of conditional quantiles, *Statistics* **35**, 139–168.

[3] Beran, R. (1981) Nonparametric regression with randomly truncated survival data, Technical Report, Departement of Statistics, Univ. of California, Berkley, CA.

[4] Berlinet, A. Gannoun, A. Matzner-Lϕber, E. (2001) Asymptotic normality of convergent estimates of conditional quantiles, *Statistics*, **35**, 139–169.

[5] Blum, J.R., Hanson, D.L., Koopmans, L.H. (1963) On the strong law of large numbers for a class of stochastic processes, *Z. Wahrsch. verw. Gebiete* **2**, 1-11.

[6] Boente, G., Fraiman, R. (1995) Asymptotic distribution of smoothers based on local means and local medians under dependence, *J. Multivariate Anal.* **54**, 77–90.

[7] Bosq, D. (1998) *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction* (Second Edition), Lecture Notes in Statistics, 110, Springer Verlag, New York.

[8] Bradley, R.C. (2007) *Introduction to strong mixing conditions;* Vol 1-3, Kendrick Press, Utah.

[9] Cai, Z.W. Roussas, G.G. (1992) Uniform strong estimation under $\alpha-$mixing, with rate, *Statist. Probab. Lett.*, **15**, 47–55.

[10] Cai, Z.W. (1998) Kernel density and hazard rate estimation for censored dependent data, *J.Multivariate Anal.* **67**, 23–34.

[11] Cai, Z.W. (2001) Estimating a distribution function for censored time series data, *J.Multivariate Anal.* **78**, 299-318.

[12] Carbonnez, A. Györfi, L., Van Meulen, E.C. (1995) Partition estimates of regression function under random consoring, *Statist. Decisions* **13**, 31–37.

[13] Chaudhuri, P., Doksum, K., Samarov, A. (1997). On average derivative quantile regression, *Ann. Statist.* **25**, 715–744.

[14] Cogburn, R. (1960) Asymptotic properties of stationary sequences, *Univ. Calif. Publ. Statist.* **3**, 99-146.

[15] Dabrowska, D. (1992) Nonparametric quantile regression with censored data, *Sankhyã Ser. A* **54**, 252–259.

[16] Dedecker, J., Doukhan, P., Lang, G., Leon, J.R., Louhichi, S., Prieur, C. (2007) *Weak Dependence: with Examples and Applications*, Lecture Notes in Statistics **190**, Springer-Verlag, New York.

[17] Doob, J.L. (1953) *Stochastic Processes,* Wiley, New York.

[18] Doukhan, P. (1994) *Mixing: Properties and examples,* Lecture Notes in Statistics **85**, Springer-Verlag, New York.

[19] Fan, J., Hu, T.C. and Troung, Y.K. (1994) Robust nonparametric function estimation, *Scand. J. Statist.* **21**, 433–446.

[20] Feigelson, E.D. Babu, G.J. (1992) *Statistical Challenges in Modern Astronomy*, Berlin Heidelberg, Springer-Verlag, New York.

[21] Ferraty, F., Rabhi, A., Vieu,P. (2005) Conditional quantiles for dependent functional data with application to the climatic *El Ninõ* phenomenon, *The Indian J.Statist.* **65**, 378–398.

[22] Ferraty, F. Vieu, P. (2006). *Nonparametric Functional Data Analysis, Theory and Practice.* Springer-Verlag, New York.

[23] Gannoun, A. (1989) *Estimation de la médiane conditionnelle,* Thèse de doctorat de l'Université de Paris VI.

[24] Gannoun, A. (1990) Estimation non paramétrique de la médiane conditionnelle médianogramme et méthode du noyau, *Publication de l'Institut de Statistique de l'université de Paris,* 11-22

[25] Gannoun, A. (1991) Estimation de la médiane conditionnelle, *C. R. Acad. Sci. Série I.*

[26] Gannoun, A., Saracco, J., Yu, K. (2003) Nonparametric prediction by conditional median and quantiles, *J.Statist. Plann. Inference* **117**, 207–223.

[27] Gannoun, A., Saracco, J., Yuan, A. and Bonnet, G.E. (2005) Nonparametric quantile regression with censored data, *Scand.J.Statist.* **32**, 527–550.

[28] Gross, S.T., Lai, T.L. (1996) Nonparametric estimation regression analysis with left truncated and censored data, *J. Amer. Statist. Assoc.* **91**, 1166–1180.

[29] Gorodetski, V.V. (1977) On the strong mixing property for linear sequences, *Theo. Prob. and its Applic.* **22**, 411–413.

[30] Gürler, U. Stute, W. Wang, J.L. (1993) Weak and strong quantile representations for randomly truncated data with applications, *Statist. Probab. Lett.*, **17**, 139–148.

[31] Hall, P., Heyde, C.C. (1980) *Martingale Limit Theory and its Application*, Academic Press, New York.

[32] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986) *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.

[33] He, S., Yang, G. (1994) Estimating a lifetime distribution under different sampling plan, In S.S. Gupta J.O. Berger (Eds.) *Statistical Decision Theory and Related Topics* (Vol. **5**, pp. 73-85), Berlin Heidelberg, Springer-Verlag, New York.

[34] He, S., Yang, G. (1998) Estimation of the truncation probability in the random truncation model, *Ann. Statist.*, **26**, 1011–1027.

[35] He, S., Yang, G. (2003) Estimation of regression parameters with left truncated data, *J.Statist. Plann. Inference* **117**, 99–122.

[36] Hounda, T. (2000) Nonparametric estimation of a conditional quantile for $\alpha-$mixing processes, *Ann. Inst.Statist.Math.* **52**, 459–470.

[37] Ibragimov, I.A. (1959) Some limit theorems for stochastic processes stationary in the strict sense, *Dokl. Akad. Nauk SSSR* **125**, 711-714.

[38] Ibragimov, I.A. (1962) Some limit theorems for stationary processes, *Theor. Probab. Appl.* **7**, 349–382.

[39] Jones, M.C., Hall, P. (1990) Mean squared error properties of kernel estimates of regression quantiles, *Statist. Prob. Lett.* **10**, 283–289.

[40] Kohler, M., Máthé, K., Pintér, M. (2002) Prediction from randomly right censored data, *J.Multivariate Anal.* **80**, 73–100.

[41] Kolmogorov, A.N. Rozanov, Yu.A. (1960) On strong mixing conditions for stationary Gaussian processes, *Theor. Probab. Appl.* **5**, 204-208.

[42] Liang H.-Y., Li D., Qi Y. (2009) Strong convergence in nonparametric regression with truncated dependent data, *J.Multivariat. Anal.* **100**, 162–174.

[43] Lecoutre, J.P., Ould-Saïd, E. (1995) Convergence of the conditional Kaplan-Meier estimate under strong mixing, *J.Statist. Plann. Inference* **44**, 359–369.

[44] Lemdani, M., Ould-Saïd, E. (2007) Asymptotic behavior of the hazard rate kernel estimator under truncated and censored data, *Comm. in Statist. Theory & Methods* **37**, 155–174.

[45] Lemdani, M. Ould Saïd, E. Poulin, N. (2005) Strong representation of the quantile function for left-truncated and dependent data, *Math. Meth. Statist.* **14**, 332–345.

[46] Lemdani, M. Ould Saïd, E. Poulin, N. (2009) Asymptotic properties of a conditional quantile estimator with randomly truncated data, *J. Multivariate Anal.* **100**, 546–559.

[47] Louani, D. Ould Saïd, E. (1999) Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis, *J. Nonparametric Statist.*, **11**, 413–442.

[48] Lynden-Bell, D. (1971) A method of allowing for known observation selection in small samples applied to 3CR quasars, *Monthly Notices Roy. Astron. Soc.*, **155**, 95–118.

[49] Masry, E. (1986) Recursive probability density estimation for weakly dependent processes, *IEEE Transactions on Information Theory* **32**, 254–267.

[50] Masry, E., Tjϕstheim, D. (1995) Nonparametric estimation and identification of nonlinear time series, *Econometric Theory* **11**, 258–289.

[51] Masry, E., Tjϕstheim, D. (1997) Additive nonlinear ARX time series and projection estimates, *Econometric Theory* **13**, 214–252.

[52] Mehra, K.L., Rao, M.S., Upadrasta, S.P. (1991) A smooth conditional quantile estimator and related applications of conditional empirical processes, *J.Multivariate Anal.* **37**, 151–179

[53] Merlevede, F., Peligrad, M. (2002) On the coupling of dependent random variables and applications, In: *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sfirensen, eds.), pp. 171-193. Birkhauser, Boston.

[54] Ould-Saïd, E. (2006) A strong uniform convergence rate of kernel conditional quantile estimator under random censorship, *Statist. Prob. Lett.* **76**, 579–586.

[55] Ould-Saïd, E., Lemdani, M. (2006) Asymptotic properties of a nonparametric regression function estimator with randomly truncated data, *Ann. Instit. Statist. Math.* **58**, 357–378.

[56] Ould-Saïd, E., Sadki, O., (2008) Prediction via the conditional quantile for right censorship model, *Far East J. Theo. Statist.* **25**, N°2, 145–179.

[57] Ould Saïd, E., Tatachak, A. (2009) Strong uniform consistency rate for the kernel mode under strong mixing hypothesis and left-truncation, *Comm. in Statist. Theory and Methods* **37**, 2735-2759.

[58] Ould Saïd, E. Yahia, D. Necir, A. (2009) A strong uniform convergence rate of a kernel conditional quantile estimator under random left-truncation and dependent data, *Elect. J. Statist.* **3**, 426-445.

[59] Peligrad, M. (1983) A note on two measures of dependence and mixing sequences. *Adv. Appl. Prob.* **15**, 461-464.

[60] Petersen, K. (1989) *Ergodic Theory*, first paperback edition. Cambridge University Press, Cambridge.

[61] Philipp, W. (1969) The central limit problem for mixing sequences of random variables, *Z. Wahrsch. verw. Gebiete* **12**, 155-171.

[62] Philipp, W. (1986) Invariance principles for independent and weakly dependent random variables, In: *Dependence in Probability and Statistics*, (E. Eberlein and M.S. Taqqu, eds.), pp. 225-268. Birkhauser, Boston.

[63] Qin,Y.S., Wu, Y. (2001) An estimator of a conditional quantile in the presence of auxiliary information, *J. Statist. Plann. Inference* **99**, 59–70.

[64] Rao, C.R. (1965) *A Linear Statistical Inference and its Applications,* Wiley, New York.

[65] Rio, E. (1994) Covariance inequalities for strongly mixing processes, *Ann. de l'I.H.P. Probab-Statist.*, **29**, 587–597.

[66] Rio, E. (2000) *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants* (In French), *Mathématiques et Applications*, **31**, Springer-Verlag, New York.

[67] Rosenblatt, M. (1956) A central limit theorem and a strong mixing condition, *Proc. Natl. Acad. Sci. USA* **42**, 43-47.

[68] Rosenblatt, M. (2000) *Gaussian and Non-Gaussian Linear Time Series and Random Fields*, Springer-Verlag, New York

[69] Roussas, G.G. (1969) Nonparametric estimation of the transition distribution function of a Markov process, *Ann. Math. Statist.* **40**, 1386–140.

[70] Roussas, G.G. Ioannides, D. (1987) Moment inequalities for mixing sequences of random variables, *Stochastic Analysis* **5**, 61–120.

[71] Samanta, M. (1989) Nonparametric estimation of conditional quantiles, *Statist. Prob. Lett.* **7**, 407–412.

[72] Schlee, W. (1982) Estimation non paramétrique du $\alpha-$quantile conditionnel, *Statistique et Analyse des Données* **1**, 32–47.

[73] Stone, C. (1977) Consistent nonparametric regression, *Ann. Statist.* **5**, 595–645.

[74] Stute, W. (1993) Almost sure representation of the product-limit estimator for truncated data, *Ann. Statit.* **21**, 146–156.

[75] Stute, W. (1996) Distributional convergence under random censorship when covariates are present, *Scand.J.Statist.* **23**, 461–471.

[76] Sun, L., Zhou, X. (2001) Survival function and density estimation for truncated dependent data, *Statit. Prob. Lett.* **52**, 47–57.

[77] Tsai, W.Y., Jewel,l N.P., Wang, M.C. (1987) A note on the product-limit estimator under right censoring and left truncation, *Biometrika* **74**, 883–886.

[78] Volkonskii,V.A., Rozanov, Yu.A. (1959) Some limit theorems for random functions I, *Theor. Probab. Appl.* **4**, 178-197.

[79] Volkonskii,V.A., Rozanov, Yu.A. (1961) Some limit theorems for random functions II, *Theor. Probab. Appl.* **6**, 186-198.

[80] Van Keilegom, I., Veraverbeke, N. (1996) Uniform strong convergence resulte for the conditional Kaplain-Meier estimator and ist quantiles, *Comm. in Statist. Theory and Methods* **25**, 2251-2265.

[81] Van Keilegom, I., Veraverbeke, N. (1998) Bootstrapping quantiles in a fixed design regression model with censored data, *J.Statist. Plann. Inference* **69**, 115–131.

[82] Van Keilegom, I., Veraverbeke, N. (2001) hazard rate estimation in nonparametric regression model with censored data, *Ann. Inst. Statist. Math.* **04**, 730–745.

[83] Wang, M.C., Jewell, N.P., Tsai, W.Y. (1986) Asymptotic properties of the product-limit estimate under random truncation, *Ann. Statist.* **14**, 1597–1605.

[84] Welsh, A.H. (1996) Robust estimation of smooth and spread functions and derivatives, *Statist. Sin.* **6**, 730–745.

[85] Withers, C.S. (1981) Conditions for linear processes to be strong mixing, *Z. Wahrsch. G.*, **57**, 477-480.

[86] Woodroofe, M. (1985) Estimating a distribution function with truncated data, *Ann. Statist.*, **13**, 163–177.

[87] Xiang, X. (1995) Deficiency of the sample quantile estimator with respect to kernel quantile estimators for censored data, *Ann. Statist.* **23**, 836–854.

[88] Xiang, X. (1996) A kernel estimator of a conditional quantile, *J.Multivariate Anal.* **59**, 206–216.

[89] Zhurbenko, I.G. (1986) *The Spectral Analysis of Time Series*, North-Holland, Amsterdam, 1986.

Abstract

In this thesis we study some asymptotic properties of the kernel conditional quantile estimator when the interest variable is subject to random left truncation. The uniform strong convergence rate of the estimator is obtained. In addition, it is shown that, under regularity conditions and suitably normalized, the kernel estimate of the conditional quantile is asymptotically normally distributed.

Our interest in conditional quantile estimation is motivated by it's robusteness, the constructing of the confidence bands and the forecasting from time series data. Our results are obtained in a more general setting (strong mixing) which includes time series modelling as a special case.

Résumé

Dans cette thèse nous étudions certaines propriétés asymptotiques de l'estimateur à noyau du quantile conditionnel lorsque la variable d'intérêt est soumise à une troncature aléatoire à gauche. La convergence uniforme presque sûre avec vitesse de l'estimateur est obtenue. En outre, il est démontré que, sous des conditions de régularité, l'estimateur à noyau du quantile conditionnel convenablement normalisé est asymptotiquement normal.

L'intérêt principal dans l'étude de l'estimation des quantiles conditionnels est sa robustesse, la construction des intervalles de confiance et la prévision à partir des données de séries chronologiques. Nos résultats sont obtenus dans un cadre général (mélangeance forte), qui inclut des modèles populaires de séries financières et économétriques comme cas particulier.