

وزارة التعليم العالي والبحث العلمي

Université Badji Mokhtar
Annaba

Badji Mokhtar University –
Annaba



جامعة باجي مختار
عنابة

Faculté des Sciences
Département de Mathématiques

THESE

Présenté en vue de l'obtention du diplôme de
Doctorat
Option : Probabilités et Statistiques

**Estimation de la fonction mode pour des données
tronquées et censurées**

Par:

CHAIB Yacine

Sous la direction de

Rapporteur: BOUTABIA Hacène

Co- Rapporteur : SADKI Ourida

Prof. Université Badji Mokhtar d'Annaba

M.C.A. Université d'Oum El Bouaghi

Devant le jury

PRESIDENTE : SEDDIK AMEUR

EXAMINATEUR : MOHDEB

EXAMINATRICE: CHADLI

EXAMINATEUR : TATACHAK

Nacira

Zaher

Asia

Abdelkader

Professeur

Professeur

MCA

MCA

Univ. Annaba

Univ. Constantine

Univ. Annaba

U.S.T.H.B.

Année : 2013

Résumé

Dans cette thèse, nous nous proposons d'étudier certaines propriétés asymptotiques de l'estimateur non paramétrique du mode simple et conditionnel, lorsque la variable d'intérêt est soumise à une censure aléatoire à droite et une troncature aléatoire à gauche.

Dans un premier temps, nous considérons une suite d'observations i.i.d. et nous construisons des estimateurs par la méthode du noyau pour la fonction de mode. Nous établissons la convergence uniforme presque sûre et la normalité asymptotique de ces estimateurs.

Dans un second temps, on généralise nos résultats au cas des données fortement mélangeantes. On établit également la convergence forte ainsi que la normalité asymptotique.

Finalement, nous considérons le mode conditionnel dans le modèle LTRC dans le cas i.i.d. et nous définissons un nouvel estimateur lissé. La convergence uniforme avec vitesse de l'estimateur est obtenue. On obtient également sous des conditions de régularité, la normalité asymptotique de l'estimateur à noyau du mode conditionnel.

Mots-Clés : Convergence uniforme forte ; Données tronquées et censurées ; Estimation à noyau ; Mélangeant forte, Fonction des modes ; Mode conditionnel ; Normalité asymptotique.

Abstract

In this thesis, we study some asymptotic properties of non parametric estimators of the simple and the conditional mode, when the interest variable is subject to random right censored and left truncation.

Firstly, we consider a sequence of i.i.d. observations. In this context, we build kernel estimators of the mode function. We state uniform strong convergence rate and asymptotic normality of kernel-based estimators of density and mode function.

Secondly, we generalize our results when the data strongly mixing. We obtain strong consistency with rates and we give an asymptotic normality.

Finally we introduce a new smooth estimator of the conditional mode function in the LTRC model in the i.i.d. case. The uniform convergence of the estimator is obtained. In addition, and under regularity conditions, the kernel estimate of the conditional mode is asymptotically normally distributed.

Keywords: Strong uniform consistency; Truncated and censored data; Kernel estimate; Strong mixing; Mode function; Conditional mode; Asymptotic normality.

Table des matières

1	Introduction	4
1.1	Motivation	4
1.2	Estimation de mode	6
1.3	Estimation du mode conditionnel	8
1.4	Les données incomplètes	11
1.5	Modèle de dépendance	12
1.6	Présentation des résultats obtenus	14
1.6.1	Mode simple et cas i.i.d.	14
1.6.2	Mode simple et cas α -mélangeant	15
1.6.3	Mode conditionnel et cas i.i.d.	16
1.7	Le plan de thèse	18
2	Données incomplètes et dépendantes	20
2.1	Données incomplètes	20
2.1.1	Les données censurées	21
2.1.2	Les données troncatures	23
2.1.3	Le modèle LTRC	24
2.2	Données dépendantes	25
2.2.1	Définitions et propriétés	25
2.2.2	Les conditions de mélange fort	27

3	La fonction de mode	30
3.1	Estimation de la fonction mode pour des données complètes . . .	31
3.2	Estimation non paramétrique de la fonction mode dans le cas des données censurées à droite	34
3.2.1	Modèle de censure	34
3.2.2	Estimation dans le modèle de censure	35
3.2.3	Le cas α -mélange	37
3.3	Estimation non paramétrique du mode dans le cas des données tronquées à gauche	38
3.3.1	Modèle de troncature	38
3.3.2	Estimation dans le modèle de troncature	40
3.3.3	Cas d' α -mélange	43
3.4	Estimation non paramétrique pour des données tronquées et cen- surées	45
3.4.1	Modèle LTRC	45
3.4.2	Estimation dans le modèle LTRC	47
4	Asymptotic properties of the mode kernel estimator under truncated and censored data	48
4.1	Introduction	48
4.2	Definition of the new estimator	50
4.3	Assumptions and main results	53
4.3.1	Consistency	54
4.3.2	Asymptotic normality	54
4.4	Auxiliary results and proofs	55
5	A nonparametric mode estimate under LTRC model and dependent data	67
5.1	Introduction	67

5.2	Definition of the estimator	70
5.3	Assumptions and main results	73
5.3.1	Consistency	74
5.3.2	Asymptotic normality	74
5.4	Auxiliary results and proofs	76
6	La fonction mode conditionnelle	93
6.1	Introduction	93
6.2	Cas de données complètes (α-mélange)	94
6.3	Cas de données censurées	95
6.4	Cas des données tronquées	96
6.5	Cas des données tronquées et censurées	98
6.5.1	Introduction	98
6.5.2	Definition of the new estimator	101
6.5.3	Assumptions and main results	104
6.5.4	Auxiliary results and proofs	107

Chapitre 1

Introduction

1.1 Motivation

L'analyse des données de survie a pour première particularité de ne concerner que des variables aléatoires positives (modélisant les durées de vie). Une conséquence de cette particularité est que la loi normale ne sera plus ici la référence en matière de distribution. Le plus souvent, toute autre loi issue de la famille exponentielle, et à support dans \mathbb{R}^+ , lui sera préférée. Une deuxième particularité de cette analyse est l'incomplétude des données différente de la troncature, qui équivaut à une perte d'information. Analysant la survenue d'un certain type d'événement, nous qualifierons de "donnée complète" un temps correspondant à l'observation de la survenue de l'événement, et de "donnée incomplète" un temps correspondant à l'absence d'observation de cet événement. Enfin, la troisième particularité de l'analyse des données de survie est la terminologie s'y rattachant. Outre les termes conçus à l'origine de cette analyse (tels que fonction de survie, fonction de risque), cette terminologie est essentiellement due au contexte épidémiologique. Concernant les modèles statistiques proprement dits, trois approches sont possibles : paramétrique, non-paramétrique et semi-paramétrique. L'approche paramétrique stipule l'appartenance de la loi de probabilité réelle des observations à une classe particulière de lois, qui dépendent d'un certain nombre (fini) de paramètres. L'avantage de cette approche est la facilitation

attendue de la phase d'estimation des paramètres, ainsi que de l'obtention d'intervalles de confiance et de la construction de tests. L'inconvénient de la méthode paramétrique est l'inadéquation pouvant exister entre le phénomène étudié et le modèle retenu. L'approche non-paramétrique ne nécessite aucune hypothèse quant à la loi de probabilité réelle des observations et c'est là son principal avantage. Il s'agit dès lors d'un problème d'estimation fonctionnelle, avec les ambiguïtés que cela implique par exemple, la fonction de survie, qui est continue, sera estimée par une fonction discontinue. L'inconvénient d'une telle approche est la nécessité de disposer d'un nombre important d'observations, le problème de l'estimation d'un paramètre fonctionnel étant délicat puisqu'il appartient à un espace de dimension infinie. L'approche semi-paramétrique est une sorte de compromis entre les deux approches précédentes. La loi de probabilité réelle des observations est supposée appartenir à une classe de lois pour partie dépendant de paramètres, et pour partie s'écrivant sous forme de fonction(s) non-paramétrique(s). Relativement récente elle est apparue au cours des années soixante-dix, cette approche est très répandue en analyse de la survie, notamment au travers du modèle de régression de Cox (1972).

Sachant que la tendance centrale d'un échantillon peut être quantifiée à partir des paramètres (moyenne, mode et médiane), la moyenne est le paramètre le plus utilisé, mais le mode, contrairement aux deux autres paramètres, possède la propriété de préserver sa définition face à une transformation monotone de la variable d'intérêt. Bien que ces trois mesures coïncident pour des distributions unimodales et symétriques, elles peuvent être totalement différentes pour des données observées. L'un des plus curieux résultats en statistique théorique classique est la relation empirique liant les trois mesures de position, mis en évidence pour des distributions continues, donnée par Pearson (1895)

$$\text{moyenne} - \text{mode} \approx 3(\text{moyenne} - \text{médiane}). \quad (1.1)$$

Le résultat est trivial pour une distribution symétrique et unimodale (gaussienne par exemple) et reste valide pour des distributions légèrement dissymétriques. A titre d'exemple, une distribution gamma de paramètre γ , a pour moyenne γ , pour mode $\gamma - 1$

et pour médiane $\gamma - \frac{1}{3} + O(\frac{1}{\gamma})$ lorsque $\gamma \rightarrow \infty$ (voir Hall (1980) et Asselin de Beauville (1978) pour plus de détails). De son côté, Fréchet (1940) a donné une borne supérieure du mode, fonction d'un majorant de la moyenne pour une variable discontinue non négative. Par exemple, pour une variable entière non négative, si sa moyenne est majorée par une constante positive a , alors son mode est majoré par $a(a + 3)/2$.

1.2 Estimation de mode

La fonction de mode simple est une caractéristique très importante en statistique et son utilité dans le domaine de la prévision a suscité l'intérêt d'un grand nombre d'auteurs. Théoriquement, le mode est lié directement à la fonction densité f , défini comme la valeur θ qui la maximise soit $f(\theta) = \sup_{t \in \mathbb{R}} f(t)$. Le problème de l'estimation du mode se produit dans de nombreux domaines d'applications. Un de ceux ayant gagné en importance dans les dernières années est celui de l'analyse de l'horloge moléculaire, qui traite du problème de l'estimation du temps de divergence entre deux espèces dans l'évolution [voir Hedges et Shah (2003)]. Cela se fait d'abord pour l'estimation du temps de divergence de deux protéines intervenant dans ces espèces. Répétée sur une grande quantité de protéines, cette approche produit une distribution des estimations de temps. Le mode empirique de cette distribution est alors pris comme estimateur pour la divergence des deux espèces. Enfin, en médecine, le mode est utilisé dans la calibration du flux sanguin [voir Wegman (1971)]. On peut distinguer deux approches principales ont été utilisées pour estimer le mode : la première est une méthode directe basée sur un algorithme qui procède d'emblée à l'estimation de θ en se basant sur ce que, dans l'échantillon, on doit observer comme une agrégation des valeurs dans le voisinage du mode. On peut citer l'estimateur de Chernoff (1964), Grenander (1965), Delenius (1965) et celui de Venter (1967), ce dernier a été repris et amélioré par Sager (1975). Nous nous intéressons à l'estimation du mode à partir de la deuxième méthode (dite indirecte), cette approche repose sur l'estimation de la densité d'abord, et la localisation du mode comme étant un maximum global ou

approximativement global, les références les plus importantes sont les premières à être adressées dans la littérature ; on cite : Parzen (1962), Wegman (1971). Plusieurs auteurs se sont intéressés à l'estimation de la fonction mode simple pour des données complètement observées (données complètes). Parzen (1962) a été l'un des premiers à considérer le problème de l'estimation du mode d'une densité de probabilité univariée. Il a montré que, sous certaines conditions, l'estimateur du mode obtenu en maximisant l'estimateur à noyau était convergent en probabilité dans le cas indépendant identiquement distribué. Notons que Yamato (1971) a obtenu les mêmes résultats dans le cas d'une densité multivariée, et que Nadaraya (1965) et Van Ryzin (1969) ont établi des résultats de convergence presque sûre dans \mathbb{R} et \mathbb{R}^d ($d > 1$) respectivement. Parzen (1962) a également établi la normalité asymptotique de θ_n , et Konakov (1973) et Samanta (1973) en ont donné une version multivariée. Les techniques de base qu'il a développées pour l'étude du mode ont été reprises par de nombreux auteurs. Eddy (1980, 1982) a amélioré le résultat de Parzen (1962) en obtenant la loi limite de $(\hat{\theta}_n - \theta)$ sous des conditions moins restrictives. Romano (1988) s'est affranchi des conditions globales de régularité sur la densité. Il a montré que la convergence presque sûre et la distribution limite de $\hat{\theta}_n$ s'obtiennent au moyen d'une hypothèse sur le comportement de f au voisinage du mode. De plus, ces deux résultats se généralisent au cas d'une suite aléatoire de fenêtres. Romano a également établi un résultat de type minimax intéressant. Grund et Hall (1995) ont étudié sa convergence en norme L^P . Vieu (1996) a obtenu une vitesse de convergence presque complète de $(\hat{\theta}_n - \theta)$. Leclerc et Loti-Viaud (2000) ont donné un majorant de la vitesse de convergence presque sûre de $\hat{\theta}_n$, en utilisant des résultats de type logarithme itéré. Abraham et al. (2004) ont étudié la normalité asymptotique du mode à partir de la densité à support dans \mathbb{R}^d et Mokkadem et Pelletier (2005) ont étudié sa déviation modérée. Nous définissons, dans ce qui suit, les principaux estimateurs du mode.

Estimateur de Parzen (1962)

L'estimateur à noyau de θ , noté $\theta_{n,P}$, introduit par Parzen (1962), est défini par

$$\theta_{n,P} = \inf \left\{ t \in \mathbb{R} : f_n(t) = \sup_{x \in \mathbb{R}} f_n(x) \right\}$$

où f_n est l'estimateur à noyau de Rosenblatt-Parzen de la densité f défini par

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - Y_i}{h_n}\right) \quad (1.2)$$

où (h_n) est une suite de réels positifs telle que $\lim h_n \rightarrow 0$ et $\lim nh_n \rightarrow +\infty$ lorsque $n \rightarrow +\infty$ et K est une fonction continue vérifiant $\lim_{|x| \rightarrow \infty} K(x) = 0$ et $\int_{\mathbb{R}} K(x) dx = 1$.

Estimateur de Chernoff (1964)

Soit a_n une suite de réels positifs décroissant vers zéro avec n , l'estimateur de Chernoff noté $\theta_{n,C}$, est défini comme étant le milieu d'un quelconque intervalle de longueur $2a_n$ contenant le maximum des observations Y_1, Y_2, \dots, Y_n .

Estimateur de Venter (1967)

Soit b_n une suite d'entiers naturels tendant vers $+\infty$ quand $n \rightarrow +\infty$. L'estimateur de Venter noté $\theta_{n,V}$, est défini comme étant le milieu du plus court intervalle contenant b_n observations parmi Y_1, Y_2, \dots, Y_n .

1.3 Estimation du mode conditionnel

Toujours dans un souci de proposer une modélisation statistique fonctionnelle pour les applications pratiques, nous nous sommes intéressés au cas où une information connexe est disponible. Lors d'études pratiques sur les durées de vie, par exemple en médecine, les chercheurs ne se contentent pas de relever uniquement la durée de vie observée. Les patients de l'échantillon bénéficieront d'un suivi médical qui permettra le relevé de nombreuses données telles les antécédents familiaux, le taux de diabète du sang, la densité des différents globules, etc... Ces données ne sont généralement pas indépendantes de la

durée de vie Y , certaines d'entre elles pouvant être des circonstances aggravantes pour la maladie considérée. La connaissance de ces données pour un certain patient peut permettre une estimation plus fiable de sa durée de vie. En statistique, ce problème peut être modélisé de la manière suivante : supposons qu'on dispose de deux variables aléatoires dépendantes X et Y , la prévision de Y sachant X se fait à travers X par une application r . Autrement dit, on cherche une fonction r telle que $r(X)$ soit une bonne approximation de Y selon un critère donné. Ainsi, le problème devient la minimisation de la fonction de risque suivante :

$$err(r) = E[l(Y - r(X))], \quad (1.3)$$

où l est une fonction de perte. Si cette fonction de perte est convexe et admet un minimum unique, alors on peut prendre la quantité

$$\hat{r}(\cdot) = \arg \min_{c \in R} E(l(Y - c) | X = x)$$

Comme approximation de Y sachant $X = x$. En conséquence, on peut dire qu'en statistique non paramétrique, les outils naturels pour faire la prévision sont les modèles liés à la distribution conditionnelle de ces variables aléatoires. Parmi ces modèles les plus sollicités citons la régression classique, la régression modale (mode conditionnel), les quantiles de régressions (la médiane conditionnelle).

Par définition, le mode conditionnel est la valeur qui maximise la densité conditionnelle. On suppose qu'il existe un compact S où le mode est unique noté θ . Ce prédicateur est obtenu, en considérant

$$\theta(x) = \arg \max_{y \in S} f(y | x) \quad (1.4)$$

où $f(y | x)$ est la densité conditionnelle de Y sachant $X = x$. En utilisant le critère (1.3), nous pouvons remarquer que le mode conditionnel est obtenu en considérant la fonction de risque

$$l(z) = \begin{cases} 0 & \text{si } z = 0 \\ 1 & \text{sinon.} \end{cases}$$

Un premier estimateur du mode conditionnel fut étudié par Collomb et al. (1987). Ces derniers ont établi la convergence uniforme de cet estimateur. Dans cet article, Collomb a donné un exemple d'un processus ϕ -mélangeant où le mode conditionnel prévoit mieux que la régression. Samanta et Thavasneswaran (1990) qui ont donné les propriétés de convergence et de normalité asymptotique dans le cadre i.i.d. alors que les conditions de convergence, dans le cas des données ergodiques données par Rosa (1993) et, dans le cas de données α -mélangeantes, par Ould-Saïd (1993, 1997). Enfin, Quintela-Del-Río et Vieu (1997) ont estimé le mode conditionnel comme étant le point annulant la dérivée d'ordre un de l'estimateur de la densité conditionnelle et établi la convergence presque complète de cet estimateur sous la condition d' α -mélange. Louani et Ould-Saïd (1999) ont établi la normalité asymptotique dans le cas de données fortement mélangeantes.

En dimension infinie, le mode conditionnel connaît, depuis peu, un intérêt croissant, malgré le peu de résultats existants. Dans le contexte des données fonctionnelles, les premiers travaux ont été réalisés par Ferraty et al. (2006). Ils ont montré, sous des conditions de régularité de la densité conditionnelle, la convergence presque complète de l'estimateur à noyau qu'ils ont introduit pour la densité conditionnelle, ainsi que celle du mode conditionnel, et établi la vitesse de convergence. Une application de leurs résultats aux données issues de l'industrie agro-alimentaire est présentée. Dans le même contexte et concernant le mode, Dabo-Niang et al. (2004) ont étudié l'estimateur à noyau du mode, pour une variable fonctionnelle à valeurs dans un espace normé de dimension infinie. Ils ont établi la convergence uniforme presque complète et spécifié les vitesses de convergence en montrant le poids de la mesure des petites boules sur ces dernières. Une application à des données spectrométriques a été présentée.

Dans ce cas, on parle de données complètes. Il n'est pas rare que les données à traiter ne soient pas complètes, dans ce cas les techniques statistiques classiques ne s'adaptent

pas correctement aux données incomplètes, car, l'inférence faite sur l'échantillon observé ne s'étend pas directement à la population mère.

1.4 Les données incomplètes

Une durée de vie est une variable aléatoire souvent positive, de manière générale c'est le temps écoulé pour passer d'un état A à un état B . Il est pourtant rare que la valeur exacte d'une durée de vie soit complètement observée, 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 plus récente. Les données censurées, qui sont bien connus modélisent des cadres d'études expérimentales, pour certains maladies, où les patients peuvent être perdus de vue suite à un déménagement ou en coré suite à un décès non-inhérent à la maladie comme un accident de la route. Dans le modèle censuré, chacune des observations de l'échantillon consiste en un couple de valeurs : (Z, δ) avec Z la durée observée et δ une variable binaire représentant la nature de la durée observée qui prend la valeur 1 s'il s'agit d'une vraie durée de vie et 0 si c'est une censure. Autrement dit, à chaque individu i ; $i = 1, 2, \dots, n$ est associé un couple de durée (Y_i, C_i) de v.a., dont seule la plus petite est observée, Y_i est la durée de survie et C_i est l'instant de censure. C'est qu'à la suite de l'article de Kaplan et Meier (1958) que les données censurées ont trouvé le poids qui est le leur dans la réalité. L'autre cas classique de données incomplètes est celui des données dites tronquées. Ce modèle est apparu tout d'abord en astronomie, mais il est observé dans plusieurs domaines comme la médecine, l'épidémiologie, la biométrie et l'économie. 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 Y est observable, cet échantillon est extrait d'un échantillon de plus grand taille N inconnue. De nombreux travaux ont été effectués sur l'analyse de données tronquées Sée par exemple Klein et Moeschberger (1997). Le plus connus sont l'estimateur de la fonction de survie

de Lynden-Bell (1971) et aussi appelé PL (Produit-Limite) car il s'obtient comme limite d'un produit. Les données censurées et les données tronquées ne sont pas l'unique type de données incomplètes. Les données censurées à droite et troncatures à gauche sont fréquemment utilisées lors d'études sur les durées de vie. 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 décédé 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.

1.5 Modèle de dépendance

Dans la réalité, il arrive souvent que les données traitées (durées de vie, temps d'attente,.....) présentent une certaine forme de dépendance. Les données dépendantes sont naturellement présentes dans beaucoup de disciplines comme la biologie, la médecine, le terme générique couvre actuellement une multitude de structures de données. Voelkel et Crowley (1984) utilisent l'approche des données dites semi-Markovienne pour établir un modèle pour des données de survie sur le cancer pour lequel ils supposent que chaque patient peut, soit reste dans un état initial, soit voir sa maladie progresser, soit répondre positivement au traitement avant de connaître une éventuelle rechute. C'est notamment le cas en ce qui concerne les virus qui se transmettent d'homme à homme. En effet, un individu côtoyant des personnes atteintes par le virus considéré aura plus de chances d'être infecté que s'il résidait dans une zone relativement épargnée par ce virus. Les données ne sont alors pas indépendantes mais on peut considérer que, asymptotiquement il y a indépendance. Si l'échantillon est composé d'individus que l'on ordonne selon leur provenance géographique, les données pour deux individus proches dans celui-ci sont dépendantes mais si les deux individus sont assez éloignés l'un de l'autre dans l'échantillon,

il y a indépendance. Mathématiquement cette dépendance est modélisée par les données. Il existe plusieurs types de mélanges qui sont défini à partir de coefficients, notés, selon les cas, α , β , ϕ , ψ et ϱ . Les données sont alors dites α -mélangeantes si le coefficient α est asymptotiquement nul. Parmi toutes ces formes de mélanges, l' α -mélange est plus faible et donc le moins restrictif. Toute suite de v.a. β , ϕ , ψ ou ϱ -mélangeante sera donc forcément α -mélangeante. Réciproquement, tout résultat énoncé pour des données α -mélangeantes sera valable pour des données soumises à une autre forme de mélange, car toutes suites du v.a's β , ϕ , ψ ou ϱ -mélangeante sera donc forcément α -mélangeante. En particulier les processus AR et ARMA, qui sont largement utilisés dans l'analyse des séries temporelles, sont α -mélangeantes avec des coefficients géométrique et exponentiel respectivement. Gorodetski (1977) et Withers (1981) dérivent les conditions pour lesquelles un processus linéaire est α -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 Tjøstheim (1990) fournissent des discussions intéressantes sur le rôle de l' α -mélange pour l'identification du modèle dans le cadre de l'analyse des séries temporelles non linéaires. Le champ d'application de telles données est très large et il est important de pouvoir faire de l'estimation dans ce cadre de données dépendantes. Afin de pouvoir établir un modèle statistique plus proche de la réalité pour certaines applications, nous nous sommes intéressés au cadre des données tronquées et censurées α -mélangeantes. Il s'agit d'un modèle récent dans le cadre duquel Liang et al. (2012) obtiennent des résultats pour l'estimation de la f.d.r. Nous énonçons dans cette thèse la convergence et la normalité asymptotique de l'estimateur de la fonction de mode. Afin de présenter au mieux ces résultats, nous donnons quelques résultats issus de la littérature pour des données α -mélangeantes. Dans le cadre des données complètes, Cai et Roussas(1992) considèrent une suite de v.a. stationnaires α -mélangeantes de f.d.r. inconnus, ils obtiennent une L.L.I. pour l'estimateur empirique de F qui converge avec une vitesse de l'ordre de $\sqrt{\frac{\log \log n}{n}}$. Dans le cas un modèle censuré, Cai(1998, 2001) a établi

la représentation forte et la normalité asymptotique de l'E.K.M., il montre la consistance forte et donne le taux de convergence et étudie le processus stochastique associé à l'E.K.M. qui converge faiblement vers un processus gaussien de structure de convergence spécifique. Sun et Zhou (2001) donnent notamment des représentations fortes pour l'estimateur de Lynden-Bell (1971) de la f.d.r.

L'étude de données α -mélangeantes dans le cadre un modèle LTRC est très récente et aucun résultat portant sur l'estimateur à noyau de la fonction mode simple dans tel cadre n'était disponible, à notre connaissance. Dans cette thèse (chapitre 5), nous établissons un résultat intéressant de type convergence uniforme et normalité asymptotique de l'estimateur de la mode simple à noyau pour des données censurées et tronquées dépendantes (α -mélangeantes) en donnant sa vitesse de convergence.

1.6 Présentation des résultats obtenus

1.6.1 Mode simple et cas i.i.d.

Dans cette partie, nous supposons pour commencer que l'échantillon que nous étudions est constitué de variables indépendantes et identiquement distribués et nous construisons un nouvel estimateur à noyau du mode simple θ est défini comme la v.a. θ_n maximisant l'estimateur à noyau de la densité $\hat{f}_n(\cdot)$ de $f(\cdot)$, c'est à dire

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

avec

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{\delta_i}{L_n(Z_i) \bar{G}_n(Z_i)} K\left(\frac{t - Z_i}{h_n}\right) \quad (1.5)$$

avec $\bar{G}_n(\cdot) = 1 - G_n(\cdot)$ où $G_n(\cdot)$ est l'estimateur de Kaplan-Meier de $G(\cdot)$ et $L_n(\cdot)$ est l'estimateur de Lynden-Bell de $L(\cdot)$, K est une densité de probabilité sur \mathbb{R} et (h_n) est une suite de réels positifs tendant vers zéro. Sous des hypothèses techniques assez géné-

rales, nous établissons les propriétés asymptotiques (la convergence uniforme, la normalité) de notre estimateur proposé.

Proposition 1.1 : Sous des hypothèses techniques sur le noyau et la fenêtre, on a

$$\sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/2}, h_n^2 \right) \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 1.1 : Sous les hypothèses de la Proposition 1.1, on a

$$\hat{\theta}_n - \theta = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 1.2 : Sous certaines hypothèses. On a

$$\sqrt{nh_n^3} (\hat{\theta}_n - \theta) \xrightarrow{L} N(0, \sigma^2)$$

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

$$\sigma^2 = \frac{\mu f(\theta)}{\bar{G}(\theta) L(\theta) [f^{(2)}(\theta)]^2} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr$$

La démonstration et le détail des conditions imposées font l'objet d'un article récemment soumis et qui est présenté dans le chapitre 4 de cette thèse.

1.6.2 Mode simple et cas α -mélangeant

Dans cette partie, nous généralisons les résultats obtenus précédemment à une suite d'observations α -mélangeantes. Ainsi, en renforçant les hypothèses sur les coefficients de mélange. Nous établissons la convergence uniforme presque sûre et la normalité asymptotique de l'estimateur à noyau de la densité et du mode.

Théorème 1.3 : Sous des hypothèses peu restrictives sur le coefficient de mélange, le noyau et la densité f , on a

$$\sup_{t \in D} |\hat{f}_n(t) - f(t)| = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/2}, h_n^2 \right) \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

$$\hat{\theta}_n - \theta = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 1.4 : Sous certaines hypothèses. On a

$$\sqrt{nh_n^3} (\hat{\theta}_n - \theta) \xrightarrow{l} N(0, \sigma^2)$$

où \xrightarrow{l} désigne la convergence en loi,

$$\sigma^2 = \frac{\mu f(\theta)}{\bar{G}(\theta) L(\theta) [f^{(2)}(\theta)]^2} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr$$

La démonstration de ces résultats et le détail des conditions imposées seront donnés dans l'article Chaib et al. (2013) qui est présenté dans le chapitre 5 de cette thèse.

1.6.3 Mode conditionnel et cas i.i.d.

Nous considérons un vecteurs X représentant une covariable de densité $l(\cdot)$. Pour tout $x \in \mathbb{R}$, on note $f(\cdot | x)$ la fonction de densité conditionnelle de Y sachant $X = x$ et on observe la suite $\{(Z_i, X_i, \delta_i) \mid i = 1, \dots, n\}$. Un estimateur à noyau du mode conditionnel $\theta(x)$ est défini comme la v.a. $\theta_n(x)$ maximisant l'estimateur à noyau de la densité conditionnelle $\hat{f}_n(\cdot | x)$ de $f(\cdot | x)$, c'est à dire

$$\hat{\theta}_n(x) = \arg \sup_{t \in \mathbb{R}} \hat{f}_n(t | x)$$

avec

$$\begin{aligned}
\hat{f}_n(t | x) &= \frac{\frac{\mu_n}{nh_n^2} \sum_{i=1}^n \frac{\delta_i}{\bar{G}_n(Z_i)L_n(Z_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Z_i}{h_n}\right)}{\frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Z_i)} K\left(\frac{x-X_i}{h_n}\right)} \\
&= \frac{\hat{f}_n(x, t)}{l_n(x)} \tag{1.6}
\end{aligned}$$

avec $G_n(\cdot)$ est l'estimateur de Kaplan-Meier de $G(\cdot)$ et $L_n(\cdot)$ est l'estimateur de Lynden-Bell de $L(\cdot)$, K et H sont des densités de probabilités sur \mathbb{R} et (h_n) est une suite de réels positifs tendant vers zéro quand $n \rightarrow \infty$. Nous établissons que $\hat{f}_n(\cdot | x)$ est consistant et nous déduisons que l'estimateur de la fonction de mode conditionnel converge presque sûrement et que convenablement standardisé, il est asymptotiquement gaussien tout en explicitant le terme de la variance.

Proposition 1.2 : Sous les conditions standards sur les noyaux et les densités, ainsi que d'autres hypothèses de régularités et techniques, on a

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{2}}, h_n^2 \right) \right\} \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 1.5 : Sous les hypothèses de la Proposition 1.2 et si la densité conditionnelle vérifie $\sup_{x \in \mathbb{R}} f^{(1)}(\theta(x) | x) < 0$, on a

$$\sup_{x \in \Omega} \left| \hat{\theta}_n(x) - \theta(x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{4}}, h_n \right) \right\} \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 1.6 : Soit $x \in \mathbb{R}$, sous certaines hypothèses. On a

$$\sqrt{nh_n^4} \left(\hat{\theta}_n(x) - \theta(x) \right) \xrightarrow{L} N(0, \sigma^2(x))$$

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

$$\sigma^2(x) = \frac{\mu f(x, \theta(x))}{\overline{G}(\theta(x)) L(\theta(x)) [f^{(0,2)}(x, \theta(x))]^2} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds$$

Ce résultat sera établi dans le chapitre 6 de cette thèse.

1.7 Le plan de thèse

Cette thèse est composée de six chapitres. Le deuxième chapitre propose une présentation des données incomplètes, puis le rappel sur les différents types de dépendances essentiellement la dépendance forte (i.e., α -mixing). Ce type de dépendance modélise beaucoup de processus en particulier les modèles ARMA ou ARCH souvent rencontrés en finance et économétrie et aussi nous donnons les versions de certaines théorèmes classiques en statistiques pour des données α -mélangeantes.

Dans le chapitre 3 nous nous intéressons à la fonction mode simple et à son estimation dans différents cadres des données i.i.d. et α -mélangeants. Nous présentons des résultats de Parzen (1962) relatifs à l'estimation de la fonction densité et la fonction mode pour des données complètes et i.i.d., puis nous présentons brièvement les travaux de Ould-Saïd et Tatachak (2009a, 2009b) sur les propriétés asymptotiques d'un estimateur à noyau pour le modèle tronqué dans le cas i.i.d. et α -mélangeant. Dans le cadre des données censurées, nous y énonçons les principaux résultats de Gannoun et Saracco (2002) sur la convergence sur l'estimateur θ_n dans le cas i.i.d. et Louani (1998) sur la normalité de l'estimateur de mode simple. Nous donnons par la suite la convergence uniforme de l'estimateur de la densité ainsi que de l'estimateur du mode simple pour des données censurées et α -mélangeantes, obtenus par Khardani et al. (2011). Pour conclure ce troisième chapitre, nous présentons les résultats obtenus par Chen et Dai (2003) sur l'asymptotiques normalité de l'estimateur à noyau du mode simple pour un modèle LTRC avec des données α -mélangeantes.

Le chapitre 4 traite du problème de l'estimation du mode simple dans le cas de données i.i.d., tronquées à gauche et censurées à droite. Nous construisons des nouveaux

estimateurs à noyau de la densité et de ses dérivées dans lesquels sont pris en compte des effets de censure et troncature au cours de nos observation. Nous établissons la consistance et la normalité asymptotique de l'estimateur de la densité, ainsi que de l'estimateur de mode simple. Ce travail a fait l'objet d'un article soumis à l'Afrika Statistika Journal.

Le cinquième chapitre est consacré à l'extension des résultats de chapitre 4 au cas des données fortement mélangeantes. Nous donnons des résultats sur la convergence forte et la normalité asymptotique d'un estimateur à noyau pour le modèle LTRC dans le cas de l' α -mélange. Ce travail a fait l'objet d'un article paru au South African Statistical Journal.

Dans le dernier chapitre, nous supposons qu'une autre source d'information est disponible au travers d'une covariable X de densité l , corrélées avec la variable d'intérêt Y la loi que nous intéresse est celle de Y sachant $X = x$. Nous présentons des résultats de Louani et Ould-Saïd (1999) qui concernent la normalité asymptotique d'un estimateur à noyau dans le cas des données complètes et fortement mélangeants, puis nous présentons d'abord les propriétés asymptotiques de l'estimateur à noyau du mode conditionnel dans le cas i.i.d. pour un modèle censuré à droite (tronqué à gauche), obtenus par Khardani et al. (2011) (Ould-Saïd et Tatachak, 2007) respectivement. Nous donnons un résultat sur la convergence uniforme presque sûre sur un compact de \mathbb{R} et la normalité asymptotique d'un estimateur à noyau pour le modèle LTRC avec des données i.i.d. Ce travail a fait l'objet d'un article soumis au Probability and Mathematical Statistics Journal.

Nous concluons ce manuscrit par une conclusion de ce travail de thèse, ainsi quelques perspectives de recherche.

Chapitre 2

Données incomplètes et dépendantes

2.1 Données incomplètes

La recherche dans les domaines tels que l'économie, médecine, biologie, épidémiologie, astronomie, fiabilité,....., n'ont a première vue rien en commun. Toute fois, dans chacun de ces domaines, on s'intéresse à des variables aléatoires généralement positives représentant la durée de temps jusqu'à l'apparition d'un certain événement, appelées durée de survie. En médecine elle est la durée de survie ou le temps de rémission d'un patient, alors qu'en industrie cette variable peut être la durée séparant deux pannes successives d'une machine. Lorsqu'on dispose d'un échantillon $\{Y_i, 1 \leq i \leq n\}$ de v.a. Y sont observées dans leur totalité nous dirons que les données sont complètement observées ou il s'agit de données complètes. Dans ce cas, un estimateur naturel de la f.d.r. F de Y est alors l'estimateur empirique F_n défini par $F_n(y) = \frac{1}{n} \sum_{i=1}^n I\{Y_i \leq y\}$. Les propriétés de $F_n(\cdot)$ sont largement étudié dans la littérature statistique. Notre travail porte sur les données incomplètes, pour lesquelles la variable d'intérêt n'est pas complètement observée pour toutes les données de l'échantillon. Nous présentons dans ce qui suit trois cas de données incomplètes; les données tronquées, censurées et les données censurées et tronquées.

2.1.1 Les données censurées

La durée de survie n'est pas toujours complètement observée, parce que pour certains individus, l'événement de début et / ou de fin n'est pas observé. On parle alors de données censurées.

Définition 2.1 : La durée Y est dite censure si la durée n'est pas intégralement observée.

Les différents types de censure

Censure de Type I : fixée

Définition 2.2 : Soit C une valeur fixée, au lieu d'observer les variables Y_1, \dots, Y_n qui nous intéressent on n'observe Y_i uniquement lorsque $Y_i \leq C$ sinon on voit uniquement que $Y_i > C$. On observe donc une variable aléatoire Z_i telle que $Z_i = \min(Y_i, C)$

Exemple 2.1 : En biologie, on peut tester l'efficacité d'une molécule sur un lot de souris (les souris vivantes au bout d'un temps μ sont sacrifiées)

Censure de Type II : attente

Définition 2.3 : On observe les durées de vie de n individus jusqu'à ce que r individus aient vu l'événement d'intérêt se produire (décès). Soient $Y_{(i)}$ et $Z_{(i)}$ les statistiques d'ordre des variables Y_i et Z_i . La date de censure est donc $Y_{(r)}$ et on observe les variables suivantes $Z_{(1)} = Y_{(1)}, Z_{(2)} = Y_{(2)}, \dots, Z_{(r)} = Y_{(r)}, Z_{(r+1)} = Y_{(r)}, \dots, Z_{(n)} = Y_{(r)}$.

Censure de Type III : aléatoire : C'est typiquement ce modèle qui est utilisé pour les essais thérapeutiques.

Censure aléatoire à droite

Définition 2.4 : La durée Y est dite censurée aléatoirement à droite si au lieu d'observer Y_1, \dots, Y_n on observe

$$Z_i = \min(Y_i, C_i) \text{ et } \delta_i = I\{Y_i \leq C_i\} \text{ pour } i = 1, \dots, n \quad (2.1)$$

où C_i est une censure aléatoire et le δ sert en fait à connaître la nature de l'observation, il indique si l'on est face à une observation réelle ($\delta = 1$) ou à une censure ($\delta = 0$)

Remarque 2.1 : La censure aléatoire à droite est la censure la plus classiquement considérée en analyse de survie.

Exemple 2.2 : Lors d'un test de fiabilité, on observe n système sur une durée C . On définit la variable aléatoire Y comme la durée de vie jusqu'à la première panne de chaque système. Cette panne ne sera pas forcément observée pour chaque élément de cet échantillon. Nous observerons alors, pendant ce temps C , $k (< n)$ durées de vie (pour $Y < C$) et $n - k$ fois C .

Censure aléatoire à gauche

Définition 2.5 : La censure à gauche correspond au cas où l'individu a déjà subi l'événement avant que l'individu sont observe. La durée Y est dite censurée aléatoirement à gauche si au lieu d'observer Y_1, \dots, Y_n on observe (Z_i, δ_i) où

$$Z_i = \max(Y_i, C_i) \text{ et } \delta_i = I\{Y_i \geq C_i\} \text{ pour } i = 1, \dots, n$$

et C_i est une censure aléatoire.

Exemple 2.3 : Une ethnologie étudie la durée d'apprentissage d'une tâche. Cette durée est une v.a. Y et C est l'âge de l'enfant. Pour les enfants qui savent déjà accomplir la tâche, C censure Y à gauche car pour eux Y est inconnu mais inférieur a C : $Y < C$.

Censure aléatoire par intervalle

Définition 2.6 : Une date est censurée par intervalle si au lieu d'observer avec certitude le temps de l'événement, la seule information disponible est qu'il a eu lieu entre deux dates connue. La durée Y est dite censurée par intervalle si au lieu d'observer

Y_1, \dots, Y_n on observe aléatoirement (Z_i, δ_i) ou

$$Z_i = \max \left[\min \left(Y_i, C_i^{(R)} \right), C_i^{(L)} \right] \quad \text{pour } i = 1, \dots, n$$

et $C_i^{(R)}$ et $C_i^{(L)}$ sont des censures aléatoires, $\delta_i = \begin{cases} 1 & \text{si } Y_i \text{ observé} \\ 0 & \text{si } Y_i \text{ censuré à droite} \\ -1 & \text{si } Y_i \text{ censuré à gauche} \end{cases}$

Exemple 2.4 : Y durée entre l'infection par le VIH et le décès par le sida.

2.1.2 Les données troncatures

Une autre situation dans laquelle les données incomplètes apparaissent est celle des données tronquées. Les troncatures diffèrent des censures au sens où elles concernent l'échantillonnage lui même. Une observation est dite tronquée si elle est conditionnelle à un autre événement. On dit que la variable T de durée de vie est tronquée si T n'est observable que sous une certaine condition dépendante de la valeur de T .

La Troncature à gauche

Définition 2.7 : Soit Y une v.a. indépendante de T , on dit qu'il y a troncature à gauche lorsque T n'est observable que si $T > Y$.

Exemple 2.5 : Si la durée de vie d'une population est étudiée à partir d'une cohorte tirée au sort dans cette population, seule la survie des sujets vivants à l'inclusion pourra être étudiée (il y a troncature à gauche car seuls les sujets ayant survécu jusqu'à la date d'inclusion dans la cohorte sont observables).

La troncature à droite

Définition 2.8 : De même, il y a troncature à droite lorsque T n'est observable que si $T < Y$.

La troncature par intervalle

Définition 2.9 : Quand une durée est tronquée à droite et à gauche, on dit qu'elle est tronquée par intervalle.

Exemple 2.6 : On rencontre ce type de troncature lors de l'étude des patients d'un registre : les patients diagnostiqués avant la mise en place du registre ou répertoriés après la consultation du registre ne seront pas inclus dans l'étude.

2.1.3 Le modèle LTRC

Définition 2.10 : On note Y le temps de survie, T la variable de troncature et C la variable de censure. Dans le cas de données tronquées à gauche et censurées à droite on observe (Z, δ, T) si $Z \geq T$, où $Z = \min(Y, W)$ et $\delta = I\{Y \leq W\}$. Lorsque $Z < T$ il n'est pas d'observation.

Exemple 2.7 : Dans le cas de la cohorte Paquid, si l'on étudie l'apparition d'une démence en fonction de l'âge nous sommes en présence de données tronquées à gauche et censurées à droite. Les données sont tronquées à gauche car les sujets ne sont inclus dans l'échantillon que s'ils ne sont pas encore déments à l'entrée dans l'étude. Comme on étudie l'apparition d'une démence en fonction de l'âge, ils ne sont à risque de développer une démence qu'à partir de l'âge où ils sont entrés (c'est à dire observés). Il y a censure à droite car certains sujets n'ont pas développé de démence à la date de point. Nous avons représenté le suivi de 4 sujets de l'étude Paquid rentrés non déments dans la cohorte. Le sujet A entre dans l'étude à 77,5 ans et est atteint de démence vers 79,6 ans. Le sujet B entre dans l'étude à 77,6 ans et à sa dernière visite à 83 ans il n'avait pas développé de démence. Le sujet C entre dans l'étude à 77,7 ans et est atteint de démence vers 81,9 ans et enfin le sujet D entre dans l'étude à 77,3 ans et à sa dernière visite à 80,4 ans il n'avait toujours pas développé de démence. En ce qui concerne l'effectif à risque calculé avec ces 4 sujets, à 77,4 ans il est de 1. En effet les sujets A, B et C ne sont pas encore entrés dans l'étude et s'ils développent une démence à cet âge ils ne pourront pas entrer dans l'étude (troncature à gauche). A 78 ans il est de 4, mais à 80 il est de 3, car le sujet A est devenu dément à 79,6 ans. A 81 ans l'effectif à risque est de 2 car le sujet D a eu

sa dernière visite à 80,4 ans et il n'avait toujours pas développé de démence (censuré à droite). Le sujet C a développé une démence à 81,9 ans, donc l'effectif à risque à 82 ans est de 1 et le sujet B n'est plus suivi après 83 ans (censuré à droite). A 83,1 ans l'effectif à risque est nul.

2.2 Données dépendantes

Des nombreux résultats de statistique ont été établis en considérant des échantillons indépendants. Cependant, dans la réalité, il arrive souvent que les données traitées présentent une certaine forme de dépendance. On peut citer les processus autorégressifs à moyenne mobile (ARMA), les processus GAREH ou les chaînes de Markov qui sont soumis à une certaine forme de dépendance mais sont asymptotiquement indépendants. Il existe plusieurs formes de dépendances, exprimées par des coefficients, notés, selon les cas, α , β , ϕ , ψ et ρ . Ces coefficients expriment l'indépendance des données selon leurs limites à l'infini (0 pour certains et 1 pour d'autres, du moins pour les coefficients qu'on présentera dans cette partie. Nous nous intéressons dans cette thèse à des phénomènes de dépendance faible. Afin de pouvoir modéliser de telles situations, plusieurs types de dépendance ont été introduits.

2.2.1 Définitions et propriétés

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

- 1- Le coefficient de mélange fort ou d' α -mélange introduit par Rosenblatt (1956)

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

- 2- Le coefficient de régularité absolue ou de β -mélange introduit par Kolmogorov (1931)

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

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

3- Le coefficient de mélange uniforme ou de Φ -mélange introduit par Ibragimov (1962)

$$\Phi(\mathcal{L}, \mathcal{L}') := \sup \left\{ |P(B \setminus A) - P(B)|; A \in \mathcal{L}, B \in \mathcal{L}' \text{ et } P(A) \neq 0 \right\}$$

4- Le coefficient de Ψ -mélange introduit par Blum et al. (1963)

$$\Psi(\mathcal{L}, \mathcal{L}') := \sup \left\{ \left| 1 - \frac{P(A \cap B)}{P(A)P(B)} \right|; A \in \mathcal{L}, B \in \mathcal{L}' \text{ et } P(A) \neq 0, P(B) \neq 0 \right\}$$

5- Le coefficient de ρ -mélange introduit par Hirschfeld (1935) et Gebelein (1941)

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

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

Propriétés

Propriété 2.1 : L'échelle de ces coefficients est définie par les inégalités suivantes

$$\begin{aligned} 0 &\leq \alpha(\mathcal{L}, \mathcal{L}') \leq \frac{1}{4} & ; & & 0 &\leq \beta(\mathcal{L}, \mathcal{L}') \leq 1 \\ 0 &\leq \Phi(\mathcal{L}, \mathcal{L}') \leq 1 & ; & & 0 &\leq \rho(\mathcal{L}, \mathcal{L}') \leq 1 \\ 0 &\leq \Psi(\mathcal{L}, \mathcal{L}') \leq \infty \end{aligned}$$

Propriété 2.2 : Chacune des égalités suivantes est équivalente à l'indépendance des tribus \mathcal{L} et \mathcal{L}'

$$\alpha(\mathcal{L}, \mathcal{L}') = 0, \beta(\mathcal{L}, \mathcal{L}') = 0, \Phi(\mathcal{L}, \mathcal{L}') = 0, \Psi(\mathcal{L}, \mathcal{L}') = 0 \text{ et } \rho(\mathcal{L}, \mathcal{L}') = 0. \quad (2.3)$$

Propriété 2.3 : Les mesures de dépendances satisfont les inégalités suivantes

- i) $2\alpha(\mathcal{L}, \mathcal{L}') \leq \beta(\mathcal{L}, \mathcal{L}') \leq \Phi(\mathcal{L}, \mathcal{L}') \leq \frac{1}{2}\Psi(\mathcal{L}, \mathcal{L}')$;
- ii) $4\alpha(\mathcal{L}, \mathcal{L}') \leq \rho(\mathcal{L}, \mathcal{L}') \leq \Psi(\mathcal{L}, \mathcal{L}')$;
- iii) $\rho(\mathcal{L}, \mathcal{L}') \leq 2[\Phi(\mathcal{L}, \mathcal{L}')]^{1/2}[\Phi(\mathcal{L}', \mathcal{L})]^{1/2} \leq 2[\Phi(\mathcal{L}, \mathcal{L}')]^{1/2}$.

2.2.2 Les conditions de mélange fort

Supposons que $c(., .)$ désigne un coefficient de mélange arbitrairement pris parmi ceux précédemment définis ($\alpha, \beta, \phi, \psi, \varrho$). Soit $\{Y_i, i \geq 1\}$ une suite de v.a. et soit \mathcal{F}_i^k la σ -algèbre engendrée par $\{Y_j, i \leq j \leq k\}$. Etant donné un entier positif l , posons pour tout k fixé, $k \geq 1$;

$$c(l) = c(A, B) \text{ où } A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+l}^\infty$$

La suite de v.a. $\{Y_i, i \geq 1\}$ est dite c -mélangeante si

$$\lim_{l \rightarrow +\infty} c(l) = 0 \quad (2.4)$$

Cette définition, mise en relation avec celles des différents coefficients de mélange, permet donc de considérer le c -mélange comme une indépendance asymptotique entre le passé et le future.

Remarques

1) Ces notions de mélange ne sont bien évidemment pas équivalentes mais il existe différentes implication entre les formes de mélanges :

$$\Psi - \text{mélange} \Rightarrow \Phi - \text{mélange} \Rightarrow \left\{ \begin{array}{l} \beta - \text{mélange} \\ \rho - \text{mélange} \end{array} \right\} \Rightarrow \alpha - \text{mélange}$$

2) Dans le cas de la stricte stationnarité de Y , on a simplement

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

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

Une inégalité exponentielle

L'outil que nous allons utiliser de manière déterminante dans les problèmes de convergence presque complète est l'inégalité exponentielle ci-dessous que l'on peut rencontrer dans Rio (2000) sous l'appellation inégalité de Fuk-Nagaev. Cette inégalité est en fait une extension au cadre de variables fortement mélangeantes de l'inégalité de Bernstein que l'on peut trouver sous diverses formes dans Hoeffding (1963) (voir aussi le Lemme 2.3.1 pour une version simplifiée). En fait, le développement des résultats en estimation non paramétrique de régression sous dépendance s'est fait en parallèle avec celui de telles inégalités. Depuis, la puissance de telles inégalités a été sérieusement améliorée (c.f. Bosq, 1993), jusqu'à l'inégalité de Rio (2000). Le lecteur trouvera dans Rio (2000) mais aussi dans Bosq (1996) plusieurs versions de cette inégalité, dont certaines sous des énoncés plus généraux que ce qui va suivre. Cependant nous en tiendrons ici à cette version simplifiée qui est largement suffisante dans notre contexte, et que l'on trouve sous cette forme dans Rio (2000, p.87, formule (6.19b)).

Lemme 2.1 Inégalité de type Fuk-Nagaev: Soit $\{U_i, i \geq 1\}$ une suite de v.a.r. centrées et de coefficient de mélange fort $\alpha(n) = O(n^{-\nu})$, $\nu > 1$, et $\forall n \in \mathbb{N}$, $1 \leq i \leq n$ $|U_i| < \infty$. Alors pour tout $r > 1$ et $\epsilon > 0$

$$P \left\{ \left| \sum_{i=1}^n U_i \right| > \epsilon \right\} \leq c \left[\left(1 + \frac{\epsilon^2}{r S_n} \right)^{-r/2} + n r^{-1} \left(\frac{2r}{\epsilon} \right)^{\nu+1} \right] \quad (2.5)$$

avec

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

Chapitre 3

La fonction de mode

Soit Y_1, \dots, Y_n une suite de v.a. de même loi qu'une v.a. Y ayant de fonction de répartition réelle F et densité continue f . On appelle mode simple de f tout point noté θ défini par

$$f(\theta) = \sup_{t \in \mathbb{R}} f(t)$$

Il s'agit d'une caractéristique de f qui joue un rôle important tant en probabilité qu'en statistique.

Exemple 3.1. Si Y suit une loi gamma, sa densité de probabilité f est

$$f(t) = \frac{1}{\sigma \Gamma(p)} \left(\frac{t - \alpha}{\sigma} \right)^{p-1} e^{-\frac{t-\alpha}{\sigma}}$$

où

$$\sigma > 0; p > 1; \alpha \leq t \leq \infty; \Gamma(p) = \int_0^{\infty} e^{-y} y^{p-1} dy.$$

Le mode θ de $f(t)$ est défini par : $\theta = (p - 1)\sigma + \alpha$.

3.1 Estimation de la fonction mode pour des données complètes

L'estimation à noyau du mode est donnée par la solution de la relation

$$\theta_n = \arg \sup_{t \in \mathbb{R}} f_n(t)$$

telle que $f_n(t)$ est l'estimateur à noyau de la densité $f(t)$ introduit par Parzen (1962) et défini pour tout $t \in \mathbb{R}$ par :

$$f_n(t) = \int_{-\infty}^{+\infty} \frac{1}{h_n} K\left(\frac{t-y}{h_n}\right) dF_n(y) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{t-Y_i}{h_n}\right) \quad (3.1)$$

où K est une fonction de probabilité continue, (h_n) est une suite de réels strictement positifs tendant vers zéro et F_n l'estimateur empirique de la fonction de répartition réelle défini par

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n I\{Y_i \leq t\}$$

Ceci nous assure de l'existence d'une variable aléatoire θ_n , telle que $f_n(\theta_n) = \sup_{t \in \mathbb{R}} f_n(t)$. Comme un tel θ_n peut ne pas être unique, considérons la fonctionnelle

$$M(g) = \inf \left\{ t : g(t) = \sup_{x \in \mathbb{R}} g(x) \right\}$$

où g est une fonction sur \mathbb{R} . Le mode empirique θ_n est alors défini de manière unique par $\theta_n = M(f_n)$ et le mode théorique θ de f est $\theta = M(f)$.

Nous rappelons dans ce paragraphe quelques résultats importants. Le théorème suivant traite de la consistance forte de θ_n .

Théorème 3.1 (Parzen, 1962) Supposons que K et (h_n) vérifient les conditions suivantes

$$\sup_{-\infty < t < +\infty} |K(t)| < \infty \quad (3.2)$$

$$\int_{-\infty}^{+\infty} |K(t)| dt < \infty \quad (3.3)$$

$$\lim_{t \rightarrow +\infty} |tK(t)| = 0 \quad (3.4)$$

$$\lim_{n \rightarrow \infty} nh_n^2 = \infty \quad (3.5)$$

et le mode de f admet la propriété suivante

$$\forall \epsilon > 0, \exists \eta > 0, \forall t, |\theta - t| \geq \epsilon \text{ implique que } |f(\theta) - f(t)| \geq \eta \quad (3.6)$$

alors pour tout $\epsilon > 0$, on a

$$P[|\theta_n - \theta| < \epsilon] \rightarrow 1 \text{ quand } n \rightarrow \infty$$

Preuve de théorème 3.1 : Il suffit de montrer que

$$f(\theta_n) \rightarrow f(\theta) \text{ en probabilité lorsque } n \rightarrow \infty$$

Maintenant on a

$$\begin{aligned} |f(\theta_n) - f(\theta)| &\leq |f(\theta_n) - f_n(\theta_n)| + |f_n(\theta_n) - f(\theta)| \\ &\leq 2 \sup_t |f_n(t) - f(t)| \end{aligned} \quad (3.7)$$

car

$$|f_n(\theta_n) - f(\theta)| = \left| \sup_t f_n(t) - \sup_t f(t) \right| \leq \sup_t |f_n(t) - f(t)|$$

d'après (3.6) et (3.7) on a le résultat.

Le résultat qui suit établit la normalité asymptotique de l'estimateur θ_n

Théorème 3.2 (Parzen, 1962) Sous les hypothèses suivantes

H1) f est de classe C^2 au voisinage de θ , avec $f^{(2)}(\theta) < 0$.

H2) $\exists \delta, 0 < \delta < 1$ tels que k vérifie

$$\int_{-\infty}^{+\infty} t^{\delta+2} |k(t)| dt < \infty$$

avec k la transformée de Fourier de K définie par

$$k(t) = \int_{-\infty}^{+\infty} e^{-ity} K(y) dy$$

H3) $\lim_{n \rightarrow \infty} nh^6 \rightarrow \infty, \quad \lim_{n \rightarrow \infty} nh_n^{2\delta+5} \rightarrow 0$

H4) La fonction caractéristique φ vérifie

$$\int_{-\infty}^{+\infty} t^{\delta+2} |\varphi(t)| dt < \infty$$

H5) Supposons que K est de classe C^1 telle que $\int_{-\infty}^{+\infty} [K^{(1)}(t)]^2 dt < \infty$.

alors, on a

$$\sqrt{nh_n^3}(\theta_n - \theta) \xrightarrow{l} N\left(0, \left\{ f(\theta) / [f^{(2)}(\theta)]^2 \right\} J\right)$$

où \xrightarrow{l} désigne la convergence en loi, avec

$$J = \int_{-\infty}^{+\infty} [K^{(1)}(t)]^2 dt = (2\pi)^{-1} \int_{\mathbb{R}} t^2 k^2(t) dt$$

3.2 Estimation non paramétrique de la fonction mode dans le cas des données censurées à droite

3.2.1 Modèle de censure

Soit Y_1, \dots, Y_n , une suite de v.a. indépendantes et identiquement distribuées de fonction de répartition continue F et de densité f . Dans certaines situations, la v.a. d'intérêt Y ne peut être observée complètement, c'est le cas dans les études médicales où le patient est perdu de vie. Cependant on observe une autre v.a. appelée censure, positive de f.d.r. G . Généralement et conformément aux études biomédicales, les v.a. C_i sont supposées être indépendantes des Y_i . Soit $\{(Z_1, \delta_1), \dots, (Z_n, \delta_n)\}$ l'échantillon réellement observé où $Z_i = Y_i \wedge C_i$ et $\delta_i = I\{Y_i \leq C_i\}$. Pour toute f.d.r. W , nous noterons par $\tau_W = \sup\{t : W(t) < 1\}$. Les Z_i sont de f.d.r. H donnée par $\bar{H}(t) = 1 - H(t) = \bar{F}(t) \bar{G}(t)$, $\forall t \geq 0$, et le cas non censure est un cas particulier du modèle de censure avec $G = 0$. Il est bien connu que l'estimateur non paramétrique \hat{F}_n de F est l'estimateur de Kaplan-Meier (1958) défini par

$$1 - \hat{F}_n(t) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{\delta_i}{n-i+1}\right)^{I\{Z_i \leq t\}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)} \end{cases} \quad (3.8)$$

où $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$ sont les statistiques d'ordre de Z_i et pour chacune des valeurs $Z_{(i)}$ la valeur δ_i est l'indicatrice correspondante. L'estimateur de Kaplan-Meier \hat{G}_n de G s'obtient de la même manière par

$$1 - \hat{G}_n(t) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_i}{n - i + 1}\right)^{I\{Z_i \leq t\}} & \text{si } t < Z_{(n)} \\ 0 & \text{si } t \geq Z_{(n)} \end{cases}$$

L'estimateur de Kaplan-Meier \hat{F}_n , de la fonction de répartition F possède beaucoup de propriétés analogues à celles de la fonction de répartition empirique. Les propriétés asymptotiques de l'E.K.M. ont été étudiées par plusieurs auteurs, Breslow et Crowley (1974) furent les premiers à traiter de la convergence de l'E.K.M. et de sa normalité asymptotique, leur résultat furent repris par Gill (1980, 1983). Pour plus de détails, nous renvoyons au livre de Shorack et Wellner (1986, p.304).

Les seuls travaux sur le mode simple en présence de données censurées où les durées de survies et les instants de censure sont deux suites de v.a. i.i.d., sont ceux de Louani (1998), Gannoun et Saracco (2002).

3.2.2 Estimation dans le modèle de censure

De manière analogue aux données complètes, la fonction mode d'une densité f est donnée par la relation

$$\theta = \arg \sup_{t \in \mathbb{R}} f(t)$$

Un estimateur à noyau du mode θ est défini comme la v.a. θ_n maximisant l'estimateur à noyau de la densité f_n de f , c-à-d

$$f_n(\theta_n) = \sup_{t \in \mathbb{R}} f_n(t)$$

où

$$f_n(t) = \frac{1}{h_n} \int_0^{\infty} K\left(\frac{t-s}{h_n}\right) dF_n(s) \quad (3.9)$$

K est une densité de probabilité sur \mathbb{R} , $(h_n)_{n \geq 0}$ est une suite de réels positifs tendant vers zéro quand $n \rightarrow \infty$, $1 - F_n$ est une estimateur de Kaplan-Meier de fonction de survie.

Le résultat qui suit établit la convergence de l'estimateur f_n et θ_n .

Théorème 3.3 (Gannoun et Saracco, 2002) : Sous des hypothèses classiques sur la densité f et le noyau $K(\cdot)$ et si $\frac{\log n}{nh_n^2} \rightarrow 0$ alors

$$\sup_{t \in \Delta} |f_n(t) - f(t)| \rightarrow 0 \quad p.s.$$

où Δ est un compact inclus dans $[0, \tau_H[$.

Corollaire 3.1 (Gannoun et Saracco, 2002) : Si $\theta \in \Delta$

$$|\theta_n - \theta| \rightarrow 0 \quad p.s.$$

Indication de la preuve du théorème 3.3 : La preuve de ce théorème est basée sur l'inégalité de Wang (2000, théorème 2.3). La deuxième technique de cette démonstration est le recouvrement du compact par des petites boules.

Le théorème suivant présente la normalité asymptotique de l'estimateur de la fonction mode simple.

Théorème 3.4 (Louani, 1998) : Sous des hypothèses sur les dérivées $K^{(1)}$ et $K^{(2)}$ et si $nh_n^5 \rightarrow +\infty$, $nh_n^7 \rightarrow 0$, alors on a

$$\sqrt{nh_n^3}(\theta_n - \theta) \xrightarrow{l} N(0, B(x))$$

où \xrightarrow{l} désigne la convergence en loi,

$$B(x) = \frac{f(\theta)}{\bar{G}(\theta) [f^{(2)}(\theta)]^2} \int_{\mathbb{R}} [K^{(1)}(x)]^2 dx$$

3.2.3 Le cas α -mélange

Dans cette partie nous allons reprendre les résultats obtenus pour Khardani et al. (2012), de convergence forte uniforme de l'estimateur à noyau du mode simple pour des données censurées à droite sous des conditions d'un certain type de dépendance. Dans le cas de censure l'estimateur de densité définit comme suit

$$\hat{f}_n(t) = \frac{1}{nh_n} \sum_{i=1}^n \frac{\delta_i}{\overline{G}_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right) \quad (3.10)$$

Proposition 3.1 (Khardani et al., 2012) : Sous des hypothèses peu restrictives sur le coefficient de mélange, le noyau K et la densité f , on a

$$\sup_{t \in \Omega} \left| \hat{f}_n(t) - f(t) \right| = O\left(\max\left(\left(\frac{\log n}{nh_n}\right)^{1/2}, h_n^2\right)\right) \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 3.5 (Khardani et al. 2012) : Sous les hypothèses de la Proposition 3.1, on a

$$\theta_n - \theta = O\left(\max\left(\left(\frac{\log n}{nh_n}\right)^{1/4}, h_n\right)\right) \text{ p.s. lorsque } n \rightarrow \infty.$$

Indication de la preuve du Proposition 3.1 : L'inégalité de tringle nous permet d'écrire :

$$\sup_{t \in \Omega} \left| \hat{f}_n(t) - f(t) \right| \leq \sup_{t \in \Omega} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| + \sup_{t \in \Omega} \left| \tilde{f}_n(t) - E\left[\tilde{f}_n(t)\right] \right| + \sup_{t \in \Omega} \left| E\left[\tilde{f}_n(t)\right] - f(t) \right|$$

où

$$\tilde{f}_n(t) = \frac{1}{nh_n} \sum_{i=1}^n \frac{\delta_i}{\overline{G}(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

La démonstration du **Proposition 3.1** repose sur les résultats suivants

$$\sup_{t \in \Omega} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| = O \left(\sqrt{\frac{\log \log n}{n}} \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

$$\sup_{t \in \Omega} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_n}} \right) \text{ p.s. lorsque } n \rightarrow \infty.$$

$$\sup_{t \in \Omega} \left| E \left[\tilde{f}_n(t) \right] - f(t) \right| = O(h_n^2) \text{ p.s. lorsque } n \rightarrow \infty.$$

3.3 Estimation non paramétrique du mode dans le cas des données tronquées à gauche

3.3.1 Modèle de troncature

Soit (Ω, A, P) un espace probabilisé, Soit Y_1, \dots, Y_N , une suite de v.a.r. d'intérêt i.i.d., de f.d.r. commune F et de densité f . Soit T_1, \dots, T_N , une suite de variables aléatoires de troncature i.i.d. de f.d.r. continue L . On suppose aussi que ces variables sont indépendantes des Y_i . La fonction d.r. jointe de Y et T est :

$$H(y, t) = P(Y_1 \leq y, T_1 \leq t) = F(y) L(t)$$

La taille de l'échantillon N est déterministe mais inconnue.

Soit $\{(Y_i, T_i), i = 1, \dots, n\}$ ($n \leq N$) l'échantillon réellement observé (i.e. $Y_i \geq T_i$). Une conséquence de la troncature, la taille de l'échantillon vraiment observé n est une v.a. distribuée selon $\beta(N, \mu)$ où $\mu = P(Y_1 \geq T_1) > 0$. Par la loi forte des grands nombres on a, lorsque N tend vers ∞

$$\mu_n := \frac{n}{N} \rightarrow \mu \text{ P p.s.}$$

Conditionnellement à la valeur de n , les données observées (Y_i, T_i) sont encore i.i.d., mais

leur f.d.r. jointe diffère de celle de (Y_i, T_i) :

$$\begin{aligned} H^*(y, t) &= P(Y_1 \leq y, T_1 \leq t \mid Y_1 \geq T_1) \\ &= \mu^{-1} \int_{-\infty}^y L(t \wedge z) dF(z) \end{aligned}$$

Pour toute f.d.r. W , nous notons par a_W et b_W respectivement, les bornes inférieure et supérieure du support de W définies respectivement par

$$a_W = \{y : W(y) > 0\} \text{ et } b_W = \{y : W(y) < 1\}$$

Pour $y \geq a_F$, notons

$$\begin{aligned} C(y) &= P(T_1 \leq y \leq Y_1 \mid Y_1 \geq T_1) \\ &= \mu^{-1} L(y) (1 - F(y^-)) \end{aligned}$$

et considérons son estimateur empirique

$$C_n(y) = \frac{1}{n} \sum_{i=1}^n I\{T_i \leq y \leq Y_i\}$$

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

$$1 - F_n(y) = \prod_{\{i: Y_i \leq y\}} \left(1 - \frac{1}{nC_n(Y_i)}\right) \text{ et } L_n(y) = \prod_{\{i: T_i > y\}} \left(1 - \frac{1}{nC_n(T_i)}\right)$$

Woodroffe (1985) étudie les propriétés asymptotiques de l'estimateur de F ainsi que celles de l'estimateur similaire pour L quand $n \rightarrow \infty$. Dans cet article il établit qu'une

condition nécessaire pour l'identifiabilité du système est que $a_L \leq a_F$ et que $b_L \leq b_F$. Il est à noter que dans le cas $a_L = a_F$, Woodrooffe utilise la condition supplémentaire :

$$\int_{a_F}^{\infty} \frac{1}{L} dF < \infty$$

3.3.2 Estimation dans le modèle de troncature

De manière analogue aux données complètes, la fonction mode d'une densité f est donnée par la relation

$$\theta = \arg \sup_{t \in \mathbb{R}} f(t)$$

Un estimateur à noyau du mode θ est défini comme la v.a. $\hat{\theta}_n$ maximisant l'estimateur à noyau de la densité \hat{f}_n de f , c-à-d

$$\hat{f}_n(\hat{\theta}_n) = \sup_{t \geq a_F} \hat{f}_n(t)$$

dans le cas de troncature l'estimateur \hat{f}_n définit comme suit

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right) \quad (3.13)$$

K est une densité de probabilité sur \mathbb{R} , (h_n) est une suite de réels positifs tendant vers zéro quand $n \rightarrow +\infty$ et L_n est un estimateur de Lyndel-Bell (1971). Dans cette partie nous allons reprendre les résultats obtenus par Ould-Saïd et Tatachak (2009a), des propriétés asymptotiques de l'estimateur à noyau du mode simple pour un modèle tronqué à gauche avec des données i.i.d.

Hypothèses et propriétés asymptotiques

Supposons que $a_L < a_F$ et $b_L < b_F$ et soit Ω un sous ensemble compact de $\Omega_0 = \{y : y \geq a_F\}$. Nous ferons les hypothèses suivantes

- H1) La fenêtre h_n satisfait $\frac{nh_n}{\log n} \rightarrow \infty$, lorsque $n \rightarrow \infty$.
- H2) Le noyau K est bornée, à support compact et trois fois dérivable telle que K , $K^{(1)}$ et $K^{(2)}$ sont bornées.
- H3) K est un noyau de second ordre.
- H4) La densité f est bornée et trois fois continument différentiable.
- H5) L'unique mode θ vérifie :

$$\forall \epsilon > 0, \exists \eta > 0 \forall y : |\theta - y| \geq \epsilon \Rightarrow |f(\theta) - f(y)| \geq \eta.$$

- H6) La fenêtre h_n satisfait $\frac{nh_n^6}{\log n} \rightarrow \infty$ et $nh_n^7 \rightarrow 0$ lorsque $n \rightarrow \infty$.

Le théorème suivant donne un résultat de convergence de l'estimateur du mode simple

Théorème 3.6 (Ould-Saïd et Tatachak 2009a) : Sous les hypothèses $H1 - H6$, on a

$$\hat{\theta}_n - \theta = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right) \quad P - p.s. \text{ lorsque } n \rightarrow \infty.$$

Schéma de la preuve : Sous l'hypothèse H5, le développement de Taylor de la fonction f , au voisinage de θ , est

$$\hat{f}_n(\hat{\theta}_n) - f(\theta) = \frac{1}{2} (\hat{\theta}_n - \theta) f^{(2)}(\bar{\theta})$$

où $\bar{\theta}$ se situe entre $\hat{\theta}_n$ et θ . Ainsi à l'aide d'arguments analytiques, on montre que

$$|\hat{\theta}_n - \theta| \leq 2 \sqrt{\frac{\sup_{t \in \Omega} |\hat{f}_n(t) - f(t)|}{|f^{(2)}(\bar{\theta})|}}$$

L'inégalité de tringle nous permet d'écrire :

$$\begin{aligned} \sup_{t \in \Omega} |\hat{f}_n(t) - f(t)| &\leq \sup_{t \in \Omega} |\hat{f}_n(t) - \tilde{f}_n(t)| + \sup_{t \in \Omega} |\tilde{f}_n(t) - E[\tilde{f}_n(t)]| + \sup_{t \in \Omega} |E[\tilde{f}_n(t)] - f(t)| \\ &= S_1 + S_2 + S_3 \end{aligned}$$

où

$$\tilde{f}_n(t) = \frac{\mu}{nh_n} \sum_{i=1}^n \frac{1}{L(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

La démonstration du Théorème 3.6 repose sur les lemmes suivants

Lemme 3.1 : Sous l'hypothèse $H4$ on a

$$S_1 = O_p\left(\frac{1}{\sqrt{n}}\right)$$

Lemme 3.2 : Sous les hypothèses $H1 - H2$ et $H4$ on a

$$S_2 = O\left(\sqrt{\frac{\log n}{nh_n}}\right)$$

Lemme 3.3 : Sous les hypothèses $H2 - H4$, on a

$$S_3 = O(h_n^2)$$

Nous énonçons, dans le prochain Théorème, la normalité asymptotique de l'estimation du mode.

Théorème 3.7 : Sous l'hypothèse $H1 - H4$ et $H6$, on a

$$\sqrt{\frac{nh_n^3 [f^{(2)}(\theta)]^2}{\sigma^2}} (\hat{\theta}_n - \theta) \xrightarrow{L} N(0, \sigma^2)$$

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

$$\sigma^2 = \frac{\mu f(\theta)}{L(\theta)} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr$$

Schéma de la preuve : On a

$$\begin{aligned}
\sqrt{nh_n^3} (\hat{\theta}_n - \theta) &= \sqrt{nh_n^3} \frac{\hat{f}_n^{(1)}(\theta) - \tilde{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&\quad + \sqrt{nh_n^3} \frac{\tilde{f}_n^{(1)}(\theta) - E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&\quad + \sqrt{nh_n^3} \frac{E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&= \frac{J_1 + J_2 + J_3}{\hat{f}_n^{(2)}(\bar{\theta}_n)}
\end{aligned}$$

La démonstration du Théorème 3.7 repose sur les résultats suivants

$$J_i \rightarrow 0 \quad (i = 1, 3)$$

$$J_2 \rightsquigarrow N(0, \sigma^2)$$

$$\hat{f}_n^{(2)}(\bar{\theta}_n) \text{ converge en probabilité vers } f^{(2)}(\theta)$$

3.3.3 Cas d' α -mélange

Cette partie porte sur l'étude asymptotique de l'estimateur à noyau du mode simple pour des données tronquées à gauche et dépendantes. Plus précisément, nous reprenons les résultats obtenus par Ould-Saïd et Tatachak (2009b), où il montre la convergence uniforme de cet estimateur.

Propriétés asymptotiques

Afin d'établir la convergence uniforme de l'estimateur $\hat{\theta}_n$ vers θ , on introduit les hypothèses suivantes

H1) K est une densité de probabilité bornée, continue au sens de Hölder d'exposant $\beta > 0$, satisfaisant

$$|u| K(u) \rightarrow 0 \text{ quand } |u| \rightarrow \infty$$

H2) K est un noyau de second ordre

H3) $\{Y_i, i \geq 1\}$ est une suite de v.a.r. et α -mélangeantes de coefficient $\alpha(n) = O(n^{-\nu})$ telle que $\nu > 5 + \frac{1}{\beta}$.

H4) $\{T_i, i \geq 1\}$ est une suite de v.a. de troncature i.i.d. et indépendante de $\{Y_i, i \geq 1\}$.

H5) L'unique mode θ vérifie :

$$\forall \epsilon > 0, \exists \eta > 0 \forall y : |\theta - y| \geq \epsilon \Rightarrow |f(\theta) - f(y)| \geq \eta.$$

H6) La densité conditionnelle jointe f_{ij}^* de (Y_i, Y_j) existe et satisfait pour toute constante c indépendante de (i, j) :

$$\sup_{u,v} |f_{ij}^*(u, v) - f^*(u) f^*(v)| \leq c < \infty$$

H7) La fenêtre h_n satisfait :

- i) $\frac{nh_n^2}{\log \log n} \rightarrow \infty$ lorsque $n \rightarrow +\infty$.
- ii) $cn^{\frac{(3-\nu)\beta}{\beta(\nu+1)+2\beta+1}+\eta} < h < c'n^{\frac{1}{1-\nu}}$ où η satisfait

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

Nous donnons dans le théorème suivant la convergence uniforme pour la fonction mode simple, établie par Ould-Saïd et Tatachak (2009b) lorsque les v.a. sont uniformément bornées et pour un taux de mélange exponentiel.

Théorème 3.8 (Ould-Saïd et Tatachak 2009b): Sous les hypothèses H1-H7, on a

$$\hat{\theta}_n - \theta = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right) \quad P - p.s. \text{ lorsque } n \rightarrow \infty.$$

La preuve du théorème en suivant le même cheminement que dans théorème 3.6 et basé sur des inégalités de type exponentiel, Fuk-Nageav (Rio 2000).

3.4 Estimation non paramétrique pour des données tronquées et censurées

3.4.1 Modèle LTRC

Considérons un vecteur aléatoire (Y, T, W) où Y est la variable aléatoire d'intérêt de fonction de répartition continue F et de densité f . T la v.a. de troncature à gauche de f.d.r. L et W la v.a. de censure à droite de f.d.r. G . Les v.a. effectivement observées sont (Z, T, δ) si $Z \geq T$, ou $Z = \min(Y, W)$ et $\delta = I\{Y \leq W\}$. Lorsque $Z < T$ il n'est pas d'observation. Soit H la f.d.r. de Z et $\mu := P[Z \geq T]$. L'indépendance nous donne $H = 1 - (1 - F)(1 - G)$. Soit N la taille de l'échantillon i.i.d. (Z_i, δ_i, T_i) et $n = \sum_{i=1}^N I\{T_i \leq Z_i\}$ la taille de l'échantillon effectivement observé (c-à-d que $T_i \leq Z_i$). Notons que n est une variable aléatoire de loi Binomial $B(N, \mu)$ et μ peut être estimé par $\frac{n}{N}$ (LGNF).

Maintenant conditionnellement à la valeur de n , (Z_i, δ_i, T_i) est encore i.i.d., mais Z et T ne sont pas indépendantes puisque $Z \geq T$. Ainsi, nous noterons P_0 la loi de probabilité absolue (de l'échantillon de taille N) tandis que P sera la loi de probabilité conditionnelle (de l'échantillon effectivement observée). Soit la fonction $C(\cdot)$ définie par :

$$C(x) = P[T \leq x \leq Z \mid Z \geq T] = \mu^{-1} L(x) (1 - H(x))$$

et considérons son estimateur empirique

$$C_n(x) = \frac{1}{n} \sum_{i=1}^n I\{T_i \leq x \leq Z_i\}$$

A la base de ce modèle, on souhaite estimer la fonction de mode

$$f(\theta) = \max_{t \in \mathbb{R}} f(t)$$

un estimateur naturel de $f(t)$ est donnée par

$$\hat{f}_n(t) = \frac{1}{h_n} \int_0^{\infty} K\left(\frac{t-s}{h_n}\right) \hat{F}_n(ds) \quad (3.14)$$

où $\hat{F}_n(\cdot)$ est un estimateur produit limite de la fonction de répartition F , K est un noyau et (h_n) est une suite de réels tendant vers 0. Soit

$$H_1(x) = P[Z \leq x, \delta = 1 \mid Z \geq T]$$

et son estimateur empirique

$$H_{1n}(x) = \frac{1}{n} \sum_{i=1}^n I\{Z_i \leq x, \delta_i = 1\}$$

L'estimateur produit-limite de F , noté \hat{F}_n est appelé estimateur TJW est défini dans Tsai et al. (1987) par

$$1 - \hat{F}_n(t) = \prod_{\{i: Z_i \leq t\}} \left(1 - \frac{1}{nC_n(Z_i)}\right)^{\delta_i} \quad (3.15)$$

Pour toute f.d.r. S , soit $a_S = \inf\{t : S(t) > 0\}$ et $b_S = \sup\{t : S(t) < 1\}$ les points extrêmes gauche et droit du support de S , respectivement.

Les propriétés asymptotiques de l'estimateur TJW ont été étudiées par plusieurs auteurs : voir e.g. Tsai et al. (1987), GU et Lai (1990), Lai et Ying(1991). Gijbels et Wang (1993) obtiennent un représentation forte de l'estimateur TJW par un somme de

v.a. i.i.d. avec un terme de reste de l'ordre $O\left(\frac{\log n}{n}\right)$ quand $a_L < a_H$. Dans le cas non censuré (c-à-d $W = +\infty$), Woodroffe (1985) montre que F n'est identifiable que si $\int_{a_F}^{\infty} \frac{1}{L} dF < \infty$ (ce que implique $a_L \leq a_F$). Pour assurer l'identifiabilité du modèle LTRC, nous supposons que les v.a. Y, T et W sont mutuellement indépendantes. La loi du logarithme itéré pour les estimateurs à noyau de la densité et de la fonction de hasard a été établie par Benrabah et Ould-Saïd (1998), Sun et Zhou (1998) obtiennent la loi asymptotique pour la déviation maximale pour l'estimateur à noyau de la densité. Dans ce qui suit, nous supposons $a_L < a_H$ et $b_L \leq b_H$.

3.4.2 Estimation dans le modèle LTRC

Dans ce paragraphe nous allons reprendre les résultats obtenus par Chen et Dai (2003), d'asymptotique normalité de l'estimateur à noyau du mode simple pour un modèle LTRC avec des données α -mélangeante.

Théorème 3.8 (Chen et Dai 2003) : Sous des hypothèses classiques sur le noyau K , la densité f et les conditions de dépendance et si

$$\lim_{n \rightarrow \infty} nh_n^5 \rightarrow +\infty, \lim_{n \rightarrow \infty} nh_n^{2r+3} \rightarrow \lambda^2 < \infty, r \geq 2.$$

$$\sqrt{nh_n^3} (\hat{\theta}_n - \theta) \xrightarrow{L} N \left((-1)^r \frac{\lambda f^{(r+1)}(\theta)}{r f^{(2)}(\theta)} B_r, \sigma^2 \right)$$

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

$$\sigma^2 = \frac{\bar{F}(\theta) f(\theta)}{[f^{(2)}(\theta)]^2 C(\theta)} \int_{\mathbb{R}} [K^{(1)}(u)]^2 du$$

Chapitre 4

Asymptotic properties of the mode kernel estimator under truncated and censored data

4.1 Introduction

Let (Y, T, W) denote random variables, where Y is the interest variable called life time variable, with density function f and distribution function (d.f.) F . T is the random left truncation time with continuous d.f. L and W denote the random right censoring time with arbitrary d.f. G . It is assumed that Y , T , W are independent and that they are non-negative. In the random LTRC model one can observe (Z, T, δ) if $Z \geq T$, where $Z = \min(Y, W)$ and $\delta = I(Y \leq W)$ is the indicator of censoring status, and when $Z < T$ nothing is observed. Clearly, if Y is independent of W , then Z has d.f. $H = 1 - (1 - F)(1 - G)$. Take $\mu = P[T \leq Z]$, it's clear if $\mu = 0$ no data can be observed therefore we suppose that $\mu > 0$. Let (Y_i, T_i, W_i) $i = 1, \dots, N$ be i.i.d. as (Y, T, W) where the population size N is fixed, but unknown. The empirical data are (Z_i, T_i, δ_i) $i = 1, \dots, n$ where n is the number of observed triplet. Let

$$C(y) = P[T \leq y \leq Z \mid T \leq Z] = \mu^{-1} L(y) (1 - F(y)) (1 - G(y)). \quad (4.1)$$

A consistent estimator of $C(y)$ is given by its empirical analogues

$$C_n(y) = \frac{1}{n} \sum_{i=1}^n I(T_i \leq y \leq Z_i).$$

For LTRC data, it is important to be able to obtain nonparametric estimates of various characteristics of the distribution function F . Tsai et al. (1987) gave the nonparametric maximum likelihood estimator of F itself, called the product-limit (PL) estimator, as

$$1 - \hat{F}_n(t) = \prod_{\{i: Z_i \leq t\}} \left(1 - \frac{1}{nC_n(Z_i)}\right)^{\delta_i} \quad (4.2)$$

\hat{F}_n reduces to the Kaplan-Meier product limit (Kaplan and Meier, 1958) estimator when $T = 0$ and to the Lynden-Bell (1971) estimator when there is no right censoring. Under independent and identically distributed (i.i.d) assumptions, the properties of $\hat{F}_n(t)$ have been studied by many authors, see Gijbels and Wang (1993), Zhou (1996) and references therein. Nonparametric estimators of the density and hazard rate for $F(t)$ have been studied by Sun and Zhou (1998), among authors. The problem of estimating the unconditional /conditional mode of the probability density has given way to a large amount of related statistics literature. In the complete-data case (no censoring, no truncation), the mode estimator for i.i.d. random variables have been studied by many authors whom we quote Parzen (1962), Romano (1988) and Vieu (1996). For the complete dependent data case, the strong consistency of the conditional mode estimator was established under a ϕ -mixing condition by Collomb et al. (1987) and their results can be applied to process forecasting. In the α -maxing case, the strong consistency over a compact set and the asymptotic normality were obtained by Ould-Saïd (1993) and Louani and Ould-Saïd (1999), respectively. In the incomplete data case, for i.i.d. random variables under random right censoring, Louani (1998) studied the asymptotic normality of the kernel

estimator of the mode, while Khardani et al. (2010) established the some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. Under random left truncation, in the i.i.d. case, the kernel mode estimator has been studied by Ould-Saïd and Tatachak (2009a).

Recent years have witnessed a renewal of interest in mode estimation. To the best of our knowledge, the problem of estimating the simple mode function under truncated and censored data has not been addressed in the statistics literature. This is the central object of interest of this chapter. Our main results establish, the strong uniform convergence rate over a compact set of the kernel density estimator which allows deduce the strong convergence rate of the kernel mode estimator and establish its asymptotic normality. The rest of the paper is organized as follows : In section 2, we define a new kernel mode estimator in LTRC model with some notations. Assumptions and main results are given in section 3. Finally, the proofs of the main results are relegated to section 4 with some auxiliary results along with their proofs.

4.2 Definition of the new estimator

As a consequence of truncation, the size of the actually observed sample, n , is a $Bin(N, \mu)$ random variable. By the strong law of large numbers (SLLN), we have

$$\mu_n := \frac{n}{N} \rightarrow \mu. \quad a.s. \quad as \quad N \rightarrow +\infty \quad (4.3)$$

We denote by a superscript (*) any d.f. that is associated to the observed sample. For any distribution function K , let $a_K = \inf \{x : K(x) > 0\}$ and $b_K = \sup \{x : K(x) < 1\}$ denote the left and right end points of its support. In the current model, as discussed by Gijbels and Wang (1993) and Zhou (1996), we assume that

$$a_L \leq a_H \quad , \quad b_L \leq b_H \quad \text{and} \quad \int_{a_H}^{\infty} \frac{dF(x)}{L^2(x)} < \infty. \quad (4.4)$$

Now, let $H_1^*(.)$ be defined by

$$H_1^*(z) = P[Z \leq z, \delta = 1 | T \leq Z] = \mu^{-1} \int_{-\infty}^z L(y) (1 - G(y^-)) dF(y) \quad (4.5)$$

where $H_1^*(z)$ is consistently estimated as

$$H_{1n}^*(z) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq z, \delta_i = 1)$$

According to (4.1) and replacing F , L and G by their respective nonparametric estimator, we can consider the estimator of μ

$$\mu_n = \frac{L_n(y) (1 - \hat{F}_n(y^-)) (1 - G_n(y))}{C_n(y)}$$

for all y when $C_n(y) > 0$, and $\hat{F}_n(y^-)$ denote the left limit of \hat{F}_n at y . Here L_n is the product-limit estimator Lynden-Bell for L

$$L_n(y) = \prod_{\{i: T_i > y\}} \left(1 - \frac{1}{nC_n(T_i)}\right). \quad (4.6)$$

and G_n is the product-limit estimator Kaplan and Meier for G

$$G_n(y) = 1 - \prod_{\{i: Y_i \leq y\}} \left(1 - \frac{1 - \delta_i}{n - (i - 1)}\right). \quad (4.7)$$

Then, define our nonparametric estimator of the mode. For the complete data is present ($n = N$), it is well known that the kernel estimator of the mode θ is defined as the random variable θ_n where

$$\theta_n = \arg \max_{t \in \mathbb{R}} f_n(t) \quad (4.8)$$

and

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

where, K is a probability density function (so-called kernel function) and (h_n) is a sequence of positive real numbers (so-called bandwidth) which goes to zero as n goes to infinity.

Under random left truncation, Ould-Saïd and Tatachak (2009a) define a new estimator $\hat{f}_n(\cdot)$ based on the n actually observed pair (Y_i, T_i) given as

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

In the right censoring model, Khardani et al.(2012) define an estimator $\hat{f}_n(\cdot)$ based on the kernel smoothing technique, given as

$$\hat{f}_n(t) = \frac{1}{nh_n} \sum_{i=1}^n \frac{\delta_i}{\overline{G}_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

where $\overline{G} = 1 - G$.

Now, we define a new estimator $\hat{f}_n(\cdot)$ in *LTRC* model based on the kernel smoothing technique, given as

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)} K\left(\frac{t - Z_i}{h_n}\right) \quad (4.9)$$

Then a natural estimator of θ is

$$\hat{\theta}_n = \arg \max_{t \in [a_H, b_H]} \hat{f}_n(t) \quad (4.10)$$

It is noticed that the estimator $\hat{\theta}_n$ is not necessarily unique and our results are valid for any satisfying chosen value (4.10). We point out that our choice can be specified by taking

$$\hat{\theta}_n = \inf \left\{ t \in [a_H, b_H] \text{ such that } \hat{f}_n(t) = \max_{y \in [a_H, b_H]} \hat{f}_n(y) \right\}.$$

Suppose that K is chosen to be twice differentiable, we get the derivative of $\hat{f}_n(\cdot)$ as

$$\hat{f}_n^{(j)}(t) = \frac{\mu_n}{nh_n^{j+1}} \sum_{i=1}^n \frac{\delta_i}{L_n(Z_i) \overline{G}_n(Z_i)} K^{(j)}\left(\frac{t - Z_i}{h_n}\right) \quad \text{for } j \geq 1$$

4.3 Assumptions and main results

In what follows, we suppose that $a_L \leq a_H$, $b_L \leq b_H$ and $D = [a, b]$ be a compact set such that $D \subset D_0 = \{t : t \in [a_H, b_H]\}$. Now introduce some assumptions, gathered below for easy reference, needed to state our results.

A1 : The bandwidth (h_n) satisfies :

(i) $\frac{nh_n^6}{\log n} \rightarrow \infty$ as $n \rightarrow \infty$.

(ii) $nh_n^7 \rightarrow 0$ as $n \rightarrow \infty$.

A2 : The density function K satisfies :

(i) K is positive valued with compact support D and there exists constant M such

that $\sup_{t \in D} K(t) < M$

(ii) K and $K^{(2)}$ are Hölderian of order β for some $\beta > 0$;

(iii) $\int_{\mathbb{R}} tK(t) dt = 0$, $\int_{\mathbb{R}} t^2K(t) dt < \infty$.

A3 : The probability density $f(\cdot)$ is differentiable up to order 3 and $f^{(2)}(\cdot)$ does not vanish.

A4 : The mode θ satisfies the following property : for any $\epsilon > 0$ and t , there exists $\eta > 0$ such that $|\theta - t| \geq \epsilon$ implies $|f(\theta) - f(t)| \geq \eta$.

Remark 4.1. (Discussion of the assumptions)

Assumption **A1**(i) is needed in the behavior study of the difference between $\hat{f}_n(t)$ and $\tilde{f}_n(t)$, while **A1**(ii) is used to prove the asymptotic normality result as well as the convergence of $\hat{f}_n^{(2)}(t)$. Assumptions **A2** and **A3** are common in nonparametric estimation. Finally, assumption **A4** stipulates the uniform uniqueness of the mode point.

4.3.1 Consistency

Our first result deals with the uniform almost sure convergence with rate of the density function estimator and is stated in **Proposition 4.1**. The uniform almost sure convergence of the mode estimator and its well is derived in **Theorem 4.1**.

Proposition 4.1. Under Assumptions **A1(i)** and **A2-A3**, we have

$$\sup_{t \in D} |\hat{f}_n(t) - f(t)| = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/2}, h_n^2 \right) \right) \text{ a.s. as } n \rightarrow \infty.$$

Theorem 4.1. Under the assumptions of **Proposition 3.2** and **A4**, we have

$$\hat{\theta}_n - \theta = O \left(\max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right) \text{ a.s. as } n \rightarrow \infty.$$

4.3.2 Asymptotic normality

Now, suppose that the density function $f(\cdot)$ is unimodal at θ . Under **A4**, we have

$$f^{(1)}(\theta) = 0 \text{ and } f^{(2)}(\theta) < 0.$$

Similarly, we have

$$\hat{f}_n^{(1)}(\hat{\theta}_n) = 0 \text{ and } \hat{f}_n^{(2)}(\hat{\theta}_n) < 0.$$

Using a Taylor expansion, we get

$$0 = \hat{f}_n^{(1)}(\hat{\theta}_n) = \hat{f}_n^{(1)}(\theta) + (\hat{\theta}_n - \theta) \hat{f}_n^{(2)}(\bar{\theta}_n)$$

where $\bar{\theta}_n$ is between $\hat{\theta}_n$ and θ , which gives that

$$\hat{\theta}_n - \theta = - \frac{\hat{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \tag{4.11}$$

if the denominator does not vanish. To establish the asymptotic normality, we show that the numerator in (4.11), suitably normalized, is asymptotically normally distributed and the denominator converges in probability to $f^{(2)}(\theta)$. The result is given in the following theorem.

Theorem 4.2. Assume that hypotheses **A1-A3** hold, then we have

$$\sqrt{nh_n^3} (\hat{\theta}_n - \theta) \xrightarrow{D} N(0, \sigma^2) \quad (4.12)$$

where \xrightarrow{D} denotes the convergence in distribution and

$$\sigma^2 = \frac{\mu f(\theta)}{\overline{G}(\theta) L(\theta) [f^{(2)}(\theta)]^2} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr$$

4.4 Auxiliary results and proofs

The first Lemma deals with the behavior of the difference between $E[\tilde{f}_n(t)]$ and $f(t)$.

Lemma 4.1. Under Assumptions **A2** and **A3**, we have

$$\sup_{t \in D} \left| E[\tilde{f}_n(t)] - f(t) \right| = O(h_n^2)$$

Proof. Adapting (4.9), we define

$$\tilde{f}_n(t) = \frac{\mu}{nh_n} \sum_{i=1}^n \frac{\delta_i}{L(Z_i) \overline{G}(Z_i)} K\left(\frac{t - Z_i}{h_n}\right)$$

Indeed, using a change of variable and a Taylor expansion, we have

$$\begin{aligned}
E \left[\tilde{f}_n(t) \right] - f(t) &= \frac{1}{h_n} E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K \left(\frac{t - Z_1}{h_n} \right) \right] - f(t) \\
&= \frac{1}{h_n} \int \frac{\mu}{L(y) \overline{G}(y)} K \left(\frac{t - y}{h_n} \right) dH_1^*(y) - f(t) \\
&= \frac{1}{h_n} \int K \left(\frac{t - y}{h_n} \right) f(y) dy - f(t) \\
&= \int K(s) \frac{(hs)^2}{2} f^{(2)}(\zeta) ds
\end{aligned}$$

with ζ is between $t - sh$ and t . Thus

$$\left| E \left[\tilde{f}_n(t) \right] - f(t) \right| \leq \frac{h^2}{2} \sup_{t \in D} \left| f^{(2)}(t) \right| \int s^2 K(s) ds \quad (4.13)$$

Under the given condition, the result holds. \square

The second Lemma deals with the behavior of the difference between $\hat{f}_n(t)$ and $\tilde{f}_n(t)$.

Lemma 4.2. Under Assumptions **A1(i)** and **A2-A3**, we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| = O \left(\sqrt{\frac{\log \log n}{nh_n^2}} \right).$$

Proof. We observe that one can have the following decomposition

$$\begin{aligned}
\left| \hat{f}_n(t) - \tilde{f}_n(t) \right| &\leq \frac{1}{nh_n} \sum_{i=1}^n \left| \delta_i K \left(\frac{t - Z_i}{h_n} \right) \right| \left| \frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right| \\
&\leq \underbrace{\frac{M}{h_n} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_{I} \frac{1}{n} \sum_{i=1}^n \delta_i
\end{aligned}$$

Since $I = O \left(\sqrt{\frac{\log \log n}{n}} \right)$ (see Lemma 3 in Sun and Zhou, 1998), in conjunction with the SLLN on the censoring law (see formula 4.28 in Deheuvels and Einmahl, 2000), we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| \leq \frac{c}{h_n} \sqrt{\frac{\log \log n}{n}}$$

we conclude the proof. \square

Lemma 4.3. Under Assumptions **A1**(i) and **A2-A3**, we have

$$\sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_n}} \right)$$

Proof. We use covering set techniques. Indeed, since D is a compact set, so it can be covered by a finite number d_n of intervals centered at t_1, \dots, t_{d_n} of length λ_n such that $\lambda_n = h_n^\gamma$, where $\gamma > \frac{\beta+1}{\beta}$, and β is the Hölder exponent. As D is bounded, there exists a constant A such that $d_n \lambda_n \leq A$. Now we put $J(t) = \arg \min_{j=1, \dots, d_n} |t - t_j|$. Thus we have the following decomposition

$$\begin{aligned} \sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| &\leq \max_{j=1, \dots, d_n} \sup_{t \in D} \left| \tilde{f}_n(t) - \tilde{f}_n(t_{J(t)}) \right| \\ &\quad + \max_{j=1, \dots, d_n} \left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| \\ &\quad + \max_{j=1, \dots, d_n} \sup_{t \in D} \left| E \left[\tilde{f}_n(t_{J(t)}) \right] - E \left[\tilde{f}_n(t) \right] \right| \\ &= I_{1n} + I_{2n} + I_{3n} \end{aligned} \tag{4.14}$$

Firstly, concerning I_{1n} and I_{3n} , with assumption **A2**, we have

$$I_{1n} \leq \frac{c}{L(a) \overline{G}(b)} \lambda_n^\beta h_n^{-1-\beta}$$

The condition upon λ_n , we get

$$I_{1n} = O(1) \quad a.s. \text{ as } n \rightarrow \infty \tag{4.15}$$

As to I_{2n} , for all $\epsilon > 0$, we have

$$P \left[\max_{j=1, \dots, d_n} \left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| > \epsilon \right] \leq d_n P \left[\left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| > \epsilon \right] \quad (4.16)$$

Now, for any $i \geq 1$, we put

$$\Delta_i(t) = \frac{\mu}{nh_n} \left[\frac{\delta_i}{L(Z_i)\bar{G}(Z_i)} K\left(\frac{t-Z_i}{h_n}\right) - E \left[\frac{\delta_1}{L(Z_1)\bar{G}(Z_1)} K\left(\frac{t-Z_1}{h_n}\right) \right] \right] \quad (4.17)$$

Under **A2**, the rv's $\Psi_i(t_{J(t)}) := nh_n \Delta_i(t_{J(t)})$ are centered and bounded by $\frac{2\mu}{L(a)\bar{G}(b)}M < \infty$. Then, applying Höeffding's inequality (see Shorack and Wellner, 1986, p. 855), we obtain

$$\begin{aligned} & P \left[\max_{j=1, \dots, d_n} \left| \sum_{i=1}^n \Delta_i(t_{J(t)}) \right| > \epsilon \right] \\ &= P \left[\max_{j=1, \dots, d_n} \left| \sum_{i=1}^n \Psi_i(t_{J(t)}) \right| > \epsilon nh_n \right] \\ &\leq \exp(-\epsilon^2 n^2 h_n^2 c) \end{aligned}$$

where c is a positive constant depending only on M , μ , $L(a)$ and $\bar{G}(b)$, so (4.16) becomes

$$\begin{aligned} & P \left[\max_{j=1, \dots, d_n} \left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| > \epsilon \right] \\ &\leq A (nh_n)^{-\gamma} n^{\gamma - c\epsilon^2 n^2 h_n^2} / \log n \end{aligned} \quad (4.18)$$

With **A1(i)**, which yields that the last term of (4.18) is the general term of a convergent series. Then, by Borel-Cantelli's Lemma, the first term of (4.18) goes to zero almost surely. Otherwise, if we replace ϵ by $\epsilon_0 \sqrt{\frac{\log n}{nh_n}}$, for some $\epsilon_0 > 0$ in all steps of the lemma, we have

$$\sqrt{\frac{nh_n}{\log n}} I_{2n} = O(1) \quad a.s. \quad (4.19)$$

In conjunction with (4.15), we complete the proof of the Lemma. \square

Proof of Proposition 4.1. Using the triangle inequality, we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| \leq \sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| + \sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| + \sup_{t \in D} \left| E \left[\tilde{f}_n(t) \right] - f(t) \right|$$

Then Lemmas **4.1**, **4.2** and **4.3** give the result. \square

Proof of Theorem 4.1. Standard argument gives us

$$\begin{aligned} \left| f(\hat{\theta}_n) - f(\theta) \right| &\leq \left| f(\hat{\theta}_n) - \hat{f}_n(\hat{\theta}_n) \right| + \left| \hat{f}_n(\hat{\theta}_n) - f(\theta) \right| \\ &\leq \sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| + \left| \hat{f}_n(\hat{\theta}_n) - f(\theta) \right| \\ &\leq 2 \sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| \end{aligned} \quad (4.20)$$

The *a.s.* uniform consistency of $\hat{\theta}_n$ follows immediately from **Proposition 4.1** and **A4**.

Now a Taylor expansion gives

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

where $\hat{\theta}_n^*$ is between $\hat{\theta}_n$ and θ . Then by (4.20), **A3** and **A4**, we have

$$\left| \hat{\theta}_n - \theta \right| \leq 2 \sqrt{\frac{\sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right|}{\left| f^{(2)}(\hat{\theta}_n^*) \right|}}$$

Hence by **Proposition 4.1** the proof is completed. \square

Proof of theorem 4.2 : Form (4.11) we have the following decomposition

$$\begin{aligned}
\sqrt{nh_n^3} (\hat{\theta}_n - \theta) &= \sqrt{nh_n^3} \frac{\hat{f}_n^{(1)}(\theta) - \tilde{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&\quad + \sqrt{nh_n^3} \frac{\tilde{f}_n^{(1)}(\theta) - E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&\quad + \sqrt{nh_n^3} \frac{E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\
&= \frac{J_1 + J_2 + J_3}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \tag{4.21}
\end{aligned}$$

So, it suffices to show that $J_i \rightarrow 0$ ($i = 1, 3$) where as J_2 is asymptotically normal and $\hat{f}_n^{(2)}(\bar{\theta}_n)$ converges in probability to $f^{(2)}(\theta)$ \square

Lemma 4.4. Under assumptions **A1-A3**, we have

$$J_1 \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned}
|J_1| &= \sqrt{nh_n^3} \left| \frac{1}{nh_n^2} \sum_{i=1}^n \delta_i K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \left(\frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right) \right| \\
&\leq \underbrace{\sqrt{nh_n} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_I \frac{1}{nh_n} \sum_{i=1}^n \left| K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \right|
\end{aligned}$$

Then for any $\epsilon > 0$, under **A2**, we have

$$P \left[\frac{1}{nh_n} \sum_{i=1}^n \left| K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \right| > \epsilon \right] \leq \frac{1}{nh_n \epsilon} \sum_{i=1}^n E \left(\left| K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \right| \right) \rightarrow \frac{f(\theta)}{\mu \epsilon} \int_{\mathbb{R}} |K^{(1)}(r)| dr < \infty$$

Since $C(a) > 0$, $I = O\left(\sqrt{\frac{\log \log n}{n}}\right)$ (see Lemma 3 in Sun and Zhou, 1998), we have $|J_1| = O\left(\sqrt{h_n \log \log n}\right)$, which goes to zero under **A1**(i, ii). \square

Lemma 4.5. Under assumptions **A1**(ii) and **A2-A3**, we have

$$J_3 \rightarrow 0 \text{ a.s. as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned} E \left[\tilde{f}_n^{(1)}(\theta) \right] &= \frac{1}{h_n^2} E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K^{(1)} \left(\frac{\theta - Z_1}{h_n} \right) \right] \\ &= \frac{1}{h_n^2} \int_{\mathbb{R}} \frac{\mu}{L(v) \overline{G}(v)} K^{(1)} \left(\frac{\theta - v}{h_n} \right) dH_1^*(v) \\ &= \frac{1}{h_n^2} \int_{\mathbb{R}} K^{(1)} \left(\frac{\theta - v}{h_n} \right) f(v) dv \end{aligned}$$

Under assumptions **A2-A3** and integration by part, we have

$$\begin{aligned} E \left[\tilde{f}_n^{(1)}(\theta) \right] &= \frac{1}{h_n} \int_{\mathbb{R}} K \left(\frac{\theta - v}{h_n} \right) f^{(1)}(v) dv \\ &= \int_{\mathbb{R}} K(r) f^{(1)}(\theta - rh_n) dr \end{aligned}$$

by a Taylor expansion of $f^{(1)}(\cdot)$ around θ , **A2** and the definition of the mode, we get

$$\sqrt{nh_n^3} E \left[\tilde{f}_n^{(1)}(\theta) \right] = \sqrt{nh_n^7} \int_{\mathbb{R}} r^2 K(r) f^{(3)}(\theta^*) dr$$

where θ^* is between θ and $\theta - rh_n$, by **A1**(ii) and **A2-A3** we conclude $J_3 \rightarrow 0$ a.s. as $n \rightarrow \infty$. \square

Lemma 4.6. Under assumptions **A1**(ii) and **A2-A3**, we have

$$\text{Var} [J_2] \longrightarrow \frac{\mu f(\theta)}{\overline{G}(\theta) L(\theta)} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr, \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned} \text{Var} [J_2] &= nh_n^3 \text{Var} \left[\tilde{f}_n^{(1)}(\theta) \right] \\ &= \frac{1}{h_n} E \left[\frac{\mu^2 \delta_1}{L^2(Z_1) \overline{G}^2(Z_1)} (K^{(1)})^2 \left(\frac{\theta - Z_1}{h_n} \right) \right] \\ &\quad - \frac{1}{h_n} \left\{ E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K^{(1)} \left(\frac{\theta - Z_1}{h_n} \right) \right] \right\}^2 \\ &= \Phi_{1,n} - \Phi_{2,n} \end{aligned}$$

On the one hand,

$$\begin{aligned} \Phi_{1,n} &= \frac{1}{h_n} \int_{\mathbb{R}} \frac{\mu^2}{L^2(y) \overline{G}^2(y)} (K^{(1)})^2 \left(\frac{\theta - y}{h_n} \right) dH_1^*(y) \\ &= \int_{\mathbb{R}} \frac{\mu}{L(\theta - rh_n) \overline{G}(\theta - rh_n)} (K^{(1)})^2(r) f(\theta - rh_n) dr \end{aligned} \tag{4.22}$$

Again, a Taylor expansion gives us

$$\Phi_{1,n} = \mu f(\theta) \int_{\mathbb{R}} \frac{1}{L(\theta - rh_n) \overline{G}(\theta - rh_n)} (K^{(1)})^2(r) dr + o(1)$$

Then, since $G(\cdot)$ and $L(\cdot)$ are continuous, we have, under **A2-A3**

$$\Phi_{1,n} = \frac{\mu f(\theta)}{\overline{G}(\theta) L(\theta)} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr + o(1)$$

On the other hand, by Lemma 4.5

$$\Phi_{2,n} = \frac{J_3^2}{n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Finally, we obtain

$$\text{Var} [J_2] \longrightarrow \frac{\mu f(\theta)}{\overline{G}(\theta) L(\theta)} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr \quad \text{as } n \rightarrow \infty.$$

This prove the Lemma. \square

Now, to prove that $\hat{f}_n^{(2)}(\bar{\theta}_n)$ converges in probability to $f^{(2)}(\theta)$, indeed, we have the following

Lemma 4.7. Under assumptions **A1**(i) and **A2-A3**, we have

$$\sup_{t \in D} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Proof. Since

$$\begin{aligned} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| &\leq \left| \hat{f}_n^{(2)}(t) - \tilde{f}_n^{(2)}(t) \right| + \left| \tilde{f}_n^{(2)}(t) - E \left[\tilde{f}_n^{(2)}(t) \right] \right| + \left| E \left[\tilde{f}_n^{(2)}(t) \right] - f^{(2)}(t) \right| \\ &= \gamma_{1,n}(t) + \gamma_{2,n}(t) + \gamma_{3,n}(t) \end{aligned}$$

The proof is completed through Lemmas **4.8-4.10**. \square

Lemma 4.8. Under assumptions **A1**(i) and **A2-A3**, we have

$$\sup_{t \in D} \gamma_{1,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned}
\sup_{t \in D} \gamma_{1,n}(t) &\leq \frac{1}{nh_n^3} \sum_{i=1}^n \left| \delta_i K^{(2)} \left(\frac{t - Z_i}{h_n} \right) \left(\frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right) \right| \\
&\leq \underbrace{\frac{1}{h_n^2} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_I \frac{1}{nh_n} \sum_{i=1}^n K^{(2)} \left(\frac{t - Z_i}{h_n} \right)
\end{aligned}$$

Analogously as in Lemma 4.4, under **A2** we get for $\epsilon > 0$

$$P \left[\frac{1}{nh_n} \sum_{i=1}^n \left| K^{(2)} \left(\frac{\theta - Z_i}{h_n} \right) \right| > \epsilon \right] \leq \frac{1}{nh_n \epsilon} \sum_{i=1}^n E \left(\left| K^{(2)} \left(\frac{\theta - Z_i}{h_n} \right) \right| \right) \rightarrow \frac{f(\theta)}{\mu \epsilon} \int_{\mathbb{R}} |K^{(2)}(r)| dr < \infty$$

In the same manner, by **A1(i)** we get the result. \square

Lemma 4.9. Under assumptions **A2** and **A3**, we have

$$\sup_{t \in D} \gamma_{3,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. Since integrating by parts twice and using a change of variable it follows as in Lemma 4.1, we obtain

$$\begin{aligned}
\sup_{t \in D} \gamma_{2,n}(t) &= \sup_{t \in D} \left| \int_{\mathbb{R}} K(r) f^{(2)}(t - rh_n) dr - f^{(2)}(t) \right| \\
&= O(h_n^2)
\end{aligned}$$

This proves the Lemma. \square

Lemma 4.10. Under assumptions **A1(i)** and **A2-A3**, we have

$$\sup_{t \in D} \gamma_{3,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. It proceeds along the same steps and by invoking the same arguments as in the proof of **Lemma 4.3**, a suitable covering of the compact set D and changing the variables $\Delta_i(t)$ into the following

$$\begin{aligned}\Gamma_i(t) &= \frac{\mu}{nh_n^3} \left[\frac{1}{L(Z_i)\overline{G}(Z_i)} K^{(2)}\left(\frac{t-Z_i}{h_n}\right) - E \left[\frac{1}{L(Z_i)\overline{G}(Z_i)} K^{(2)}\left(\frac{t-Z_i}{h_n}\right) \right] \right] \\ &= (nh_n^3)^{-1} V_i(t)\end{aligned}$$

This completes the proof of the Lemma.

In order to achieve the proof, we observe that we have, for n large enough

$$\left| \hat{f}_n^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta) \right| \leq \sup_{t \in D} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| + |f^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta)|$$

By lemma **4.7** and assumption **A3**, we get the result. \square

Now the final step to prove **Theorem 3.4** is to show the Berry-Esséen condition for J_2 (see Chow and Teicher, 1997, p.322). For that, put

$$J_2 = \sum_{i=1}^n \Upsilon_{i,n}(\theta)$$

where

$$\Upsilon_{i,n}(\theta) = \frac{\mu}{\sqrt{nh_n}} \left\{ \frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K^{(1)}\left(\frac{\theta-Z_i}{h_n}\right) - E \left[\frac{\delta_1}{L(Z_1)\overline{G}(Z_1)} K^{(1)}\left(\frac{\theta-Z_1}{h_n}\right) \right] \right\}$$

Then, we have

Lemma 4.11. Under assumptions **A1-A3**, we have

$$\sum_{i=1}^n E [|\Upsilon_{i,n}(\theta)|^3] < \infty$$

Proof. Applying the C_r -inequality (see Loève (1963), p.155), we have

$$E [|\Upsilon_{i,n}(\theta)|^3] \leq 4(nh_n)^{-\frac{3}{2}} \left\{ E \left[\left| \frac{\mu\delta_1}{L(Z_1)\overline{G}(Z_1)} K^{(1)} \left(\frac{\theta - Z_1}{h_n} \right) \right|^3 \right] + \left| E \left[\frac{\mu\delta_1}{L(Z_1)\overline{G}(Z_1)} K^{(1)} \left(\frac{\theta - Z_1}{h_n} \right) \right] \right|^3 \right\} \quad (4.23)$$

Both expectation terms in (4.23) being finite under **A1-A3**, we get

$$\sum_{i=1}^n E [|\Upsilon_{in}(\theta)|^3] = o(nh_n^3)^{-\frac{1}{2}} = o(1)$$

This completes the proof of **Lemma 4.11** and there for that of **Theorem 4.2** □

Chapitre 5

A nonparametric mode estimate under LTRC model and dependent data

Abstract

In this chapter, we consider the estimation of the mode function when the data are subject to random left truncation and right censorship. Under suitable conditions, we drive almost sure (a.s.) convergence and asymptotic normality for a kernel estimator of the mode, when observed data exhibit α -mixing dependence.

Keywords : Asymptotic normality ; Kernel estimator ; Simple mode ; Strong-mixing ; Uniform almost sure convergence ; Truncated and Censored data.

5.1 Introduction

Let (Y, T, W) denote random variables, where Y is the interest variable called lifetime variable, with density function f and distribution function (d.f.) F . T is the random left truncation time with continuous d.f. L and W denote the random right censoring time with arbitrary d.f. G . In the random left truncated and right censored (LTRC) model one

can observe (Z, T, δ) if $Z \geq T$, where $Z = \min(Y, W)$ and $\delta = I(Y \leq W)$ is the indicator of censoring status, and when $Z < T$ nothing is observed. Clearly, if Y is independent of W , then Z has d.f. $H = 1 - (1 - F)(1 - G)$. Taking into account the truncation and the censorship effects, we denote by $\{(Z_i, T_i, \delta_i), i = 1, \dots, n\}$ the actually observed sample (*i.e.* $Z_i \geq T_i, \forall i$) and suppose that $\mu = P[T \leq Z] > 0$.

Let

$$C(y) = P[T \leq y \leq Z \mid T \leq Z] = \mu^{-1}L(y)(1 - H(y^-)). \quad (5.1)$$

and define their empirical estimators as

$$C_n(y) = \frac{1}{n} \sum_{i=1}^n I(T_i \leq y \leq Z_i).$$

For LTRC data, it is important to be able to obtain nonparametric estimators of the various characteristics of the distribution function, F . Tsai et al. (1987) gave the nonparametric maximum likelihood estimator of F itself, called the product-limit (PL) estimator, as

$$1 - \hat{F}_n(t) = \prod_{\{i: Z_i \leq t\}} \left(1 - \frac{1}{nC_n(Z_i)}\right)^{\delta_i} \quad (5.2)$$

\hat{F}_n reduces to the Kaplan-Meier product limit (Kaplan and Meier, 1958) estimator when $T = 0$ and to the Lynden-Bell (1971) estimator when there is no right censoring. Under independent and identically distributed (i.i.d) assumptions, the properties of $\hat{F}_n(\cdot)$ have been studied by many authors. See Gijbels and Wang (1993), Zhou (1996) and references therein. Nonparametric estimators of the density and hazard rate for $F(\cdot)$ have been studied by Sun and Zhou (1998), among others. The problem of estimating the unconditional/conditional mode of the probability density has given way to a large amount of related statistics literature. In the complete-data case (no censoring, no truncation), the mode estimator for i.i.d. random variables have been studied by many authors whom

we quote Parzen (1962), Romano (1988) and Vieu (1996). For the complete dependent data case, the strong consistency of the conditional mode estimator was established under a ϕ -mixing condition by Collomb et al. (1987) and their results can be applied to process forecasting. In the α -maxing case, the strong consistency over a compact set and the asymptotic normality were obtained by Ould-Saïd (1993) and Louani and Ould-Saïd (1999), respectively. In the incomplete data case, for i.i.d. random variables under random right censoring, Louani (1998) studied the asymptotic normality of the kernel estimator of the mode, while Khardani et al. (2010) established the some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. These authors also studied the strong uniform consistency of the mode estimator for censored time series (Khardani et al, 2012). Under random left truncation (RLT model), in the i.i.d. case, Ould-Saïd and Tatachak (2009a) studied the nonparametric estimation of the simple mode under left-truncation model. They also studied the strong consistency rate for the kernel mode estimator under strong mixing hypothesis and left truncation (2009b).

The aim of this chapter is to establish a strong uniform convergence rate and asymptotic normality for a kernel mode estimator with randomly left truncated and right censored data under α -mixing condition. We begin by recalling the definition of the strong property, for this we introduce the following notations.

Let $\{U_i, i \geq 1\}$ denote a sequence of random variables and $\mathcal{F}_i^k(U)$ denote the σ -algebra of events generated by $\{U_j, i \leq j \leq k\}$.

Given a positive integer n , set :

$$\alpha(n) = \sup \left\{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^k(U), B \in \mathcal{F}_{k+n}^\infty(U), k \in \mathbb{N}^* \right\} \quad (5.3)$$

The sequence is said to be α -mixing if the mixing coefficient $\alpha(n) \rightarrow 0$ as $n \rightarrow \infty$. The α -mixing condition is the weakest among mixing conditions known in the literature. Many stochastic processes satisfy the α -mixing condition; the ARMA processes are

geometrically strongly mixing, i.e., there exists $\varrho \in (0, 1)$ such that, $\alpha(n) = O(\varrho^{-n})$ (see, e.g., Jones 1978).

The rest of the paper is organized as follows : In section 2, we define a new kernel mode estimator in LTRC model with some notations. Assumptions and main results are given in section 3. Finally, the proofs of the main results are relegated to section 4 with some auxiliary results.

5.2 Definition of the estimator

For any distribution function (d.f.) K , let $a_K = \inf \{x : K(x) > 0\}$ and $b_K = \sup \{x : K(x) < 1\}$ denote the left and right end points of its support. In the current model, as discussed by Gijbels and Wang (1993) and Zhou (1996), we assume that

$$a_L \leq a_H \quad , \quad b_L \leq b_H \quad \text{and} \quad \int_{a_H}^{\infty} \frac{dF(x)}{L^2(x)} < \infty. \quad (5.4)$$

We denote by a superscript (*) any d.f. that is associated to the observed sample.

Now, let $H_1^*(\cdot)$ be defined as

$$H_1^*(z) = P[Z \leq z, \delta = 1 \mid T \leq Z] = \mu^{-1} \int_{-\infty}^z L(y) (1 - G(y^-)) dF(y) \quad (5.5)$$

where $H_1^*(z)$ is consistently estimated as

$$H_{1n}^*(z) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq z, \delta_i = 1)$$

According to (5.1) and by replacing F , L and G by their respective nonparametric estimator, we can consider the estimator of μ

$$\mu_n = \frac{L_n(y) \left(1 - \hat{F}_n(y^-)\right) (1 - G_n(y))}{C_n(y)}$$

For all y when $C_n(y) > 0$, and $\hat{F}_n(y^-)$ denote the left limit of \hat{F}_n at y . Here, L_n is the product-limit estimator Lynden-Bell for L

$$L_n(y) = \prod_{\{i: T_i > y\}} \left(1 - \frac{1}{nC_n(T_i)}\right). \quad (5.6)$$

And G_n is the product-limit estimator Kaplan and Meier for G

$$G_n(y) = 1 - \prod_{\{i: Y_i \leq y\}} \left(1 - \frac{1 - \delta_i}{n - (i - 1)}\right). \quad (5.7)$$

For the complete data, it is well known that the kernel estimator of the mode θ is defined as the random variable θ_n

$$\theta_n = \arg \max_{t \in \mathbb{R}} f_n(t) \quad (5.8)$$

where

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

where, K is a probability density function (so-called kernel function) and (h_n) is a sequence of positive real numbers (so-called bandwidth) which goes to zero as n goes to infinity.

Under *RTL* model, Ould-Saïd and Tatachak (2009a) define a new estimator $\hat{f}_n(\cdot)$ based on the n actually observed pair (Y_i, T_i) given as

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

In the right censoring model, Khardani et al. (2012) define an estimator $\hat{f}_n(\cdot)$ based on the kernel smoothing technique, given by

$$\hat{f}_n(t) = \frac{1}{nh_n} \sum_{i=1}^n \frac{\delta_i}{\overline{G}_n(Y_i)} K\left(\frac{t - Y_i}{h_n}\right)$$

where $\overline{G} = 1 - G$.

Now, we define a new estimator $\hat{f}_n(\cdot)$ in *LTRC* model based on the kernel smoothing technique, given as

$$\hat{f}_n(t) = \frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)} K\left(\frac{t - Z_i}{h_n}\right) \quad (5.9)$$

Then a natural estimator of θ is

$$\hat{\theta}_n = \arg \max_{t \in [a_H, b_H]} \hat{f}_n(t) \quad (5.10)$$

It is noticed that the estimator $\hat{\theta}_n$ is not necessarily unique and our results are valid for any chosen value satisfying (5.10). We point out that our choice can be specified by taking

$$\hat{\theta}_n = \inf \left\{ t \in [a_H, b_H] \text{ such that } \hat{f}_n(t) = \max_{y \in [a_H, b_H]} \hat{f}_n(y) \right\}.$$

Suppose that K is chosen to be twice differentiable, we get the derivative of $\hat{f}_n(\cdot)$ as

$$\hat{f}_n^{(j)}(t) = \frac{\mu_n}{nh_n^{j+1}} \sum_{i=1}^n \frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)} K^{(j)}\left(\frac{t - Z_i}{h_n}\right) \quad \text{for } j \geq 1$$

Adapting (5.9), we define

$$\tilde{f}_n(t) = \frac{\mu}{nh_n} \sum_{i=1}^n \frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K\left(\frac{t - Z_i}{h_n}\right)$$

5.3 Assumptions and main results

In what follows, we suppose that $a_L \leq a_H$, $b_L \leq b_H$ and $D = [a, b]$ be a compact set such that $D \subset D_0 = \{t : t \in [a_H, b_H]\}$.

Now, some assumptions needed to state our results, are introduced and gathered below for easy reference.

A1 : The bandwidth h_n satisfies :

$$(i) \quad \sqrt{\frac{\log \log n}{n}} = O(h_n^3);$$

$$(ii) \quad \lim_{n \rightarrow \infty} \frac{nh_n^{(\nu+4)/(\nu-4)}}{(\log n)^{\frac{\nu+1}{\nu-4}} (\log \log n)^{\frac{6}{\nu-4}}} \rightarrow \infty \quad \text{as } n \rightarrow \infty \quad (\text{where } \nu \text{ is as in } \mathbf{A3}).$$

A2 : The density function K satisfies :

(i) K is three times differentiable with compact support on \mathbb{R} ;

(ii) K and $K^{(2)}$ are Lipschitz continuous functions;

$$(iii) \quad \int_{\mathbb{R}} tK(t) dt = 0, \quad \int_{\mathbb{R}} t^2 K(t) dt < \infty.$$

A3 : $(Y_i)_{i \geq 1}$ is a stationary α -mixing sequence of rv's, with coefficient $\alpha(n) = O(n^{-\nu})$ for some $\nu > 4$.

A4 : The probability density $f(\cdot)$ is differentiable up to order 3 and $f^{(2)}(\cdot)$ does not vanish.

A5 : The joint density $f_{1j}(\cdot, \cdot)$ of (Y_1, Y_j) exists and satisfies for any constant $M > 0$

$$\sup_{r,t} |f_{1j}(r, t) - f(r)f(t)| \leq M < \infty$$

A6 : The mode θ satisfies the following property : for any $\epsilon > 0$ and t , there exists a $\eta > 0$ such that $|\theta - t| \geq \epsilon$ implies $|f(\theta) - f(t)| \geq \eta$.

The following assumptions intervene in the asymptotic normality results

B1 : The bandwidth (h_n) satisfies $nh_n^7 \rightarrow 0$ as $n \rightarrow \infty$.

B2 : Let (M_n) and (N_n) be subsequences of (n) tending to infinity such that :

$$(i) \quad M_n + N_n \leq n; \quad \frac{r_n M_n}{n} \rightarrow 1 \quad \text{and} \quad \frac{r_n N_n}{n} \rightarrow 0;$$

$$(ii) \quad M_n (nh_n)^{-1/2} \rightarrow 0;$$

(iii) $r_n \alpha(N_n) \rightarrow 0$, as $n \rightarrow \infty$, where (r_n) is the largest positive integer for which

$$r_n(M_n + N_n) \leq n.$$

Remark 5.1. (Discussion of the assumptions)

Assumption **A1**(i) is needed in the behavior study of the difference between $\hat{f}_n(\cdot)$ and $\tilde{f}_n(t)$, while **A1**(ii) is gives a condition for the bandwidth which allows to estimate the covariance term. Assumptions **A2** and **A4** are common in nonparametric estimation. Assumption **A3** specifies the considered models as well as the rate for the mixing coefficient. Assumption **A5** is needed for covariance calculus. Assumption **A6** stipulates the uniform uniqueness of the mode function. Assumption **B1** is used to prove the asymptotic normality result as well as the convergence of $\hat{f}_n^{(2)}(\cdot)$. Finally, assumption **B2** deals with real sequences that take part in establishing our results (see Louani and Ould-Saïd 1999).

Remark 5.2. Assumption **A2**(i) implies that the kernels K , $K^{(1)}$ and $K^{(2)}$ are bounded by constants M , M' and M'' respectively.

5.3.1 Consistency

In this subsection, we prove the consistency of our estimator and give a rate of convergence.

Proposition 5.1. Under assumptions **A1-A5**, we have

$$\sup_{t \in D} |\hat{f}_n(t) - f(t)| = O \left\{ \max \left(\left(\frac{\log n}{nh_n} \right)^{1/2}, h_n^2 \right) \right\} \text{ a.s. as } n \rightarrow \infty.$$

Theorem 5.1. Under the assumptions of **Proposition 5.1** and **A6**, we have

$$\hat{\theta}_n - \theta = O \left\{ \max \left(\left(\frac{\log n}{nh_n} \right)^{1/4}, h_n \right) \right\} \text{ a.s. as } n \rightarrow \infty.$$

5.3.2 Asymptotic normality

Now, suppose that the density function $f(\cdot)$ is unimodal at θ . Under **A6**, we have

$$f^{(1)}(\theta) = 0 \text{ and } f^{(2)}(\theta) < 0.$$

Similarly, we have

$$\hat{f}_n^{(1)}(\hat{\theta}_n) = 0 \text{ and } \hat{f}_n^{(2)}(\hat{\theta}_n) < 0.$$

Using a Taylor expansion, we get

$$0 = \hat{f}_n^{(1)}(\hat{\theta}_n) = \hat{f}_n^{(1)}(\theta) + (\hat{\theta}_n - \theta) \hat{f}_n^{(2)}(\bar{\theta}_n)$$

where $\bar{\theta}_n$ is between $\hat{\theta}_n$ and θ , which gives that

$$\hat{\theta}_n - \theta = -\frac{\hat{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \tag{5.11}$$

if the denominator does not vanish.

To establish the asymptotic normality, we show that the numerator in (5.11), suitably normalised, is asymptotically normally distributed and that the denominator converges in probability to $f^{(2)}(\theta)$. The result is given in the following theorem :

Theorem 5.2. Assume that assumptions **A1(i)**, **A2-A5** and **B1-B2** hold, then we have

$$\sqrt{nh_n^3} (\hat{\theta}_n - \theta) \xrightarrow{D} N(0, \sigma^2) \tag{5.12}$$

where \xrightarrow{D} denotes the convergence in distribution and

$$\sigma^2 = \frac{\mu f(\theta)}{\bar{G}(\theta) L(\theta) [f^{(2)}(\theta)]^2} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr.$$

5.4 Auxiliary results and proofs

Some auxiliary results and notations are needed to be proved. The first Lemma deals with the behavior of the difference between $E \left[\tilde{f}_n(t) \right]$ and $f(t)$.

Lemma 5.1. Under assumptions **A2** and **A4**, we have

$$\sup_{t \in D} \left| E \left[\tilde{f}_n(t) \right] - f(t) \right| = O(h_n^2)$$

Proof. Indeed, using a change of variable and a Taylor expansion, we have

$$\begin{aligned} E \left[\tilde{f}_n(t) \right] - f(t) &= \frac{1}{h_n} E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K \left(\frac{t - Z_1}{h_n} \right) \right] - f(t) \\ &= \frac{1}{h_n} \int \frac{\mu}{L(y) \overline{G}(y)} K \left(\frac{t - y}{h_n} \right) dH_1^*(y) - f(t) \\ &= \frac{1}{h_n} \int K \left(\frac{t - y}{h_n} \right) f(y) dy - f(t) \\ &= \int K(s) \frac{(hs)^2}{2} f^{(2)}(\zeta) ds \end{aligned}$$

with ζ is between $t - sh$ and t . Thus

$$\left| E \left[\tilde{f}_n(t) \right] - f(t) \right| \leq \frac{h^2}{2} \sup_{t \in D} \left| f^{(2)}(t) \right| \int s^2 K(s) ds \quad (5.13)$$

Under the given condition, the result holds. \square

The second Lemma deals the with behavior of the difference between $\hat{f}_n(t)$ and $\tilde{f}_n(t)$.

Lemma 5.2. Under assumptions **A1(i)**, **A2** and **A4**, for n large enough we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| = O(h_n^2).$$

Proof. We have the following decomposition

$$\begin{aligned}
\left| \hat{f}_n(t) - \tilde{f}_n(t) \right| &\leq \frac{1}{nh_n} \sum_{i=1}^n \delta_i K \left(\frac{t - Z_i}{h_n} \right) \left| \frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right| \\
&\leq \underbrace{\frac{M}{h_n} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_{I} \frac{1}{n} \sum_{i=1}^n \delta_i \\
&\leq \frac{M}{h_n} \left[\sup_{t \in D} \left| \frac{\hat{F}_n(t) - F(t)}{C_n(t)} \right| + \sup_{t \in D} \left| \frac{1 - F(t)}{C_n(t)C(t)} (C_n(t) - C(t)) \right| \right] \frac{1}{n} \sum_{i=1}^n \delta_i \\
&\leq \frac{M}{h_n} \left[\frac{1}{C_n(a)} \underbrace{\sup_{t \in D} |\hat{F}_n(t) - F(t)|}_I + \frac{1 - F(a)}{C_n(a)C(a)} \underbrace{\sup_{t \in D} |C_n(t) - C(t)|}_J \right] \frac{1}{n} \sum_{i=1}^n \delta_i
\end{aligned}$$

Since $C(a) > 0$, $I = O\left(\sqrt{\frac{\log \log n}{n}}\right)$ (see Lemma 2.1 in Chen and Dai 2003) and $J = O\left(\sqrt{\frac{\log \log n}{n}}\right)$ (see Theorem 3.2 in Cai and Roussas 1992). In conjunction with the SLLN on the censoring law (see Theorem 3.2 in Cai and Roussas 1992), we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| \leq \frac{c}{h_n} \sqrt{\frac{\log \log n}{n}} \quad a.s.$$

where c is a positive constant. With assumption **A1**(i) permits concluding the proof. \square

Lemma 5.3. Under assumptions **A1-A5**, we have

$$\sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_n}} \right)$$

Proof. We use covering set techniques. Indeed, since D is a compact set, so it can be covered by a finite number d_n of intervals centered at t_j of length λ_n such that $\lambda_n = (n^{-1}h_n^3)^{\frac{1}{2}}$. As D is bounded, there exists a constant A such that $d_n\lambda_n \leq A$. Now we put $J(t) = \arg \min_{j=1, \dots, d_n} |t - t_j|$. Thus we have the following decomposition

$$\begin{aligned}
\sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| &\leq \max_{j=1, \dots, d_n} \sup_{t \in D} \left| \tilde{f}_n(t) - \tilde{f}_n(t_{J(t)}) \right| \\
&\quad + \max_{j=1, \dots, d_n} \left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| \\
&\quad + \max_{j=1, \dots, d_n} \sup_{t \in D} \left| E \left[\tilde{f}_n(t_{J(t)}) \right] - E \left[\tilde{f}_n(t) \right] \right| \\
&= I_{1n} + I_{2n} + I_{3n}
\end{aligned} \tag{5.14}$$

Firstly, concerning I_{1n} and I_{3n} , with assumption **A2**, we have

$$I_{1n} \leq \frac{cM}{L(a)\overline{G}(b)} \lambda_n h_n^{-2}$$

The condition upon λ_n , we get

$$I_{1n} = O\left(\frac{1}{\sqrt{nh_n}}\right) \quad a.s. \text{ as } n \rightarrow \infty \tag{5.15}$$

As to I_{2n} , for all $\epsilon > 0$, we have

$$P \left[\max_{j=1, \dots, d_n} \left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| > \epsilon \right] \leq d_n P \left[\left| \tilde{f}_n(t_{J(t)}) - E \left[\tilde{f}_n(t_{J(t)}) \right] \right| > \epsilon \right] \tag{5.16}$$

Now, for any $i \geq 1$, we put

$$\Delta_i(t) = \frac{\mu}{nh_n} \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K\left(\frac{t-Z_i}{h_n}\right) - E \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K\left(\frac{t-Z_i}{h_n}\right) \right] \right] \tag{5.17}$$

Under **A2**, the rv's $U_i(t_{J(t)}) := nh_n \Delta_i(t_{J(t)})$ are centered and bounded by $\frac{2\mu}{L(a)\overline{G}(b)} M < \infty$. The use of the well-known Fuk-Nagaev's inequality (see Rio 2000, p. 87, 6.19b), allows us to get, for all $\epsilon > 0$ and $r > 1$

$$\begin{aligned}
& P \left[\max_{j=1, \dots, d_n} \left| \sum_{i=1}^n \Delta_i (t_{J(t)}) \right| > \epsilon \right] \\
&= P \left[\max_{j=1, \dots, d_n} \left| \sum_{i=1}^n U_i (t_{J(t)}) \right| > \epsilon n h_n \right] \\
&\leq A \lambda_n^{-1} \left[\frac{n}{r} \left(\frac{2r}{n h_n \epsilon} \right)^{\nu+1} + \left(1 + \frac{n^2 h_n^2 \epsilon^2}{r S_n} \right)^{-r/2} \right] \\
&= Q_{1n} + Q_{2n}
\end{aligned}$$

where

$$S_n = \sum_{i=1}^n \sum_{j=1}^n |\text{cov}(U_i, U_j)|. \quad (5.18)$$

Now, we have to study the asymptotic behavior of (5.18). First, we calculate

$$S_n = n \text{Var}(U_1) + \sum_{i \neq j} |\text{cov}(U_i, U_j)| = n \text{Var}(U_1) + S_n^* \quad (5.19)$$

By using **A2**, the conditional expectation properties and a change of variables, we get

$$\begin{aligned}
n \text{Var}(U_1) &\leq n E \left[\frac{\mu^2 \delta_1}{L^2(Z_1) \overline{G}^2(Z_1)} K^2 \left(\frac{t_{J(t)} - Z_1}{h_n} \right) \right] \\
&\leq \frac{n h_n \mu}{L(a) \overline{G}(b)} \int_{\mathbb{R}} K^2(s) f(t_{J(t)} - s h_n) ds \\
&= O(n h_n). \quad (5.20)
\end{aligned}$$

Then, from (5.18), using again the conditional expectation, we get under **A3**

$$\begin{aligned}
\text{cov}(U_i, U_j) &= E(U_i U_j) \\
&= \int \int K\left(\frac{t_{J(t)} - r}{h_n}\right) K\left(\frac{t_{J(t)} - s}{h_n}\right) \\
&\quad \times f_{1j}(r, s) dr ds \\
&\quad - \int K\left(\frac{t_{J(t)} - r}{h_n}\right) f(r) dr \\
&\quad \times \int K\left(\frac{t_{J(t)} - s}{h_n}\right) f(s) ds
\end{aligned}$$

Then

$$\begin{aligned}
|\text{cov}(U_i, U_j)| &\leq h_n^2 \int \int K(m_1) K(m_2) \\
&\quad \times |f_{1,j}(t_{J(t)} - h_n m_1, t_{J(t)} - h_n m_2) \\
&\quad - f(t_{J(t)} - h_n m_1) f(t_{J(t)} - h_n m_2)| dm_1 dm_2
\end{aligned}$$

Using assumption **A5**, it follows that

$$|\text{cov}(U_i, U_j)| = O(h_n^2). \quad (5.21)$$

Now, following Masry (1986), we define the sets

$$E_1 = \{(i, j) \text{ such that } 1 \leq |i - j| \leq \varphi_n\} \text{ and } E_2 = \{(i, j) \text{ such that } \varphi_n + 1 \leq |i - j| \leq n - 1\}$$

where $\varphi_n \rightarrow \infty$ as $n \rightarrow \infty$, we can write

$$S_n^* = \mathcal{F}_{1,n} + \mathcal{F}_{2,n}$$

where $\mathcal{F}_{1,n}$ and $\mathcal{F}_{2,n}$ be the sums of covariances over E_1 and E_2 respectively. First, by

applying the upper bound (5.21) we get

$$\mathcal{F}_{1,n} = O(nh_n^2\varphi_n)$$

For the second term, we use the modified Davydov covariance inequality for mixing processes (see Rio, 2000, formula 1.12a, page 10),

$$\forall i \neq j, \quad |\text{cov}(U_i, U_j)| \leq c\alpha(|i - j|) \quad (5.22)$$

Then, we get for assumption **A3**

$$\mathcal{F}_{2,n} = O(n\varphi_n^{1-\nu})$$

Choosing $\varphi_n = h_n^{-2/\nu}$ permits us to get

$$S_n^* = O\left(nh_n^{(\nu-1)\frac{2}{\nu}}\right) \quad (5.23)$$

Finally, under **A3**, combined to (5.20) and (5.23) yields

$$S_n = O(nh_n) \quad (5.24)$$

Therefore, by putting

$$\epsilon = \epsilon_0 \sqrt{\frac{\log n}{nh_n}} \quad ; \quad r = c \log n (\log \log n)^{1/\nu} \quad (5.25)$$

we get

$$Q_{1n} = O\left(\lambda_n^{-1} n^{(1-\nu)/2} h_n^{-(\nu+1)/2} (\log n)^{(\nu-1)/2} (\log \log n) \epsilon_0^{-(\nu+1)}\right)$$

Under **A1(ii)**, yields $Q_{1n} = O\left(n^{-1} \log^{-1} n (\log \log^{-2} n)\right)$, which is the general term of a convergent Bertrand's series. Using a Taylor expansion of $\log(x+1)$ and (5.25), we have

$$\begin{aligned}
Q_{2n} &= A_2 n^{1/2} h_n^{-3/2} \exp \left[-\frac{r}{2} \log \left(1 + \frac{nh_n \epsilon_0^2 \log n}{r S_n} \right) \right] \\
&\leq c h_n^{-3/2} n^{\frac{1-\epsilon_0^2}{2}}
\end{aligned} \tag{5.26}$$

In the same way, we can choose ϵ_0 such that Q_{2n} is the general term of convergent series. Thus, $\sum_{n \geq 1} (Q_{1n} + Q_{2n}) < \infty$, and the result is then a direct application of Borel-Cantelli's Lemma. \square

Proof of Proposition 5.1. Using the triangle inequality, we have

$$\sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| \leq \sup_{t \in D} \left| \hat{f}_n(t) - \tilde{f}_n(t) \right| + \sup_{t \in D} \left| \tilde{f}_n(t) - E \left[\tilde{f}_n(t) \right] \right| + \sup_{t \in D} \left| E \left[\tilde{f}_n(t) \right] - f(t) \right|$$

Then Lemmas **5.1**, **5.2** and **5.3** give the result. \square

Proof of Theorem 5.1. Standard argument gives us

$$\begin{aligned}
\left| f(\hat{\theta}_n) - f(\theta) \right| &\leq \left| f(\hat{\theta}_n) - \hat{f}_n(\hat{\theta}_n) \right| + \left| \hat{f}_n(\hat{\theta}_n) - f(\theta) \right| \\
&\leq \sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right| + \left| \hat{f}_n(\hat{\theta}_n) - f(\theta) \right| \\
&\leq 2 \sup_{t \in D} \left| \hat{f}_n(t) - f(t) \right|
\end{aligned} \tag{5.27}$$

The *a.s.* uniform consistency of $\hat{\theta}_n$ follows then immediately from **Proposition 5.1** and **A6**. Now a Taylor expansion gives

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

where $\hat{\theta}_n^*$ is between $\hat{\theta}_n$ and θ . Then by (5.27) and **A4**, we have

$$|\hat{\theta}_n - \theta| \leq 2 \sqrt{\frac{\sup_{t \in D} |\hat{f}_n(t) - f(t)|}{|f^{(2)}(\hat{\theta}_n^*)|}}$$

The result is then a direct consequence of **Proposition 5.1**. \square

Proof of Theorem 5.2. Form (5.11) we have the following decomposition

$$\begin{aligned} \sqrt{nh_n^3} (\hat{\theta}_n - \theta) &= \sqrt{nh_n^3} \frac{\hat{f}_n^{(1)}(\theta) - \tilde{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\ &\quad + \sqrt{nh_n^3} \frac{\tilde{f}_n^{(1)}(\theta) - E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\ &\quad + \sqrt{nh_n^3} \frac{E[\tilde{f}_n^{(1)}(\theta)]}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \\ &= \frac{J_1 + J_2 + J_3}{\hat{f}_n^{(2)}(\bar{\theta}_n)} \end{aligned} \tag{5.28}$$

To prove the results, we establish that J_1 and J_3 are negligible and J_2 is asymptotically normal and $\hat{f}_n^{(2)}(\bar{\theta}_n) \rightarrow f^{(2)}(\theta)$. \square

Lemma 5.4. Under assumptions **A1(i)**, **A2-A4** and **B1**, we have

$$J_1 \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Proof. We have

$$\begin{aligned} |J_1| &= \sqrt{nh_n^3} \left| \frac{1}{nh_n^2} \sum_{i=1}^n \delta_i K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \left(\frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right) \right| \\ &\leq \sqrt{nh_n} \left[\underbrace{\frac{1}{C_n(a)} \sup_{t \in D} |\hat{F}_n(t) - F(t)|}_I + \frac{1 - F(a)}{C_n(a)C(a)} \underbrace{\sup_{t \in D} |C_n(t) - C(t)|}_J \right] \frac{1}{nh_n} \sum_{i=1}^n \delta_i \left| K^{(1)}\left(\frac{\theta - Z_i}{h_n}\right) \right| \end{aligned}$$

Then for any $\epsilon > 0$, under **A2**, we have

$$P \left[\frac{1}{nh_n} \sum_{i=1}^n \left| K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) \right| > \epsilon \right] \leq \frac{1}{nh_n \epsilon} \sum_{i=1}^n E \left(\left| K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) \right| \right) \rightarrow \frac{f(\theta)}{\mu \epsilon} \int_{\mathbb{R}} |K^{(1)}(r)| dr$$

Since $C(a) > 0$, $I = O\left(\sqrt{\frac{\log \log n}{n}}\right)$ (see Lemma 2.1 in Chen and Dai, 2003) and $J = O\left(\sqrt{\frac{\log \log n}{n}}\right)$ (see Theorem 3.2 in Cai and Roussas, 1992), we have $|J_1| = O\left(\sqrt{h_n \log \log n}\right)$, which goes to zero under **A1**(i) and **B1**. \square

Lemma 5.5. Under assumptions **A2**, **A4** and **B1**, we have

$$J_3 \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Proof. Integrating by part and a change of variables, we have

$$\begin{aligned} E \left[\tilde{f}_n^{(1)}(\theta) \right] &= \frac{1}{h_n^2} E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K^{(1)} \left(\frac{\theta - Z_1}{h_n} \right) \right] \\ &= \frac{1}{h_n^2} \int_{\mathbb{R}} K^{(1)} \left(\frac{\theta - v}{h_n} \right) f(v) dv \\ &= \int_{\mathbb{R}} K(r) f^{(1)}(\theta - rh_n) dr \end{aligned}$$

By a Taylor expansion of $f^{(1)}(\cdot)$ around θ , **A2** and the definition of the mode, we get

$$\sqrt{nh_n^3} E \left[\tilde{f}_n^{(1)}(\theta) \right] = \sqrt{nh_n^7} \int_{\mathbb{R}} r^2 K(r) f^{(3)}(\theta^*) dr$$

where θ^* is between θ and $\theta - rh_n$, by **A2**, **A4** and **B1**, we conclude that $J_3 \rightarrow 0$ a.s. as $n \rightarrow \infty$. \square

Lemma 5.6. Under assumptions **A2-A5** and **B2** we have

$$J_2 \xrightarrow{D} N(0, \sigma^2)$$

Proof. In order to establish the asymptotic normality for sums of dependent rv's, we use Doob's small-block and large-block technique (see Doob 1953, pp. 228-232). Partition the set $\{1, 2, \dots, n\}$ into $2r_n + 1$ subsets with large blocks of size M_n and small blocks of size N_n , where (M_n) , (N_n) and (r_n) are sequences of integer numbers described in assumption **B2**. Put

$$\Gamma_i = \frac{\mu}{\sqrt{h_n}} \left\{ \frac{\delta_i}{L(Z_i) \overline{G}(Z_i)} K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) - E \left[\frac{\delta_i}{L(Z_i) \overline{G}(Z_i)} K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) \right] \right\}$$

Then

$$J_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Gamma_i = \frac{1}{\sqrt{n}} [T_{1,n} + T_{2,n} + T_{3,n}]$$

where

$$T_{1,n} = \sum_{m=1}^{r_n} L_m, \quad T_{2,n} = \sum_{m=1}^{r_n} L'_m \quad \text{and} \quad T_{3,n} = \sum_{i=(M_n+N_n)r_n+1}^n \Gamma_i$$

Also set

$$L_m = \sum_{i=(m-1)(M_n+N_n)+1}^{(m-1)(M_n+N_n)+M_n} \Gamma_i \quad \text{and} \quad L'_m = \sum_{i=(m-1)(M_n+N_n)+M_n+1}^{m(M_n+N_n)} \Gamma_i$$

We first show that $\frac{1}{n} (E [T_{2,n}^2] + E [T_{3,n}^2]) \rightarrow 0$, as $n \rightarrow \infty$. Next, we state that $\frac{1}{\sqrt{n}} (T_{1,n})$ converges in distribution to a normal variable with 0 mean and a variance given explicitly.

Lemma 5.7. Under the assumptions of **Lemma 5.6**, we have

$$\frac{1}{n} (E [T_{2,n}^2] + E [T_{3,n}^2]) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

Proof. Obviously

$$\begin{aligned}
\frac{1}{n} E [T_{2,n}^2] &= \frac{1}{n} \sum_{m=1}^{r_n} \sum_{i=(m-1)(M_n+N_n)+M_n+1}^{m(M_n+N_n)} E [\Gamma_i^2] \\
&\quad + \frac{2}{n} \sum_{m=1}^{r_n} \sum_{(m-1)(M_n+N_n)+M_n+1 \leq i < j \leq m(M_n+N_n)} cov (\Gamma_i, \Gamma_j) \\
&\quad + \frac{2}{n} \sum_{1 \leq i < j \leq r_n} cov (L'_i, L'_j) \\
&= \Lambda_{1n} + \Lambda_{2n} + \Lambda_{3n}
\end{aligned}$$

First, we have

$$\begin{aligned}
E [\Gamma_i^2] &= \frac{1}{h_n} E \left[\frac{\mu^2 \delta_i}{L^2 (Z_i) \overline{G}^2 (Z_i)} (K^{(1)})^2 \left(\frac{\theta - Z_i}{h_n} \right) \right] \\
&\quad - \frac{1}{h_n} \left\{ E \left[\frac{\mu \delta_i}{L (Z_i) \overline{G} (Z_i)} K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) \right] \right\}^2 \\
&= \int_{\mathbb{R}} \frac{\mu}{L (\theta - sh_n) \overline{G} (\theta - sh_n)} (K^{(1)})^2 (s) f (\theta - sh_n) ds \\
&\quad - h_n \left(\int_{\mathbb{R}} (K^{(1)}) (s) f (\theta - sh_n) ds \right)^2
\end{aligned}$$

Finally, we obtain

$$E [\Gamma_i^2] \longrightarrow \frac{\mu f (\theta)}{\overline{G} (\theta) L (\theta)} \int_{\mathbb{R}^2} [K^{(1)} (s)]^2 ds \text{ as } n \rightarrow \infty \quad (5.29)$$

which yields that $\Lambda_{1n} = O \left(\frac{r_n N_n}{n} \right) = O (1)$ from **B2**(i). Since

$$|\Lambda_{kn}| \leq \frac{2}{n} \sum_{1 \leq i < j \leq n} |cov (\Gamma_i, \Gamma_j)|, \quad k = 2, 3 \quad (5.30)$$

we prove that the right-hand side term in (5.30) tends to 0 as n tends infinity. We use a

technique developed by Masry (1986). We define the sets

$$\mathfrak{R}_1 = \{(i, j) \text{ such that } 1 \leq |i - j| \leq \rho_n\} \text{ and } \mathfrak{R}_2 = \{(i, j) \text{ such that } \rho_n + 1 \leq |i - j| \leq n - 1\}$$

where ρ_n is a sequence of integers such that $\rho_n \rightarrow \infty$, $\rho_n h_n \rightarrow 0$ as $n \rightarrow \infty$ and for some $\delta \in (0, 1)$, we have

$$h_n^{-\delta} \sum_{l=\rho_n}^{\infty} (\alpha(l))^\delta \rightarrow 0 \text{ as } n \rightarrow \infty \quad (5.31)$$

We can write

$$\frac{1}{n} \sum_{1 \leq i < j \leq n} |\text{cov}(\Gamma_i, \Gamma_j)| = \frac{1}{n} \mathfrak{S}_{1,n} + \frac{1}{n} \mathfrak{S}_{2,n} \quad (5.32)$$

where $\mathfrak{S}_{1,n}$ and $\mathfrak{S}_{2,n}$ be the sums of covariances over \mathfrak{R}_1 and \mathfrak{R}_2 respectively. Furthermore, for $i < j$, we observe that

$$\begin{aligned} \text{cov}(\Gamma_i, \Gamma_j) &= E(\Gamma_i \Gamma_j) \\ &= \frac{1}{h_n} \int \int K^{(1)}\left(\frac{\theta - r}{h_n}\right) K^{(1)}\left(\frac{\theta - s}{h_n}\right) \\ &\quad f_{1,j}(r, s) dr ds \\ &\quad - \frac{1}{h_n} \int K^{(1)}\left(\frac{\theta - r}{h_n}\right) f(r) dr \\ &\quad \times \int K^{(1)}\left(\frac{\theta - s}{h_n}\right) f(s) ds \end{aligned}$$

Then

$$\begin{aligned}
|cov(\Gamma_i, \Gamma_j)| &\leq h_n \int \int K^{(1)}(m_1) K^{(1)}(m_2) \\
&\quad \times |f_{1,j}(\theta - h_n m_1, \theta - h_n m_2) \\
&\quad - f(\theta - h_n m_1) f(\theta - h_n m_2)| dm_1 dm_2
\end{aligned}$$

Using assumption **A5**, it follows that

$$|cov(\Gamma_i, \Gamma_j)| = O(h_n).$$

Hence

$$\frac{1}{n} \mathfrak{S}_{1,n} = O(h_n \rho_n) \quad (5.33)$$

For the second term of (5.32), let use a version of the moment inequality due to Rio (1994, p. 588 – 590). Let p, q, γ be integer numbers such that $\frac{1}{p} + \frac{1}{q} + \frac{1}{\gamma} = 1$, it follows that

$$\frac{1}{n} \mathfrak{S}_{2,n} \leq \frac{c}{n} \sum_{(i,j) \in \mathfrak{R}_2} 2^{1+\frac{1}{\gamma}} (\alpha(|i-j|))^{\frac{1}{\gamma}} (E|\Gamma_i|^p)^{\frac{1}{p}} (E|\Gamma_i|^q)^{\frac{1}{q}}$$

Moreover, under **A1**, we get for n large enough.

$$\begin{aligned}
E|\Gamma_i|^p &\leq \frac{2^{p-1}}{h_n^{P/2}} E \left[\left| \frac{\mu \delta_i}{L(Z_i) \overline{G}(Z_i)} K^{(1)} \left(\frac{\theta - Z_i}{h_n} \right) \right|^p \right] \\
&\leq \frac{2^{p-1} \mu^{p-1} h_n}{h_n^{P/2} L^{p-1}(a) \overline{G}^{p-1}(b)} \int_{\mathbb{R}} |K^{(1)}(s)|^p f(\theta - sh_n) ds \\
&= O(h_n^{1-p/2}) \quad (5.34)
\end{aligned}$$

Hence, using (5.34), it follows that

$$\begin{aligned}
\frac{1}{n} \mathfrak{S}_{2,n} &\leq \frac{c}{n} \sum_{(i,j) \in \mathfrak{R}_2} 2^{1+\frac{1}{\gamma}} (\alpha(|i-j|))^{\frac{1}{\gamma}} (h_n^{1-p/2})^{\frac{1}{p}} (h_n^{1-q/2})^{\frac{1}{q}} \\
&\leq c (h_n^{1-p/2})^{\frac{1}{p}} (h_n^{1-q/2})^{\frac{1}{q}} \sum_{l=\rho_n}^{\infty} (\alpha(l))^{\frac{1}{\gamma}} \\
&\leq ch_n^{-\frac{1}{\gamma}} \sum_{l=\rho_n}^{\infty} (\alpha(l))^{\frac{1}{\gamma}}
\end{aligned} \tag{5.35}$$

Therefore, by (5.31), we get

$$\frac{1}{n} \mathfrak{S}_{2,n} \rightarrow 0 \text{ as } n \rightarrow \infty \tag{5.36}$$

Thus, from (5.33) and (5.36), we obtain

$$\frac{1}{n} \sum_{1 \leq i < j \leq n} |\text{cov}(\Gamma_i, \Gamma_j)| \rightarrow 0 \text{ as } n \rightarrow \infty \tag{5.37}$$

As to $n^{-1}E[T_{3,n}^2]$, from (5.29), we have

$$\begin{aligned}
n^{-1}E[T_{3,n}^2] &= n^{-1}E \left[\left(\sum_{i=(M_n+N_n)r_n+1}^n \Gamma_i \right)^2 \right] \\
&\leq n^{-1} \sum_{i=(M_n+N_n)r_n+1}^n E[\Gamma_i^2] + 2n^{-1} \sum_{1 \leq i < j \leq n} |\text{cov}(\Gamma_i, \Gamma_j)|
\end{aligned}$$

Therefore, (5.37) and assumption **B2** give the result. \square

Lemma 5.8. Under assumptions **A2-A5** and **B2**, we have

$$\frac{1}{n} \text{Var}(T_{1,n}) \longrightarrow \frac{\mu f(\theta)}{\bar{G}(\theta) L(\theta)} \int_{\mathbb{R}} [K^{(1)}(r)]^2 dr, \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned}
\frac{1}{n}Var(T_{1,n}) &= \frac{1}{n} \left(\sum_{m=1}^{r_n} Var(L_m) \right) + \frac{2}{n} \sum_{1 \leq i < j \leq r_n} cov(L_i, L_j) \\
&\leq \frac{1}{n} \sum_{m=1}^{r_n} \sum_{i=(m-1)(M_n+N_n)+1}^{(m-1)(M_n+N_n)+M_n} E[\Gamma_i^2] + \frac{4}{n} \sum_{1 \leq i < j \leq n} cov(\Gamma_i, \Gamma_j)
\end{aligned}$$

Since $\frac{r_n M_n}{n} \rightarrow 1$ and from (5.29) and (5.37) the result can be achieved. \square

Lemma 5.9. Under the assumptions of **Lemma 5.6**, we have

$$\frac{1}{\sqrt{n}} T_{1,n} \xrightarrow{D} N(0, \sigma^2)$$

Proof. The proof is based on two steps. Firstly, it will be shown that the characteristic functions of $\frac{1}{\sqrt{n}} T_{1,n}$ minus the product of the characteristic functions of L_m , $m = 1, \dots, r_n$ converges to 0. Secondly, it is proved that the distribution determined by the product of the characteristic function of L_m , $m = 1, \dots, r_n$ is asymptotically, the distribution $N(0, \sigma^2)$.

For the first step, according to the Lemma of Volonskii and Rosanov (see Appendix in Liang et al. 2010). Note that L_m is $F_{i_m}^{j_m}$ measurable with $i_m = (m-1)(M_n + N_n) + 1$ and $j_m = (m-1)(M_n + N_n) + M_n$. Then we have

$$\left| E \left[\prod_{m=1}^{r_n} \exp \left(itn^{-\frac{1}{2}} L_m \right) \right] - \prod_{m=1}^{r_n} E \left[\exp \left(itn^{-\frac{1}{2}} L_m \right) \right] \right| \leq 16r_n \alpha (N_n + 1)$$

which tends to zero by **B2(iii)**.

Now we have to show that the standard Lindeberg-Feller conditions (see Loève, 1963 p. 280) for asymptotic normality of $\frac{1}{\sqrt{n}} T_{1,n}$ under independence are satisfied. So, we have to establish

$$\frac{1}{n} \sum_{m=1}^{r_n} E \left[L_m^2 I(|L_m| > \epsilon \sigma \sqrt{n}) \right] \rightarrow 0$$

It suffices to show that, for every $\epsilon > 0$, the set $\{|L_m| > \epsilon \sigma \sqrt{n}\}$ is empty. Since

$$\frac{1}{\sqrt{n}} |L_m| \leq \frac{M_n c}{\sqrt{nh_n}}$$

which leads to the fact that the set $\{|L_m| > \epsilon \sigma \sqrt{n}\}$ is an empty set under assumption **B2(ii)**. Then this completes the proof of **Lemma 5.9**. \square

Now, to prove that $\hat{f}_n^{(2)}(\bar{\theta}_n) \rightarrow f^{(2)}(\theta)$, we have the following :

Lemma 5.10. Under assumptions **A1(i)** and **A2-A5**, we have

$$\sup_{t \in D} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Proof. Since

$$\begin{aligned} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| &\leq \left| \hat{f}_n^{(2)}(t) - \tilde{f}_n^{(2)}(t) \right| + \left| E \left[\tilde{f}_n^{(2)}(t) \right] - f^{(2)}(t) \right| + \left| \tilde{f}_n^{(2)}(t) - E \left[\tilde{f}_n^{(2)}(t) \right] \right| \\ &= \gamma_{1,n}(t) + \gamma_{2,n}(t) + \gamma_{3,n}(t) \end{aligned}$$

such that

$$\sup_{t \in D} \gamma_{1,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty \tag{5.38}$$

$$\sup_{t \in D} \gamma_{2,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty \tag{5.39}$$

$$\sup_{t \in D} \gamma_{3,n}(t) \rightarrow 0 \text{ as } n \rightarrow \infty \tag{5.40}$$

Using the same arguments as in the proof of **Lemma 5.2**, we get **(5.38)**.

Since integrating by parts twice and using a change of variable, as followed in **Lemma 5.1**, we obtain

$$\begin{aligned} \sup_{t \in D} \gamma_{2,n}(t) &= \sup_{t \in D} \left| \int_{\mathbb{R}} K(r) f^{(2)}(t - rh_n) dr - f^{(2)}(t) \right| \\ &= O(h_n^2) \end{aligned}$$

we get **(5.39)**. To prove **(5.40)**, we use the same framework as in the proof of **Lemma 5.3**, and a suitable covering of the compact set D and changing the variables $\Delta_i(t)$ into the following

$$\begin{aligned} \Pi_i(t) &= \frac{\mu}{nh_n^3} \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K^{(2)}\left(\frac{t-Z_i}{h_n}\right) - E \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K^{(2)}\left(\frac{t-Z_i}{h_n}\right) \right] \right] \\ &= (nh_n^3)^{-1} V_i(t) \end{aligned}$$

and if the bandwidth h_n satisfies :

$$\lim \frac{nh_n^{3(\nu+6)/(\nu-6)}}{(\log n)^{\frac{\nu+1}{\nu-6}} (\log \log n)^{\frac{6}{\nu-6}}} \rightarrow \infty$$

we get the result. \square

We observe that we have, for n large enough

$$\left| \hat{f}_n^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta) \right| \leq \sup_{t \in D} \left| \hat{f}_n^{(2)}(t) - f^{(2)}(t) \right| + |f^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta)|$$

By **Lemma 5.10** and assumption **A3**, the proof of **Theorem 5.2** is completed. \square

Chapitre 6

La fonction mode conditionnelle

6.1 Introduction

Considérons une suite (X_n, Y_n) stationnaire de vecteurs aléatoires définis dans le même espace probabilisé (Ω, \mathcal{F}, P) à valeur dans $\mathbb{R}^d \times \mathbb{R}$. Soit $F(\cdot, \cdot)$ la f.d.r. continue et f la densité de probabilité du couple (X_1, Y_1) . Pour tout $x \in \mathbb{R}^d$, la fonction de densité conditionnelle de Y_1 sachant $X_1 = x$ noté par $f(\cdot | x)$. Le mode conditionnel de Y_1 sachant $X_1 = x$, noté par $\theta(x)$, est défini par

$$f(\theta(x) | x) = \max_{t \in \mathbb{R}} f(t | x)$$

Nous nous intéressons ici au problème de l'estimation du mode conditionnel θ qui est un outil de prévision alternatif à la méthode de régression (voir Collomb et al. 1987). Un estimateur du mode conditionnel $\theta(x)$ est défini comme la v.a. $\theta_n(x)$ maximisant l'estimateur de la densité conditionnelle $\hat{f}_n(\cdot | x)$ de $f(\cdot | x)$, c-à-d

$$\hat{f}_n(\theta_n(x) | x) = \max_{t \in \mathbb{R}} \hat{f}_n(t | x) \tag{6.1}$$

La littérature sur l'estimation du mode conditionnel est très abondante lorsque la variable explicative X est à valeurs dans un espace de dimension fini et de densité $l(\cdot)$.

Dans le cas i.i.d., Samanta et Thavaneswaran (1990) établissant la convergence presque sûre et la normalité asymptotique de l'estimateur à noyau. Dans le cas α -mélangeantes, Louani et Ould-Saïd (1999) ont établi la normalité asymptotique dans le cas des données fortement mélangeantes.

6.2 Cas de données complètes (α -mélange)

L'estimateur à noyau du mode conditionnel est l'estimateur introduit par Louani et Ould-Saïd (1999) et défini pour tout $x \in \mathbb{R}^d$:

$$\hat{f}_n(\theta_n(x) | x) = \max_{t \in \mathbb{R}} \hat{f}_n(t | x)$$

où

$$\hat{f}_n(t | x) = \frac{\frac{1}{nh_n^{d+1}} \sum_{i=1}^n K_d\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Y_i}{h_n}\right)}{\frac{1}{nh_n^d} \sum_{i=1}^n K_d\left(\frac{x-X_i}{h_n}\right)} \quad (6.2)$$

K_d et H sont des densités de probabilités sur \mathbb{R}^d et \mathbb{R} respectivement, (h_n) est une suite de réels positifs tendant vers zéro quand $n \rightarrow \infty$. Le processus (X_n, Y_n) satisfait certaines conditions de dépendance. Le théorème suivant donne un résultat de normalité de l'estimateur du mode conditionnel dans le cas α -mélange.

Théorème 6.1 (Louani et Ould-Saïd 1999) : Sous les hypothèses, en particulier sur les noyaux K_d et H , la fenêtre h_n , les densités de probabilités $f(.,.)$ et $l(.)$ et sur la structure de dépendance. On a

$$\sqrt{\frac{nh_n^{d+3}}{V(x, \theta(x))}} f^{(0,2)}(x, \theta(x)) (\theta_n(x) - \theta(x)) \xrightarrow{l} N(0, 1)$$

où \xrightarrow{l} désigne la convergence en loi,

$$V(x, \theta(x)) = f(x, \theta(x)) \int_{\mathbb{R}^{d+1}} [K_d(r) H^{(1)}(s)]^2 dr ds$$

6.3 Cas de données censurées

Soit Y une variable aléatoire positive de fonction de répartition F et de densité f , représentant une durée de vie d'un individu censurée à droite par une v.a. positive C de fonction de répartition G . Lors d'une étude expérimentale, on relève de plus d'autres données relatives à l'état de l'individu, comme par exemple la pression artérielle, le taux de cholestérol, etc....., pour une étude médicale, dans un tel cas, on regroupe toutes les mesures annexes dans un vecteur X de \mathbb{R} représentant une covariable de densité $l(\cdot)$. Notons par $f(\cdot, \cdot)$ la densité de probabilité du couple (X, Y) et $f(\cdot | x)$ la densité conditionnelle de Y sachant $X = x$. Sous l'hypothèse que Y et C sont conditionnellement indépendants sachant X et on observe les v.a. $Z_i = Y_i \wedge C_i$ et $\delta_i = I\{Y_i \leq C_i\}$ et X_i pour toute $i \geq 1$ tels que les (Z_i, δ_i, X_i) sont i.i.d.

Un estimateur à noyau du mode conditionnel $\theta(x)$ est défini comme la v.a. $\theta_n(x)$ maximisant l'estimateur à noyau de la densité conditionnelle $\hat{f}_n(\cdot | x)$ de $f(\cdot | x)$, c-à-d

$$\theta_n(x) = \arg \sup_{t \in \mathbb{R}} \hat{f}_n(t | x)$$

avec

$$\begin{aligned} \hat{f}_n(t | x) &= \frac{\frac{1}{nh_n^2} \sum_{i=1}^n \frac{\delta_i}{G_n(Z_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Z_i}{h_n}\right)}{\frac{1}{nh_n^2} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)} \\ &= \frac{\hat{f}_n(x, t)}{l_n(x)} \end{aligned} \quad (6.3)$$

où $G_n(\cdot)$ est l'estimateur de Kaplan-Meier de $G(\cdot)$, $l_n(\cdot)$ estimateur à noyau de $l(\cdot)$, K et

H sont des densités de probabilités sur \mathbb{R} et (h_n) est une suite de réels positifs tendant vers zéro quand $n \rightarrow \infty$. Nous donnons dans les théorèmes suivants les propriétés asymptotiques de l'estimateur à noyau du mode conditionnel, lorsque la variable d'intérêt est assujettie à une censure à droite, établie par Khardani et al. (2010) dans le cas i.i.d.

Théorème 6.2 (Khardani et al. 2010) : Sous des hypothèses de type VC-classes sur le noyau et des conditions techniques standard sur la fenêtre, on a

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{2}}, h_n^2 \right) \right\} \text{ p.s. lorsque } n \rightarrow \infty.$$

$$\sup_{x \in \Omega} |\theta_n(x) - \theta(x)| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{4}}, h_n \right) \right\} \text{ p.s. lorsque } n \rightarrow \infty.$$

Théorème 6.3 (Khardani et al. 2010) : Soit $x \in \mathbb{R}$, sous certaines hypothèses. On a

$$\sqrt{nh_n^4} (\theta_n(x) - \theta(x)) \xrightarrow{l} N(0, \sigma^2(x))$$

où \xrightarrow{l} désigne la convergence en loi,

$$\sigma^2(x) = \frac{f(x, \theta(x))}{\overline{G}(\theta(x)) [f^{(0,2)}(x, \theta(x))]^2} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds$$

6.4 Cas des données tronquées

Soit $\{Y_i, 1 \leq i \leq N\}$ une suite de variables aléatoires d'intérêt, i.i.d., de fonction de répartition inconnu F . Soit $\{X_i, 1 \leq i \leq N\}$ un échantillon correspondant de vecteurs aléatoires de covariables prenant leurs valeurs dans \mathbb{R} de fonction de répartition $V(\cdot)$ et une densité $v(\cdot)$. Les variables aléatoires Y_i sont considérées comme les durées de vie des objets d'étude et sont supposées être soumises à une troncature à gauche. Nous notons T_1, \dots, T_N la suite de v.a. de troncature, i.i.d., de f.d.r. continue L . La variable T sera supposée indépendante du vecteur (X, Y) . Soit $(X_1, Y_1, T_1), \dots, (X_n, Y_n, T_n)$ ($n \leq N$) l'échantillon observé (i.e. $Y_i \geq T_i$), la vraie taille n de l'échantillon observé est une variable aléatoire

distribuée selon la loi Binomial de paramètres N et μ où $\mu = P[Y_1 \geq T_1]$.

Le mode conditionnel $\theta(x)$ de Y sachant $X = x$ est donné par

$$\theta(x) = \arg \sup_{t \in \mathbb{R}} f(t | x)$$

où $f(t | x) = \frac{f(x,t)}{v(x)}$ est la densité conditionnelle de Y sachant $X = x$, et $f(.,.)$ est la densité de couple (X, Y) . Un estimateur à noyau du mode conditionnel $\theta(x)$ est défini comme la v.a. $\theta_n(x)$ solution de l'équation

$$\theta_n(x) = \arg \sup_{t \in \mathbb{R}} \hat{f}_n(t | x)$$

où $\hat{f}_n(., | x)$ est un estimateur de $f(., | x)$ défini par

$$\begin{aligned} \hat{f}_n(t | x) &= \frac{\frac{\mu_n}{nh_n^2} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Y_i}{h_n}\right)}{\frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{x-X_i}{h_n}\right)} \\ &= \frac{\hat{f}_n(x, t)}{v_n(x)} \end{aligned} \tag{6.4}$$

où $L_n(.)$ est l'estimateur produit-limite obtenu par Lynden-Bell (1971) de $L(.)$, K et H sont des densités de probabilités sur \mathbb{R} et (h_n) est une suite de réels positifs tendant vers zéro quand $n \rightarrow \infty$. Nous donnons dans le théorème suivant la convergence uniforme presque sûre de l'estimateur à noyau du mode conditionnel pour un modèle tronqué à gauche puis le théorème de la normalité asymptotique, ces résultats établis par Ould-Saïd et Tatachak (2007) dans le cas i.i.d.

Théorème 6.4 (Ould-Saïd et Tatachak 2007) : Sous des hypothèses de type VC-classes sur le noyau et des conditions techniques standards sur la densité conjointe et la

densité marginale, la fenêtre. On a

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^4} \right)^{\frac{1}{2}}, h_n \right) \right\} \quad P. p.s. \text{ lorsque } n \rightarrow \infty.$$

Corollaire 6.1 (Ould-Saïd et Tatachak 2007) : Sous les hypothèses de le théorème (6.4) avec la condition d'unicité de mode, on a

$$\sup_{x \in \Omega} |\theta_n(x) - \theta(x)| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^4} \right)^{\frac{1}{4}}, h_n^{\frac{1}{2}} \right) \right\} \quad P.p.s. \text{ lorsque } n \rightarrow \infty.$$

Théorème 6.5 (Ould-Saïd et Tatachak 2007) : Soit $x \in \mathbb{R}$, sous certaines hypothèses. On a

$$\sqrt{nh_n^4} (\theta_n(x) - \theta(x)) \xrightarrow{l} N(0, \sigma^2(x))$$

où \xrightarrow{l} désigne la convergence en loi,

$$\sigma^2(x) = \frac{\mu f(x, \theta(x))}{L(\theta(x)) [f^{(0,2)}(x, \theta(x))]^2} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds$$

Dans cette partie nous donnons maintenant des résultats soumis à parution avec Madame Sedki et Monsieur Boutabia. Nous présentons ici, dans la version soumise, les parties de cet article où sont donnés les résultats et leurs preuves.

6.5 Cas des données tronquées et censurées

6.5.1 Introduction

Let (Y, T, W) denote random variables, where Y is the interest variable called lifetime variable, with density function f and distribution function (d.f.) F . T is the random left truncation time with continuous d.f. L and W denote the random right censoring time

with arbitrary d.f. G . It is assumed that Y , T , W are independent and that they are non-negative. In the random LTRC model one can observe (Z, T, δ) if $Z \geq T$, where $Z = \min(Y, W)$ and $\delta = I(Y \leq W)$ is the indicator of censoring status, and when $Z < T$ nothing is observed. Clearly, if Y is independent of W , then Z has d.f. $H = 1 - (1 - F)(1 - G)$. Take $\mu = P[T \leq Z]$, it's clear if $\mu = 0$ no data can be observed therefore we suppose that $\mu > 0$. Let (Y_i, T_i, W_i) $i = 1, \dots, N$ be i.i.d. as (Y, T, W) where the population size N is fixed, but unknown. The empirical data are (Z_i, T_i, δ_i) $i = 1, \dots, n$ where n is the number of observed triplet. Let

$$C(y) = P[T \leq y \leq Z \mid T \leq Z] = \mu^{-1} L(y) (1 - F(y)) (1 - G(y)). \quad (6.5)$$

and its empirical estimator

$$C_n(y) = \frac{1}{n} \sum_{i=1}^n I(T_i \leq y \leq Z_i)$$

where $I(A)$ denotes the indicator function of the set A .

For LTRC data, it is important to be able to obtain nonparametric estimates of various characteristics of the distribution function F . Tsai et al. (1987) gave the nonparametric maximum likelihood estimator of F itself, called the product-limit (PL) estimator, as

$$1 - \hat{F}_n(t) = \prod_{\{i: Z_i \leq t\}} \left(1 - \frac{1}{nC_n(Z_i)} \right)^{\delta_i} \quad (6.6)$$

\hat{F}_n reduces to the Kaplan-Meier product limit (Kaplan-Meier, 1958) estimator when $T = 0$ and to the Lynden-Bell (1971) estimator when there is no right censoring. There is a vast literature to the study of the PL estimator. In particular, the asymptotic properties of the PLE under independent and identically distributed (i.i.d) assumptions, have been studied by many authors see Gijbels and Wang (1993), Zhou (1996) and references therein. Nonparametric estimators of the density and hazard rate for $F(\cdot)$ have been studied by Sun and Zhou (1998).

Let X be a real valued random variable of covariates related with Y . For any x denote by $f(t | x) = \frac{f(t, x)}{l(x)}$, the conditional probability density function of Y given $X = x$, where $f(t, x)$ is the joint probability density function of (Y, X) and $l(x)$ is the marginal density of X with respect of the Lebesgue measure. Then denote by $(X_i, Z_i, T_i, \delta_i)$ $i = 1, \dots, n$, a stationary random sample from (X, Z, T, δ) which one observes $(Z_i \geq T_i, \forall i)$. We assume that $f(\cdot | x)$ has a unique mode $\theta(x)$ which is given by

$$\theta(x) = \arg \sup_{t \in \mathbb{R}} f(t | x)$$

The problem of estimating the unconditional /conditional mode of the probability density has given way to a large amount of related statistics literature, and number of distinguished papers deal with this topic. In the complete data case (no censoring, no truncation), the mode estimator for i.i.d. random variables has been studied by many authors among whom we quote Parzen (1962), Romano (1988) and Vieu (1996). For the complete dependent data case, the strong consistency of the conditional mode estimator was established under a ϕ -mixing condition by Collomb et al. (1987) and their results can be applied to process forecasting. In the α -maxing case, the strong consistency over a compact set and the asymptotic normality were obtained by Ould-Saïd (1993) and Louani and Ould-Saïd (1999), respectively.

In the incomplete data case, for i.i.d. random variables under random right censoring, Louani (1998) studied the asymptotic normality of the kernel estimator of the mode, while Khardani et al. (2010) established, some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. These authors also studied uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series (Khardani et al, 2011). Under left-truncated model, in the i.i.d. case, the kernel mode estimate has been studied by Ould-Saïd and Tatachak (2009). They also studied the asymptotic properties of the kernel estimator of the conditional mode for the left-truncated model (Ould-Saïd and Tatachak, 2007). To the best of our knowledge, the problem of estimating the conditional mode function under truncated and censored

data has not been addressed in the statistics literature.

Our main results establish, the strong uniform convergence rate, over a compact set, of the kernel conditional density estimator which allows deduce the strong convergence rate of the kernel conditional mode estimator and establish its asymptotic normality. The rest of the paper is organized as follows. In section 2, we define a new kernel conditional mode estimator in LTRC model with some notations. Assumptions and main results are given in section 3. Finally, the proofs of the main results are relegated to section 4 with some auxiliary results and their proofs.

6.5.2 Definition of the new estimator

As a consequence of truncation, the size of the actually observed sample, n , is a $Bin(N, \mu)$ random variable. By the strong law of large numbers (SLLN), we have

$$\mu_n := \frac{n}{N} \rightarrow \mu. \quad a.s. \quad as \quad N \rightarrow +\infty \quad (6.7)$$

We denote by a superscript (*) any d.f. that is associated to the observed sample. For any distribution function K , let $a_K = \inf \{x : K(x) > 0\}$ and $b_K = \sup \{x : K(x) < 1\}$ denote the left and right end points of its support. As in the current model, as discussed by Gijbels and Wang (1993) and Zhou (1996), we assume that

$$a_L \leq a_H \quad , \quad b_L \leq b_H \quad \text{and} \quad \int_{a_H}^{\infty} \frac{dF(x)}{L^2(x)} < \infty \quad (6.8)$$

Now, let $H_1^*(\cdot)$ be defined by

$$H_1^*(z) = P[Z \leq z, \delta = 1 \mid T \leq Z] = \mu^{-1} \int_{-\infty}^z L(y) (1 - G(y^-)) dF(y) \quad (6.9)$$

$H_1^*(z)$ is consistently estimated by

$$H_{1n}^*(z) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq z, \delta_i = 1)$$

Indeed, according to (6.5) and replacing F , L and G by their respective nonparametric estimator, we can consider the estimator of μ , namely

$$\mu_n = \frac{L_n(y) \left(1 - \hat{F}_n(y^-)\right) (1 - G_n(y))}{C_n(y)}$$

for all y such that $C_n(y) > 0$, where $\hat{F}_n(y^-)$ denote the left limit of \hat{F}_n at y . Here L_n is the product-limit estimator Lynden-Bell for L

$$L_n(y) = \prod_{\{i: T_i > y\}} \left(1 - \frac{1}{nC_n(T_i)}\right) \quad (6.10)$$

and G_n is the product limit estimator Kaplan and Meier for G

$$G_n(y) = 1 - \prod_{\{i: Y_i \leq y\}} \left(1 - \frac{1 - \delta_i}{n - (i - 1)}\right) \quad (6.11)$$

We now define our nonparametric estimator of the conditional mode. For the complete data is present ($n = N$) it is well known that the kernel estimator of the conditional mode $\theta(x)$ is defined as the r.v. $\hat{\theta}_n(x)$ maximizing the kernel estimator $\hat{f}_n(t | x)$ of $f(t | x)$, that is

$$\hat{f}_n(\hat{\theta}_n(x) | x) = \max_{t \in \mathbb{R}} \hat{f}_n(t | x) \quad (6.12)$$

where

$$\hat{f}_n(t | x) = \frac{\frac{1}{nh_n^2} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) H\left(\frac{t - Y_i}{h_n}\right)}{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)}$$

K and H are a probability density functions, (h_n) is a sequence of positive real numbers

(so-called bandwidth) which goes to zero as n goes to infinity.

Under left-truncated model, Ould-Saïd and Tatachak (2007) define a new estimator $\hat{f}_n(\cdot | x)$ based on the n actually observed pair (Y_i, X_i) given by

$$\hat{f}_n(t | x) = \frac{\frac{\mu_n}{nh_n^2} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Y_i}{h_n}\right)}{\frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{x-X_i}{h_n}\right)}$$

In the right censoring model ; Khardani et al. (2010) define an estimator $\hat{f}_n(\cdot | x)$ based on the kernel smoothing technique, given by

$$\hat{f}_n(t | x) = \frac{\frac{1}{nh_n^2} \sum_{i=1}^n \frac{\delta_i}{\bar{G}_n(Y_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Y_i}{h_n}\right)}{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}$$

where $\bar{G} = 1 - G$.

Now, we define a new estimator $\hat{f}_n(\cdot | x)$ in *LTRC* model based on the kernel smoothing technique, given by

$$\begin{aligned} \hat{f}_n(t | x) &= \frac{\frac{\mu_n}{nh_n^2} \sum_{i=1}^n \frac{\delta_i}{\bar{G}_n(Z_i)L_n(Z_i)} K\left(\frac{x-X_i}{h_n}\right) H\left(\frac{t-Z_i}{h_n}\right)}{\frac{\mu_n}{nh_n} \sum_{i=1}^n \frac{1}{L_n(Z_i)} K\left(\frac{x-X_i}{h_n}\right)} \\ &= \frac{\hat{f}_n(x, t)}{l_n(x)} \end{aligned} \quad (6.13)$$

Remark that the estimator $\hat{\theta}_n(x)$ is not necessarily unique and our results are valid for any chosen value satisfying (6.12). We point out that we can specify our choice by taking

$$\hat{\theta}_n(x) = \inf \left\{ y \in \mathbb{R} : \hat{f}_n(y | x) = \max_{t \in \mathbb{R}} \hat{f}_n(t | x) \right\}. \quad (6.14)$$

We establish, the strong uniform convergence rate, over a compact set, of the estimator

(6.13) which allows deduce the strong convergence rate of the kernel conditional mode estimator (6.14) and establish its asymptotic normality.

In order to state the asymptotic normality we need some notations; for any function $\varphi(\cdot, \cdot)$ and $(i, j) \in \mathbb{N}^2$ denote $\varphi^{(i+j)}$ the $(i+j)$ th-order derivative of φ define by

$$\varphi^{(i,j)}(x, t) = \frac{\partial^{(i+j)}}{\partial x^i \partial t^j} \varphi(x, t)$$

and for $j \geq 1$

$$\hat{f}_n^{(0,j)}(x, t) = \frac{\partial^j \hat{f}_n(x, t)}{\partial t^j} = \frac{\mu_n}{nh_n^{2+j}} \sum_{i=1}^n \frac{\delta_i}{\bar{G}_n(Z_i) L_n(Z_i)} K\left(\frac{x - X_i}{h_n}\right) H^{(j)}\left(\frac{t - Z_i}{h_n}\right)$$

6.5.3 Assumptions and main results

In what follows, we suppose that $a_L \leq a_H$, $b_L \leq b_H$ and $D = [a, b]$ be a compact set such that $D \subset D_0 = \{t : t \in [a_H, b_H]\}$ and $\Omega_0 = \{x \in \mathbb{R} : l(x) > 0\}$, let $\Omega \subset \Omega_0$ be a compact set of \mathbb{R} and $\gamma = \inf_{x \in \Omega} l(x) > 0$. We introduce some assumptions below for easy reference needed to state our results.

A1 : The bandwidth (h_n) satisfies :

- (i) $\frac{nh_n^6}{\log n} \rightarrow \infty$ as $n \rightarrow \infty$;
- (ii) $nh_n^8 \rightarrow 0$ as $n \rightarrow \infty$.

A2 : The kernel K satisfies :

- (i) K is a positive valued, bounded, and there exists constant m such that

$$\sup_{t \in D} K(t) = m;$$

- (ii) It is Lipschitz continuous function ;

$$(iii) \int_{\mathbb{R}} tK(t) = 0; \int_{\mathbb{R}} t^2K(t) = 0 \text{ and } \int_{\mathbb{R}} t^3K(t) < \infty.$$

A3 : The density function H satisfies :

- (i) $H(\cdot)$ is three times differentiable with compact support on \mathbb{R} ;
- (ii) $H(\cdot)$ and $H^{(2)}(\cdot)$ satisfies the Lipschitz condition ;

(iii) $\int_{\mathbb{R}} tH(t) = 0$, $\int_{\mathbb{R}} t^2H(t) = 0$; $\lim_{|t| \rightarrow \infty} H(t) = 0$ and $\int_{\mathbb{R}} t^3H(t) < \infty$.

A4 : The marginal density $l(\cdot)$ satisfies the Lipschitz condition and $l(x) > \gamma_0$ for all $x \in \Omega$ for some $\gamma_0 > 0$.

A5 : $f(t | x)$ is twice differentiable, uniformly continuous in t and the second derivative $f^{(2)}(\cdot | x)$ is continuous.

A6 : The joint pdf $f(\cdot, \cdot)$ is bounded and differentiable up to order 3 and

$\sup_{x \in \Omega} \sup_{t \in D} |f^{(i,j)}(x, t)| < \infty$ for $i + j \leq 3$.

A7 : The function $\theta(\cdot)$ satisfies the following property : for any $\epsilon > 0$

and $\eta(x)$, $\exists \lambda > 0$ such that

$\sup_{x \in \Omega} |\theta(x) - \eta(x)| \geq \epsilon \Rightarrow \sup_{x \in \Omega} |f(\theta(x) | x) - f(\eta(x) | x)| \geq \lambda$

A8 : The function $f^{(0,2)}(\cdot, \cdot)$ is continuous and does not vanish.

Remark 6.1. (Discussion of the assumptions)

Assumption **A1**(i) is needed in the behavior study of the difference between $\hat{f}_n(\cdot, \cdot)$ and $\tilde{f}_n(\cdot, \cdot)$, while **A1**(ii) is used to prove the asymptotic normality result as well as the convergence of $\hat{f}_n^{(0,2)}(\cdot, \cdot)$. Assumptions **A2-A3** and **A6** are common in nonparametric estimation for independent case. Assumption **A4** intervenes in the convergence of kernel density estimator $l_n(\cdot)$. Assumption **A5** and **A7** stipulates the existence and the uniform uniqueness of the conditional mode function. Finally, assumption **A8** intervenes in the convergence of $\hat{f}_n^{(0,2)}(\cdot, \cdot)$.

Remark 6.2. Assumption **A3**(i) implies that the kernels H , $H^{(1)}$ and $H^{(2)}$ are bounded by constants M , M' and M'' respectively.

Consistency

Our first result deals with the uniform almost sure convergence with rate of the conditional density function estimator and is stated in **Proposition 6.1**. The uniform almost sure convergence of the conditional mode estimator and its well be derived in **Theorem 6.6**.

Proposition 6.1. Under assumptions **A1**(i) and **A2-A6**, we have

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{2}}, h_n^2 \right) \right\} \text{ a.s. as } n \rightarrow \infty$$

Theorem 6.6. Under the assumptions of **Proposition 6.1** and **A7**, we have

$$\sup_{x \in \Omega} \left| \hat{\theta}_n(x) - \theta(x) \right| = O \left\{ \max \left(\left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{4}}, h_n \right) \right\} \text{ a.s. as } n \rightarrow \infty$$

Asymptotic normality

We now consider the problem of asymptotic normality, suppose that the density function $f(\cdot | x)$ is unimodal at $\theta(x)$, then by assumptions **A5** and **A8**, we have

$$f^{(1)}(\theta(x) | x) = 0 \text{ and } f^{(2)}(\theta(x) | x) < 0$$

Similarly, we have

$$\hat{f}_n^{(1)}(\hat{\theta}_n(x) | x) = 0 \text{ and } \hat{f}_n^{(2)}(\hat{\theta}_n(x) | x) < 0$$

A Taylor series expansion of $\hat{f}_n^{(1)}(\cdot | x)$ in the neighborhood of $\theta(x)$, gives

$$0 = \hat{f}_n^{(1)}(\hat{\theta}_n(x) | x) = \hat{f}_n^{(1)}(\theta(x) | x) + \left(\hat{\theta}_n(x) - \theta(x) \right) \hat{f}_n^{(2)}(\bar{\theta}_n(x) | x)$$

where $\bar{\theta}_n(x)$ is between $\hat{\theta}_n(x)$ and $\theta(x)$, which gives that

$$\hat{\theta}_n(x) - \theta(x) = - \frac{\hat{f}_n^{(0,1)}(x, \theta(x))}{\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))} \tag{6.15}$$

if the denominator does not vanish.

To establish the asymptotic normality, we show that the numerator in (6.15), suitably normalized, is asymptotically normally distributed and that the denominator converges in probability to $f^{(0,2)}(x, \theta(x))$. The result is given in the following theorem.

Theorem 6.7. Let $x \in \mathbb{R}$ and suppose that assumptions **A1-A6** and **A8** hold. We have

$$\sqrt{nh_n^4} \left(\hat{\theta}_n(x) - \theta(x) \right) \xrightarrow{D} N(0, \sigma^2(x))$$

where \xrightarrow{D} denotes the convergence in distribution,

$$\sigma^2(x) = \frac{\mu f(x, \theta(x))}{\overline{G}(\theta(x)) L(\theta(x)) [f^{(0,2)}(x, \theta(x))]^2} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds$$

6.5.4 Auxiliary results and proofs

The proof of our main results is split up into several Lemmas. The first Lemma deals with the behavior of the difference between $E[\tilde{f}_n(x, t)]$ and $f(x, t)$.

Lemma 6.1. Under Assumptions **A2-A4** and **A6**, we have

$$\sup_{x \in \Omega} \sup_{t \in D} \left| E[\tilde{f}_n(x, t)] - f(x, t) \right| = O(h_n^2)$$

Proof. Indeed, integrating by parts twice and using a change of variable, we have

$$\begin{aligned} E[\tilde{f}_n(x, t)] &= \frac{1}{h_n^2} E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K \left(\frac{x - X_1}{h_n} \right) H \left(\frac{t - Z_1}{h_n} \right) \right] \\ &= \frac{1}{h_n^3} \int_{\mathbb{R}} \int_{\mathbb{R}} K \left(\frac{x - r}{h_n} \right) H^{(1)} \left(\frac{t - s}{h_n} \right) F(s | r) l(r) dr ds \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} K(\beta) H(\alpha) f(x - \beta h_n, t - \alpha h_n) d\beta d\alpha \end{aligned}$$

Finally a Taylor expansion gives, under **A2** and **A3**

$$E \left[\tilde{f}_n(x, t) \right] - f(x, t) = \frac{h_n^2}{2} \left[\int_{\mathbb{R}} \int_{\mathbb{R}} K(\beta) H(\alpha) \left[\beta^2 \frac{\partial^2 f}{\partial x^2}(x^*, t^*) + \alpha\beta \frac{\partial^2 f}{\partial x \partial t}(x^*, t^*) + \alpha^2 \frac{\partial^2 f}{\partial t^2}(x^*, t^*) \right] d\beta d\alpha \right]$$

where $x^* \in [x - \beta h_n, x]$ and $t^* \in [t - \alpha h_n, t]$. Considering assumption **A6**, the integrals are finite and we get result. \square

The second Lemma deals the behavior of the difference between $\hat{f}_n(x, t)$ and $\tilde{f}_n(x, t)$.

Lemma 6.2 : Under assumption **A1(i)**, **A2** and **A3**, we have

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(x, t) - \tilde{f}_n(x, t) \right| = O \left(\frac{1}{h_n^2} \sqrt{\frac{\log \log n}{n}} \right)$$

Proof. We observe that one can have the following decomposition

$$\begin{aligned} \left| \hat{f}_n(x, t) - \tilde{f}_n(x, t) \right| &\leq \frac{1}{nh_n^2} \sum_{i=1}^n \left| \delta_i K \left(\frac{x - X_i}{h_n} \right) H \left(\frac{t - Z_i}{h_n} \right) \right| \left| \frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right| \\ &\leq \underbrace{\frac{mM}{h_n^2} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_{I} \frac{1}{n} \sum_{i=1}^n \delta_i \end{aligned}$$

Since $I = O \left(\sqrt{\frac{\log \log n}{n}} \right)$ (see Lemma 3 in Sun and Zhou, 1998), in conjunction with the SLLN on the censoring law (see formula 4.28 in Deheuvelds and Einmahl, 2000), we have

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(x, t) - \tilde{f}_n(x, t) \right| \leq \frac{c}{h_n^2} \sqrt{\frac{\log \log n}{n}} \text{ as } n \rightarrow \infty$$

we concludes the proofs. \square

Lemma 6.3. Under Assumptions **A1(i)**, **A2** and **A3**, we have

$$\sup_{x \in \Omega} \sup_{t \in D} \left| \tilde{f}_n(x, t) - E \left[\tilde{f}_n(x, t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_n^2}} \right)$$

Proof. We use covering set technique. Indeed, since Ω and D are compact sets, then they can be covered by a finite numbers s_n and d_n of intervals centered at x_1, \dots, x_{s_n} and t_1, \dots, t_{d_n} of length a_n and b_n respectively such that $a_n = b_n = n^{-\frac{1}{2}} h_n^2$. As Ω and D are bounded, there exist two constant A_1 and A_2 such that $s_n a_n \leq A_1$ and $d_n b_n \leq A_2$. Now put $k(x) = \arg \min_{k=1, \dots, s_n} |x - x_k|$ and $J(t) = \arg \min_{j=1, \dots, d_n} |t - t_j|$. Thus we have the following decomposition

$$\begin{aligned} \sup_{x \in \Omega} \sup_{t \in D} \left| \tilde{f}_n(x, t) - E \left[\tilde{f}_n(x, t) \right] \right| &\leq \max_j \sup_{x \in \Omega} \sup_{t \in D} \left| \tilde{f}_n(x, t) - \tilde{f}_n(x, t_{J(t)}) \right| + \\ &\max_j \sup_{x \in \Omega} \sup_{t \in D} \left| E \left(\tilde{f}_n(x, t_{J(t)}) \right) - E \left(\tilde{f}_n(x, t) \right) \right| + \\ &\max_k \max_j \sup_{x \in \Omega} \left| \tilde{f}_n(x, t_{J(t)}) - \tilde{f}_n(x_{k(x)}, t_{J(t)}) \right| + \\ &\max_k \max_j \sup_{x \in \Omega} \left| E \left(\tilde{f}_n(x_{k(x)}, t_{J(t)}) \right) - E \left(\tilde{f}_n(x, t_{J(t)}) \right) \right| + \\ &\max_k \max_j \left| \tilde{f}_n(x_{k(x)}, t_{J(t)}) - E \left[\tilde{f}_n(x_{k(x)}, t_{J(t)}) \right] \right| \\ &= I_{1n} + I_{2n} + I_{3n} + I_{4n} + I_{5n} \end{aligned}$$

Firstly, concerning I_{1n} and I_{2n} , assumptions **A2** and **A3** yields

$$I_{1n} \leq \frac{c\mu m}{L(a) \overline{G}(b)} \frac{b_n}{h_n^3}$$

The condition upon b_n , we get

$$I_{1n} = O\left(\frac{1}{\sqrt{nh_n^2}}\right) \text{ a.s.} \quad (6.16)$$

Similarly, we obtain for I_{3n} and I_{4n}

$$I_{3n} \leq \frac{c\mu M}{L(a)\overline{G}(b)} \frac{a_n}{h_n^3}$$

The condition upon a_n , we get

$$I_{3n} = O\left(\frac{1}{\sqrt{nh_n^2}}\right) \text{ a.s.} \quad (6.17)$$

As to I_{5n} , for all $\epsilon > 0$ we have

$$\begin{aligned} & P \left[\max_{k=1, \dots, s_n} \max_{j=1, \dots, d_n} \left| \tilde{f}_n(x_{k(x)}, t_{J(t)}) - E \left[\tilde{f}_n(x_{k(x)}, t_{J(t)}) \right] \right| > \epsilon \right] \\ & \leq s_n d_n P \left[\left| \tilde{f}_n(x_{k(x)}, t_{J(t)}) - E \left[\tilde{f}_n(x_{k(x)}, t_{J(t)}) \right] \right| > \epsilon \right] \end{aligned} \quad (6.18)$$

Now, for any $i \geq 1$, setting

$$\Psi_i(x, t) = \frac{\mu}{nh_n^2} \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K\left(\frac{x - X_i}{h_n}\right) H\left(\frac{t - Z_i}{h_n}\right) - E \left[\frac{\delta_1}{L(Z_1)\overline{G}(Z_1)} K\left(\frac{x - X_1}{h_n}\right) H\left(\frac{t - Z}{h_n}\right) \right] \right]$$

The rv's $U_i = nh_n^2 \Psi_i(x_{k(x)}, t_{J(t)})$ are centered and bounded by $\frac{2\mu M m}{L(a)\overline{G}(b)} < \infty$, then, applying Höeffding's inequality (see Shorack and Wellner, 1986, p. 855) yields

$$\begin{aligned} & P \left[\max_k \max_j \left| \sum_{i=1}^n \Psi_i(x_{k(x)}, t_{J(t)}) \right| > \epsilon \right] \\ & = P \left[\max_k \max_j \left| \sum_{i=1}^n U_i \right| > nh_n^2 \epsilon \right] \leq \exp(-\epsilon^2 n^2 h_n^4 c) \end{aligned}$$

where c is a positive constant depending only on M , m , μ and $L(a)$, $\overline{G}(b)$ then (6.18) becomes

$$\begin{aligned}
& P \left[\max_k \max_j \left| \tilde{f}_n(x_{k(x)}, t_{J(t)}) - E \left[\tilde{f}_n(x_{k(x)}, t_{J(t)}) \right] \right| > \epsilon \right] \\
& \leq A_1 A_2 (nh_n)^{-4} n^{5 - \epsilon^2 n^2 h_n^4 c / \log n}
\end{aligned} \tag{6.19}$$

By A1(i), which yields that the last term of (6.19) is the general term of a convergent series, then by Borel-Cantelli's Lemma, the first term of (6.19) goes to zero almost surely. Otherwise, if we replace ϵ by $\epsilon_0 \sqrt{\frac{\log n}{nh_n^2}}$, for some $\epsilon_0 > 0$ in all steps of the Lemma, we have

$$\sqrt{\frac{nh_n^2}{\log n}} I_{5n} = O(1) \quad a.s. \tag{6.20}$$

which to gather with (6.16)-(6.17) and (6.20) gives the result. \square

The following Lemma in which they state a rate of convergence for μ_n under censored and truncated data.

Lemma 6.4. Under Assumption **A1**(i), we have

$$|\mu_n - \mu| = O \left(\sqrt{\frac{\log \log n}{n}} \right)$$

Proof : First, observe that

$$\begin{aligned}
\mu_n - \mu &= \frac{L_n(y) (1 - \hat{F}_n(y)) (1 - G_n(y))}{C_n(y)} - \frac{L(y) (1 - F(y)) (1 - G(y))}{C(y)} \\
&= L_n(y) (1 - G_n(y)) \underbrace{\left[\frac{(1 - \hat{F}_n(y))}{C_n(y)} - \frac{(1 - F(y))}{C(y)} \right]}_{J_{1n}} \\
&\quad + \frac{(1 - F(y))}{C(y)} \left\{ (1 - G_n(y)) \underbrace{[L_n(y) - L(y)]}_{J_{2n}} + L(y) \underbrace{[G(y) - G_n(y)]}_{J_{3n}} \right\}
\end{aligned}$$

On the one hand, from Lemma 3 in Sun and Zhou (1998), we have

$$\sup_{y \in D} |J_{1n}| = O\left(\sqrt{\frac{\log \log n}{n}}\right) \text{ as } n \rightarrow \infty \quad (6.21)$$

On the other hand, using Remark 6 in Woodroffe (1985), we have

$$\sup_{y \in D} |J_{2n}| = O\left(\sqrt{\frac{1}{n}}\right) \text{ as } n \rightarrow \infty \quad (6.22)$$

To end the proof, from LIL on the censoring law (see formula 4.28 in Deheuvels and Einmahl, 2000), we have

$$\sup_{y \in D} |J_{3n}| = O\left(\sqrt{\frac{\log \log n}{n}}\right) \text{ as } n \rightarrow \infty \quad (6.23)$$

Hence, for the continuity of F and eqnarrys (6.21) – (6.23) we obtain the result \square

The following Lemma gives the uniform convergence with rate of the estimator $l_n(x)$ defined in (6.13)

Lemma 6.5. Under Assumption **A1**(i), **A2** and **A4**, we have

$$\sup_{x \in \Omega} |l_n(x) - l(x)| = O\left(\max\left(\left(\frac{\log n}{nh_n}\right)^{\frac{1}{2}}; h_n^2\right)\right) \text{ a.s. as } n \rightarrow \infty$$

Proof. Adapting (6.13), we define

$$\tilde{l}_n(x) = \frac{\mu}{nh_n} \sum_{i=1}^n \frac{1}{L(Z_i)} K\left(\frac{x - X_i}{h_n}\right)$$

We have

$$\begin{aligned} \sup_{x \in \Omega} |l_n(x) - l(x)| &\leq \sup_{x \in \Omega} |l_n(x) - \tilde{l}_n(x)| + \sup_{x \in \Omega} |\tilde{l}_n(x) - E[\tilde{l}_n(x)]| \\ &\quad + \sup_{x \in \Omega} |E[\tilde{l}_n(x)] - l(x)| \\ &= \varphi_{1n} + \varphi_{2n} + \varphi_{3n} \end{aligned} \quad (6.24)$$

For the first term φ_{1n} , we have

$$\left| l_n(x) - \tilde{l}_n(x) \right| \leq \left\{ \frac{|\mu_n - \mu|}{L_n(a)} + \frac{\mu}{L_n(a)L(a)} \sup_y |L_n(y) - L(y)| \right\} \frac{1}{nh_n} \sum_{i=1}^n \left| K\left(\frac{x - X_i}{h_n}\right) \right|$$

Then for any $\epsilon > 0$, we have

$$\begin{aligned} P \left[\frac{1}{nh_n} \sum_{i=1}^n \left| K\left(\frac{x - X_i}{h_n}\right) \right| > \epsilon \right] &\leq \frac{1}{nh_n \epsilon} \sum_{i=1}^n E \left(\left| K\left(\frac{x - X_i}{h_n}\right) \right| \right) \\ &= \frac{1}{\epsilon} \int_{\mathbb{R}} |K(r)| l(x - rh_n) dr \rightarrow \frac{l(x)}{\epsilon} \int_{\mathbb{R}} |K(r)| dr \end{aligned}$$

Since $L(a) > 0$, from **lemma 6.4** $|\mu_n - \mu| = O\left(\sqrt{\frac{\log \log n}{n}}\right)$, $\sup_y |L_n(y) - L(y)| = O\left(\frac{1}{\sqrt{n}}\right)$ (see Woodroffe, 1985). Then we get

$$\sup_{x \in \Omega} \left| l_n(x) - \tilde{l}_n(x) \right| = O\left(\sqrt{\frac{\log \log n}{n}}\right) \quad (6.25)$$

In addition, by using analogous framework as in **lemma 6.3**, we get

$$\varphi_{2n} = O\left(\left(\frac{\log n}{nh_n}\right)^{\frac{1}{2}}\right) \quad \text{as } n \rightarrow \infty \quad (6.26)$$

Finally, a change of variable and a Taylor expansion, we get, under **A2** and **A4**

$$E \left[\tilde{l}_n(x) \right] - l(x) = \frac{h_n^2}{2} \int s^2 K(s) l''(\bar{x}) ds$$

with $\bar{x} \in [x - sh_n, x]$, which yields that

$$\varphi_{3n} = O(h_n^2) \quad \text{as } n \rightarrow \infty \quad (6.27)$$

Combining (6.25) – (6.27) permits to conclude the proof. \square

Proof of Proposition 6.1 : Using the triangular inequality, we have

$$\begin{aligned}
\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| &\leq \frac{1}{\inf_{x \in \Omega} l_n(x)} \sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(x, t) - \tilde{f}_n(x, t) \right| \\
&+ \frac{1}{\inf_{x \in \Omega} l_n(x)} \sup_{x \in \Omega} \sup_{t \in D} \left| E \left[\tilde{f}_n(x, t) \right] - f(x, t) \right| \\
&+ \frac{1}{\inf_{x \in \Omega} l_n(x)} \sup_{x \in \Omega} \sup_{t \in D} \left| \tilde{f}_n(x, t) - E \left[\tilde{f}_n(x, t) \right] \right| \\
&+ \frac{1}{\inf_{x \in \Omega} l_n(x)} \gamma^{-1} \sup_{x \in \Omega} \sup_{t \in D} |f(x, t)| \sup_{x \in \Omega} |l_n(x) - l(x)|
\end{aligned}$$

Lemmas **6.1-6.5** permits to conclude the proof. \square

Proof of Theorem 6.6 : We have

$$\begin{aligned}
\sup_{x \in \Omega} \left| f \left(\hat{\theta}_n(x) | x \right) - f \left(\theta(x) | x \right) \right| &\leq \sup_{x \in \Omega} \left| f \left(\hat{\theta}_n(x) | x \right) - \hat{f}_n \left(\hat{\theta}_n(x) | x \right) \right| \\
&+ \sup_{x \in \Omega} \left| \hat{f}_n \left(\hat{\theta}_n(x) | x \right) - f \left(\theta(x) | x \right) \right| \\
&\leq \sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right| + \sup_{x \in \Omega} \left| \sup_{t \in D} \hat{f}_n(t | x) - \sup_{t \in D} f(t | x) \right| \\
&\leq 2 \sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right|
\end{aligned}$$

The a.s. uniform consistency of $\hat{\theta}_n(x)$ follows then immediately from **Proposition 6.1** and assumption **A7**, and the continuity of $f(\cdot | x)$. Now expansion of $f(\cdot | x)$ in a neighborhood of $\theta(x)$ gives

$$f \left(\hat{\theta}_n(x) | x \right) - f \left(\theta(x) | x \right) = \frac{1}{2} \left(\hat{\theta}_n(x) - \theta(x) \right)^2 f^{(2)} \left(\bar{\theta}(x) | x \right) \quad (6.28)$$

where $\bar{\theta}(x)$ is between $\hat{\theta}_n(x)$ and $\theta(x)$. Then by (6.28) and **A5**, we have

$$\sup_{x \in \Omega} \left| \hat{\theta}_n(x) - \theta(x) \right| \leq 2 \sqrt{\frac{\sup_{x \in \Omega} \sup_{t \in D} \left| \hat{f}_n(t | x) - f(t | x) \right|}{\left| f^{(2)}(\bar{\theta}(x) | x) \right|}}$$

Hence, by **Proposition 3.1**, the proof of **Theorem 3.2** is complete. \square

Proof of Theorem 6.7. From (6.15) we have the following decomposition

$$\begin{aligned} \sqrt{nh_n^4} \left(\hat{\theta}_n(x) - \theta(x) \right) &= \sqrt{nh_n^4} \frac{\hat{f}_n^{(0,1)}(x, \theta(x)) - \tilde{f}_n^{(0,1)}(x, \theta(x))}{\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))} \\ &\quad + \sqrt{nh_n^4} \frac{\tilde{f}_n^{(0,1)}(x, \theta(x)) - E \left[\tilde{f}_n^{(0,1)}(x, \theta(x)) \right]}{\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))} \\ &\quad + \sqrt{nh_n^4} \frac{E \left[\tilde{f}_n^{(0,1)}(x, \theta(x)) \right]}{\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))} \\ &= \frac{S_1 + S_2 + S_3}{\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))} \end{aligned} \tag{6.29}$$

So, it suffices to show that $\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x))$ converges in probability to $f^{(0,2)}(x, \theta(x))$

and $S_i \rightarrow 0$ ($i = 1, 3$) where as S_2 is asymptotically normal. \square

Now, to prove that $\hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x)) \rightarrow f^{(0,2)}(x, \theta(x))$, indeed, we have the following

Lemma 6.6. Under assumptions **A1(i)**, **A2-A3** and **A6**, we have

$$\sup_{t \in D} \left| \hat{f}_n^{(0,2)}(x, t) - f^{(0,2)}(x, t) \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Proof. Since

$$\begin{aligned} \left| \hat{f}_n^{(0,2)}(x, t) - f^{(0,2)}(x, t) \right| &\leq \left| \hat{f}_n^{(0,2)}(x, t) - \tilde{f}_n^{(0,2)}(x, t) \right| + \left| \tilde{f}_n^{(0,2)}(x, t) - f^{(0,2)}(x, t) \right| \\ &= \gamma_{1,n}(x, t) + \gamma_{2,n}(x, t) \end{aligned}$$

The proof is completed through **Lemma 6.7** and **Lemma 6.8**. \square

Lemma 6.7. Under assumptions **A1(i)**, **A2-A3** and **A6**, we have

$$\sup_{t \in D} \gamma_{1,n}(x, t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned} \sup_{t \in D} \gamma_{1,n}(x, t) &\leq \sup_{t \in D} \frac{1}{nh_n^4} \sum_{i=1}^n \left| \delta_i K \left(\frac{x - X_i}{h_n} \right) H^{(2)} \left(\frac{t - Z_i}{h_n} \right) \left(\frac{(1 - \hat{F}_n(Z_i))}{C_n(Z_i)} - \frac{(1 - F(Z_i))}{C(Z_i)} \right) \right| \\ &\leq \underbrace{\frac{M''}{h_n^3} \sup_{t \in D} \left| \frac{(1 - \hat{F}_n(t))}{C_n(t)} - \frac{(1 - F(t))}{C(t)} \right|}_I \frac{1}{nh_n} \sum_{i=1}^n \left| K \left(\frac{x - X_i}{h_n} \right) \right| \end{aligned}$$

Then for any $\epsilon > 0$, we have

$$\begin{aligned} P \left[\frac{1}{nh_n} \sum_{i=1}^n \left| K \left(\frac{x - X_i}{h_n} \right) \right| > \epsilon \right] &\leq \frac{1}{nh_n \epsilon} \sum_{i=1}^n E \left(\left| K \left(\frac{x - X_i}{h_n} \right) \right| \right) \\ &= \frac{1}{\epsilon} \int_{\mathbb{R}} |K(r)| l(x - rh_n) dr \rightarrow \frac{l(x)}{\epsilon} \int_{\mathbb{R}} |K(r)| dr \end{aligned}$$

Since $I = O \left(\sqrt{\frac{\log \log n}{n}} \right)$ (see Lemma 3 in Sun and Zhou, 1998), we have $\sup_{t \in D} \gamma_{1,n}(x, t) =$

$O \left(\frac{1}{h_n^3} \sqrt{\frac{\log \log n}{n}} \right)$, which goes to zero under **A1(i)** \square

Lemma 6.8. Under assumptions **A1(i)**, **A2-A3** and **A6**, we have

$$\sup_{t \in D} \gamma_{2,n}(x, t) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Proof. Since

$$\begin{aligned} \sup_{t \in D} \gamma_{2,n}(x, t) &\leq \sup_{t \in D} \left| E \left[\tilde{f}_n^{(0,2)}(x, t) \right] - f^{(0,2)}(x, t) \right| + \sup_{t \in D} \left| \tilde{f}_n^{(0,2)}(x, t) - E \left[\tilde{f}_n^{(0,2)}(x, t) \right] \right| \\ &= \beta_{1,n} + \beta_{2,n} \end{aligned}$$

On the one hand, integrating by parts twice and using a change of variable, it follows as in **Lemma 6.1**

$$\begin{aligned} \beta_{1,n} &= \sup_{t \in D} \left| \int_{\mathbb{R}} \int_{\mathbb{R}} K(r) H(s) f^{(0,2)}(x - rh_n, t - sh_n) dr ds - f^{(0,2)}(x, t) \right| \\ &= O(h_n^2) \end{aligned}$$

On the other hand, it proceeds along the same steps and by invoking the same arguments as in the proof of **Lemma 6.3**, a suitable covering of the compact set D (that is using interval with length $b_n = n^{-\frac{1}{2}}h_n^4$) and changing the variables $\Psi_i(\cdot, \cdot)$ into the following :

$$\begin{aligned} \Gamma_i(x, t_{J(t)}) &= \frac{\mu}{nh_n^4} \left[\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)} K\left(\frac{x-X_i}{h_n}\right) H^{(2)}\left(\frac{t_{J(t)}-Z_i}{h_n}\right) - E \left[\frac{\delta_1}{L(Z_1)\overline{G}(Z_1)} K\left(\frac{x-X_1}{h_n}\right) H^{(2)}\left(\frac{t_{J(t)}-Z_1}{h_n}\right) \right] \right] \\ &= (nh_n^4)^{-1} V_i(x, t_{J(t)}) \end{aligned}$$

to get

$$\beta_{2,n} = O \left\{ \left(\frac{\log n}{nh_n^2} \right)^{\frac{1}{2}} \right\} \text{ as } n \rightarrow \infty$$

which gives the proof of the Lemma. \square

In order to achieve the proof, we observe that we have, for n large enough

$$\left| \hat{f}_n^{(0,2)}(x, \bar{\theta}_n(x)) - f^{(0,2)}(x, \theta(x)) \right| \leq \sup_{t \in D} \left| \hat{f}_n^{(0,2)}(x, t) - f^{(0,2)}(x, t) \right| + \left| f^{(0,2)}(x, \bar{\theta}_n(x)) - f^{(0,2)}(x, \theta(x)) \right|$$

By **Lemma 6.6** and assumption **A8**, we get the result \square

Lemma 6.9. Under assumptions **A1-A3** and **A6**, we have

$S_1 \rightarrow 0$ as $n \rightarrow \infty$

Proof. We have

$$\begin{aligned} |S_1| &= \sqrt{nh_n^4} \left| \frac{1}{nh_n^3} \sum_{i=1}^n \delta_i K \left(\frac{x - X_i}{h_n} \right) H^{(1)} \left(\frac{\theta(x) - Z_i}{h_n} \right) \left(\frac{1 - \hat{F}_n(Z_i)}{C_n(Z_i)} - \frac{1 - F(Z_i)}{C(Z_i)} \right) \right| \\ &\leq \frac{\sqrt{nh_n^4}}{h_n} \underbrace{\sup_{t \in D} \left| \frac{1 - \hat{F}_n(t)}{C_n(t)} - \frac{1 - F(t)}{C(t)} \right|}_I \frac{1}{nh_n^2} \sum_{i=1}^n \delta_i K \left(\frac{x - X_i}{h_n} \right) H^{(1)} \left(\frac{\theta(x) - Z_i}{h_n} \right) \end{aligned}$$

Analogously as in Lemma 6.7, we have $|S_1| = O\left(\sqrt{h_n^2 \log \log n}\right)$, which goes to zero under **A1**(i, ii). \square

Lemma 6.10. Under assumptions **A1**(ii), **A2-A3** and **A6**, we have

$S_3 \rightarrow 0$ as $n \rightarrow \infty$

Proof. We have

$$\begin{aligned} E \left[\tilde{f}_n^{(0,1)}(x, \theta(x)) \right] &= \frac{1}{h_n^3} E \left[\frac{\mu \delta_1}{L(Z_1) \bar{G}(Z_1)} K \left(\frac{x - X_1}{h_n} \right) H^{(1)} \left(\frac{\theta(x) - Z_1}{h_n} \right) \right] \\ &= \frac{1}{h_n^3} E \left[K \left(\frac{x - X_1}{h_n} \right) E \left[\frac{\mu \delta_1}{L(Z_1) \bar{G}(Z_1)} H^{(1)} \left(\frac{\theta(x) - Z_1}{h_n} \right) \mid X_1 \right] \right] \\ &= \frac{1}{h_n^3} \int_{\mathbb{R}} \int_{\mathbb{R}} K \left(\frac{x - u}{h_n} \right) H^{(1)} \left(\frac{\theta(x) - v}{h_n} \right) f(u, v) dudv \end{aligned}$$

Integrating by parts with respect to v and using conditions **A2-A3** and **A6** we obtain, by a Taylor expansion

$$\begin{aligned}
E \left[\tilde{f}_n^{(0,1)}(x, \theta(x)) \right] &= \int_{\mathbb{R}} \int_{\mathbb{R}} K(r) H(s) f^{(0,1)}(x - rh_n, \theta(x) - sh_n) dr ds \\
&= O(h_n^2)
\end{aligned}$$

thus by **A1(ii)**, we conclude $S_3 \rightarrow 0$ as $n \rightarrow \infty$. \square

Lemma 6.11. Under assumptions **A2-A3** and **A6**, we have

$$\text{Var} [S_2] \longrightarrow \frac{\mu f(x, \theta(x))}{\overline{G}(\theta(x)) L(\theta(x))} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds, \text{ as } n \rightarrow \infty$$

Proof. We have

$$\begin{aligned}
\text{Var} [S_2] &= nh_n^4 \text{Var} \left[\tilde{f}_n^{(0,1)}(x, \theta(x)) \right] \\
&= \frac{1}{h_n^2} E \left[\frac{\mu^2 \delta_1}{L^2(Z_1) \overline{G}^2(Z_1)} K^2 \left(\frac{x - X_1}{h_n} \right) (H^{(1)})^2 \left(\frac{\theta(x) - Z_1}{h_n} \right) \right] \\
&\quad - \frac{1}{h_n^2} \left\{ E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K \left(\frac{x - X_1}{h_n} \right) H^{(1)} \left(\frac{\theta(x) - Z_1}{h_n} \right) \right] \right\}^2 \\
&= J_{1,n} - J_{2,n} \tag{6.30}
\end{aligned}$$

on the one hand,

$$\begin{aligned}
J_{1,n} &= \frac{1}{h_n^2} E \left[K^2 \left(\frac{x - X_1}{h_n} \right) E \left[\frac{\mu^2 \delta_1}{L^2(Z_1) \overline{G}^2(Z_1)} (H^{(1)})^2 \left(\frac{\theta(x) - Z_1}{h_n} \right) \mid X_1 \right] \right] \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\mu}{L(\theta(x) - sh_n) \overline{G}(\theta(x) - sh_n)} (H^{(1)})^2(s) K^2(r) f(x - rh_n, \theta(x) - sh_n) dr ds
\end{aligned}$$

Again, a Taylor expansion gives us

$$J_{1,n} = \mu f(x, \theta(x)) \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{(H^{(1)})^2(s) K^2(r)}{L(\theta(x) - sh_n) \overline{G}(\theta(x) - sh_n)} dr ds + o(1)$$

Then, since $G(\cdot)$ and $L(\cdot)$ are continuous, we have, under **A2** and **A3**

$$J_{1,n} = \frac{\mu f(x, \theta(x))}{L(\theta(x)) \overline{G}(\theta(x))} \int_{\mathbb{R}} \int_{\mathbb{R}} [(H^{(1)}(s) K(r))^2] dr ds + o(1)$$

On the other hand, by **Lemma 6.10**

$$J_{2,n} = \frac{S_3^2}{n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Finally, we obtain

$$\text{Var}[S_2] \rightarrow \frac{\mu f(x, \theta(x))}{\overline{G}(\theta(x)) L(\theta(x))} \int_{\mathbb{R}^2} [K(r) H^{(1)}(s)]^2 dr ds \text{ as } n \rightarrow \infty.$$

This proves the Lemma. \square

Now the final step to prove **Theorem 6.7** is to show the Berry-Esséen condition for S_2 (see Chow and Teicher 1997, p. 322). For that, put

$$S_2 = \sum_{i=1}^n \Delta_{i,n}(x)$$

where

$$\Delta_{i,n}(x) = \frac{1}{\sqrt{nh_n^2}} \left\{ \frac{\mu \delta_i}{L(Z_i) \overline{G}(Z_i)} K\left(\frac{x - X_i}{h_n}\right) H^{(1)}\left(\frac{\theta(x) - Z_i}{h_n}\right) - E \left[\frac{\mu \delta_1}{L(Z_1) \overline{G}(Z_1)} K\left(\frac{x - X_1}{h_n}\right) H^{(1)}\left(\frac{\theta(x) - Z_1}{h_n}\right) \right] \right\}$$

Then, we have

Lemma 6.12. Under assumptions **A1-A3** and **A6**, we have

$$\sum_{i=1}^n E[|\Delta_{i,n}(x)|^3] < \infty$$

Proof. Applying the C_r -inequality (see Loève, 1963, p.155), we have

$$\begin{aligned}
E [|\Delta_{i,n}(x)|^3] &\leq 4 (nh_n^2)^{-3/2} \left\{ E \left[\left| \frac{\mu\delta_1}{L(Z_1)\overline{G}(Z_1)} K\left(\frac{x-X_1}{h_n}\right) H^{(1)}\left(\frac{\theta(x)-Z_1}{h_n}\right) \right|^3 \right] \right. \\
&\quad \left. + E \left[\left| E \left[\frac{\mu\delta_1}{L(Z_1)\overline{G}(Z_1)} K\left(\frac{x-X_1}{h_n}\right) H^{(1)}\left(\frac{\theta(x)-Z_1}{h_n}\right) \right] \right|^3 \right] \right\} \quad (6.31)
\end{aligned}$$

Both expectation terms in (6.31) being finite under **A1-A3** and **A6**, we get

$$\sum_{i=1}^n E [|\Delta_{i,n}(x)|^3] = o(nh_n^6)^{-\frac{1}{2}} = o(1)$$

This completes the proof of **Lemma 6.12**, and therefore that of **Theorem 6.7**.

Conclusion et perspectives

La modélisation statistique par le biais de données incomplètes est largement employée lors d'études pratiques sur les durées de vie. Les formes de données incomplètes, le plus couramment utilisées sont les données censurées et tronquées. Ce type de données a des caractéristiques très spécifiques dont on doit tenir compte dans leur modélisation. Dans cette thèse, nous avons proposé des estimateurs pour les densités, ainsi que pour le mode, simple et conditionnel, ces derniers restent un sujet très abordé de nos jours.

Nous nous sommes intéressés, dans un premier temps, à la fonction de mode simple. Nous avons construit, pour ce modèle de LTRC dans le cas i.i.d., un nouvel estimateur lissé de la densité ainsi que pour le mode simple. Nous avons établi sa convergence uniforme presque sûre sur un compact et sa normalité asymptotique, en donnant leurs vitesses de convergence.

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

Nous avons considéré les données censurées, tronquées et α -mélangeantes pour établir, sous une hypothèse sur le coefficient de mélange, la consistance de l'estimateur du mode simple, ainsi, que pour la fonction densité. Nous avons aussi montré, la normalité asymptotique de l'estimateur à noyau de mode tout en précisant le terme de variance.

Au regard des champs d'applications et pour obtenir une modélisation plus fine, nous avons, dans un second temps, envisagé de travailler sur l'estimation de la fonction mode lorsqu'on dispose d'informations supplémentaires regroupées dans une variable explicative X . Nous sommes revenus au modèle LTRC dans le cas i.i.d., sous ce modèle nous construisons un nouvel estimateur lissé de la fonction de densité conditionnelle. Nous avons établi sa convergence uniforme presque sûre sur un compact et sa normalité asymptotique en donnant leur vitesse de convergence.

En perspectives de recherche, nous présentons plusieurs projets se situant directement en continuité des études effectuées dans le cadre de cette thèse. L'objectif est naturellement d'établir des propriétés asymptotiques. Nous donnons quelques détails ci-après.

1- Sur les modèles conditionnels

Les résultats établis au chapitre 6, on peut également envisager d'étendre au cas des données tronquées et censurées les autres modèles non paramétriques conditionnels, par exemple la fonction de hasard conditionnelle, la fonction de répartition conditionnelle et le quantile conditionnel.

2- Cas d'un modèle LTRC en situation de dépendance

A notre connaissance, c'est la situation la moins développée dans la littérature et les résultats de ce chapitre de cette thèse ouvrent la porte à de nombreuses extensions. Concernant le cadre du α -mélange, les résultats obtenus précédemment devraient se généraliser à modèles conditionnels.

3- Cas des données fonctionnelles

Nous envisageons de traiter l'estimation du mode conditionnelle dans le cas où la variable explicative est fonctionnelle. Le cas où les données sont complètes est traité par Ezzahrioui et Ould-Saïd (2005, 2006), nous nous pencherons sur le cas des données censurées et tronquées.

4- Cas des données spatiales

Étendre dans le cas de censure et troncature, la régression pour des données spatiales.

Bibliographie

- [1] Abraham, C., Biau, G. et Cadre, B. (2004). On the asymptotic properties of a simple estimate of the mode. *ESAIM, Probab. Statist.*, 8 : 1-11.
- [2] Asselin de Beauville, J. P. (1978). Estimation non paramétrique de la densité et du mode : Exemple de la distribution gamma. *Revue de Statist. Appl.*, 26 : 47-70
- [3] Auestad, B. et Tjøtheim, D. (1990). Identification of nonlinear time series : First order characterisation and order determination. *Biometrika*, 77 : 669-687
- [4] Bailey, K.R. (1979). The general maximum likelihood approach to the Cox regression model. Ph.D. Dissertation, University of Chicago, Illinois.
- [5] Begum, J., Hall, W., Huang, W., et Wellner, J. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.*, 11 : 432-452.
- [6] Benrabah, W., Ould-Said, E. (1998). On the law of the logarithm for density and hazard rate estimation from truncated and censored data. Technical Report, 44, Univ. du Littoral Côte d'Opale.
- [7] Berlinet, A., Gannoun, A. et Matzner-Løber, E. (1998). Normalité asymptotique d'estimateurs convergent du mode conditionnel. *Canad. J.Statist.*, 26 : 365-380.
- [8] Bickel, P., Klaassen, C., Ritov, T., et Wellner, J. (1987). Efficient and adaptive inference in semi parametric models. John Hopkins University Press, Baltimore.
- [9] Bickel, D.R. et Frühwirth, R. (2006). On a fast robust estimator of the mode : Comparaisons on to other robust estimators with applications. *Computat. Statist. Data Anal.*, 50 : 3500-3530.

- [10] 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.
- [11] Bosq, D. (1993). Bernstein type large deviations inequalities for partial sums of strong mixing processes. *Statistics*, 24, 59-70.
- [12] Bosq, D. (1996). Limits theorems for Banach-valued autoregressive processes. Application to real continuous time processes. *Bull. Belg. Math. Soc. Simon. Stevin.* 3(5) : 537-555.
- [13] Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes : Estimation and Prediction (Second Edition)*. Lecture Notes in Statistics, 110 : Springer-Verlag, Berlin.
- [14] Breslow, N. et Crowley, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *Ann. Statist.*, 2 : 437-453.
- [15] Cai, Z. (1998). Asymptotic properties of Kaplan-Meier estimator for censored dependent data. *Statist. Probab. Lett.*, 37 : 381-389.
- [16] Cai, Z. (2001). Estimating a Distribution Function for Censored Time Series Data, *J. Multivariate Anal.*, 78 : 299-318.
- [17] Cai, Z. et Roussas, G. G. (1992). Uniform strong estimation under α -mixing, with rates, *Statist. Probab. Lett.*, 15 : 47-55.
- [18] Chen, Q., Dai, Y.(2003). Kernel estimation of higher derivatives of density and hazard rate function for truncated and censored dependent data. *J. Acta. Math. Scientia* 23B, 477-486.
- [19] Chernoff, H., (1964). Estimation of the mode. *Ann. Instit. Statist. Math.*, 16 : 31-41.
- [20] Chow, Y.S. , Teicher, H. (1997). *Probability theory. Independence, interchangeability, martingales*. Springer. New York.
- [21] Collomb, G., Härdle, W. et Hassani, S. (1987). A note on prediction via estimation of the conditional mode function. *J. Statist. Plann. and Inf.*, 15 : 227-236.

- [22] Cox, D.R. (1972). Regression models and left-tables (with discussion). *J. Roy. Statist. Soc.B*, 34, 187-220.
- [23] Dabo-Niang, S., Ferraty, F. et Vieu, P. (2004). Estimation du mode dans un espace vectoriel semi-normé. *C. R. Acad. Sci. Paris Sér. I*, 339 : 659-662.
- [24] Dalenius, T. (1965). The mode of a neglected statistical parameter. *J. Roy. Statist. Soc. Ser. A.*, 128 : 110-117.
- [25] Deheuvels, P. , Einmahl, J.H.J. (2000). Functional limit laws for the in increments of Kaplan-Meier product limit processes and applications. *J Ann. Probab.* 28, 1301-1335.
- [26] Doob, J.L. (1953). *Stochastic Processes*. New York. Wiley.
- [27] Doukhan, P. (1994). *Mixing : Properties and Examples*. Lecture Notes in Statistics., Springer-Verlag, Berlin.
- [28] Eddy, W.F. (1980). Optimum kernel estimators of the mode. *Ann. Statist.*, 8 : 870-882.
- [29] Eddy, W.F. (1982). The asymptotic distributions of kernel estimators of the mode. *Z. Wahrsch.Verw. Gebiete*, 59 : 279-290.
- [30] Ezzahrioui, M., Ould-Saïd, E. (2005). Asymptotic normality of nonparametric estimators of the conditional mode for functional data. Technical report, No.249, LMPA, Univ. Littoral Côte d'Opale.
- [31] Ezzahrioui, M., Ould-Saïd, E. (2006). On the asymptotic properties of a nonparametric estimator of the conditional mode for functional dependent data. Preprint, LMPA No 277, Univ. du Littoral Côte d'Opale.
- [32] Ferraty, F., Laksaci, A. et Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. Stoch. Proc.*, 9 : 47-76.
- [33] Fréchet, M. (1940). The general relation between the mean and the mode for a discontinuous variate. *Ann. Math. Statist.*, 18 : 290-293.

- [34] Gannoun, A. et Saracco, J. (2002). A new proof of strong consistency of kernel estimation of density function and mode under random censorship. *Statist. Probab. Lett.*, 59 : 61-66.
- [35] Gebelein, H. (1941). Das statistische problem der korrelation als variations und eigenwert problem und sein zusammenhang mit der ausgleichsrechnung. *Z. Angew. Math. Mech.* 21 : 364-379.
- [36] Gijbels, I., Wang, J.L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *J Multivariate Anal.*, 47, 210-229.
- [37] Grenander, U. (1965). Some direct estimates of the mode. *Ann. Math. Statist.*, 131-138.
- [38] Grund, B. et Hall, P. (1995). On the minimisation of L_p error in mode estimation. *Ann. Statist.*, 23 : 2264-2284.
- [39] Gorodetski, V.V. (1977) On the strong mixing property for linear sequences, *Theo. Prob. and its Applic.* 22, 411–413.
- [40] Gu M.G., Lai T.L. (1990) Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann Probab* 18 :160–189
- [41] Hall, P. (1980). On the limiting behaviour of the mode and median of a sum of independent random variables. *Ann. Statist.* 8 : 419-430.
- [42] Hedges, S. et Shah, P. (2003). Comparison of mode estimation methods and application in molecular clock analysis. *BMC Bioinformatics*, 4 : 31-41.
- [43] Hirschfeld, H.O. (1935). A connection between correlation and contingency. *Math. Proceed. of the Cambridge Philosophical Soc*, 31 : 520-524.
- [44] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58, 15-30.

- [45] Ibragimov, I.A. (1962). Some limit theorems from stationary processes. *Theory Proc. Appl.*, 7 : 349-382.
- [46] Jones, D.A. (1978). Nonlinear autoregressive processes. *Proc. Roy. Soc. London A.*, 360, 7–95.
- [47] Kaplan, E.L. et Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.*, 53 : 457-481.
- [48] Khardani S, Lemdani M, Ould-Saïd E (2011) Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series. *J. Statist. Plann. Inference.* 141, 3426-3436
- [49] Khardani, S., Lemdani, M., Ould-Saïd, E. (2010). Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. *J Korean Statist. Soc.* 39, 455–469.
- [50] Khardani S, Lemdani M, Ould-Saïd E (2012) On the strong uniform consistency of the mode estimator for censored times series, *J. Metrika* 75 229–241.
- [51] Klein J.P., Moeschberger, M.L. (1997). *Survival analysis : techniques for censored and truncated data.* Springer-Verlag, New York.
- [52] Konakov V.D., On the asymptotic normality of the mode of multidimensional distributions, *Theory of Probab. Appl.* 19 (1973) 794–799.
- [53] Lai TL, Ying Z. (1991) Estimating a distribution with truncated and censored data. *Ann Stat* 19 :417–442.
- [54] Leclerc, J. et Pierre-Loti-Viau, D. (2000). Vitesse de convergence presque sûre de l'estimateur à noyau du mode. *C.R. Acad. Sci. Paris, Sér. I*, 331 : 637-640.
- [55] Liang, H.Y, De Uña-Álvarez, J. (2010). Asymptotic normality for estimator of conditional mode under left-truncated and dependent observations. *Journal of Metrika*, 72, 1-19.

- [56] Liang, H.Y, De Uña-Álvarez, J, Iglesias-Pérez, C. (2012). Asymptotic properties of conditional distribution estimator with truncated, censored and dependent data, *Test*,
- [57] Loève, M. (1963).. *Probability theory*. Springer-Verlag. New York.
- [58] Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronom Soc.*, 155, 95–118.
- [59] Louani, D. (1998). On the asymptotic normality of the kernel estimators of the density function and its derivatives under random censoring. *Comm. Statist. Theory & Meth.*, 27 : 2909-2924.
- [60] Louani, D. et Ould Saïd, E. (1999). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric Statist.*, 11 : 413-442.
- [61] Masry, E. (1986). Recursive probability density estimation for weakly dependent process. *IEEE. Trans. Inform. Theory*, 32 : 254-267.
- [62] Mokkadem, A., Pelletier, M. (2005). Moderate deviations for the kernel mode estimator and some applications. *J. Statist. Plann. Inf.*, 135 : 276-299.
- [63] Nadaraya E.A. (1965) , On non-parametric estimates of density functions and regression curves, *Theory Probab. Appl.* 10 186–190.
- [64] Ould Saïd, E. (1993). Estimation non paramétrique du mode conditionnel. Application à la prévision. *C. R. Acad. Sci. Paris Sér. I*, 316 : 943-947.
- [65] Ould Saïd, E.(1997). A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. Statist.*, 24 : 231-239.
- [66] Ould-Saïd, E., Tatachak, A. (2007). Asymptotic properties of the kernel estimator of the conditional mode the left-truncated model. *C. R. Acad. Sci. Paris. Ser. I* 344 651-656.

- [67] Ould-Saïd, E., Tatachak, A. (2009a). On the nonparametric estimation of the simple mode under random left-truncation model. *Romanian Journal of Pure and Applied Mathematics*, 54, 243-266.
- [68] Ould-Saïd, E., Tatachak, A. (2009b). Strong consistency rate for the kernel mode under strong mixing hypothesis and left truncation. *J. Comm. Statist. Theory Methods*, 38, 1154-1169.
- [69] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33 : 1065-1076.
- [70] Pearson, K. (1895). Contributions to the mathematical theory of evolution, II : skew variation in homogeneous material. *Philosophical transactions of the royal society of London array*, 186 : 343-414.
- [71] Quintela-Del-Río, A. et Vieu, P. (1997). A nonparametric conditional mode estimate. *J. Nonparametric Statist.*, 8 : 253-266.
- [72] Rio, E. (1994). Covariance inequalities for strongly mixing processes. *Ann. de l'I.H.P. Probab..Statist.*, 29, 587–597.
- [73] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, Mathématiques et applications. Springer. New-York.
- [74] Romano, J. (1988). On weak convergence and optimality of kernel density estimates of the mode. *Ann. Statist.*, 16 : 629-647.
- [75] Rosa, A.C. (1993). *Prévision robuste sous une hypothèse ergodique*. Thèse de Doctorat à l'Université de Toulouse I.
- [76] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.*, 42 : 43-47.
- [77] Sager, T. (1975). Consistency in nonparametric estimation of the mode. *Ann. Stat.* 3 :698-706.
- [78] Samanta, M., (1973). Nonparametric estimation of the mode of a multivariate density. *South African Statistical Journal*, 7, 109–117.

- [79] Samanta, M., and Thavaneswaran, A. (1990). Nonparametric estimation of the conditional mode. *Comm. Statist. Theory and Methods.*, 16 : 4515-4524.
- [80] Shorack, G.R. and Wellner, J.A. (1986). *Empirical processes with applications to statistics.* Wiley, New-York.
- [81] Sun, L., Zhou, Y. (1998). Sequential confidence bands for densities under truncated and censored data. *Statist. Probab. Lett.*, 40, 31-41.
- [82] Sun, L., Zhou, X. (2001). Survival function and density estimation for truncated dependent data. *Statist. Probab. Lett.* 52 :47-57.
- [83] Tsai, W.Y., Jewell, N.P., Wang, M.C. (1987). A note on the product limit estimator under right censoring and left truncation. *J. Biometrika*, 74, 883-886.
- [84] Van Ryzin J. (1969) , On strong consistency of density estimates, *Ann. Math. Statist.* 40 1765–1772.
- [85] Venter, J. (1967). On estimation of the mode. *Ann. Math. Stat.*, 38 : 1446-1455.
- [86] Vieu, P. (1996). A note on density mode function. *Statist. Probab. Lett.*, 26 : 297-307.
- [87] Voelkel, J. et Crowley, J. (1984). Nonparametric inference for a class of semi-Markov process with censored observations. *Ann. Statist.*, 12 : 142-160.
- [88] Wegman, E.J. (1971). A note on the estimation of the mode. *Ann. Math. Statist.*, 42 : 1909-1915.
- [89] Withers, C.S. (1981) Conditions for linear processes to be strong mixing, *Z. Wahrsch. G.*, 57, 477-480.
- [90] Woodroffe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.*, 13 : 163-177.
- [91] Yamato H.,(1971) Sequential estimation of a continuous probability density function and mode, *Bull. Math. Statist.* 14 1–12.

- [92] Yoshihara, K.I. (2004). Weakly dependent stochastic sequences and their applications. Vol XIV : Recent topics on weak and strong limit theorems. Sanseidoco., Ltd., Chiyoda-Ku.
- [93] Zhou, Y. (1996). A note on the TJW product-limit estimator for truncated and censored data *Statist. J. Probab. Lett.*, 26, 381-387.