

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



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

N° d'ordre :

N° de série :

THÈSE

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

« Etude de modèles semi et non paramétriques pour
des données censurées »

Par Mohamed BOUKÉLOUA

OPTION

Modélisation Stochastique

Devant le jury :

Présidente	M ^{me} N. Nemouchi	Prof.	Université frères Mentouri
Directrice de thèse	M ^{me} F. Messaci	Prof.	Université frères Mentouri
Co-directeur de thèse	M. A. Keziou	Prof.	Université de Reims Champagne-Ardenne
Examinatrice	M ^{me} O. Sadki	Prof.	U.S.T.H.B., Alger
Examineur	M. A. Laksaci	Prof.	Université D. Liabbes Sidi Belabbes
Examinatrice	M ^{me} S. Belaloui	M.C. A.	Université frères Mentouri

Soutenue le :

A la mémoire de ma mère

A mon père

A ma sœur Samira

A toute la famille

Remerciements

J'aimerais tout d'abord remercier ma directrice de thèse, madame Fatiha Messaci pour sa confiance, son orientation, son ouverture envers mes idées et son soutien permanent durant ces années de travail sous sa direction. J'ai eu l'immense chance d'être encadré par elle depuis la licence, et ce sera pour moi un réel plaisir de continuer de travailler avec elle.

Je tiens également à remercier mon co-directeur de thèse, monsieur Amor Keziou pour sa patience, son orientation et ses conseils durant mon séjour à l'université de Reims Champagne-Ardenne. J'espère que nous continuons de développer les thèmes de recherche que nous avons entamés.

J'adresse mes vifs remerciements à madame Nahima Nemouchi d'avoir accepté de présider le jury, et pour tous les bons moments que nous avons partagés au cours de mes années de thèse.

Merci infiniment à madame O. Sadki, à monsieur A. Laksaci et à madame S. Belaloui, de me faire l'honneur de faire partie du jury, et pour le temps accordé à la lecture attentive de la thèse.

Enfin, je remercie monsieur Nadji Rahmania pour son accueil chaleureux à l'université de Lille 1, et tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.

Articles

1. Kitouni, A., Boukeloua, M. et Messaci, F. (2015). Rate of strong consistency for nonparametric estimators based on twice censored data. *Statistics and Probability Letters*, 96, 255–261.
2. Boukeloua, M. (2015). Rates of mean square convergence of density and failure rate estimators under twice censoring. *Statistics and Probability Letters*, 106, 121–128.
3. Boukeloua, M. et Messaci, F. (2016). Asymptotic normality of kernel estimators based upon incomplete data. *Journal of Nonparametric Statistics*, 28(3), 469–486.

Communications

1. Boukeloua, M., Kitouni, A., et Messaci, F. Taux de convergence presque complète d'estimateurs non paramétriques basés sur des données doublement censurées. JSB2014. Biskra, mai 2014.
2. Boukeloua, M., et Messaci, F. Convergence en moyenne quadratique d'estimateurs de la densité et du taux de hasard dans un modèle de censure mixte. CMA2014. Tlemcen, mai 2014.
3. Boukeloua, M., et Messaci, F. Estimation de la fonction de densité et du taux de hasard dans un modèle de censure mixte. CISAA2014. Constantine, 30 novembre et 1 décembre 2014.
4. Boukeloua, M., et Messaci, F. Normalité asymptotique d'estimateurs à noyau de la densité et du taux de hasard pour des données censurées. JDS SFDS 2015. Lille, juin 2015.

Table des matières

Introduction	3
1 Estimation non paramétrique dans un modèle de censure à droite	9
1.1 Le cas des données complètes	10
1.1.1 Estimation de la fonction de répartition	10
1.1.2 Estimateur à noyau de la densité	11
1.2 Le cas des données censurées à droite	14
1.2.1 L'estimateur de Kaplan-Meier	14
1.2.2 Estimation de la densité	17
1.2.3 Estimation du taux de hasard	18
2 Estimation non paramétrique pour des données doublement censurées	20
2.1 Le modèle de Turnbull (1974)	21
2.1.1 Les estimateurs self-consistants	21
2.1.2 Estimation de la densité et du taux de hasard	23
2.2 Le modèle de Patilea et Rolin (2006)	24
2.2.1 L'estimateur de Patilea et Rolin	25
2.2.2 Estimation de la densité et du taux de hasard	26

3	Normalité asymptotique des estimateurs à noyau basés sur des données censurées	28
3.1	Le résultat principal	29
3.2	Application à des cas particuliers	36
3.2.1	Le modèle de censure double	36
3.2.2	Le modèle de censure mixte	41
3.3	Etude de simulation	46
3.3.1	Simulation des estimateurs étudiés	47
3.3.2	La normalité asymptotique	51
4	Convergence en moyenne quadratique d'estimateurs à noyau dans le cas de la censure mixte	57
4.1	Un Lemme préliminaire	58
4.2	Résultats et preuves	61
4.3	Utilisation des noyaux d'ordre infini	69
4.4	Etude de simulation	72
5	Utilisation des divergences dans l'étude de modèles semi-paramétriques avec des données censurées	77
5.1	Introduction	77
5.2	La mesure empirique de Kaplan-Meier	80
5.3	Les estimateurs du minimum de divergence	82
5.4	Les propriétés asymptotiques des estimateurs étudiés	89
5.5	Etude de simulation	99
	Perspectives	103
	Bibliographie	104

Introduction

L'objet de la statistique non paramétrique est d'étudier des phénomènes dont la loi de probabilité nous est inconnue sans supposer que cette loi appartient à une des familles de modèles théoriques dépendant d'un nombre fini de paramètres (comme le modèle gaussien ou exponentiel) qui relève de la statistique paramétrique. En fait, dans certaines situations concrètes, rien ne permet de faire un choix raisonnable. Bien que la fonction de répartition détermine entièrement la loi d'une variable aléatoire, elle ne donne pas d'information concernant la symétrie, la multimodalité ou l'aplatissement de la loi. C'est la fonction de densité qui permet de mieux visualiser la loi, d'où l'intérêt de son estimation. Il existe plusieurs méthodes pour estimer la densité et nous nous intéressons dans cette thèse à l'estimation par la méthode du noyau. L'estimateur à noyau de la densité a été introduit par Rosenblatt (1956), il s'écrit comme un produit de convolution entre un noyau K convenablement rééchelonné et la fonction de répartition empirique. Cet estimateur a été largement étudié par la suite. Nous citons à titre d'exemple les travaux de Parzen (1962) qui a montré sa convergence en moyenne quadratique, Rosenblatt (1971) qui a montré sa normalité asymptotique et Silverman (1978) qui a montré sa convergence presque sûre. En dehors du fait qu'il présente lui même de l'intérêt, il peut être utilisé pour estimer la fonction du taux de hasard, fonction très employée dans l'analyse de survie et en fiabilité. L'estimation non paramétrique du taux de hasard a été introduite pour la première

fois par Watson et Leadbetter (1964).

Mais dans la pratique, il arrive qu'un phénomène de censure empêche l'observation complète de la variable d'intérêt. Par exemple quand on s'intéresse au temps de survie à une maladie grave, la fixation du temps de l'étude va introduire une censure à droite. En effet, à la fin de l'étude, il est possible que certains malades soient encore vivants (heureusement pour eux !). Mais le statisticien ne disposera que de l'information partielle que leurs temps de survie dépassent les valeurs observées. Les estimateurs précédents ne sont plus valables dans un tel modèle, aussi Kaplan et Meier (1958) ont proposé un nouvel estimateur de la fonction de survie (complément à 1 de la fonction de répartition). En se basant sur cet estimateur, Földes et al. (1981) ont introduit un estimateur à noyau de la densité qui généralise l'estimateur de Rosenblatt. Ils ont remplacé dans l'expression de ce dernier, la fonction de répartition empirique par l'estimateur de Kaplan-Meier. L'estimateur de Földes et al. (1981) a fait l'objet de plusieurs travaux comme ceux de Mielniczuk (1986) et Diehl et Stute (1988) qui se sont intéressés à sa convergence presque sûre et en probabilité et à sa normalité asymptotique et Kagba (2004) qui a montré sa convergence en moyenne quadratique. Quant à l'estimation du taux de hasard, elle a été considérée par plusieurs auteurs dans le cadre de la censure à droite, comme Nelson (1969, 1972), Aalen (1978) et Xiang (1994).

Il est vrai que la censure à droite est la plus fréquente dans la pratique, mais un phénomène de censure à gauche peut aussi intervenir dans certaines situations. Les données censurées à gauche correspondent au cas où nous n'observons pas la variable d'intérêt mais nous savons seulement qu'elle est inférieure à la variable de censure. Généralement, la censure à gauche s'accompagne de la censure à droite ce qui donne des données doublement censurées. Parmi les modèles qui traitent ce genre de données deux ont été largement étudiés dans la littérature statistique, à savoir le modèle de Turnbull

(1974) connu par le modèle de censure double et le modèle de Patilea et Rolin (2006) que nous nommons par le modèle de censure mixte. Le premier modèle suppose que la variable de censure à gauche est plus petite que la variable de censure à droite. Ce modèle est très présent dans la pratique, nous renvoyons le lecteur à Turnbull (1974) et Chang et Yang (1987) pour des exemples de son application. Dans ce cadre, Turnbull (1974) a proposé des estimateurs pour les fonctions de survie des variables latentes, donnés par un système d'équations intégrales connues par les équations de self-consistance. Ensuite, Ren (1997) a proposé un estimateur de la densité généralisant les estimateurs à noyau connus dans les cas des données complètes et censurées à droite. En effet, il a remplacé la fonction de répartition empirique (ou l'estimateur de Kaplan-Meier) par l'estimateur de Turnbull. Ren (1997) a montré la convergence presque sûre et la normalité asymptotique pour cet estimateur ainsi que pour un estimateur du taux de hasard qu'il a lui même introduit.

Contrairement au modèle de Turnbull (1974), le modèle de Patilea et Rolin (2006) suppose que les variables latentes sont indépendantes. Des exemples d'application de ce modèle peuvent être trouvés dans Morales et al. (1991) et Patilea et Rolin (2006). Patilea et Rolin (2006) ont proposé un estimateur produit-limite de la fonction de survie qui généralise, comme l'estimateur de Turnbull, l'estimateur de Kaplan-Meier. Shen (2011, 2012) a proposé deux estimateurs alternatifs. Ensuite, Kitouni et al. (2015) ont utilisé l'estimateur de Patilea et Rolin pour introduire, par analogie aux cas précédents, des estimateurs à noyau de la densité et du taux de hasard pour lesquels ils ont prouvé la convergence presque complète uniforme avec taux.

Dans cette thèse, nous définissons un cadre général de censure qui englobe tous les modèles précédents et nous montrons la normalité asymptotique des estimateurs à noyau de la densité et du taux de hasard introduits. Cette étude unifiée est alors appliquée à trois cas de censure. Le premier, à savoir la censure à droite, a permis d'améliorer légèrement (en allégeant un peu les

hypothèses) le résultat établi en 1986 par Mielniczuk. Le second cas, consacré à la censure double a également permis d'améliorer, en termes d'hypothèses moins contraignantes, un résultat déjà connu et dû à Ren en 1997. Le dernier cas, qui concerne la censure mixte, est un résultat établi pour la première fois dans le cadre de cette thèse. Une étude de simulation a permis d'illustrer, d'une part, les bonnes performances des estimateurs étudiés pour une taille finie de l'échantillon (100) et a pu vérifier, d'autre part, la normalité asymptotique des estimateurs en question. Tout cela a été testé sur des modèles très utiles en pratique : la loi de Weibull et la loi de Berthelon qui contrairement à la première tient compte aussi bien des défaillances accidentelles que celles dues au vieillissement et qui a été introduite en 2001.

Ensuite, nous montrons la convergence en moyenne quadratique des mêmes estimateurs dans le cas de la censure mixte. Quoique la censure complique le calcul de la variance, des idées développées pour l'établissement de la convergence presque complète dans l'article de Kitouni et al. (2015) ont été exploitées avec succès. Dans la première partie, le noyau utilisé était d'ordre fini, mais il est connu que pour des densités infiniment dérivables les noyaux d'ordre infini sont plus adéquats. C'est la raison pour laquelle, des taux de convergence en moyenne quadratique ont été donnés dans ce cas et se sont avérés, sans surprise, meilleurs que les premiers. Finalement, une étude de simulation a permis d'appuyer les résultats théoriques pour le modèle de censure mixte et de laisser penser à la possibilité d'avoir des résultats théoriques semblables pour le modèle de censure double.

Dans une deuxième partie de la thèse, nous nous intéressons à des modèles semi-paramétriques définis par un nombre fini de contraintes linéaires avec un paramètre inconnu. Ce genre de modèles est très présent dans la pratique et particulièrement en économétrie. La méthode la plus utilisée pour traiter ces modèles est la méthode de la vraisemblance empirique (EL) introduite par Owen (1988, 1990). Plusieurs auteurs ont travaillé dans ce cadre comme Qin

et Lawless (1994) et Corcoran (1998). Pour les données censurées à droite, il existe plusieurs généralisations de la méthode EL. Nous citons par exemple les travaux de Wang et Jing (2001), Ren (2001) et Pan et Zhou (2002). Par ailleurs, Broniatowski et Keziou (2012) ont utilisé la théorie des divergences développée dans Broniatowski et Keziou (2006) pour traiter les modèles de contraintes linéaires dans le cas des données complètes. Leur approche généralise la méthode EL. Dans la présente thèse, nous initions l'extension de leurs résultats au cas des données censurées à droite. Nous construisons des estimateurs pour le paramètre d'intérêt et nous étudions leurs propriétés asymptotiques. Nous comparons également notre approche avec la méthode EL de Wang et Jing (2001) à l'aide d'une étude de simulation. Les résultats de cette dernière conjugués avec des résultats de convergence montrent que notre approche améliore la méthode EL qu'ils proposent.

Organisation de la thèse

Cette thèse est organisée comme suit.

Chapitre 1 : Nous y présentons le modèle de censure à droite ainsi que les estimateurs de la fonction de survie, de la densité et du taux de hasard pour ce modèle. Nous rappelons également quelques résultats de convergence de ces estimateurs que nous utilisons dans les chapitres qui suivent.

Chapitre 2 : Il est consacré aux modèles de censure double et mixte. Nous introduisons les estimateurs d'intérêt (des fonctions de répartition, de densité et de hasard) pour tous ces modèles et nous rappelons quelques résultats concernant leurs propriétés asymptotiques.

Chapitre 3 : Nous y considérons un cadre général de censure qui englobe tous les cas traités dans les chapitres précédents. Nous commençons par établir la normalité asymptotique d'un estimateur à noyau

que nous introduisons, ensuite nous appliquons ce résultat pour déduire la normalité asymptotique des estimateurs de la densité et du taux de hasard pour les modèles de censure double, de censure mixte et de censure à droite. Nous terminons par une étude de simulation qui a pour but d'illustrer les performances des estimateurs étudiés ainsi que leur comportement gaussien.

Chapitre 4 : Nous y montrons la convergence en moyenne quadratique, avec taux, des estimateurs de la densité et du taux de hasard dans le cas de la censure mixte. Nous présentons également une étude de simulation afin de conforter les résultats théoriques obtenus.

Chapitre 5 : Il est voué à l'étude des modèles semi-paramétriques vérifiant des contraintes linéaires, avec des données censurées à droite. Nous appliquons la théorie des divergences et de la dualité pour proposer des estimateurs pour le paramètre d'intérêt. Nous terminons par l'illustration des performances des méthodes proposées par une étude de simulation.

Chapitre 1

Estimation non paramétrique dans un modèle de censure à droite

Soit X une v.a.r. définie sur un espace de probabilité (Ω, \mathcal{A}, P) , de fonction de répartition F et de densité de probabilité f . Pour estimer ces caractéristiques fonctionnelles de la loi de X , nous nous basons sur un échantillon de données qui peuvent être complètes ou bien censurées à droite et/ou à gauche. Le but de ce chapitre est de rappeler les principales méthodes et résultats de l'estimation non paramétrique pour des données censurées à droite. Mais comme ces méthodes sont généralement inspirées de celles connues dans le cas des données complètes, nous commençons d'abord par regarder ce cas.

1.1 Le cas des données complètes

1.1.1 Estimation de la fonction de répartition

Dans ce cas on dispose d'un échantillon de n v.a.i.i.d. X_1, \dots, X_n de même loi que X . Un estimateur naturel de F est la fonction de répartition empirique, définie pour tout $x \in \mathbb{R}$ par :

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

L'étude asymptotique de cet estimateur a beaucoup suscité l'intérêt des statisticiens. Glivenko (1933) et Cantelli (1933) ont montré sa convergence presque sûre uniforme. La vitesse de cette convergence a été précisée par les lois du logarithme itéré unidimensionnelle de Chung (1949) et multidimensionnelle de Kiefer (1961). Un autre résultat fondamental sur la loi asymptotique de F_n^C est le test de Kolmogorov-Smirnov (voir Kolmogorov (1933); Smirnov (1939)).

Dans la suite nous allons rappeler quelques bornes exponentielles pour F_n^C , que nous utiliserons ultérieurement.

Dvoretzky et al. (1956) ont donnée la borne suivante appelée la borne DKW.

Théorème 1. *Il existe une constante $D < +\infty$ tel que pour tout $u > 0$, nous avons*

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

Massart (1990) a montré que ce Théorème est vrai pour $D = 2$ et que cette constante est la meilleure que l'on peut obtenir. En effet, Kolmogorov (1933) a montré que si F est continue, on a

$$\lim_{n \rightarrow \infty} P \left(\sqrt{n} \sup_{x \in \mathbb{R}} |F_n^C(x) - F(x)| > u \right) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2 u^2),$$

et lorsque u augmente le membre droit est équivalent à $2 \exp(-2u^2)$.

Si $X = (X^{(1)}, X^{(2)}, \dots, X^{(m)})$ est à valeurs dans \mathbb{R}^m , sa fonction de répartition empirique est définie pour tout $x = (x^{(1)}, x^{(2)}, \dots, x^{(m)}) \in \mathbb{R}^m$ comme suit

$$F_{n,m}^C(x) := \frac{1}{n} \sum_{i=1}^n 1_{\{X_i^{(1)} \leq x^{(1)}, X_i^{(2)} \leq x^{(2)}, \dots, X_i^{(m)} \leq x^{(m)}\}}.$$

Kiefer (1961) a donné la borne exponentielle suivante pour $F_{n,m}^C$.

Théorème 2. *Pour tous m et $\varepsilon > 0$, il existe une constante $D(\varepsilon, m)$ telle que pour tous F , n et $u \geq 0$, nous avons*

$$P \left(\sqrt{n} \sup_{x \in \mathbb{R}} |F_{n,m}^C(x) - F(x)| \geq u \right) \leq D \exp(-(2 - \varepsilon)u^2).$$

Ces bornes exponentielles permettent d'obtenir, de façon directe, la convergence presque complète avec un taux de convergence de l'ordre de $\sqrt{\log n/n}$ (voir Kitouni et al. (2015)). Quant à la convergence faible, elle est donnée par le Théorème de Donsker suivant.

Théorème 3. *Le processus empirique $\sqrt{n}(F_n^C - F)$ converge faiblement vers un processus gaussien centré de fonction de covariance donnée par*

$$\Gamma(s, t) := F(s)(1 - F(t)) \text{ pour } s \leq t.$$

Démonstration.

Voir Pollard (1984) page 97. □

1.1.2 Estimateur à noyau de la densité

Bien que la fonction de répartition détermine entièrement la loi d'une variable aléatoire, elle ne donne pas d'information concernant la symétrie,

la multimodalité ou l'aplatissement de la loi. C'est la fonction de densité qui permet de mieux visualiser la loi, d'où l'intérêt du problème de son estimation. Une solution intuitive à ce problème a été proposée par Rosenblatt (1956). Pour $h > 0$ assez petit on a :

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

en remplaçant F par son estimateur F_n^C , nous obtenons l'estimateur suivant de f :

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

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

Ceci mène à généraliser cet estimateur comme suit

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

où $K : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction intégrable, telle que $\int K(u)du = 1$, c'est l'estimateur à noyau de f . La fonction K est le noyau et le paramètre h est la fenêtre.

Voici quelques exemples de noyaux usuels :

- $K(u) = \frac{1}{2}1_{\{|u| \leq 1\}}$: noyau rectangulaire,
- $K(u) = (1 - |u|)1_{\{|u| \leq 1\}}$: noyau triangulaire,
- $K(u) = \frac{3}{4}(1 - u^2)1_{\{|u| \leq 1\}}$: noyau d'Epanechnikov ou parabolique,
- $K(u) = \frac{15}{16}(1 - u^2)^2 1_{\{|u| \leq 1\}}$: noyau quadratique,
- $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)$: noyau gaussien,
- $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-|u|/\sqrt{2}) \sin(|u|/\sqrt{2} + \pi/4)$: noyau de Silverman.

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

L'estimateur f_n^C a été largement étudié dans la littérature. Sa convergence uniforme faible et forte a été considérée par plusieurs auteurs comme Schuster (1969), Van Ryzin (1969), Rosenblatt (1971) et Silverman (1978). La loi du logarithme itéré a été établie par Deheuvels (1991). Par ailleurs, Parzen (1962) a montré sa convergence en moyenne quadratique en appliquant le théorème de Bochner suivant.

Théorème 4. *Soient g et K deux fonctions intégrables, avec K bornée et $|z|K(z) \xrightarrow{|z| \rightarrow \infty} 0$, et posons : $g_n(x) := \frac{1}{h_n} \int K(\frac{z}{h_n})g(x-z) dz$, où $(h_n)_{n \geq 1}$ est une suite de nombres réels strictement positifs qui converge vers 0.*

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

Démonstration.

Voir Bochner (1955). □

En effet, en supposant (en plus des hypothèses du théorème de Bochner) que $nh_n \xrightarrow{n \rightarrow \infty} \infty$, nous obtenons en vertu du théorème de Bochner

$$\begin{aligned} Ef_n^C(x) &= \frac{1}{h_n} EK \left(\frac{x-X}{h_n} \right) \\ &= \frac{1}{h_n} \int K \left(\frac{x-y}{h_n} \right) f(y) dy \\ &= \frac{1}{h_n} \int K \left(\frac{z}{h_n} \right) f(x-z) dz \xrightarrow{n \rightarrow \infty} f(x), \end{aligned}$$

et

$$\begin{aligned} Var f_n^C(x) &= \frac{1}{nh_n^2} Var K \left(\frac{x-X}{h_n} \right) \\ &\leq \frac{1}{nh_n^2} EK^2 \left(\frac{x-X}{h_n} \right) \\ &= \frac{1}{nh_n} \left[\frac{1}{h_n} \int K^2 \left(\frac{z}{h_n} \right) f(x-z) dz \right] \xrightarrow{n \rightarrow \infty} 0 \times f(x) \int K^2(z) dz = 0. \end{aligned}$$

1.2 Le cas des données censurées à droite

Dans la pratique, il n'est pas toujours possible de disposer d'un échantillon de données complètes. Une variable de censure R peut empêcher l'observation de la vraie variable d'intérêt et ne nous fournit alors qu'une information partielle sur elle. Il existe plusieurs types de censure dont la censure à droite, qui est l'objet d'étude dans cette section. Nous parlons de donnée censurée à droite lorsque nous observons la censure R et non pas la variable d'intérêt X , et que nous savons que $X > R$. Donc dans un tel modèle, nous n'observons la variable d'intérêt que lorsqu'elle est inférieure à la variable de censure, et par conséquent, les observations sont des répliques du couple $(Z := X \wedge R, \Delta := 1_{\{X \leq R\}})$ où Δ est l'indicateur de censure valant 1 quand l'observation est complète et 0 quand elle est censurée. Ce modèle est le plus fréquent en pratique, il est par exemple adapté au cas où l'événement d'intérêt est le temps de survie à une maladie et où la date de fin de l'étude est préalablement fixée ; les patients vivants à la fin de l'étude fournissent des données censurées à droite. La censure n'est pas nécessairement fixe, elle peut être aléatoire, c'est le cas par exemple d'un individu perdu de vue ou mort dans un accident au cours de l'étude.

1.2.1 L'estimateur de Kaplan-Meier

Dans toute la suite, pour toute variable aléatoire U , nous notons F_U sa fonction de répartition et $S_U := 1 - F_U$ sa fonction de survie. De plus, on définit le point initial du support de U , noté I_U , par : $I_U := \inf\{t \in \mathbb{R}/F_U(t) > 0\}$, et le point terminal du support de U , noté T_U par : $T_U := \sup\{t \in \mathbb{R}/F_U(t) < 1\}$. I_U et T_U possèdent les propriétés suivantes :

i) $I_U \leq U \leq T_U$ p.s.

ii) Si V est une v.a.r. indépendante de U , on a : $I_{U \wedge V} = I_U \wedge I_V$, $I_{U \vee V} = I_U \vee I_V$, $T_{U \wedge V} = T_U \wedge T_V$ et $T_{U \vee V} = T_U \vee T_V$.

Nous supposons que les variables aléatoires X et R sont positives et indépendantes. A partir de l'échantillon de n couple de v.a.i.i.d. $(Z_i := X_i \wedge R_i, \Delta_i := 1_{\{X_i \leq R_i\}})_{1 \leq i \leq n}$, Kaplan et Meier (1958) ont construit l'estimateur produit-limite suivant pour F .

$$F_n^R(t) := 1 - \prod_{j/Z'_j \leq t} \left(1 - \frac{M(Z'_j)}{U(Z'_j)} \right), \quad \forall t \geq 0,$$

où

$(Z'_j)_{1 \leq j \leq m}$ ($m \leq n$) sont les valeurs distinctes des $(Z_i)_{1 \leq i \leq n}$ rangées dans l'ordre croissant,

$M(Z'_j) := \sum_{i=1}^n \Delta_i 1_{\{Z_i = Z'_j\}}$ est le nombre de morts exactes à l'instant Z'_j

et $U(Z'_j) := \sum_{i=1}^n 1_{\{Z_i \geq Z'_j\}}$ est le nombre d'individus à risque juste avant l'instant Z'_j .

Cet estimateur coïncide avec la fonction de répartition empirique quand il n'y a pas de données censurées. Il est donc naturel que les statisticiens se soient intéressés à étendre les résultats connus pour la fonction de répartition empirique au cas de l'estimateur de Kaplan-Meier. La convergence presque sûre uniforme et la loi du logarithme itéré ont été montrées respectivement par Winter et al. (1978) et Földes et Rejtő (1981a). Ensuite, Stute et Wang (1993) ont montré la loi forte des grands nombres dans le cas de la censure à droite, qui entraîne, entre autres, la convergence presque sûre de F_n^R . Quant à la convergence presque complète, elle a été montrée par Földes et al. (1980), avec un taux de convergence de l'ordre de $\sqrt{\log n / \sqrt{n}}$. Puis en imposant la continuité de F et de F_R , Földes et Rejtő (1981b) ont amélioré le taux de convergence qui est passé à l'ordre de $\sqrt{\log n / n}$. Ensuite, Kitouni et al. (2015) ont montré qu'on peut se passer de cette hypothèse de continuité en utilisant la borne exponentielle suivante.

Théorème 5. *Il existe une constante absolue D telle que, pour tout réel positif u ,*

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

Démonstration.

Voir Bitouzé et al. (1999). □

En ce qui concerne la convergence faible, Breslow et Crowley (1974) ont donné le résultat suivant.

Théorème 6. *Soit $T < \infty$ vérifiant $F_Z(T) < 1$ et supposons que F et F_R sont continues. Donc la fonction aléatoire $\sqrt{n}(F_n^R - F)$, pour $0 < t < T$, converge faiblement vers un processus gaussien centré, de fonction de covariance donnée par*

$$\Gamma(s, t) := S_X(s)S_X(t) \int_0^s \frac{dF}{S_X S_Z}, \quad s \leq t.$$

Gill (1983) a montré qu'on peut se passer de la continuité de F_R .

Nous énonçons également le résultat suivant, de convergence uniforme en probabilité, que nous utilisons dans la suite.

Théorème 7. *Supposons que $F(T_Z) < 1$ et soit $p \in]0, \frac{1}{2}[$. Nous avons*

$$\sup_{t \leq T_Z} |F_n^R(t) - F(t)| = O_P(n^{-p})$$

si et seulement si

$$\limsup_{t \rightarrow T_Z} \frac{\left(\int_t^{T_Z} S_R dF\right)^{1-p}}{S_R(t)} < \infty.$$

Démonstration.

Voir Chen et Lo (1997) Théorème 3.1. page 1064. □

1.2.2 Estimation de la densité

En s'inspirant de (1.1), Földes et al. (1981) ont proposé d'estimer f dans ce cas de censure à droite par

$$f_n^R(t) := \frac{1}{h_n} \int K\left(\frac{t-y}{h_n}\right) dF_n^R(y). \quad (1.2)$$

Földes et al. (1981) ont montré dans le même article la convergence presque complète ponctuelle et uniforme de f_n^R . Ensuite, plusieurs auteurs se sont intéressés à cet estimateur. Nous citons à titre d'exemple les travaux de Diehl et Stute (1988) et Karunamuni et Yang (1991) qui ont donné des taux de convergence uniforme presque sûre et en probabilité, Marron et Padgett (1987) qui ont proposé une méthode de validation croisée pour le choix de la fenêtre et Gannoun et Saracco (2002) qui ont proposé une nouvelle preuve de la convergence uniforme presque sûre de f_n^R .

Dans la suite, nous allons détailler quelques résultats de normalité asymptotique et de convergence en moyenne quadratique de f_n^R , car ce sont les modes de convergence d'intérêt dans cette thèse.

Concernant la normalité asymptotique, Mielniczuk (1986) a donné le résultat suivant.

Théorème 8. *Soit $x \in \mathbb{R}^+$ tel que $f(x)S_R(x) > 0$, sous les hypothèses*

- *f admet une dérivée seconde bornée au voisinage de x ,*
- *R admet une densité de probabilité continue au point x ,*
- *K est une densité bornée, paire et à support dans $[-1, 1]$,*
- *$h_n = o(n^{-1/3})$ et $nh_n \xrightarrow{n \rightarrow \infty} \infty$,*

nous avons

$$\sqrt{nh_n}(f_n^R(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_R(x)} \int K^2(y) dy\right).$$

(où $\xrightarrow{\mathcal{D}}$ désigne la convergence en loi).

Démonstration.

Voir Mielniczuk (1986) corollaire 3-ii) page 772. \square

Louani (1998) a montré un résultat plus général qui concerne la normalité asymptotique de f_n^R ainsi que ses dérivés.

Quant à la convergence en moyenne quadratique, Kagba (2004) a montré le résultat suivant.

Théorème 9. *Soit $x \in \mathbb{R}^+$, sous les hypothèses*

- *f et F_R sont deux fois différentiables et ayant des dérivées d'ordre 3 bornées au voisinage de x ,*
- *K est une densité paire à support compact et à variation bornée,*
- *$h_n \xrightarrow[n \rightarrow \infty]{} 0$ et $nh_n \xrightarrow[n \rightarrow \infty]{} \infty$,*

nous avons

$$\begin{aligned} MSE(f_n^R(x)) &= \left[\frac{f''(x)}{4} \left(\int y^2 K(y) dy \right)^2 \right] h_n^4 + \left[\frac{f(x)}{S_R(x)} \int K^2(y) dy \right] \times \frac{1}{nh_n} \\ &+ \frac{f^2(x)}{n} \left[\int_0^x \frac{f(r)}{S_X^2(r)S_R(r)} dr - \frac{1}{S_X(x)S_R(x)} \right] + O(h_n^6) + O(n^{-1}h_n) + o((nh_n)^{-1}). \end{aligned}$$

Démonstration. Voir Kagba (2004) Théorème 2.1. page 30. \square

1.2.3 Estimation du taux de hasard

Le taux de hasard de X est défini par

$$\lambda(t) := \begin{cases} \frac{f(t)}{S_X(t)} & \text{si } S_X(t) \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

Si X est le temps écoulé avant la survenue d'un certain événement, alors $\lambda(t)$ représente la probabilité que cet événement survienne à l'instant t sachant qu'il n'est pas arrivé avant. Par exemple si X est le temps de survie à une maladie grave, $\lambda(t)$ représente la probabilité qu'un individu meurt à l'instant t sachant qu'il était vivant juste avant. En fiabilité, $\lambda(t)$ représente la

probabilité qu'un appareil tombe en panne à l'instant t sachant qu'il fonctionnait juste avant ; on l'appelle dans ce cas taux de panne ou bien taux de défaillance.

Plus généralement, si X est une v.a.r. (pas nécessairement absolument continue), la mesure de hasard de X est définie par

$$d\Lambda(t) := \frac{dF(t)}{S_X(t^-)},$$

où $h(t^-)$ désigne la limite à gauche (lorsqu'elle existe) de h au point t pour toute fonction h .

On définit également la fonction de hasard cumulé de X par

$$\Lambda(t) := \int_0^t d\Lambda(s).$$

Si X est absolument continue alors

$$d\Lambda(t) = \lambda(t)dt \text{ et } \Lambda(t) = \int_0^t \lambda(s) ds.$$

Le taux de hasard et la fonction de hasard cumulé caractérisent la loi de probabilité de X , car elles sont liées à f et S_X par les relations suivantes

$$S_X(t) = \exp(-\Lambda(t)),$$

$$f(t) = \lambda(t) \exp(-\Lambda(t))$$

$$\text{et } \lambda(t) = (-\log S_X(t))'$$

et ceci pour tout t tel que $S_X(t) \neq 0$.

D'après la définition de λ , il s'avère naturel de l'estimer par

$$\lambda_n^R(t) := \frac{f_n^R(t)}{1 - F_n^R(t)} 1_{\{F_n^R(t) \neq 1\}}.$$

Plusieurs auteurs ont étudié cet estimateur ainsi que d'autres types d'estimateurs de λ . Nous citons par exemple les travaux de Földes et al. (1981), Tanner et Wong (1983), Diehl et Stute (1988), Muller et Wang (1994) et Xiang (1994).

Chapitre 2

Estimation non paramétrique pour des données doublement censurées

On parle de censure à gauche lorsque la variable d'intérêt n'est observée que si elle est supérieure à une certaine variable de censure. Dans la pratique, ce mode de censure accompagne généralement la censure à droite ce qui donne des échantillons censurés à droite et à gauche au même temps. Il existe plusieurs modèles non paramétriques qui traitent ce genre de données, comme les modèles de Peto (1973), Turnbull (1974), Samuelsen (1989), Huang (1999) et Patilea et Rolin (2001, 2004, 2006). Nous allons nous intéresser à deux de ces modèles qui sont très utilisés dans la littérature, à savoir le modèle de Turnbull (1974) connu par le modèle de censure double et le modèle de Patilea et Rolin (2006) que nous appelons par le modèle de censure mixte.

2.1 Le modèle de Turnbull (1974)

Dans ce modèle, la variable d'intérêt X est censurée à droite par une variable R et $\min(X, R)$ est lui-même censuré à gauche par une variable L . De plus, X est indépendante du couple (L, R) , $P(0 \leq L \leq R) = 1$ et nous n'observons que le couple (Z, Δ) où $Z := \max(\min(X, R), L)$ et

$$\Delta := \begin{cases} 0 & \text{si } L < X \leq R, \\ 1 & \text{si } X > R, \\ 2 & \text{si } X \leq L. \end{cases}$$

Δ prend la valeur 0 lorsque $Z = X$ (donnée complète) et la valeur 1 lorsque $Z = R$ (donnée censurée à droite) et la valeur 2 lorsque $Z = L$ (donnée censurée à gauche).

Un exemple pratique qui correspond à ce modèle est l'étude de la précocité des enfants kényens menée par Leiderman et al. (1973). Ils disposaient d'un échantillon de 65 enfants nés entre le 1 juillet et le 31 décembre 1969, et ils commençaient en janvier 1970 à vérifier chaque mois si les enfants ont appris à accomplir certaines tâches standard. La variable d'intérêt est donc l'âge d'apprentissage de ces tâches. Les enfants qui n'arrivent pas à accomplir une tâche à la fin de l'étude représentent des données censurées à droite, et les enfants qui maîtrisent une tâche avant d'entrer dans l'étude représentent des données censurées à gauche.

2.1.1 Les estimateurs self-consistants

Turnbull (1974) a proposé des estimateurs non paramétriques pour S_X , S_R et S_L notés respectivement par S_n^D , $S_R^{(n)}$ et $S_L^{(n)}$. L'inconvénient de ces estimateurs est le fait qu'il n'existe pas de formules explicites connues pour eux, mais ils sont définis comme les solutions d'équations intégrales appelées les équations de self-consistance. Ces équations sont données dans Ren (1997)

pour tout $t \geq 0$ par

$$S_n^D(t) = Q^{(n)}(t) - \int_{\{u \leq t\}} \frac{S_n^D(t)}{S_n^D(u)} dQ_1^{(n)}(u) + \int_{\{t < u\}} \frac{1 - S_n^D(t)}{1 - S_n^D(u)} dQ_2^{(n)}(u),$$

$$S_R^{(n)}(t) = 1 + \int_{\{u \leq t\}} \frac{dQ_1^{(n)}(u)}{S_n^D(u)}, \quad t < B_n,$$

$$S_L^{(n)}(t) = - \int_{\{t < u\}} \frac{dQ_2^{(n)}(u)}{1 - S_n^D(u)}, \quad t \geq A_n,$$

où

$$Q_j^{(n)}(t) := \frac{1}{n} \sum_{i=1}^n 1_{\{Z_i > t, \Delta_i = j\}}, \quad j \in \{0, 1, 2\},$$

$$Q^{(n)} := \sum_{j=0}^2 Q_j^{(n)},$$

$$A_n := \min\{Z_i / S_n^D(Z_i) < 1\}$$

$$\text{et } B_n := \max\{Z_i / S_n^D(Z_i^-) > 0\}.$$

Les propriétés asymptotiques de S_n^D (convergence presque sûre, convergence faible, efficacité asymptotique,...) ont fait l'objet de divers travaux, parmi lesquels ceux de Tsai et Crowley (1985), Chang et Yang (1987), Chang (1990), Gu et Zhang (1993) et Bih-Sheue et Cheun-Der (2004).

Nous allons expliciter deux de ces résultats dont nous aurons besoin dans la suite. Pour cela considérons les hypothèses suivantes.

D1 $\forall t \geq 0, S_R(t) - S_L(t) > 0.$

D2 S_X, S_R et S_L sont des fonctions continues de t pour $t \geq 0$, et $0 < S_X(t) < 1$ pour $t > 0$.

D3 $S_X(0) = S_R(0) = 1$ et $\lim_{u \rightarrow \infty} S_X(u) = \lim_{u \rightarrow \infty} S_R(u) = \lim_{u \rightarrow \infty} S_L(u) = 0$.

D4 il existe α et $\beta, 0 < \alpha < \beta < \infty$, tels que $P(L \in]0, \alpha]) = 0$ et $P(L \leq \beta) = 1$.

Théorème 10. *Sous les hypothèses **D1** et **D2** nous avons*

$$\sup_{t \geq 0} |S_n^D(t) - S_X(t)| \xrightarrow{p.s.} 0,$$

$$\sup_{t \geq 0} |S_R^{(n)}(t) - S_R(t)| \xrightarrow{p.s.} 0$$

et $\sup_{t \geq 0} |S_L^{(n)}(t) - S_L(t)| \xrightarrow{p.s.} 0.$

Démonstration.

Voir Chang et Yang (1987) Théorème 4.2 page 1546. □

Notons $D[0, \beta]$ l'espace des fonctions réelles définies sur $[0, \beta]$, continues à droite et ayant des limites à gauche en tout point de $[0, \beta]$ (les fonctions cadlag). Le résultat suivant concerne la convergence faible des estimateurs $S_n^D, S_R^{(n)}$ et $S_L^{(n)}$.

Théorème 11. *Sous les hypothèses **D1–D4**, le processus $u^{(n)} = \sqrt{n}(S_n^D - S_X, S_R^{(n)} - S_R, S_L^{(n)} - S_L)$ converge faiblement vers un processus gaussien sur $D[0, \beta] \times D[0, \beta] \times D[0, \beta]$.*

Démonstration.

Voir Chang (1990) Théorème 3.1. page 399. □

2.1.2 Estimation de la densité et du taux de hasard

Par analogie aux cas des données complètes et censurées à droite, Ren (1997) a proposé d'estimer la densité f de X par

$$f_n^D(t) := \frac{1}{h_n} \int K\left(\frac{t-y}{h_n}\right) dF_n^D(y), \quad (2.1)$$

où $F_n^D := 1 - S_n^D$.

Ren (1997) a montré la convergence presque sûre uniforme et la normalité asymptotique de f_n^D . Dans le chapitre suivant, nous montrons un résultat de normalité asymptotique pour l'estimateur à noyau de la densité dans un cadre général de censure. Notre résultat entraîne, entre autres, la normalité asymptotique de f_n^D sous des conditions légèrement différentes de ceux imposées par Ren (1997). Rappelons alors le résultat de Ren (1997) pour pouvoir le comparer au nôtre.

Théorème 12. Soit $x \in \mathbb{R}^+$, sous les hypothèses **D1–D4** et si

- f est bornée, $f(x) > 0$ et au voisinage de x , la dérivée seconde de $f(S_R - S_L)$ existe et elle est bornée,
- K est une densité paire, bornée et à support dans $[-1, 1]$,
- $nh_n^3 \xrightarrow[n \rightarrow \infty]{} 0$ et $nh_n \xrightarrow[n \rightarrow \infty]{} \infty$,

nous avons

$$\sqrt{nh_n}(f_n^D(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_R(x) - S_L(x)} \int K^2(y) dy\right).$$

Démonstration.

Voir Ren (1997) Théorème 2-(ii) page 35. □

Ren (1997) a également étudié l'estimateur suivant du taux de hasard λ

$$\lambda_n^D(t) := \frac{f_n^D(t)}{S_n^D(t)},$$

il a montré que sous les mêmes hypothèses du Théorème 12 ci-dessus

$$\sqrt{nh_n}(\lambda_n^D(x) - \lambda(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_X^2(x)(S_R(x) - S_L(x))} \int K^2(y) dy\right).$$

(Voir Ren (1997) Corollaire page 35).

2.2 Le modèle de Patilea et Rolin (2006)

Patilea et Rolin (2006) ont étudié deux modèles de censure mixte. Nous nous intéressons au modèle I où la variable d'intérêt X est censurée à droite par une variable R , $\min(X, R)$ est censuré à gauche par une variable L et les variables X , L et R sont positives et indépendantes. L'observation disponible consiste en n répliques indépendantes du couple (Z, Δ) où $Z :=$

$\max(\min(X, R), L)$ et Δ est l'indicateur de censure défini par

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

Un exemple d'application de ce modèle donné par Patilea et Rolin (2006) consiste à considérer un système de trois composants C_1 , C_2 et C_3 , avec C_1 et C_2 en série et C_3 en parallèle avec le système composé de C_1 et C_2 . Les variables X , R et L représentent les temps de survie de C_1 , C_2 et C_3 respectivement, et nous pouvons déterminer quel composant tombe en panne au même temps que le système. Donc au lieu d'observer X , on peut seulement observer le couple (Z, Δ) .

2.2.1 L'estimateur de Patilea et Rolin

Patilea et Rolin (2006) ont proposé un estimateur produit limite de F s'écrivant en supposant qu'il n'y a pas d'ex-æquo parmi les X_i , ce qui est naturel dans notre situation d'existence de la densité, comme suit

$$F_n^T(t) = 1 - S_n^T(t) := 1 - \prod_{k: Z_k \leq t} \left\{ 1 - \frac{1_{\{\Delta_k=0\}}}{n \left(F_L^{(n)}(Z_k^-) - F_Z^{(n)}(Z_k^-) \right)} \right\},$$

où $F_Z^{(n)}(t) := \frac{1}{n} \sum_{i=1}^n 1_{\{Z_i \leq t\}}$ est la version empirique de F_Z et

$$F_L^{(n)}(t) := \prod_{j: Z'_j > t} \left(1 - \frac{\sum_{i=1}^n 1_{\{Z_i=Z'_j, \Delta_i=2\}}}{n F_Z^{(n)}(Z'_j)} \right),$$

$((Z'_j)_{1 \leq j \leq m} \ (m \leq n)$ étant les valeurs distinctes des $(Z_i)_{1 \leq i \leq n}$) est l'estimateur produit limite de la fonction de répartition de L qui est censurée à gauche par $\min(X, R)$. Cet estimateur peut être déduit de celui de Kaplan-Meier en inversant le temps.

Patilea et Rolin (2006) ont montré la convergence presque sûre uniforme de $F_n^T(t)$, la vitesse de cette convergence a été précisée par la loi du logarithme itéré de Messaci et Nemouchi (2011, 2013). Kitouni et al. (2015) ont donné un taux de la convergence presque complète uniforme, de l'ordre de $\sqrt{\log n/n}$. Notons $H_k(t) := P(Z \leq t, \Delta = k)$, $k \in \{0, 1, 2\}$ les sous loi de Z et $I_{H_0} := \inf\{t \in \mathbb{R}/H_0(t) > 0\}$ le point initial du support de H_0 . Dans la suite nous aurons besoin des deux résultats suivants de la convergence faible.

Théorème 13. *Sous les hypothèses*

T1 $I_L \leq I_X$ et $T_X \leq T_R$,

T2 $\int_{\{u > I_{H_0}\}} \frac{dH_2(u)}{(F_Z(u))^2} < \infty$,

nous avons

- i) $\sqrt{n}(F_L^{(n)} - F_L)$ converge faiblement vers un processus gaussien centré dans $D[I_{H_0}, \infty]$,
- ii) $\sqrt{n}(S_n^T - S_X)$ converge faiblement vers un processus gaussien centré dans $D[0, \tau]$, où τ est tel que $I_{H_0} < \tau$ et $H_0(\tau) + H_1(\tau) < 1$.

Démonstration.

i) Voir Patilea et Rolin (2006) Lemme 7.2. page 935.

ii) Voir Patilea et Rolin (2006) Théorème 7.3. page 937.

Remarquons que $F_{n2}^I([0, t]) - F_2^I([0, t])$ (resp. $F_{nT}^I([0, t]) - F_T^I([0, t])$) de Patilea et Rolin (2006) n'est autre que $F_L^{(n)}(t) - F_L(t)$ (resp. $S_X(t) - S_n^T(t)$) pour $t > I_X$ (resp. sous l'hypothèse **T1**). \square

2.2.2 Estimation de la densité et du taux de hasard

Par analogie aux cas précédents, Kitouni et al. (2015) ont proposé l'estimateur suivant pour f

$$f_n^T(t) := \frac{1}{h_n} \int K\left(\frac{t-y}{h_n}\right) dF_n^T(y) \quad (2.2)$$

et ils ont montré sa convergence presque complète énoncée comme suit.

Théorème 14. (Kitouni, Boukeloua et Messaci 2015).

Soit C un compact inclus dans $[0, \min(T_X, T_R)[$, sous les hypothèses

- $\max(I_L, I_R) < I_X$,
- $h_n \rightarrow 0$ et $nh_n^2 / \log n \rightarrow \infty$,
- K est une fonction continue à droite, à variation bornée, à support compact et telle que $\int K(u) du = 1$,

nous avons

- i) S'il existe un entier $r \geq 2$ telle que f est r fois continûment différentiable autour de C et $\int u^j K(u) du = 0, \forall 1 \leq j \leq r - 1$, alors

$$\sup_{x \in C} |f_n^T(x) - f(x)| = O_{p.co.} \left(h_n^r + \frac{1}{h_n} \sqrt{\frac{\log n}{n}} \right).$$

- ii) S'il existe $\alpha, \beta, \varepsilon \in \mathbb{R}_+^*$ tels que $\forall x \in C, \forall y \in]x - \varepsilon, x + \varepsilon[, |f(x) - f(y)| \leq \alpha|x - y|^\beta$, alors

$$\sup_{x \in C} |f_n^T(x) - f(x)| = O_{p.co.} \left(h_n^\beta + \frac{1}{h_n} \sqrt{\frac{\log n}{n}} \right).$$

Kitouni et al. (2015) ont également montré des résultats similaires pour l'estimateur du taux de hasard suivant

$$\lambda_n^T(t) := \frac{f_n^T(t)}{S_n^T(t) + u_n},$$

où $u_n > 0$ est le terme général d'une suite qui converge vers zéro et qui sert à éviter la division par zéro.

Dans le chapitre suivant, nous montrons la normalité asymptotique de f_n^T et d'un estimateur du taux de hasard légèrement différent de λ_n^T .

Chapitre 3

Normalité asymptotique des estimateurs à noyau basés sur des données censurées

Dans les chapitres précédents, nous avons présenté des estimateurs à noyau de la densité pour des types différents de données censurées. Dans ce chapitre, nous considérons un cadre général de censure, qui englobe tous les cas précédents et nous proposons un estimateur de la densité qui généralise les estimateurs que nous avons déjà introduits. Ensuite, nous montrons la normalité asymptotique de cet estimateur qui va entraîner celle des estimateurs de la densité dans les cas de la censure à droite, la censure double et la censure mixte. Nous en déduisons également la normalité asymptotique des estimateurs du taux de hasard que nous avons introduits dans ces mêmes cas. Ce travail a fait l'objet d'une publication dans la revue "Journal of Nonparametric Statistics" (Boukeloua et Messaci (2016)).

3.1 Le résultat principal

Soit X une durée d'intérêt de fonction de répartition F et de densité de probabilité f . On suppose que X peut être censurée, donc au lieu d'observer un échantillon de X , nous disposons d'un échantillon $(Z_i, \Delta_i)_{1 \leq i \leq n}$ de n variables aléatoires indépendantes et de même loi que le couple (Z, Δ) , où $Z = X$ si et seulement si l'indicateur de censure Δ prend la valeur 0 (observation complète). En s'inspirant des estimateurs f_n^C , f_n^R , f_n^D et f_n^T (voir (1.1), (1.2), (2.1) et (2.2) respectivement), nous estimons f par

$$f_n(t) := \frac{1}{h_n} \int K\left(\frac{t-y}{h_n}\right) dF_n(y), \quad (3.1)$$

où F_n est un estimateur de F , continu à droite, à variation bornée et vérifiant d'autres conditions que nous précisons dans la suite (voir l'hypothèse **H4**), K est le noyau et h_n est la fenêtre.

Posons $\psi(t) := P(\Delta = 0 | X = t)$, nous avons besoin des hypothèses suivantes pour énoncer la normalité asymptotique de f_n en un point fixé x .

H1 f est dérivable au point x .

H2 ψ est continue au point x .

H3 $\exists a > 0$ tel que $\inf_{t \in [x-a, x+a]} \psi(t) > 0$.

H4 F_n est une fonction en escalier ayant des éventuels sauts uniquement

en $(Z_i)_{1 \leq i \leq n}$ et telle que $\sqrt{n} \max_{i: Z_i \in [x-a, x+a]} \left| n \Delta F_n(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| =$

$O_P(1)$ et $\sum_{i=1}^n \Delta F_n(Z_i) 1_{\{\Delta_i \neq 0\}} = o_P\left(\sqrt{h_n/n}\right)$.

Si nous disposons d'un échantillon de données complètes $(X_i)_{1 \leq i \leq n}$ de X alors $f_n(x) = f_n^C(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)$. Puisque $E\left(K\left(\frac{x-Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)}\right) = E\left(K\left(\frac{x-X_i}{h_n}\right)\right)$, il semble naturel de remplacer $K\left(\frac{x-X_i}{h_n}\right)$ par $K\left(\frac{x-Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)}$ quand X est censurée. Nous introduisons alors

$$\tilde{f}_n(x) := \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x-Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)}, \quad (3.2)$$

qui va être asymptotiquement normal (comme démontré dans la preuve du Théorème 15). L'hypothèse **H4** va permettre à l'estimateur défini dans (3.1) d'être proche en probabilité de \tilde{f}_n avec un taux adéquat.

Comme on va le montrer dans la section suivante, les hypothèses **H3** et **H4** sont vérifiées pour les modèles de censure à droite, de censure double et de censure mixte sous des conditions appropriées. Nous supposons en plus, quelques hypothèses standard, dans le cadre de la statistique non paramétrique, concernant K et h_n

H5 K est une densité bornée et à support compact sur $[-b, b]$.

H6 $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$, lorsque $n \rightarrow \infty$.

Notre résultat principal s'énonce comme suit.

Théorème 15. (*Boukeloua et Messaci 2016*) *Sous **H1-H6**, nous avons*

$$\sqrt{nh_n}(f_n(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{\psi(x)} \int K^2(z) dz\right).$$

Démonstration.

Rappelons que

$$\tilde{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)}.$$

Notons que sous **H5** et **H6**, on peut prendre $Z_i \in [x - a, x + a]$ pour n assez grand, donc $\psi(Z_i) \neq 0$ en vertu de **H3**.

Considérons maintenant la décomposition suivante

$$f_n(x) - f(x) = (f_n(x) - \tilde{f}_n(x)) + (\tilde{f}_n(x) - E\tilde{f}_n(x)) + (E\tilde{f}_n(x) - f(x)).$$

Nous allons étudier chacun de ces termes à part.

Etape 1 : Etude du terme $\tilde{f}_n(x) - E\tilde{f}_n(x)$.

Nous avons $\tilde{f}_n(x) - E\tilde{f}_n(x) = \sum_{i=1}^n \xi_{n,i}$, où

$$\xi_{n,i} := \frac{1}{nh_n} K\left(\frac{x - Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} - \frac{1}{nh_n} E\left(K\left(\frac{x - Z_i}{h_n}\right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)}\right). \quad (3.3)$$

Nous allons appliquer le Théorème de Lindeberg suivant.

Théorème 16. Soient $\xi_{n,1}, \xi_{n,2}, \dots, \xi_{n,n}$ des v.a.r. indépendantes, centrées et de carrés intégrables et notons

$$s_n^2 := \sum_{i=1}^n \text{Var}(\xi_{n,i}),$$

si pour tout $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{i=1}^n E \left(\xi_{n,i}^2 1_{\{|\xi_{n,i}| > \varepsilon s_n\}} \right) = 0, \quad (3.4)$$

alors

$$\frac{1}{s_n} \sum_{i=1}^n \xi_{n,i} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Démonstration.

Voir Billingsley (1995) Théorème 27.2. page 359. □

Les $\xi_{n,i}$ définies dans (3.3) sont bien indépendantes, centrées et de carrés intégrables, il reste à montrer la condition de Lindeberg (3.4) qui s'écrit dans ce cas comme suit

$$\lim_{n \rightarrow \infty} \frac{1}{\text{Var}(\tilde{f}_n(x))} \sum_{i=1}^n E \left(\xi_{n,i}^2 1_{\{|\xi_{n,i}| > \varepsilon \sqrt{\text{Var}(\tilde{f}_n(x))}\}} \right) = 0. \quad (3.5)$$

Pour n suffisamment grand, nous avons

$$|\xi_{n,i}| \leq \frac{2\|K\|_\infty}{\inf_{t \in [x-a, x+a]} \psi(t)} \times \frac{1}{nh_n},$$

où $\|K\|_\infty$ est la norme uniforme de K sur \mathbb{R} .

Donc

$$\begin{aligned}
& \frac{1}{\text{Var}(\tilde{f}_n(x))} \sum_{i=1}^n E \left(\xi_{n,i}^2 1_{\{|\xi_{n,i}| > \varepsilon \sqrt{\text{Var}(\tilde{f}_n(x))}\}} \right) \\
&= \frac{n}{\text{Var}(\tilde{f}_n(x))} E \left(\xi_{n,1}^2 1_{\{|\xi_{n,1}| > \varepsilon \sqrt{\text{Var}(\tilde{f}_n(x))}\}} \right) \\
&\leq \frac{1}{\text{Var}(\tilde{f}_n(x))} \times \frac{4\|K\|_\infty^2}{(\inf_{t \in [x-a, x+a]} \psi(t))^2} \times \frac{1}{nh_n^2} \times P \left(|\xi_{n,1}| > \varepsilon \sqrt{\text{Var}(\tilde{f}_n(x))} \right) \\
&\leq \frac{1}{\text{Var}(\tilde{f}_n(x))} \times \frac{4\|K\|_\infty^2}{(\inf_{t \in [x-a, x+a]} \psi(t))^2} \times \frac{1}{nh_n^2} \times \frac{E\xi_{n,1}^2}{\varepsilon^2 \text{Var}(\tilde{f}_n(x))},
\end{aligned}$$

en vertu de l'inégalité de Chebyshev–Bienaymé (voir inégalité (5.32) page 80 de Billingsley (1995)).

D'où

$$\begin{aligned}
& \frac{1}{\text{Var}(\tilde{f}_n(x))} \sum_{i=1}^n E \left(\xi_{n,i}^2 1_{\{|\xi_{n,i}| > \varepsilon \sqrt{\text{Var}(\tilde{f}_n(x))}\}} \right) \\
&\leq \frac{1}{\text{Var}(\tilde{f}_n(x))} \times \frac{4\|K\|_\infty^2}{(\inf_{t \in [x-a, x+a]} \psi(t))^2} \times \frac{1}{nh_n^2} \times \frac{E\xi_{n,1}^2}{\varepsilon^2 n E\xi_{n,1}^2} \\
&\leq \frac{4\|K\|_\infty^2}{\varepsilon^2 (\inf_{t \in [x-a, x+a]} \psi(t))^2} \times \frac{1}{n^2 h_n^2 \text{Var}(\tilde{f}_n(x))} \\
&= \frac{4\|K\|_\infty^2}{\varepsilon^2 (\inf_{t \in [x-a, x+a]} \psi(t))^2} \times \frac{1}{nh_n \times \frac{1}{h_n} \text{Var} \left(K \left(\frac{x-Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right)}. \tag{3.6}
\end{aligned}$$

Il nous reste à calculer

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} \text{Var} \left(K \left(\frac{x-Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right).$$

Pour cela calculons la limite suivante

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1}{h_n} E \left(K \left(\frac{x - Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right)^2 \\
&= \lim_{n \rightarrow \infty} \frac{1}{h_n} E \left(E \left(K^2 \left(\frac{x - X_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi^2(X_1)} \middle| X_1 \right) \right) \\
&= \lim_{n \rightarrow \infty} \frac{1}{h_n} E \left(K^2 \left(\frac{x - X_1}{h_n} \right) \frac{1}{\psi^2(X_1)} E(1_{\{\Delta_1=0\}} | X_1) \right) \\
&= \lim_{n \rightarrow \infty} \frac{1}{h_n} \int K^2 \left(\frac{x - u}{h_n} \right) \varphi(u) du, \text{ où } \varphi(u) = \frac{f(u)}{\psi(u)} 1_{[x-a, x+a]}(u) \\
&= \lim_{n \rightarrow \infty} \frac{1}{h_n} \int K^2 \left(\frac{z}{h_n} \right) \varphi(x - z) dz \text{ (changement de variable } z = x - u) \\
&= \frac{f(x)}{\psi(x)} \int K^2(z) dz, \tag{3.7}
\end{aligned}$$

grâce au Théorème de Bochner (voir Théorème 4) qui est applicable sous les hypothèses **H1–H3**, **H5** et **H6**.

Par un argument similaire, nous obtenons également

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} E \left(K \left(\frac{x - Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right) = f(x),$$

ce qui donne

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} \left(E \left(K \left(\frac{x - Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right) \right)^2 = 0.$$

En combinant ceci à (3.7) nous obtenons

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} Var \left(K \left(\frac{x - Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right) = \frac{f(x)}{\psi(x)} \int K^2(z) dz. \tag{3.8}$$

La condition de Lindeberg (3.5) découle facilement de (3.6), (3.8) et **H6**.

Nous en déduisons alors que

$$\frac{\tilde{f}_n(x) - E\tilde{f}_n(x)}{\sqrt{Var(\tilde{f}_n(x))}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

et que

$$\sqrt{nh_n} \left(\tilde{f}_n(x) - E\tilde{f}_n(x) \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{f(x)}{\psi(x)} \int K^2(z) dz \right), \quad (3.9)$$

vu (3.8).

Etape 2 : Etude du terme $E\tilde{f}_n(x) - f(x)$.

En procédant comme dans (3.7) nous obtenons

$$\begin{aligned} E\tilde{f}_n(x) &= \frac{1}{h_n} E \left(K \left(\frac{x - Z_1}{h_n} \right) \frac{1_{\{\Delta_1=0\}}}{\psi(Z_1)} \right) \\ &= \frac{1}{h_n} \int K \left(\frac{x - u}{h_n} \right) f(u) du \\ &= \int_{-b}^b K(z) f(x - h_n z) dz \quad \left(\text{changement de variable } z = \frac{x - u}{h_n} \right) \end{aligned} \quad (3.10)$$

donc l'application de La formule de Taylor-Young donne

$$\begin{aligned} E\tilde{f}_n(x) - f(x) &= \int_{-b}^b K(z) (f(x - h_n z) - f(x)) dz \\ &= \int_{-b}^b K(z) (-h_n z f'(x) + h_n z \eta(h_n z)) dz, \end{aligned} \quad (3.11)$$

où η est une fonction de \mathbb{R} dans \mathbb{R} vérifiant $\lim_{u \rightarrow 0} \eta(u) = 0$, donc $\exists \alpha > 0$, $\forall u \in \mathbb{R} : |u| < \alpha \Rightarrow |\eta(u)| < 1$ et puisque h_n tend vers 0, nous avons pour n assez grand $|h_n z| \leq b h_n < \alpha$ donc $|\eta(h_n z)| < 1$ et par conséquent

$$|E\tilde{f}_n(x) - f(x)| \leq b(f'(x) + 1)h_n = O(h_n).$$

Il vient alors de l'hypothèse **H6**

$$\lim_{n \rightarrow \infty} \sqrt{nh_n} (E\tilde{f}_n(x) - f(x)) = 0. \quad (3.12)$$

Etape 3 : Il reste à traiter le terme $f_n(x) - \tilde{f}_n(x)$.

Il est facile de voir que pour n suffisamment grand, nous avons

$$\begin{aligned}
\sqrt{nh_n} \left| f_n(x) - \tilde{f}_n(x) \right| &= \sqrt{nh_n} \left| \frac{1}{h_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \Delta F_n(Z_i) \right. \\
&\quad \left. - \frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \\
&\leq \sqrt{nh_n} \left| \frac{1}{h_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \Delta F_n(Z_i) 1_{\{\Delta_i=0\}} \right. \\
&\quad \left. - \frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \\
&\quad + \frac{\sqrt{nh_n}}{h_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \Delta F_n(Z_i) 1_{\{\Delta_i \neq 0\}} \\
&\leq \frac{\sqrt{nh_n}}{nh_n} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \left| n \Delta F_n(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \\
&\quad + \sqrt{\frac{n}{h_n}} \sum_{i=1}^n K \left(\frac{x - Z_i}{h_n} \right) \Delta F_n(Z_i) 1_{\{\Delta_i \neq 0\}} \\
&\leq \left\{ \sqrt{nh_n} \max_{i: Z_i \in [x-a, x+a]} \left| n \Delta F_n(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \right\} \times \\
&\quad \frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right) + \|K\|_\infty \sqrt{\frac{n}{h_n}} \sum_{i=1}^n \Delta F_n(Z_i) 1_{\{\Delta_i \neq 0\}}.
\end{aligned}$$

Sous **H4** et **H6** nous avons

$$\sqrt{nh_n} \max_{i: Z_i \in [x-a, x+a]} \left| n \Delta F_n(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \xrightarrow{P} 0,$$

et

$$\sqrt{\frac{n}{h_n}} \sum_{i=1}^n \Delta F_n(Z_i) 1_{\{\Delta_i \neq 0\}} \xrightarrow{P} 0.$$

De plus, Parzen (1962) a montré que

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

D'où

$$\sqrt{nh_n} \left| f_n(x) - \tilde{f}_n(x) \right| \xrightarrow{P} 0.$$

En combinant ceci à (3.9) et (3.12), nous obtenons le résultat voulu. \square

3.2 Application à des cas particuliers

Dans cette section nous allons appliquer le résultat du Théorème 15 pour établir la normalité asymptotique de l'estimateur à noyau de la densité dans les modèles de censure double et de censure mixte, ensuite nous déduisons le résultat pour le modèle de censure à droite.

Nous sommes dans la situation où X est censurée à droite par une v.a.r. R et $\min(X, R)$ est censuré à gauche par une v.a.r. L . En d'autres termes $Z := \max(\min(X, R), L)$. Dans toute cette section, nous supposons (sans perte de généralité) qu'il n'y a pas d'ex-æquo parmi les $(X_i)_{1 \leq i \leq n}$.

3.2.1 Le modèle de censure double

On se place dans le modèle de Turnbull (1974) décrit dans la section 2.1, autrement dit la durée de survie X est indépendante du couple (L, R) , $P(0 \leq L \leq R) = 1$ et $\Delta := 1_{\{X > R\}} + 2 \times 1_{\{X \leq L\}}$. Dans cette situation, nous avons pour tout borélien A de \mathbb{R}

$$\begin{aligned} \int_{\{X \in A\}} 1_{\{\Delta=0\}} dP &= \int_A \int 1_{\{l < t \leq r\}} dP_{L,R}(l, r) dP_X(t) \\ &= \int_A P(L < t \leq R) dP_X(t) \\ &= \int_A (S_R(t^-) - S_L(t^-)) dP_X(t). \end{aligned}$$

Donc $\psi(t) = S_R(t^-) - S_L(t^-)$.

Rappelons les équations de self-consistance des estimateurs S_n^D , $S_R^{(n)}$ et $S_L^{(n)}$

de S_X , S_R et S_L respectivement, donnés à la section 2.1.

$$S_n^D(t) = Q^{(n)}(t) - \int_{\{u \leq t\}} \frac{S_n^D(t)}{S_n^D(u)} dQ_1^{(n)}(u) + \int_{\{t < u\}} \frac{1 - S_n^D(t)}{1 - S_n^D(u)} dQ_2^{(n)}(u),$$

$$S_R^{(n)}(t) = 1 + \int_{\{u \leq t\}} \frac{dQ_1^{(n)}(u)}{S_n^D(u)}, \quad t < B_n,$$

$$S_L^{(n)}(t) = - \int_{\{t < u\}} \frac{dQ_2^{(n)}(u)}{1 - S_n^D(u)}, \quad t \geq A_n,$$

où

$$Q_j^{(n)}(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z_i > t, \Delta_i = j\}}, \quad j \in \{0, 1, 2\},$$

$$Q^{(n)} = \sum_{j=0}^2 Q_j^{(n)},$$

$$A_n = \min\{Z_i / S_n^D(Z_i) < 1\}$$

$$\text{et } B_n = \max\{Z_i / S_n^D(Z_i^-) > 0\}.$$

Nous imposons les conditions suivantes comme dans Ren (1997).

$$S_R^{(n)}(0) = 1, \quad \lim_{u \rightarrow \infty} S_L^{(n)}(u) = 0 \quad \text{et} \quad S_n^D(t) = \begin{cases} 1 & \text{si } t < \min_{1 \leq i \leq n} Z_i, \\ 0 & \text{si } t \geq \max_{1 \leq i \leq n} Z_i. \end{cases}$$

Sous cette dernière hypothèse A_n et B_n sont respectivement le plus petit et le plus grand points de discontinuité de S_n^D (voir Ren (1997) page 30).

En remplaçant dans (3.1) F_n par $F_n^D = 1 - S_n^D$, nous retrouvons l'estimateur (2.1) introduit par Ren (1997) qui a imposé, en particulier, les hypothèses suivantes :

D1 $\forall t \geq 0, S_R(t) - S_L(t) > 0.$

D2 S_X, S_R et S_L sont des fonctions continues de t pour $t \geq 0$, et $0 < S_X(t) < 1$ pour $t > 0.$

D3 $S_X(0) = S_R(0) = 1$ et $\lim_{u \rightarrow \infty} S_X(u) = \lim_{u \rightarrow \infty} S_R(u) = \lim_{u \rightarrow \infty} S_L(u) = 0.$

D4 il existe α et $\beta, 0 < \alpha < \beta < \infty$, tels que $P(L \in]0, \alpha]) = 0$ et $P(L \leq \beta) = 1.$

Ces conditions nous permettent d'appliquer le Théorème 11 dans la preuve du corollaire suivant qui donne la normalité asymptotique de f_n^D et λ_n^D , où λ_n^D est l'estimateur de λ donné par

$$\lambda_n^D(x) := \frac{f_n^D(x)}{S_n^D(x)} 1_{\{S_n^D(x) \neq 0\}}.$$

Corollaire 1. (*Boukeloua et Messaci 2016*) Soit $x \in]I_X, T_X[$, sous **H1**, **H5**, **H6** et **D1–D4**, nous avons

$$\begin{aligned} i) \quad & \sqrt{nh_n}(f_n^D(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_R(x) - S_L(x)} \int K^2(z) dz\right). \\ ii) \quad & \sqrt{nh_n}(\lambda_n^D(x) - \lambda(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_X^2(x)(S_R(x) - S_L(x))} \int K^2(z) dz\right). \end{aligned}$$

Ren (1997) a donné les mêmes résultats sous des conditions différentes sur f (voir Théorème 12). Notons de plus que la variance asymptotique de λ_n^D est plus grande que celle de f_n^D , donc λ_n^D est asymptotiquement plus variant que f_n^D .

Démonstration.

- i) Afin d'appliquer le Théorème 15, nous devons vérifier les hypothèses **H2**, **H3** et **H4**.
- L'hypothèse **H2** découle directement de **D2**.
 - **D2** implique que pour tout $T > 0$, $\exists t_0 \in [0, T]$ tel que $\inf_{t \in [0, T]} (S_R(t) - S_L(t)) = S_R(t_0) - S_L(t_0) > 0$ (grâce à **D1**) ce qui donne **H3**.
 - Il reste à vérifier **H4**.

Posons $T := x + a$. En vertu des hypothèses **D1** et **D2**, du Lemme 3.1. (page 35) et de la relation (2.26) (page 33 de Ren (1997)), pour n assez grand, F_n^D est une fonction en escalier vérifiant pour tout $Z_i \in [0, T]$

$$\Delta F_n^D(Z_i) = \begin{cases} \frac{1}{n(S_R^{(n)}(A_n) - S_L^{(n)}(A_n))} & \text{si } Z_i = A_n, \\ \frac{1_{\{\Delta_i=0\}}}{n(S_R^{(n)}(Z_i) - S_L^{(n)}(Z_i))} & \text{si } Z_i \in]A_n, B_n[, \\ \frac{1}{n(S_R^{(n)}(B_n) - S_L^{(n)}(B_n))} & \text{si } Z_i = B_n \text{ et } B_n \leq T, \\ \frac{1_{\{\Delta_i=0\}}}{n(S_R^{(n)}(C_n) - S_L^{(n)}(C_n))} & \text{si } Z_i = C_n \text{ et } B_n > T, \end{cases} \quad (3.13)$$

avec la probabilité 1, où $C_n := \max\{Z_i/Z_i \leq \min(B_n, T)\}$.

Donc pour tout $Z_i \leq T$ nous avons

$$\begin{aligned}
\sqrt{n} \left| n\Delta F_n^D(Z_i)1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| &= \sqrt{n} \left| \frac{1_{\{\Delta_i=0\}}}{S_R^{(n)}(Z_i) - S_L^{(n)}(Z_i)} - \frac{1_{\{\Delta_i=0\}}}{S_R(Z_i) - S_L(Z_i)} \right| \\
&\leq \sqrt{n} \sup_{t \in [0, T]} \left| \frac{1}{S_R^{(n)}(t) - S_L^{(n)}(t)} - \frac{1}{S_R(t) - S_L(t)} \right| \\
&\leq \frac{\sqrt{n} \sup_{t \in [0, T]} |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))|}{\inf_{t \in [0, T]} |S_R^{(n)}(t) - S_L^{(n)}(t)| \inf_{t \in [0, T]} (S_R(t) - S_L(t))}.
\end{aligned} \tag{3.14}$$

Sous les hypothèses **D1–D4**, le Théorème 11 est applicable et il donne

$$\sqrt{n} \sup_{t \in [0, T]} |S_R^{(n)}(t) - S_R(t)| = O_P(1), \tag{3.15}$$

et

$$\sqrt{n} \sup_{t \in [0, T]} |S_L^{(n)}(t) - S_L(t)| = O_P(1). \tag{3.16}$$

De plus pour $\varepsilon_1 \in]0, \inf_{t \in [0, T]} (S_R(t) - S_L(t))$ nous avons

$$\inf_{t \in [0, T]} |S_R^{(n)}(t) - S_L^{(n)}(t)| < \varepsilon_1 \Rightarrow \exists t_0 \in [0, T], |S_R^{(n)}(t_0) - S_L^{(n)}(t_0)| < \varepsilon_1$$

ce qui implique

$$\begin{aligned}
\sup_{t \in [0, T]} |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))| &\geq (S_R(t_0) - S_L(t_0)) - \\
&|S_R^{(n)}(t_0) - S_L^{(n)}(t_0)| \\
&\geq \inf_{t \in [0, T]} (S_R(t) - S_L(t)) - \varepsilon_1 \\
&=: \eta_1 > 0,
\end{aligned}$$

d'où

$$\begin{aligned}
&P \left(\inf_{t \in [0, T]} |S_R^{(n)}(t) - S_L^{(n)}(t)| < \varepsilon_1 \right) \\
&\leq P \left(\sup_{t \in [0, T]} |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))| > \eta_1 \right). \tag{3.17}
\end{aligned}$$

En combinant ceci à (3.14), (3.15) et (3.16) nous obtenons

$$\sqrt{n} \max_{i: Z_i \in [0, T]} \left| n \Delta F_n^D(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| = O_P(1). \quad (3.18)$$

D'autre part, d'après la relation (3.13) nous avons

$$\sqrt{\frac{n}{h_n}} \sum_{i=1}^n \Delta F_n^D(Z_i) 1_{\{\Delta_i \neq 0\}} \leq \frac{2}{\sqrt{nh_n} \inf_{t \in [0, T]} |S_R^{(n)}(t) - S_L^{(n)}(t)|}.$$

Or pour tout $t \in [0, T]$ nous avons

$$\begin{aligned} |S_R^{(n)}(t) - S_L^{(n)}(t)| &\geq S_R(t) - S_L(t) - |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))| \\ &\geq \inf_{t \in [0, T]} (S_R(t) - S_L(t)) - \sup_{t \in [0, T]} |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))|, \end{aligned}$$

et comme $\sup_{t \in [0, T]} |(S_R^{(n)}(t) - S_L^{(n)}(t)) - (S_R(t) - S_L(t))| \xrightarrow{p.s.} 0$ (grâce au Théorème 10 qui est applicable sous **D1** et **D2**), nous en déduisons que pour n assez grand

$$\inf_{t \in [0, T]} |S_R^{(n)}(t) - S_L^{(n)}(t)| \geq \frac{\inf_{t \in [0, T]} (S_R(t) - S_L(t))}{2},$$

et par conséquent

$$\sqrt{\frac{n}{h_n}} \sum_{i=1}^n \Delta F_n^D(Z_i) 1_{\{\Delta_i \neq 0\}} \leq \frac{4}{\sqrt{nh_n} \inf_{t \in [0, T]} (S_R(t) - S_L(t))}.$$

D'où

$$\sum_{i=1}^n \Delta F_n^D(Z_i) 1_{\{\Delta_i \neq 0\}} = o_{p.s.} \left(\sqrt{\frac{h_n}{n}} \right), \quad (3.19)$$

en vertu de **H6**.

Les relations (3.18) et (3.19) montrent que **H4** est vérifiée ainsi que le résultat visé.

ii) D'après le point *i*), nous avons

$$\sqrt{nh_n} (f_n^D(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{f(x)}{S_R(x) - S_L(x)} \int K^2(z) dz \right)$$

et le Théorème 11 implique que

$$\sqrt{nh_n}(S_n^D(x) - S_X(x)) \xrightarrow{P} 0,$$

nous en déduisons alors que

$$\sqrt{nh_n}((f_n^D(x), S_n^D(x)) - \theta) \xrightarrow{\mathcal{D}} (Y, 0),$$

où $\theta := (f(x), S_X(x))$ et Y est une v.a.r. de loi $\mathcal{N}\left(0, \frac{f(x)}{S_R(x) - S_L(x)} \int K^2(z) dz\right)$.
Donc l'application de la méthode delta pour $r_n := \sqrt{nh_n}$ ($r_n \rightarrow +\infty$ grâce à **H6**) et

$$\begin{aligned} \Phi : \quad \mathbb{R}_+ \times \mathbb{R}_+ &\longrightarrow \mathbb{R}_+ \\ (x_1, x_2) &\longrightarrow \frac{x_1}{x_2} \mathbf{1}_{\{x_2 \neq 0\}}, \end{aligned}$$

donne

$$\sqrt{nh_n}(\lambda_n^D(x) - \lambda(x)) \xrightarrow{\mathcal{D}} \Phi'_\theta(Y, 0),$$

où Φ'_θ est la différentielle de Φ en θ (Φ est différentiable au point θ car $S_X(x) \neq 0$). Le calcul donne $\Phi'_\theta(Y, 0) = \frac{Y}{S_X(x)}$, ce qui implique le résultat visé. □

3.2.2 Le modèle de censure mixte

On se place dans le modèle I de Patilea et Rolin (2006) décrit dans la section 2.2. En d'autres termes, nous supposons que les variables X , L et R sont positives et indépendantes et que $\Delta := \mathbf{1}_{\{L < R < X\}} + 2 \times \mathbf{1}_{\{\min(X, R) \leq L\}}$. Rappelons que l'estimateur de Patilea et Rolin (2006) peut s'écrire, en supposant qu'il n'y a pas d'ex-æquo parmi les X_i , comme suit

$$F_n^T(t) = 1 - S_n^T(t) = 1 - \prod_{k: Z_k \leq t} \left\{ 1 - \frac{\mathbf{1}_{\{\Delta_k=0\}}}{n \left(F_L^{(n)}(Z_k^-) - F_Z^{(n)}(Z_k^-) \right)} \right\},$$

où $F_Z^{(n)}(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{Z_i \leq t\}}$ est la version empirique de F_Z et

$$F_L^{(n)}(t) = \prod_{j: Z'_j > t} \left(1 - \frac{\sum_{i=1}^n 1_{\{Z_i = Z'_j, \Delta_i = 2\}}}{n F_Z^{(n)}(Z'_j)} \right),$$

$((Z'_j)_{1 \leq j \leq m} \ (m \leq n)$ étant les valeurs distinctes des $(Z_i)_{1 \leq i \leq n}$) est l'estimateur produit limite de F_L .

En remplaçant cette fois-ci dans (3.1) F_n par F_n^T , nous retrouvons l'estimateur f_n^T de f .

Par ailleurs, l'indépendance des variables latentes permet d'écrire pour tout borélien A de \mathbb{R}

$$\begin{aligned} \int_{\{X \in A\}} 1_{\{\Delta=0\}} dP &= \int_A \int 1_{\{l < t \leq r\}} dP_{L,R}(l, r) dP_X(t) \\ &= \int_A P(L < t \leq R) dP_X(t) \\ &= \int_A F_L(t^-) S_R(t^-) dP_X(t) \end{aligned}$$

Donc $\psi(t) = F_L(t^-) S_R(t^-)$.

Notons par ailleurs $H_k(t) := P(Z \leq t, \Delta = k)$, $k \in \{0, 1, 2\}$ les sous loi de Z et $I_{H_0} := \inf\{t \in \mathbb{R} / H_0(t) > 0\}$ le point initial du support de H_0 .

Afin de déduire la normalité asymptotique de f_n^T du Théorème 15, nous avons besoin des conditions suivantes.

T1 $I_L \leq I_X$ et $T_X \leq T_R$.

T2 $\int_{\{u > I_{H_0}\}} \frac{dH_2(u)}{(F_Z(u))^2} < \infty$.

T1 assure l'identification du modèle étudié et elle est introduite dans Patilea et Rolin (2006). C'est une hypothèse naturelle puisque L est une variable de censure à gauche et R est une variable de censure à droite. Quant à **T2**, elle correspond à la condition (11) de Patilea et Rolin (2006) (voir page 935). Notons qu'elle est moins restrictive que la condition $I_L \vee I_R < I_X$ supposée dans Messaci et Nemouchi (2011, 2013) et dans Kitouni et al. (2015). En effet, cette dernière hypothèse combinée à l'indépendance de X , L et R implique

que $I_Z = (I_X \wedge I_R) \vee I_L < I_X \leq I_{H_0}$ (car $H_0(t) = P(X \leq t, \Delta = 0) \leq P(X \leq t) = F(t)$) et par conséquent $F_Z(I_{H_0}) \neq 0$ et donc **T2** est vérifiée.

Le corollaire suivant donne la normalité asymptotique de f_n^T ainsi que celle de l'estimateur suivant de λ .

$$\gamma_n^T(x) := \frac{f_n^T(x)}{S_n^T(x)} \mathbf{1}_{\{S_n^T(x) \neq 0\}}.$$

Corollaire 2. (*Boukeloua et Messaci 2016*) Soit $x \in]I_X, T_X[$, sous les hypothèses **H1**, **H2**, **H5**, **H6**, **T1** et **T2**, nous obtenons

- i) $\sqrt{nh_n}(f_n^T(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{F_L(x)S_R(x)} \int K^2(z) dz\right)$.
- ii) $\sqrt{nh_n}(\gamma_n^T(x) - \lambda(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{f(x)}{S_X^2(x)F_L(x)S_R(x)} \int K^2(z) dz\right)$.

Comme dans le cas de la censure double, nous remarquons que γ_n^T est asymptotiquement plus variant que f_n^T car sa variance asymptotique est plus grande.

Démonstration.

- i) Nous allons vérifier que **H3** et **H4** sont satisfaites pour pouvoir appliquer le Théorème 15.

D'abord, en choisissant $a \in]0, \min(x - I_X, T_X - x)[$ et en posant $\theta_1 := x - a$ et $\theta_2 := x + a$, nous obtenons $\inf_{t \in [\theta_1, \theta_2]} \psi(t) = F_L(\theta_1^-)S_R(\theta_2^-)$, donc **H3** vient directement de **T1**.

Maintenant, il reste à vérifier **H4**.

Puisque $\Delta_i \neq 0 \Rightarrow \Delta F_n^T(Z_i) = 0$, nous avons $\sum_{i=1}^n \Delta F_n^T(Z_i) \mathbf{1}_{\{\Delta_i \neq 0\}} = 0$.

De plus, en remarquant que $F_L - F_Z = F_L S_X S_R$ (car $F_Z = F_L F_{X \wedge R} =$

$F_L(1 - S_X S_R)$), nous obtenons pour tout $Z_i \in [\theta_1, \theta_2]$

$$\begin{aligned}
& \sqrt{n} \left| n \Delta F_n^T(Z_i) 1_{\{\Delta_i=0\}} - \frac{1_{\{\Delta_i=0\}}}{\psi(Z_i)} \right| \\
&= \sqrt{n} \left| \frac{S_n^T(Z_i^-) 1_{\{\Delta_i=0\}}}{F_L^{(n)}(Z_i^-) - F_Z^{(n)}(Z_i^-)} - \frac{S_X(Z_i) 1_{\{\Delta_i=0\}}}{S_X(Z_i) F_L(Z_i^-) S_R(Z_i^-)} \right| \\
&= \sqrt{n} \left| \frac{S_n^T(Z_i^-) 1_{\{\Delta_i=0\}}}{F_L^{(n)}(Z_i^-) - F_Z^{(n)}(Z_i^-)} - \frac{S_X(Z_i) 1_{\{\Delta_i=0\}}}{F_L(Z_i^-) - F_Z(Z_i^-)} \right| \\
&= \sqrt{n} \left| \frac{S_n^T(Z_i^-) (F_L(Z_i^-) - F_Z(Z_i^-)) + S_X(Z_i) (F_L(Z_i^-) - F_Z(Z_i^-))}{(F_L(Z_i^-) - F_Z(Z_i^-)) (F_L^{(n)}(Z_i^-) - F_Z^{(n)}(Z_i^-))} \right. \\
&\quad \left. - \frac{S_X(Z_i) (F_L(Z_i^-) - F_Z(Z_i^-)) + S_X(Z_i) (F_L^{(n)}(Z_i^-) - F_Z^{(n)}(Z_i^-))}{(F_L(Z_i^-) - F_Z(Z_i^-)) (F_L^{(n)}(Z_i^-) - F_Z^{(n)}(Z_i^-))} \right| 1_{\{\Delta_i=0\}} \\
&\leq \frac{\sqrt{n} \sup_{t \in [I_{H_0}, \theta_2]} |S_n^T(t^-) - S_X(t)|}{F_L(\theta_1^-) S_R(\theta_2^-) S_X(\theta_2) \inf_{t \in [\max(\theta_1, I_{H_0}), \theta_2]} |F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)|} \\
&\quad + \frac{\sqrt{n} \sup_{t \in [I_{H_0}, \theta_2]} |(F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)) - (F_L(t^-) - F_Z(t^-))|}{F_L(\theta_1^-) S_R(\theta_2^-) S_X(\theta_2) \inf_{t \in [\max(\theta_1, I_{H_0}), \theta_2]} |F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)|}.
\end{aligned} \tag{3.20}$$

(Notons que lorsque $\Delta = 0$, nous avons $Z \geq I_{H_0}$ *p.s.* car $H_0(I_{H_0}^-) = 0$).

D'une part,

- en vertu du Théorème de Donsker (voir Théorème 3), nous avons

$$\sqrt{n} \sup_{t \in [I_{H_0}, \theta_2]} |F_Z^{(n)}(t^-) - F_Z(t^-)| = O_P(1). \tag{3.21}$$

- sous **T1** et **T2**, le Théorème 13 donne

$$\sqrt{n} \sup_{t \in [I_{H_0}, \theta_2]} |F_L^{(n)}(t^-) - F_L(t^-)| = O_P(1) \tag{3.22}$$

et

$$\sqrt{n} \sup_{t \in [I_{H_0}, \theta_2]} |S_n^T(t^-) - S_X(t)| = O_P(1). \tag{3.23}$$

D'autre part,

on peut montrer comme dans (3.17) que pour $\varepsilon_2 \in]0, F_L(\theta_1^-)S_R(\theta_2^-)S_X(\theta_2)[$ il existe $\eta_2 > 0$ tel que

$$P \left(\inf_{t \in [\max(\theta_1, I_{H_0}), \theta_2]} |F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)| < \varepsilon_2 \right) \leq \\ P \left(\sup_{t \in [\max(\theta_1, I_{H_0}), \theta_2]} |(F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)) - (F_L(t^-) - F_Z(t^-))| > \eta_2 \right).$$

Ceci combiné à (3.20), (3.21), (3.22) et (3.23) montre que **H4** est vérifiée.

ii) Comme dans le Corollaire 1-ii), ce point découle du point i) et du Théorème 13 par application de la méthode delta.

□

Maintenant en supposant que $L = 0$ p.s., nous nous ramenons au modèle de censure à droite et nous trouvons que $\psi(t) = S_R(t^-)$. De plus, dans ce cas les estimateurs F_n^T , f_n^T et λ_n^T coïncident avec F_n^R , f_n^R et λ_n^R respectivement. Notons que l'hypothèse $I_L \leq I_X$ (resp. **T2**) est nécessairement vérifiée car $I_L = 0$ (resp. $H_2 \equiv 0$).

Ainsi, nous obtenons aisément le résultat suivant.

Corollaire 3. (Boukeloua et Messaci 2016) Soit $x \in]I_X, T_X[$, en supposant que $T_X \leq T_R$, nous obtenons sous les hypothèses **H1**, **H2**, **H5** et **H6**

$$i) \sqrt{nh_n}(f_n^R(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{f(x)}{S_R(x)} \int K^2(z) dz \right). \\ ii) \sqrt{nh_n}(\lambda_n^R(x) - \lambda(x)) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{f(x)}{S_X^2(x)S_R(x)} \int K^2(z) dz \right).$$

Rappelons que Mielniczuk (1986) donne le même résultat que i) sous des conditions plus fortes (voir Théorème 8).

3.3 Etude de simulation

Dans le but de montrer les performances des estimateurs que nous avons étudiés sur des échantillons de tailles finies, nous allons mené une étude de simulation. D'abord, nous présentons la qualité d'estimation en comparant la courbe de chaque estimateur, basé sur un échantillon d'une loi connue avec la courbe de la fonction théorique correspondante. Ensuite, nous illustrons le comportement gaussien des mêmes estimateurs en appliquant des tests numériques et graphiques.

Les modèles théoriques utilisés dans cette étude sont basés sur deux lois de probabilité connues dans l'analyse de survie, à savoir la loi de Weibull et la loi de Bertholon. La loi de Weibull (notée $\mathcal{W}(\eta, \beta)$) est caractérisée par la densité de probabilité suivante :

$$f(x) := \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\eta}\right)^\beta\right) 1_{]0, +\infty[}(x),$$

où $\eta > 0$ et $\beta > 0$ sont les paramètres de la loi ; η est dit le paramètre d'échelle, et β est dit le paramètre de forme. La particularité de cette loi est que son taux de hasard peut être croissant ou décroissant, suivant la valeur de β :

- Si $\beta < 1$: Le taux de hasard décroît.
- Si $\beta = 1$: Le taux de hasard est constant.
- Si $\beta > 1$: Le taux de hasard croît.

Donc cette loi modélise parfaitement les défaillances d'un appareil qui vieillit, car le taux de hasard (ou de panne) d'un appareil est généralement décroissant durant la période de rodage, ensuite il devient constant pendant la période de vie utile et il termine croissant dans la période de vieillissement. Mais la loi de Weibull ne permet pas de tenir compte des défaillances accidentelles que court un appareil, ces derniers sont modélisés par la loi exponentielle qui,

elle même, ne tient pas compte des défaillances dues au vieillissement. Pour pallier à ces insuffisances, dans Bertholon (2001) est proposée une nouvelle loi de probabilité qui permet de modéliser à la fois les défaillances accidentelles et les défaillances dues au vieillissement, et elle est définie comme suit : Une variable aléatoire B est dite de loi de Bertholon si $B = \min(E, W)$, où E est une v.a.r. de loi exponentielle de paramètre η_0 et W est une v.a.r. de loi $\mathcal{W}(\eta_1, \beta)$. Donc la fonction de survie de B est donnée par

$$S_B(x) = S_E(x)S_W(x) = \begin{cases} \exp\left(-\frac{1}{\eta_0}x - \left(\frac{x}{\eta_1}\right)^\beta\right) & \text{si } x > 0, \\ 1 & \text{sinon,} \end{cases}$$

et on obtient facilement la densité et le taux de hasard de B

$$f_B(x) := \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x}{\eta_1}\right)^{\beta-1}\right) \exp\left(-\frac{1}{\eta_0}x - \left(\frac{x}{\eta_1}\right)^\beta\right) 1_{]0, +\infty[}(x)$$

et

$$\lambda_B(x) := \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x}{\eta_1}\right)^{\beta-1}\right) 1_{]0, +\infty[}(x).$$

On note $\mathcal{B}(\eta_0, \eta_1, \beta)$ la loi de Bertholon de paramètres η_0 , η_1 et β .

3.3.1 Simulation des estimateurs étudiés

Pour tracer les courbes des estimateurs étudiés, nous générons 200 échantillons de taille 100 des variables latentes et nous calculons l'estimateur de la densité (ou du taux de hasard) donné par chaque échantillon. Ensuite, nous calculons la valeur approximative de l'erreur quadratique intégrée (ISE) qui correspond à chaque échantillon. Finalement, nous sélectionnons la courbe qui correspond au quantile d'ordre 95% des approximations de l'ISE et nous traçons avec la courbe de la densité (ou du taux de hasard) théorique. Nous faisons la même chose avec la médiane et le quantile d'ordre 5% de l'ISE et nous traçons les courbes correspondantes dans le même graphe, dans le but

d'illustrer la variabilité dans l'estimation à travers les échantillons générés. Dans ce paragraphe, nous utilisons le noyau gaussien et une fenêtre optimale obtenue en minimisant l'ISE de l'estimateur en question parmi un ensemble de valeurs de la fenêtre.

Le cas de la censure double

Dans le premier modèle que nous considérons $X \sim \mathcal{W}(10, 4)$, $L \sim \mathcal{W}(6, 4)$ et $R := L + V$ où V est une v.a.r. de loi $\mathcal{W}(22, 1)$. Pour ce modèle, le taux de censure à droite (resp. à gauche) est de 15% (resp. 11%). Dans le deuxième modèle, $X \sim \mathcal{B}(\frac{200}{3}, 12, 4)$, $L \sim \mathcal{B}(\frac{100}{3}, 6.5, 4)$ et $R := L + (1.75) \times W$ où W est une v.a.r. de loi $\mathcal{W}(5, 50)$. Dans ce cas, le taux de censure à droite (resp. à gauche) est de 16.3% (resp. 13.7%). Notons que pour les deux modèles, le choix de R satisfait à l'hypothèse $L \leq R$ p.s. et X est générée indépendamment du couple (L, R) . Nous obtenons la figure 3.1 pour le premier modèle et la figure 3.2 pour le deuxième. Dans chaque figure, le graphique gauche représente l'estimateur de la densité et le graphique droit représente l'estimateur du taux de hasard.

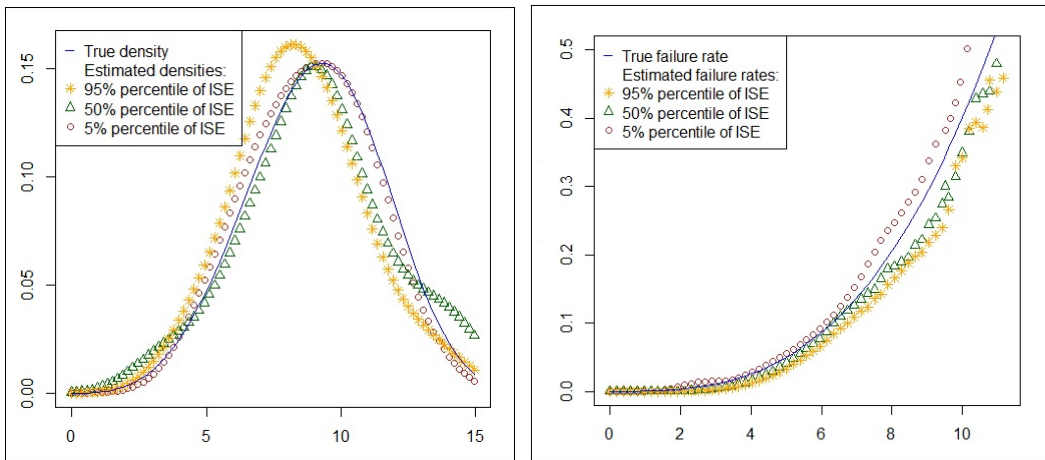


FIGURE 3.1: Représentation des estimateurs étudiés pour le premier modèle de censure double.

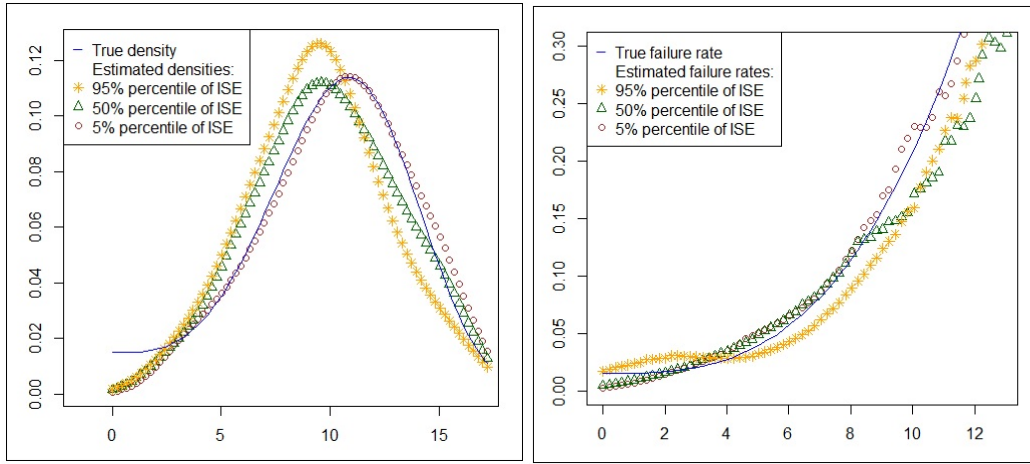


FIGURE 3.2: Représentation des estimateurs étudiés pour le deuxième modèle de censure double.

Le cas de la censure mixte

Dans le premier modèle $X \sim \mathcal{W}(2, 2)$, $R \sim \mathcal{W}(4, 2)$ et $L \sim \mathcal{W}(0.6, 2)$. Le taux de censure à droite (resp. à gauche) est de 18% (resp. 10%). Dans le deuxième modèle $X \sim \mathcal{B}(\frac{25}{3}, 2, 2)$, $R \sim \mathcal{B}(20, 4, 2)$ et $L \sim \mathcal{B}(2.5, 0.5, 3)$. Le taux de censure à droite (resp. à gauche) est de 19% (resp. 12%). Dans chaque modèle les variables X , R et L sont générées de manière indépendante. Nous obtenons la figure 3.3 pour le premier modèle et la figure 3.4 pour le deuxième. Comme précédemment, le graphique gauche (resp. droit) représente l'estimateur de la densité (resp. du taux de hasard).

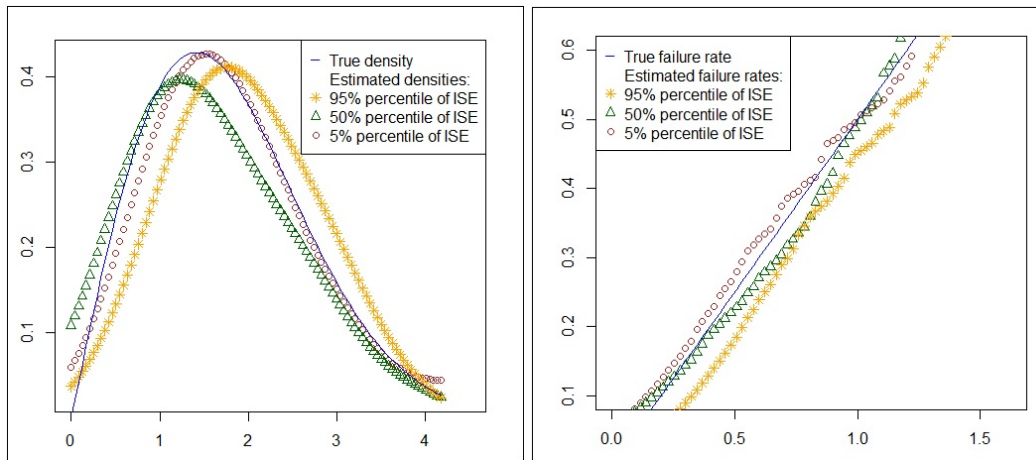


FIGURE 3.3: Représentation des estimateurs étudiés pour le premier modèle de censure mixte.

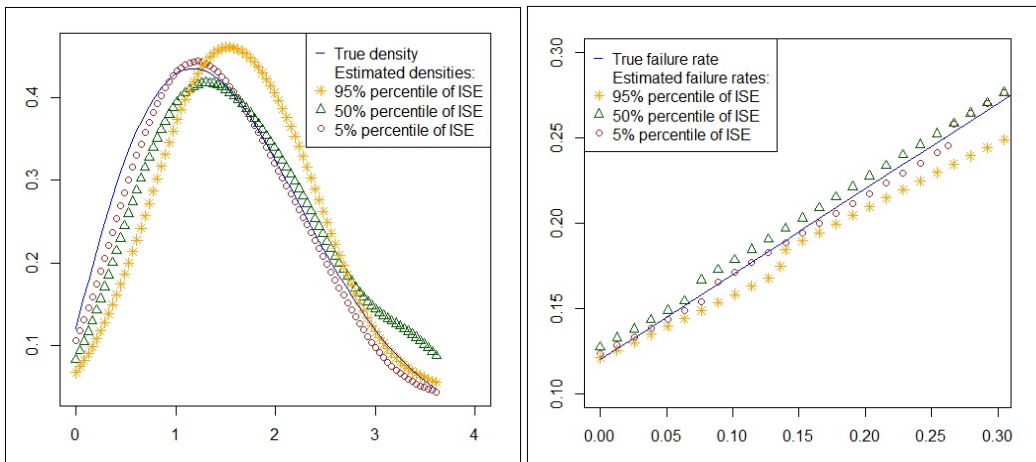


FIGURE 3.4: Représentation des estimateurs étudiés pour le deuxième modèle de censure mixte.

Nous remarquons que pour un taux de censure avoisinant les 30% (qui est un taux acceptable), les estimateurs étudiés donnent, pour tous les modèles, une bonne estimation des fonctions théoriques correspondantes. Sans surprise, la qualité d'estimation s'améliore lorsque l'ISE décroît.

3.3.2 La normalité asymptotique

En un point fixé x qui va être spécifié dans chaque cas, nous calculons 200 valeurs de l'estimateur à noyau de la densité (ou du taux de hasard) basé sur des échantillons de taille 100. Ensuite, nous centrons et nous normalisons les valeurs obtenues en retranchons la vraie valeur de la densité (ou du taux de hasard) et en divisant par l'écart-type de ces valeurs. Finalement, nous traçons l'histogramme et l'estimateur à noyau de la densité de l'échantillon obtenu, que nous comparons avec la courbe de la densité de la loi normale centrée réduite. Nous utilisons également le test graphique Q-Q plot correspondant à cet échantillon pour confirmer la convergence de sa distribution vers la loi normale centrée réduite.

Pour calculer nos estimateurs, nous utilisons le noyau gaussien et une fenêtre optimale obtenue en minimisant la distance entre la valeur de l'estimateur et la valeur de la fonction théorique qui lui correspond, parmi un ensemble de valeurs de la fenêtre.

Nous considérons les mêmes modèles du paragraphe précédent. Pour le premier modèle de censure double, nous obtenons, en $x = 13$, la figure 3.5 pour l'estimateur de la densité et la figure 3.6 pour l'estimateur du taux de hasard. Quant au deuxième modèle de censure double, nous obtenons, en $x = 8$, la figure 3.7 (resp. 3.8) pour l'estimateur de la densité (resp. du taux de hasard). En ce qui concerne la censure mixte, nous obtenons, dans le premier modèle la figure 3.9 (resp. 3.10) pour l'estimateur de la densité (resp. du taux de hasard), avec $x = 1$. Dans le deuxième modèle, nous obtenons, pour $x = 3$, les figures 3.11 et 3.12 pour les estimateurs de la densité et du taux de hasard respectivement.

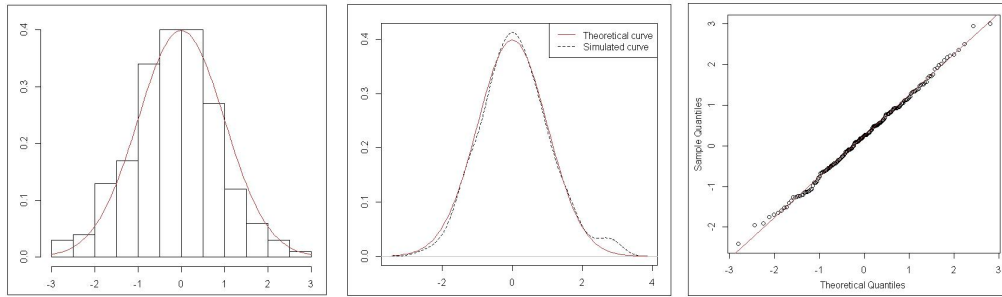


FIGURE 3.5: Normalité de l'estimateur de la densité dans le premier modèle de censure double.

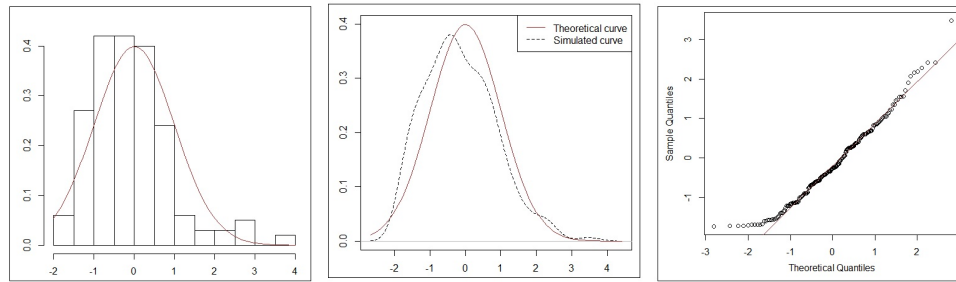


FIGURE 3.6: Normalité de l'estimateur du taux de hasard dans le premier modèle de censure double.

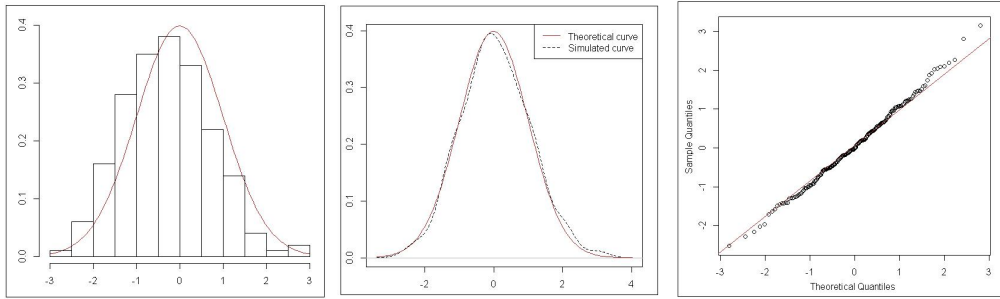


FIGURE 3.7: Normalité de l'estimateur de la densité dans le deuxième modèle de censure double.

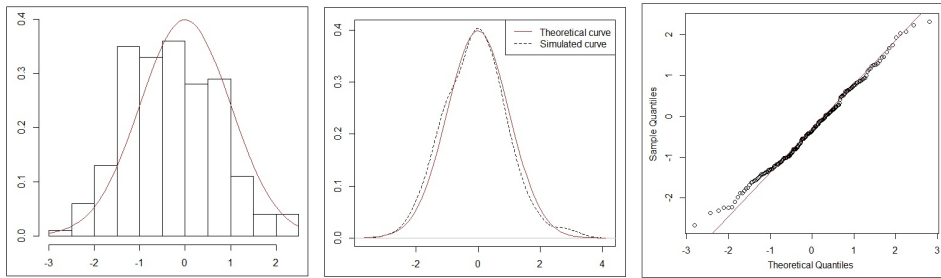


FIGURE 3.8: Normalité de l'estimateur du taux de hasard dans le deuxième modèle de censure double.

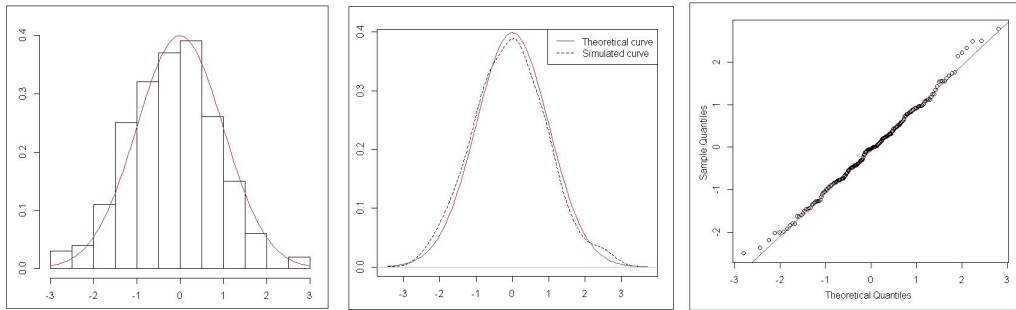


FIGURE 3.9: Normalité de l'estimateur de la densité dans le premier modèle de censure mixte.

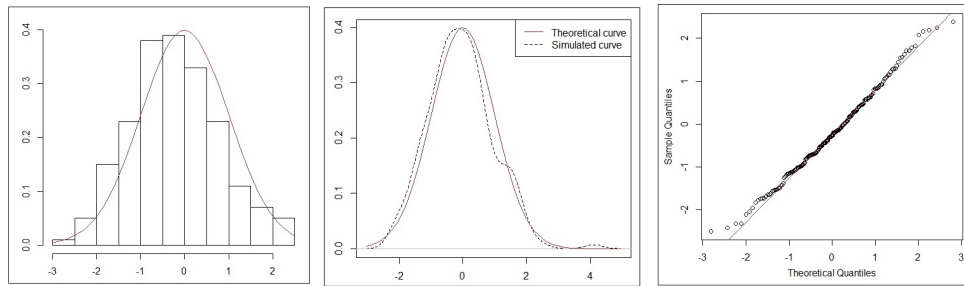


FIGURE 3.10: Normalité de l'estimateur du taux de hasard dans le premier modèle de censure mixte.

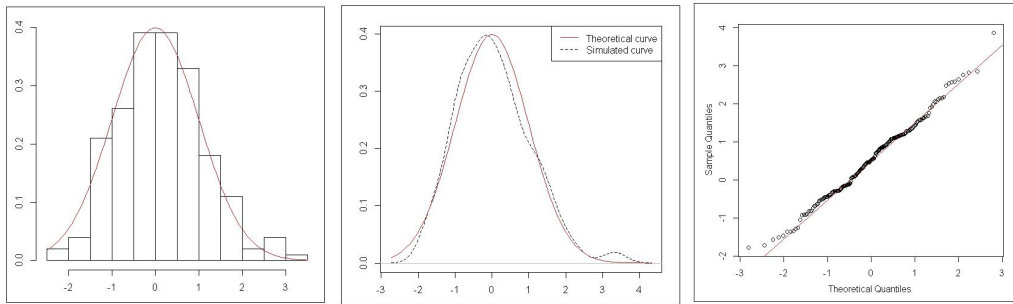


FIGURE 3.11: Normalité de l'estimateur de la densité dans le deuxième modèle de censure mixte.

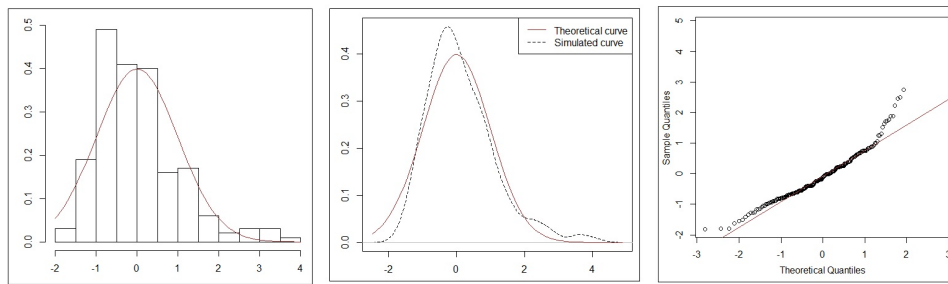


FIGURE 3.12: Normalité de l'estimateur du taux de hasard dans le deuxième modèle de censure mixte.

Pour environ 30% de censure, les figures obtenues montrent que le comportement des lois asymptotiques étudiées est proche à celui de la loi normale centrée réduite. Ceci est confirmé par les tests de Kolmogorov-Smirnov et de Shapiro-Wilk qui acceptent la normalité avec les valeurs-p données dans le tableau suivant. Notons que les estimateurs du taux de hasard présentent plus de variabilité par rapport aux estimateurs de la densité, ce qui est conforme avec les résultats théoriques des corollaires 1 et 2.

La figure	La valeur-p (Kolmogorov-Smirnov)	La valeur-p (Shapiro-Wilk)
Figure 3.5	0.781	0.804
Figure 3.6	0.153	0.406
Figure 3.7	0.781	0.911
Figure 3.8	0.372	0.420
Figure 3.9	0.860	0.776
Figure 3.10	0.367	0.380
Figure 3.11	0.465	0.560
Figure 3.12	0.165	0.570

TABLE 3.1: Les valeurs-p des tests de Kolmogorov-Smirnov et de Shapiro-Wilk.

Chapitre 4

Convergence en moyenne quadratique d'estimateurs à noyau dans le cas de la censure mixte

Dans ce chapitre, nous nous plaçons dans le cadre du modèle I de Patilea et Rolin (2006) et nous gardons les notations précédemment utilisées dans ce contexte. Notre but est d'étudier l'erreur quadratique moyenne des estimateurs à noyau de la densité f_n^T et du taux de hasard λ_n^T que nous avons déjà introduits. Suivant la régularité de f , nous montrons que ces estimateurs convergent en moyenne quadratique et nous précisons la vitesse de cette convergence. Les sections 4.1 et 4.2 de ce chapitre ont fait l'objet d'une publication dans la revue "Statistics and Probability Letters" (Boukeloua (2015)).

4.1 Un Lemme préliminaire

Pour établir la convergence en moyenne quadratique de f_n^T et λ_n^T , nous avons besoin des résultats asymptotiques suivants :

- En vertu de la borne DKW (voir Théorème 1), nous avons pour tout $c > 0$

$$P \left(\sup_{t \in \mathbb{R}} |F_Z^{(n)}(t) - F_Z(t)| > \sqrt{\frac{c \log n}{2n}} \right) = O(n^{-c}). \quad (4.1)$$

- Puisque $H_0(t) = P(X \leq t, L - X < 0, X - R \leq 0)$, on peut lui appliquer le Théorème 2 qui donne pour tous $u \geq 0$ et $\varepsilon > 0$

$$P \left(\sqrt{n} \sup_{t \in \mathbb{R}} |H_0^{(n)}(t) - H_0(t)| > u \right) \leq D(\varepsilon) e^{-(2-\varepsilon)u^2},$$

où $D(\varepsilon)$ est une constante qui ne dépend que de ε . Donc pour tout $c > 0$, nous obtenons en choisissant $u = \sqrt{c \log n}$ et $\varepsilon < 1$

$$P \left(\sup_{t \in \mathbb{R}} |H_0^{(n)}(t) - H_0(t)| > \sqrt{\frac{c \log n}{n}} \right) \leq D(\varepsilon) e^{-(2-\varepsilon)c \log n} = O(n^{-c}). \quad (4.2)$$

- En adaptant le Théorème 5 au cas de la censure à gauche, nous obtenons pour tout $c > 0$ et $\tau > \min(I_X, I_R)$

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

où k est une constante. En remarquant que $\exp \left\{ k \sqrt{\frac{(c+1) \log n}{2}} \right\} \leq n$ pour n assez grand, nous obtenons

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

et puisque

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

il s'ensuit que

$$P \left(\sup_{t \geq \tau} |F_L^{(n)}(t) - F_L(t)| > \frac{1}{1 - S_X(\tau)S_R(\tau)} \sqrt{\frac{(c+1) \log n}{2n}} \right) = O(n^{-c}). \quad (4.3)$$

Nous avons également besoin de l'hypothèse d'identification suivante supposée dans Messaci et Nemouchi (2011, 2013) et dans Kitouni et al. (2015).

$$\mathbf{A1} \max(I_L, I_R) < I_X.$$

Puisque L est une variable de censure à gauche, la condition $I_L < I_X$ est naturelle. De plus, nous avons également besoin d'estimer F_L et l'hypothèse $I_R < I_X$ est introduite pour pouvoir appliquer la formule (4.3) sur tout le support de la variable d'intérêt X .

Le Lemme suivant est le moteur des preuves de nos résultats visés.

Lemme 1. *Sous A1, nous avons pour tout $c > 0$, $\theta < \min(T_X, T_R)$ et $\theta' > I_Z$*

i) Il existe $a > 0$ tel que

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

ii) Il existe $a' > 0$ tel que

$$P \left(\sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t)}{F_L^{(n)}(t) - F_Z^{(n)}(t)} - \frac{S_X(t)}{F_L(t) - F_Z(t)} \right| > a' \sqrt{\frac{\log n}{n}} \right) = O(n^{-c}).$$

Démonstration.

i) Nous enlevons la preuve vu qu'elle découle de (4.1), (4.2) et (4.3) de la même façon que le Théorème 1 dans Kitouni et al. (2015).

ii) Puisque $F_L - F_Z = F_L S_X S_R$ (déjà vu), nous avons pour tout $t \in [\theta', \theta]$

$$\begin{aligned}
\left| \frac{S_n^T(t)}{F_L^{(n)}(t) - F_Z^{(n)}(t)} - \frac{S_X(t)}{F_L(t) - F_Z(t)} \right| &= \left| \frac{S_n^T(t)(F_L(t) - F_Z(t)) - S_X(t)(F_L^{(n)}(t) - F_Z^{(n)}(t))}{(F_L^{(n)}(t) - F_Z^{(n)}(t))(F_L(t) - F_Z(t))} \right. \\
&\quad \left. + \frac{S_X(t)(F_L(t) - F_Z(t)) - S_X(t)(F_L(t) - F_Z(t))}{(F_L^{(n)}(t) - F_Z^{(n)}(t))(F_L(t) - F_Z(t))} \right| \\
&\leq \frac{|S_n^T(t) - S_X(t)|}{|F_L^{(n)}(t) - F_Z^{(n)}(t)|} \\
&\quad + \frac{S_X(t)|F_L^{(n)}(t) - F_Z^{(n)}(t) - (F_L(t) - F_Z(t))|}{(F_L(t) - F_Z(t))|F_L^{(n)}(t) - F_Z^{(n)}(t)|} \\
&\leq \frac{|S_n^T(t) - S_X(t)|}{|F_L^{(n)}(t) - F_Z^{(n)}(t)|} \\
&\quad + \frac{|F_L^{(n)}(t) - F_Z^{(n)}(t) - (F_L(t) - F_Z(t))|}{(F_L(t)S_X(t)S_R(t))|F_L^{(n)}(t) - F_Z^{(n)}(t)|} \\
&\leq \frac{\sup_{\theta' \leq u \leq \theta} |S_n^T(u) - S_X(u)|}{\inf_{\theta' \leq u \leq \theta} |F_L^{(n)}(u) - F_Z^{(n)}(u)|} \\
&\quad + \frac{\sup_{\theta' \leq u \leq \theta} |F_L^{(n)}(u) - F_Z^{(n)}(u) - (F_L(u) - F_Z(u))|}{(F_L(\theta')S_X(\theta)S_R(\theta)) \inf_{\theta' \leq u \leq \theta} |F_L^{(n)}(u) - F_Z^{(n)}(u)|}.
\end{aligned} \tag{4.4}$$

Et pour $\varepsilon_0 \in]0, F_L(\theta')S_X(\theta)S_R(\theta)[$, on peut montrer comme dans (3.17) que

$$\begin{aligned}
&P \left(\inf_{\theta' \leq u \leq \theta} |F_L^{(n)}(u) - F_Z^{(n)}(u)| < \varepsilon_0 \right) \\
&\leq P \left(\sup_{\theta' \leq u \leq \theta} |F_L(u) - F_L^{(n)}(u) + F_Z^{(n)}(u) - F_Z(u)| > \varepsilon \right),
\end{aligned}$$

où $\varepsilon = F_L(\theta')S_X(\theta)S_R(\theta) - \varepsilon_0$. Ceci combiné avec (4.1), (4.3), (4.4) et le point i) entraîne le résultat visé.

□

4.2 Résultats et preuves

Maintenant, nous sommes en mesure d'établir la convergence en moyenne quadratique, avec taux, de f_n^T (donné à la relation 2.2) et λ_n^T en un point fixé $x \in]I_X, \min(T_X, T_R)[$, où rappelons le

$$\lambda_n^T(x) := \frac{f_n^T(x)}{S_n^T(x) + u_n}, \quad (4.5)$$

où $u_n > 0$ est le terme général d'une suite qui converge vers zéro et qui sert à éviter la division par zéro. Notons que Földes et al. (1981) ont étudié un estimateur analogue pour des données censurées à droite, avec $u_n = 1/n$.

Considérons les conditions suivantes habituellement supposées dans le cadre de l'estimation non paramétrique.

A2 f est continue au point x .

A3 Il existe un entier $r \geq 1$ tel que f est r fois dérivable au point x et $\int u^j K(u) du = 0, \forall 1 \leq j \leq r - 1$.

A4 $\exists \alpha, \beta, \varepsilon > 0, \forall y \in]x - \varepsilon, x + \varepsilon[, |f(x) - f(y)| \leq \alpha|x - y|^\beta$.

A5 $h_n = O\left(\frac{1}{\log n}\right)$ et $nh_n \rightarrow \infty$.

A6 K est une fonction bornée à support compact et telle que $\int K(u) du = 1$.

Remarquons que les hypothèses **A3** et **A4** sont plus restrictives que **A2**; elles nous permettent d'obtenir les taux de la convergence en moyenne quadratique de f_n .

Théorème 17. (Boukeloua 2015) Pour tout $x \in]I_X, \min(T_X, T_R)[$ nous avons

i) Sous **A1, A2, A5** et **A6**,

$$\lim_{n \rightarrow \infty} MSE(f_n^T(x)) = 0.$$

ii) Sous **A1, A3, A5** et **A6**,

$$MSE(f_n^T(x)) = O\left(h_n^{2r} + \frac{1}{nh_n}\right).$$

iii) Sous **A1** et **A4–A6**,

$$MSE(f_n^T(x)) = O\left(h_n^{2\beta} + \frac{1}{nh_n}\right).$$

En posant $r = 2$ dans *ii*), nous retrouvons le résultat de Rosenblatt (1956) dans le cas des données complètes, où il suppose que f est trois fois dérivable en x . Pour les données censurées à droite, Kagba (2004) donne le même résultat en supposant que f et F_R admettent des dérivées d'ordre trois bornées au voisinage de x (voir Théorème 9).

Démonstration.

Dans ce cadre de censure mixte, \tilde{f}_n défini dans (3.2) s'écrit comme suit

$$\begin{aligned}\tilde{f}_n(x) &= \frac{1}{nh_n} \sum_{j=1}^n K\left(\frac{x - Z_j}{h_n}\right) \frac{1_{\{\Delta_j=0\}}}{F_L(Z_j^-)S_R(Z_j^-)} \\ &= \frac{1}{nh_n} \sum_{j=1}^n K\left(\frac{x - Z_j}{h_n}\right) \frac{S_X(Z_j)1_{\{\Delta_j=0\}}}{F_L(Z_j^-) - F_Z(Z_j^-)}.\end{aligned}$$

Remarquons là aussi que grâce à **A5** et **A6**, nous pouvons prendre, pour n assez grand, $Z_j \in [\theta', \theta]$, où $\theta' \in]I_X, x[$ et $\theta \in]x, \min(T_X, T_R)[$. Donc, $F_L(Z_j^-) - F_Z(Z_j^-) = F_L(Z_j^-)S_X(Z_j)S_R(Z_j^-) \neq 0$ en vertu de **A1**.

De plus, en se basant sur la relation

$$MSE(f_n^T(x)) \leq 2E(f_n^T(x) - \tilde{f}_n(x))^2 + 4Var(\tilde{f}_n(x)) + 4(E\tilde{f}_n(x) - f(x))^2,$$

il suffit de traiter les étapes suivantes.

Étape 1 : Etude du terme $E(f_n^T(x) - \tilde{f}_n(x))^2$.

Soit $c \geq 4$ une constante, selon le Lemme 1, il existe $a' > 0$ vérifiant la

relation ii) du Lemme et nous pouvons écrire

$$\begin{aligned}
E(f_n^T(x) - \tilde{f}_n(x))^2 &= \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| \leq a' \sqrt{\frac{\log n}{n}} \right\} (f_n^T(x) - \tilde{f}_n(x))^2 dP \\
&\quad + \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| > a' \sqrt{\frac{\log n}{n}} \right\} (f_n^T(x) - \tilde{f}_n(x))^2 dP \\
&=: I_{n,1}(x) + I_{n,2}(x). \tag{4.6}
\end{aligned}$$

En tenant compte du fait que $Z_j \in [\theta', \theta]$ pour n assez grand, nous obtenons

$$\begin{aligned}
I_{n,1}(x) &= \frac{1}{n^2 h_n^2} \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| \leq a' \sqrt{\frac{\log n}{n}} \right\} \left(\sum_{j=1}^n K \left(\frac{x - Z_j}{h_n} \right) \right. \\
&\quad \left. \left(\frac{S_n^T(Z_j^-)}{F_L^{(n)}(Z_j^-) - F_Z^{(n)}(Z_j^-)} - \frac{S_X(Z_j)}{F_L(Z_j^-) - F_Z(Z_j^-)} \right) 1_{\{\Delta_j=0\}} \right)^2 dP \\
&\leq \frac{1}{n^2 h_n^2} \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| \leq a' \sqrt{\frac{\log n}{n}} \right\} \left(\sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} \right. \right. \\
&\quad \left. \left. - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| \right)^2 \times \left(\sum_{j=1}^n K \left(\frac{x - Z_j}{h_n} \right) 1_{\{\Delta_j=0\}} \right)^2 dP \\
&\leq \frac{(a')^2 \log n}{n^3 h_n^2} E \left(\sum_{j=1}^n K \left(\frac{x - X_j}{h_n} \right) 1_{\{\Delta_j=0\}} \right)^2 \\
&= \frac{(a')^2 \log n}{n^3 h_n^2} \sum_{j=1}^n E \left(K^2 \left(\frac{x - X_j}{h_n} \right) 1_{\{\Delta_j=0\}} \right) \\
&\quad + \frac{(a')^2 \log n}{n^3 h_n^2} \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} E \left(K \left(\frac{x - X_j}{h_n} \right) 1_{\{\Delta_j=0\}} \right) E \left(K \left(\frac{x - X_i}{h_n} \right) 1_{\{\Delta_i=0\}} \right) \\
&\leq \frac{(a')^2 \|K\|_\infty \log n}{n^2 h_n} \times \left(\frac{1}{h_n} E \left| K \left(\frac{x - X_1}{h_n} \right) \right| \right) \\
&\quad + \frac{(a')^2 (n^2 - n) \log n}{n^3} \times \left(\frac{1}{h_n} E \left| K \left(\frac{x - X_1}{h_n} \right) \right| \right)^2,
\end{aligned}$$

où, rappelons-le, $\|K\|_\infty = \sup_{u \in \mathbb{R}} |K(u)|$.

Maintenant nous devons traiter le terme $\frac{1}{h_n} E \left| K \left(\frac{x - X_1}{h_n} \right) \right|$.

En utilisant le changement de variable $z = x - u$, nous obtenons

$$\frac{1}{h_n} E \left| K \left(\frac{x - X_1}{h_n} \right) \right| = \frac{1}{h_n} \int \left| K \left(\frac{x - u}{h_n} \right) \right| f(u) du = \frac{1}{h_n} \int \left| K \left(\frac{z}{h_n} \right) \right| f(x - z) dz.$$

Sous les hypothèses **A2**, **A5** et **A6**, le Théorème de Bochner (voir Théorème 4) est applicable et il donne

$$\lim_{n \rightarrow \infty} \frac{1}{h_n} E \left| K \left(\frac{x - X_1}{h_n} \right) \right| = f(x) \int |K(z)| dz, \quad (4.7)$$

d'où

$$I_{n,1}(x) = O \left(\frac{1}{nh_n} \right), \quad (4.8)$$

en vertu de **A5** et **A6**.

Par ailleurs nous avons

$$\begin{aligned} I_{n,2}(x) \leq & \frac{1}{n^2 h_n^2} \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| > a' \sqrt{\frac{\log n}{n}} \right\} \left[\sum_{j=1}^n \left(\frac{S_n^T(Z_j^-)}{F_L^{(n)}(Z_j^-) - F_Z^{(n)}(Z_j^-)} \right. \right. \\ & \left. \left. + \frac{S_X(Z_j)}{F_L(Z_j^-) - F_Z(Z_j^-)} \right) \times \left| K \left(\frac{x - Z_j}{h_n} \right) \right| 1_{\{\Delta_j=0\}} \right]^2 dP. \end{aligned}$$

Puisque pour tout $j \in \{1, \dots, n\}$ tel que $\Delta_j = 0$, $\frac{S_n^T(Z_j^-)}{n(F_L^{(n)}(Z_j^-) - F_Z^{(n)}(Z_j^-))} =$

$\Delta F_n^T(Z_j) \leq 1$, nous obtenons

$$\begin{aligned}
I_{n,2}(x) &\leq \frac{\left(n + \frac{1}{F_L(\theta^-)S_R(\theta^-)}\right)^2}{n^2 h_n^2} \\
&\times \int \left\{ \sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| > a' \sqrt{\frac{\log n}{n}} \right\} \left(\sum_{j=1}^n \left| K \left(\frac{x - Z_j}{h_n} \right) \right| 1_{\{\Delta_j=0\}} \right)^2 dP \\
&\leq \frac{\|K\|_\infty^2 \left(n + \frac{1}{F_L(\theta^-)S_R(\theta^-)}\right)^2}{h_n^2} \\
&\times P \left(\sup_{\theta' \leq t \leq \theta} \left| \frac{S_n^T(t^-)}{F_L^{(n)}(t^-) - F_Z^{(n)}(t^-)} - \frac{S_X(t)}{F_L(t^-) - F_Z(t^-)} \right| > a' \sqrt{\frac{\log n}{n}} \right) \\
&= O \left(\frac{1}{nh_n} \right),
\end{aligned}$$

grâce à **A5**, ii) du Lemme 1 et le fait que $c \geq 4$. En combinant ceci à (4.6) et (4.8), nous obtenons

$$E(f_n^T(x) - \tilde{f}_n(x))^2 = O \left(\frac{1}{nh_n} \right). \quad (4.9)$$

Etape 2 : Etude de la variance de $\tilde{f}_n(x)$.

L'indépendance et l'équidistribution des paires $(Z_j, \Delta_j)_{1 \leq j \leq n}$ permettent d'écrire

$$\begin{aligned}
Var(\tilde{f}_n(x)) &= \frac{1}{n^2 h_n^2} \sum_{j=1}^n Var \left(K \left(\frac{x - Z_j}{h_n} \right) \frac{S_X(Z_j) 1_{\{\Delta_j=0, \theta' \leq Z_j \leq \theta\}}}{F_L(Z_j^-) - F_Z(Z_j^-)} \right) \\
&\leq \frac{1}{nh_n^2} E \left(\frac{K \left(\frac{x - Z_1}{h_n} \right) 1_{\{\Delta_1=0, \theta' \leq Z_1 \leq \theta\}}}{F_L(Z_1^-) S_R(Z_1^-)} \right)^2 \\
&\leq \frac{1}{nh_n} \times \left\{ \frac{\|K\|_\infty}{F_L(\theta')^2 S_R(\theta^-)^2} E \left(\frac{1}{h_n} \left| K \left(\frac{x - X_1}{h_n} \right) \right| \right) \right\} \\
&= O \left(\frac{1}{nh_n} \right), \quad (4.10)
\end{aligned}$$

En vertu de (4.7) et de l'hypothèse **A6**.

Etape 3 : Il reste à traiter le terme $(E\tilde{f}_n(x) - f(x))^2$.

Selon (3.10) nous avons

$$E\tilde{f}_n(x) = \frac{1}{h_n} \int K\left(\frac{x-u}{h_n}\right) f(u) du.$$

Donc

— Sous les hypothèses **A2**, **A5** et **A6**, le Théorème de Bochner implique

$$\lim_{n \rightarrow \infty} E\tilde{f}_n(x) = \lim_{n \rightarrow \infty} \frac{1}{h_n} \int K\left(\frac{z}{h_n}\right) f(x-z) dz = f(x).$$

— Sous les hypothèses **A3** et **A6**, le changement de variable $z = \frac{x-u}{h_n}$ et l'application de la formule de Taylor-Young donnent

$$\begin{aligned} E\tilde{f}_n(x) - f(x) &= \int K(z) (f(x-zh_n) - f(x)) dz \\ &= \int K(z) \left(\sum_{k=1}^r (-1)^k z^k h_n^k \frac{f^{(k)}(x)}{k!} + z^r h_n^r \eta(zh_n) \right) dz \\ &= (-1)^r h_n^r \frac{f^{(r)}(x)}{r!} \int z^r K(z) dz + h_n^r \int z^r \eta(zh_n) K(z) dz, \end{aligned} \tag{4.11}$$

où η est une fonction de \mathbb{R} dans \mathbb{R} vérifiant $\lim_{u \rightarrow 0} \eta(u) = 0$, donc sous **A5** et **A6** on peut montrer comme dans (3.11) que η est bornée indépendamment de z et n , pour n assez grand. Donc

$$E\tilde{f}_n(x) - f(x) = O(h_n^r),$$

en vertu de **A6**.

— Sous les hypothèses **A4–A6**, nous avons pour n assez grand $x - zh_n \in]x - \varepsilon, x + \varepsilon[$, donc

$$\begin{aligned} |E\tilde{f}_n(x) - f(x)| &\leq \int |K(z) (f(x-zh_n) - f(x))| dz \\ &\leq \alpha \int |K(z)| |z|^\beta h_n^\beta dz \\ &= O(h_n^\beta). \end{aligned}$$

□

Le résultat suivant, obtenu du Théorème 17, concerne l'erreur quadratique moyenne de λ_n^T , pour cela nous avons besoin de l'hypothèse suivante

A'5 : **A5** est vérifiée et il existe $d > 0$ tel que $\frac{1}{h_n u_n} = O(n^d)$.

Théorème 18. (Boukeloua 2015) Soit $x \in]I_X, \min(T_X, T_R)[$

i) Sous **A1**, **A2**, **A'5** et **A6**, nous avons

$$\lim_{n \rightarrow \infty} MSE(\lambda_n^T(x)) = 0.$$

ii) Sous **A1**, **A3**, **A'5** et **A6** et avec un choix de $u_n = O\left(h_n^r + \frac{1}{\sqrt{nh_n}}\right)$, nous obtenons

$$MSE(\lambda_n^T(x)) = O\left(h_n^{2r} + \frac{1}{nh_n}\right).$$

iii) Sous **A1**, **A4**, **A'5** et **A6** et avec un choix de $u_n = O\left(h_n^\beta + \frac{1}{\sqrt{nh_n}}\right)$, nous obtenons

$$MSE(\lambda_n^T(x)) = O\left(h_n^{2\beta} + \frac{1}{nh_n}\right).$$

Démonstration.

Nous avons

$$\begin{aligned} MSE(\lambda_n^T(x)) &= \int_{\left\{S_n^T(x) + u_n < \frac{S_X(x)}{2}\right\}} (\lambda_n^T(x) - \lambda(x))^2 dP \\ &\quad + \int_{\left\{S_n^T(x) + u_n \geq \frac{S_X(x)}{2}\right\}} (\lambda_n^T(x) - \lambda(x))^2 dP \\ &=: J_{n,1}(x) + J_{n,2}(x). \end{aligned} \tag{4.12}$$

Il suffit donc d'étudier ces deux termes séparément.

Sous **A6** on a

$$|\lambda_n^T(x)| \leq \frac{\|K\|_\infty}{h_n(S_n^T(x) + u_n)} \leq \frac{\|K\|_\infty}{h_n u_n}.$$

En choisissant $c \geq 2d + 1$ (d étant la constante spécifiée dans l'hypothèse **A'5**) et en appliquant le Lemme 1, il existe $a > 0$ tel que d'une part

$$\begin{aligned} J_{n,1}(x) &\leq 2 \int_{\{S_n^T(x)+u_n < \frac{S_X(x)}{2}\}} ((\lambda_n^T(x))^2 + \lambda(x)^2) dP \\ &\leq 2 \left(\frac{\|K\|_\infty^2}{h_n^2 u_n^2} + \lambda(x)^2 \right) P \left(S_n^T(x) + u_n < \frac{S_X(x)}{2} \right). \end{aligned}$$

Or

$$S_n^T(x) + u_n < \frac{S_X(x)}{2} \Rightarrow |S_n^T(x) - S_X(x)| = S_X(x) - S_n^T(x) > \frac{S_X(x)}{2} > a \sqrt{\frac{\log n}{n}},$$

pour n assez grand.

Nous en déduisons alors que

$$J_{n,1}(x) \leq 2 \left(\frac{\|K\|_\infty^2}{h_n^2 u_n^2} + \lambda(x)^2 \right) P \left(|S_n^T(x) - S_X(x)| > a \sqrt{\frac{\log n}{n}} \right) = O(n^{2d-c}), \quad (4.13)$$

en vertu du Lemme 1 et de l'hypothèse **A'5**.

D'autre part, la décomposition

$$\lambda_n^T(x) - \lambda(x) = \frac{f_n^T(x) - f(x)}{S_n^T(x) + u_n} + (S_X(x) - S_n^T(x) - u_n) \frac{\lambda(x)}{S_n^T(x) + u_n}$$

implique que

$$\begin{aligned} J_{n,2}(x) &\leq 2 \int_{\{S_n^T(x)+u_n \geq \frac{S_X(x)}{2}\}} \frac{(f_n^T(x) - f(x))^2}{(S_n^T(x) + u_n)^2} dP \\ &\quad + 2\lambda(x)^2 \int_{\{S_n^T(x)+u_n \geq \frac{S_X(x)}{2}\}} \frac{(S_X(x) - S_n^T(x) - u_n)^2}{(S_n^T(x) + u_n)^2} dP \\ &\leq \frac{8}{S_X(x)^2} MSE(f_n^T(x)) + \frac{16\lambda(x)^2}{S_X(x)^2} MSE(S_n^T(x)) + \frac{16\lambda(x)^2}{S_X(x)^2} u_n^2. \end{aligned} \quad (4.14)$$

Grâce au Théorème 17, il reste à traiter $MSE(S_n^T(x))$.

$$\begin{aligned}
MSE(S_n^T(x)) &= \int_{\{|S_n^T(x) - S_X(x)| \leq a\sqrt{\frac{\log n}{n}}\}} (S_n^T(x) - S_X(x))^2 dP \\
&\quad + \int_{\{|S_n^T(x) - S_X(x)| > a\sqrt{\frac{\log n}{n}}\}} (S_n^T(x) - S_X(x))^2 dP \\
&\leq a^2 \frac{\log n}{n} + 4 \times P\left(|S_n^T(x) - S_X(x)| > a\sqrt{\frac{\log n}{n}}\right) \\
&= O\left(\frac{\log n}{n}\right),
\end{aligned}$$

en vertu du Lemme 1. En combinant ceci avec (4.12), (4.13), (4.14), l'hypothèse **A'5** et le Théorème 17, nous obtenons les résultats visés. \square

4.3 Utilisation des noyaux d'ordre infini

Dans la section précédente, nous avons utilisé un noyau K d'ordre fini r , où r est un entier qui quantifie la régularité de f , car nous avons supposé que f est r fois dérivable (voir l'hypothèse **A3**). Dans ce cas, nous avons obtenu un taux de convergence en moyenne quadratique de f_n^T de l'ordre de $h_n^{2r} + 1/(nh_n)$. Mais lorsque f est infiniment dérivable, la relation (4.11) montre que l'utilisation d'un noyau d'ordre fini r n'est pas recommandée, car le taux de convergence de $E\tilde{f}_n(x)$ vers $f(x)$ reste toujours $O(h_n^r)$, même si on peut développer f en une série infinie. Donc dans ce cas, le fait que f soit infiniment dérivable ne permet pas d'améliorer le taux de convergence à cause du choix du noyau. Pour cela, Devroye (1992) a introduit des noyaux d'ordre infini vérifiant

$$\int u^j K(u) du = 0, \quad \forall j \in \mathbb{N}^*,$$

et définis par

$$K(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} k(t) e^{-itx} dt,$$

où

$$k(t) = \begin{cases} 1 & \text{si } |t| \leq 1, \\ g(|t|) & \text{sinon,} \end{cases}$$

avec g une fonction continue, de carré intégrable, majorée en valeur absolue par 1 et telle que $g(1) = g(-1) = 1$.

Ce type de noyaux est donc plus adéquat pour estimer les densités infiniment dérivables. Dans la suite, nous donnons des taux de convergence de l'estimateur f_n^T basé sur ce type de noyaux. Le degré de régularité de f sera quantifié par le taux de convergence vers 0, de sa transformée de Fourier, que nous notons Φ . Ceci est précisé dans les hypothèses suivantes.

A7 $\exists p > 0$ tel que $\int_{-\infty}^{+\infty} |t|^p |\Phi(t)| dt < \infty$.

A8 $\exists q, D > 0$ tels que $|\Phi(t)| \leq D e^{-q|t|}, \forall t \in \mathbb{R}$.

A9 Φ est à support compact.

Théorème 19. Soit $x \in]I_X, \min(T_X, T_R)[$,

i) Sous les hypothèses **A1** et **A5–A7** et pour un choix de $h_n \sim a n^{-\beta}$ avec $a > 0$ et $\beta = (2p + 1)^{-1}$, nous obtenons

$$MSE(f_n^T(x)) = O\left(n^{\frac{-2p}{2p+1}}\right),$$

ii) Sous les hypothèses **A1, A5, A6** et **A8**, et pour un choix de $h_n \sim \frac{1}{a \log n}$ avec $a > \frac{1}{2q}$, nous obtenons

$$MSE(f_n^T(x)) = O\left(\frac{\log n}{n}\right),$$

iii) Sous les hypothèses **A1, A5, A6** et **A9**, nous obtenons

$$MSE(f_n^T(x)) = O\left(\frac{1}{nh_n}\right).$$

Remarquons que ces taux de convergence sont meilleurs que ceux du Théorème 17. Berg (2007) a montré des résultats similaires pour des données censurées à droite.

Démonstration. Commençons par traiter la variance de $f_n^T(x)$.

Nous avons

$$\begin{aligned}
\text{Var}(f_n^T(x)) &\leq 2E \left(f_n^T(x) - \tilde{f}_n(x) \right)^2 + 4E \left(\tilde{f}_n(x) - E\tilde{f}_n(x) \right)^2 \\
&\quad + 4(E\tilde{f}_n(x) - Ef_n^T(x))^2 \\
&\leq 6E \left(f_n^T(x) - \tilde{f}_n(x) \right)^2 + 4E \left(\tilde{f}_n(x) - E\tilde{f}_n(x) \right)^2 \\
&= O \left(\frac{1}{nh_n} \right), \tag{4.15}
\end{aligned}$$

en vertu de (4.9) et (4.10).

Passons maintenant au biais de $f_n^T(x)$, noté $b(f_n^T(x))$.

En procédant comme dans le théorème 7 de Berg (2007), nous pouvons montrer que

— Sous **A7**, nous avons

$$b(f_n^T(x)) = o \left(n^{\frac{-p}{2p+1}} \right),$$

— Sous **A8**, nous avons

$$b(f_n^T(x)) = O \left(\frac{1}{\sqrt{n}} \right),$$

— Sous **A9**, nous avons

$$b(f_n^T(x)) = 0,$$

en combinant ceci avec (4.15), nous obtenons les résultats visés. \square

En suivant la même démarche du Théorème 18, nous obtenons les résultats suivants pour $\lambda_n^T(x)$.

Théorème 20. Soit $x \in]I_X, \min(T_X, T_R)[$,

i) Sous les hypothèses **A1, A'5, A6** et **A7** et pour un choix de $h_n \sim an^{-\beta}$ avec $a > 0$ et $\beta = (2p + 1)^{-1}$ et de $u_n = O \left(n^{\frac{-p}{2p+1}} \right)$, nous obtenons

$$MSE(\lambda_n^T(x)) = O \left(n^{\frac{-2p}{2p+1}} \right),$$

ii) Sous les hypothèses **A1**, **A'5**, **A6** et **A8**, et pour un choix de $h_n \sim \frac{1}{a \log n}$ avec $a > \frac{1}{2q}$ et de $u_n = O\left(\sqrt{\frac{\log n}{n}}\right)$, nous obtenons

$$MSE(\lambda_n^T(x)) = O\left(\frac{\log n}{n}\right),$$

iii) Sous les hypothèses **A1**, **A'5**, **A6** et **A9** et pour un choix de $u_n = O\left(\frac{1}{\sqrt{nh_n}}\right)$, nous obtenons

$$MSE(\lambda_n^T(x)) = O\left(\frac{1}{nh_n}\right).$$

4.4 Etude de simulation

Dans le but de conforter les résultats théoriques prouvés dans ce chapitre, nous allons présenter une étude de simulation. Dans cette étude, nous évaluons le biais, la variance et l'erreur quadratique moyenne des estimateurs étudiés pour quelques valeurs fixées de x . Pour cela, nous calculons 200 valeurs de chaque estimateur, auxquelles nous retranchons la valeur de la fonction théorique correspondante. Nous calculons ensuite la moyenne, la variance et la moyenne des carrés de l'échantillon que nous obtenons, qui approchent respectivement le biais, la variance et l'erreur quadratique moyenne de l'estimateur considéré.

Pour calculer nos estimateurs, nous utilisons le noyau gaussien et une fenêtre optimale obtenue en minimisant la distance entre la valeur de l'estimateur et la valeur de la fonction théorique qui lui correspond, parmi un ensemble de valeurs de la fenêtre. Afin de montrer l'effet de la taille d'échantillon sur la qualité de l'estimation, nous utilisons à chaque fois deux tailles d'échantillon $n = 100$ et $n = 250$.

Nous allons considérer les deux types de censure : la censure mixte et la censure double et ceci pour guider l'intuition sur la possibilité d'étendre les

résultats prouvés dans ce chapitre au cas de la censure double.

Comme dans le chapitre précédent, nous considérons deux modèles pour chaque type de censure. Dans le premier modèle de censure mixte, $X \sim \mathcal{W}(3, 3)$, $R \sim \mathcal{W}(4.5, 4)$ et $L \sim \mathcal{W}(1.7, 3)$. Pour ce modèle les taux de censure à droite et à gauche sont respectivement de 16% et 17%. Dans le deuxième modèle, $X \sim \mathcal{B}(10, 2, 4)$, $R \sim \mathcal{B}(20, 5, 1.7)$ et $L \sim \mathcal{B}(2.5, 0.5, 0.2)$. Dans ce cas, le taux de censure à droite (resp. à gauche) est de 16.4% (resp. 18.1%). Quant aux modèles de censure double, nous les considérons comme suit : dans le premier modèle $X \sim \mathcal{W}(12, 4)$, $L \sim \mathcal{W}(7, 4)$ et $R := L + V$ où V est une v.a.r. de loi $\mathcal{W}(22, 1)$. Le taux de censure à droite (resp. à gauche) est de 18.4% (resp. 10.3%). Dans le deuxième modèle, $X \sim \mathcal{B}(100, 10, 5)$, $L \sim \mathcal{B}(\frac{100}{3}, 7, 5)$ et $R := L + 3 \times W$, où W est une v.a. de loi $\mathcal{W}(2, 25)$. Le taux de censure à droite (resp. à gauche) est de 17.2% (resp. 16.7%). Les résultats que nous avons obtenus sont présentés dans les tableaux suivants.

	$x = 1$		$x = 3$		$x = 4$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	8.02×10^{-3}	1.94×10^{-2}	-2.97×10^{-2}	-1.42×10^{-2}	1.6×10^{-2}	1.47×10^{-2}
Variance	7.79×10^{-4}	2.12×10^{-4}	1.08×10^{-3}	6.48×10^{-4}	2.55×10^{-4}	1.15×10^{-4}
MSE	8.39×10^{-4}	5.85×10^{-4}	1.96×10^{-3}	8.48×10^{-4}	5.11×10^{-4}	3.3×10^{-4}

TABLE 4.1: L'estimateur de la densité dans le premier modèle de censure mixte.

	$x = 0.5$		$x = 1.5$		$x = 3$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	6.62×10^{-3}	4.06×10^{-4}	-2.88×10^{-2}	-1.51×10^{-2}	-7.75×10^{-2}	-4.68×10^{-2}
Variance	3.82×10^{-4}	2.6×10^{-4}	5.71×10^{-4}	3.38×10^{-4}	2.15×10^{-2}	9.29×10^{-3}
MSE	4.24×10^{-4}	2.6×10^{-4}	1.4×10^{-3}	5.66×10^{-4}	2.74×10^{-2}	1.14×10^{-2}

TABLE 4.2: L'estimateur du taux de hasard dans le premier modèle de censure mixte.

	$x = 0.5$		$x = 1.5$		$x = 2.5$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	3.11×10^{-2}	1.49×10^{-2}	-3.03×10^{-2}	-2.18×10^{-2}	-6.06×10^{-3}	-5.6×10^{-3}
Variance	8.32×10^{-4}	6.77×10^{-4}	4.81×10^{-3}	2.91×10^{-3}	1.75×10^{-4}	7.43×10^{-5}
MSE	1.79×10^{-3}	8.96×10^{-4}	5.71×10^{-3}	3.37×10^{-3}	2.11×10^{-4}	1.05×10^{-4}

TABLE 4.3: L'estimateur de la densité dans le deuxième modèle de censure mixte.

	$x = 0.2$		$x = 0.5$		$x = 1$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	7.3×10^{-3}	1.35×10^{-2}	2.95×10^{-2}	2.12×10^{-2}	1.65×10^{-2}	2.03×10^{-3}
Variance	5.39×10^{-4}	2.05×10^{-4}	1.37×10^{-3}	8.14×10^{-4}	8.34×10^{-4}	2.8×10^{-4}
MSE	5.9×10^{-4}	3.86×10^{-4}	2.24×10^{-3}	1.26×10^{-3}	1.1×10^{-3}	2.83×10^{-4}

TABLE 4.4: L'estimateur du taux de hasard dans le deuxième modèle de censure mixte.

	$x = 5$		$x = 10$		$x = 13$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	3.4×10^{-3}	3.12×10^{-3}	-1.35×10^{-3}	-1.03×10^{-2}	-8.65×10^{-3}	-5.36×10^{-3}
Variance	2.77×10^{-5}	2.14×10^{-5}	1.25×10^{-4}	7.44×10^{-5}	1.73×10^{-4}	9.32×10^{-5}
MSE	3.91×10^{-5}	3.1×10^{-5}	3.08×10^{-4}	1.8×10^{-4}	2.48×10^{-4}	1.21×10^{-4}

TABLE 4.5: L'estimateur de la densité dans le premier modèle de censure double.

	$x = 2$		$x = 4$		$x = 7$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	-9.49×10^{-5}	-5.63×10^{-4}	3.06×10^{-3}	2.29×10^{-3}	-2.54×10^{-4}	-2.13×10^{-3}
Variance	4.41×10^{-6}	3.22×10^{-6}	3.59×10^{-5}	2.37×10^{-5}	1.48×10^{-4}	5.31×10^{-5}
MSE	4.4×10^{-6}	3.52×10^{-6}	4.51×10^{-5}	2.88×10^{-5}	1.47×10^{-4}	5.74×10^{-5}

TABLE 4.6: L'estimateur du taux de hasard dans le premier modèle de censure double.

	$x = 5$		$x = 10$		$x = 13$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	1.08×10^{-2}	5.74×10^{-3}	-1.2×10^{-2}	7.83×10^{-3}	1.75×10^{-3}	-1.63×10^{-3}
Variance	1.09×10^{-4}	8.57×10^{-5}	3.97×10^{-4}	2.84×10^{-4}	1.05×10^{-4}	4.85×10^{-5}
MSE	2.25×10^{-4}	1.18×10^{-4}	5.22×10^{-4}	3.44×10^{-4}	1.07×10^{-4}	5.09×10^{-5}

TABLE 4.7: L'estimateur de la densité dans le deuxième modèle de censure double.

	$x = 2$		$x = 5$		$x = 8$	
	$n = 100$	$n = 250$	$n = 100$	$n = 250$	$n = 100$	$n = 250$
Biais	8.17×10^{-4}	3.75×10^{-3}	8.36×10^{-3}	1.67×10^{-3}	-1.93×10^{-2}	-1.11×10^{-2}
Variance	5.37×10^{-5}	1.43×10^{-5}	2.01×10^{-4}	1.41×10^{-4}	5.81×10^{-4}	3.22×10^{-4}
MSE	5.41×10^{-5}	2.83×10^{-5}	2.7×10^{-4}	1.43×10^{-4}	9.5×10^{-4}	4.43×10^{-4}

TABLE 4.8: L'estimateur du taux de hasard dans le deuxième modèle de censure double.

Pour les tailles d'échantillon $n = 100$ et $n = 250$, l'erreur quadratique moyenne des estimateurs étudiés est assez petite, ce qui confirme leur performance. Sans surprise, la qualité de l'estimation s'améliore lorsque la taille d'échantillon augmente. Grosso-modo, les erreurs quadratiques moyennes des estimateurs étudiés dans le cas de la censure double sont plus petites que celles de la censure mixte, ce qui nous mène à penser à la possibilité de prouver des résultats similaires à ceux obtenus dans ce chapitre pour le modèle de censure double.

Chapitre 5

Utilisation des divergences dans l'étude de modèles semi-paramétriques avec des données censurées

5.1 Introduction

Soit X une variable aléatoire positive de fonction de répartition inconnue F . Nous nous intéressons à des modèles semi-paramétriques, définis par des contraintes de moments, de la forme

$$E(g(X, \theta)) = 0, \tag{5.1}$$

où $g := (g_1, \dots, g_l)^\top \in \mathbb{R}^l$ est une fonction connue définie sur $\mathbb{R} \times \Theta$ et $\theta \in \Theta \subset \mathbb{R}^d$ est un paramètre vectoriel. Pour les données complètes, ces modèles ont été largement étudiés dans la littérature. Nous citons par exemple les travaux de Qin et Lawless (1994), qui ont utilisé la méthode de la vraisemblance empirique (EL) (voir Owen (1988, 1990, 2001)), pour proposer des

estimateurs de θ et F , et pour construire des régions de confiance pour θ et des tests relatifs au paramètre. Par ailleurs, Corcoran (1998) a considéré le modèle (5.1) avec $g(x, \theta) = x - \theta$ (i.e. θ est la moyenne de X) et a introduit deux classes de statistiques généralisant la méthode EL. Les modèles (5.1) peuvent être traités par la méthode des moments généralisée (GMM) introduite par Hansen (1982), mais dans ce cas, le biais des estimateurs obtenus croît, en général, avec le nombre de contraintes (voir Newey et Smith (2004)). Comme alternative aux estimateurs GMM, Smith (1997) a introduit la classe d'estimateurs de vraisemblance empirique généralisée (GEL). Les propriétés de ces estimateurs ont été étudiées dans Newey et Smith (2004) et comparés aux estimateurs GMM. Smith (2011) et Chang et al. (2015) ont étudié le modèle (5.1) pour des données dépendantes.

Pour les données censurées à droite, la méthode la plus utilisée pour traiter le modèle (5.1) est la méthode EL. Plusieurs auteurs ont travaillé dans ce domaine comme Wang et Jing (2001) qui ont étudié le modèle (5.1) avec $g(x, \theta) = \xi(x) - \theta$ (où ξ est une fonction connue). Ils ont appliqué la méthode EL pour construire des intervalles de confiance pour θ . Ren (2001) a considéré le modèle (5.1) avec $g(x, \theta) = x - \theta$ et a proposé une méthode du rapport de vraisemblance empirique qui donne des intervalles de confiance pour θ , qui est, dans ce cas, la moyenne de X . Pan et Zhou (2002) ont proposé une méthode EL en termes de la fonction de hasard cumulée. Par ailleurs, He et al. (2015) ont utilisé certaines fonctions d'influence pour améliorer les intervalles de confiance obtenus par la méthode EL. Zheng et al. (2014) ont proposé une autre méthode pour améliorer la méthode EL, appelée la vraisemblance empirique ajustée. Shen et al. (2016) ont construit des intervalles de confiance pour θ dans le cas des données doublement censurées, avec un ou deux échantillons. De plus, plusieurs auteurs ont étudié les problèmes à deux échantillons et les problèmes de régression liés au modèle (5.1) comme Wang et Wang (2001), Fang et al. (2013), He et Liang (2014) et Wu et al.

(2015).

Par ailleurs, la théorie des φ -divergences et de la dualité développée dans Broniatowski et Keziou (2006) a été largement utilisée dans l'inférence statistique, voir Keziou (2003) et Broniatowski et Keziou (2009) pour les problèmes paramétriques et Keziou et Leoni-Aubin (2008) pour un problème de rapport de densité avec deux échantillons. Broniatowski et Keziou (2012) ont étudié le modèle (5.1) avec des données complètes en utilisant les φ -divergences. Ils ont proposé des estimateurs de θ et ont étudié leurs propriétés asymptotiques. De plus, ils ont construit des régions de confiance pour θ ainsi que des tests pour le modèle et pour θ . Ils ont donné la distribution asymptotique des statistiques de ces tests aussi bien sous l'hypothèse nulle que sous les alternatives, même lorsque ces dernières contiennent une mauvaise spécification. L'obtention des distributions statistiques sous les alternatives leur a permis de déduire des approximations de la fonction de puissance et de la taille d'échantillon qui assure une puissance voulue pour une alternative donnée. Dans ce chapitre, nous initions l'extension de leurs résultats au cas des données censurées à droite. Nous construisons des estimateurs pour θ basés sur la mesure empirique de Kaplan-Meier et nous étudions leurs propriétés asymptotiques. Pour cela, nous appliquons un théorème central limite (TCL) dans le cas des données censurées (voir Zhou (2015)) et une loi faible uniforme des grands nombres (LFUGN) que nous prouvons dans le présent chapitre. Les résultats théoriques que nous obtenons ont plusieurs applications comme celle donnée dans l'exemple suivant.

Exemple 1. Considérons le modèle (5.1) avec $g(x, \theta) = (x - \theta, x^2 - h(\theta))^T$, où h est une fonction connue. Dans ce cas, la vraie valeur de θ est la moyenne de X . Ces modèles contiennent toutes les distributions pour lesquelles le moment d'ordre deux peut être écrit comme fonction du moment d'ordre un. Plusieurs lois usuelles dans l'analyse de survie appartiennent à cette classe de distributions, comme la loi exponentielle et la loi de Rayleigh. Le fait de

considérer le modèle semi-paramétrique (5.1), au lieu d'un modèle complètement paramétrique, mène à des estimateurs plus robustes de la moyenne de X .

Nous renvoyons le lecteur à He et al. (2015) pour plus d'exemples d'application du modèle (5.1) pour les considérations de l'analyse de survie.

La méthode EL a été adaptée au cas des données censurées par Wang et Jing (2001). Ils ont remplacé, dans la formule de la vraisemblance empirique, le $i^{\text{ème}}$ élément de l'échantillon X_i , par une quantité observable et qui a la même espérance. Par ailleurs, Broniatowski et Keziou (2012) ont donné un autre point de vue de la vraisemblance empirique. En effet, utiliser la méthode EL revient à projeter la mesure empirique de l'échantillon sur le modèle, en choisissant la divergence de Kullback-Leibler modifiée pour évaluer les distances entre les mesures. Dans ce travail, nous suivons ce même point de vue et nous estimons θ en projetant la mesure empirique de Kaplan-Meier sur le modèle. Ceci donne un estimateur différent de celui de Wang et Jing (2001), mais une comparaison par une étude de simulation a montré que notre estimateur est plus performant.

Ce chapitre est organisé comme suit. La section 5.2 décrit la mesure empirique de Kaplan-Meier. Dans la section 5.3, nous introduisons les estimateurs du minimum de divergences empiriques. Dans la section 5.4, nous donnons les propriétés asymptotiques des estimateurs proposés. La section 5.5 présente des résultats de simulation.

5.2 La mesure empirique de Kaplan-Meier

Soit X une durée de survie de fonction de répartition continue F . Nous supposons que X est censurée à droite par une v.a. positive R , indépendante de X . Donc au lieu d'observer X , nous disposons d'un échantillon de n v.a. indépendantes et de même loi que le couple (Z, Δ) , où $Z := \min(X, R)$ et

$\Delta := 1_{\{X \leq R\}}$.

Rappelons que dans ce cas F peut être estimée par l'estimateur de Kaplan-Meier suivant

$$F_n^R(x) = 1 - S_n(x) := 1 - \prod_{i/Z'_i \leq x} \left(1 - \frac{D(Z'_i)}{U(Z'_i)}\right),$$

où $(Z'_i)_{1 \leq i \leq m}$ ($m \leq n$) sont les valeurs distinctes des $(Z_i)_{1 \leq i \leq n}$, $D(Z'_i) := \sum_{j=1}^n \Delta_j 1_{\{Z_j = Z'_i\}}$ et $U(Z'_i) := \sum_{j=1}^n 1_{\{Z_j \geq Z'_i\}}$. La mesure empirique de la fonction de répartition F_n^R peut être écrite comme suit

$$\mathbb{P}_n^R(\cdot) := \sum_{i=1}^m \Delta F_n^R(Z'_i) \delta_{Z'_i}(\cdot) = \sum_{i=1}^n \frac{\Delta F_n^R(Z_i)}{D(Z_i)} \delta_{Z_i}(\cdot),$$

avec la convention $\frac{0}{0} := 0$, $\delta_x(\cdot)$ désigne la mesure de Dirac au point x , pour tout x .

Soit $S_R^{(n)}$ l'estimateur produit-limite, de S_R , donné par

$$S_R^{(n)}(x) := \prod_{i/Z'_i \leq x} \left(1 - \frac{\bar{D}(Z'_i)}{U(Z'_i)}\right),$$

où $\bar{D}(Z'_i) := \sum_{j=1}^n (1 - \Delta_j) 1_{\{Z_j = Z'_i\}}$.

Puisque pour tout x

$$S_n(x) S_R^{(n)}(x) = \frac{1}{n} \sum_{j=1}^n 1_{\{Z_j > x\}} =: \bar{H}_n(x),$$

nous obtenons

$$\begin{aligned} \Delta F_n^R(Z_i) &= S_n(Z_i^-) \frac{D(Z_i)}{U(Z_i)} \\ &= S_n(Z_i^-) \frac{D(Z_i)}{n \bar{H}_n(Z_i^-)} \\ &= \frac{D(Z_i)}{n S_R^{(n)}(Z_i^-)}. \end{aligned}$$

D'où

$$\frac{\Delta F_n^R(Z_i)}{D(Z_i)} = \frac{\Delta_i}{nS_R^{(n)}(Z_i^-)}$$

et

$$\mathbb{P}_n^R(\cdot) = \sum_{i=1}^n \frac{\Delta_i}{nS_R^{(n)}(Z_i^-)} \delta_{Z_i}(\cdot).$$

5.3 Les estimateurs du minimum de divergence

Notons M l'espace de toutes les mesures finies et signées (m.f.s.) sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et considérons le modèle statistique suivant.

$$\mathcal{M} := \bigcup_{\theta \in \Theta} \mathcal{M}_\theta,$$

où

$$\mathcal{M}_\theta := \left\{ Q \in M \text{ telle que } \int_{\mathbb{R}} dQ(x) = 1 \text{ and } \int_{\mathbb{R}} g(x, \theta) dQ(x) = 0 \right\},$$

$g := (g_1, \dots, g_l)^\top \in \mathbb{R}^l$ étant une fonction vectorielle de $x \in \mathbb{R}$ et d'un paramètre $\theta \in \Theta \subset \mathbb{R}^d$.

Nous notons θ_V , si elle existe, la valeur du paramètre telle que (tq) $P_X \in \mathcal{M}_{\theta_V}$ (où P_X est la loi de probabilité de X). Nous allons construire des estimateurs de θ_V , basés sur la notion des divergences et de la dualité, que nous rappelons dans la suite.

Soit φ une fonction convexe définie de \mathbb{R} dans $[0, +\infty]$ telle que son domaine $\text{dom}_\varphi := \{x \in \mathbb{R} \text{ tq } \varphi(x) < \infty\}$ est un intervalle de bornes $a_\varphi < 1 < b_\varphi$ (qui peut être borné ou non, ouvert ou non). Nous supposons que $\varphi(1) = 0$ et que φ est fermée (i.e., si a_φ ou b_φ est fini alors $\varphi(x) \rightarrow \varphi(a_\varphi)$, quand $x \downarrow a_\varphi$, et $\varphi(x) \rightarrow \varphi(b_\varphi)$, quand $x \uparrow b_\varphi$). Pour toute mesure de probabilité (m.p.) P sur l'espace probablisable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et pour toute m.f.s. $Q \in M$, la φ -divergence entre Q et P , lorsque Q est absolument continue par rapport

à (a.c.p.r.à) P , est définie par

$$D_\varphi(Q, P) := \int_{\mathbb{R}} \varphi \left(\frac{dQ}{dP}(x) \right) dP(x),$$

où $dQ/dP(\cdot)$ est la dérivée de Radon-Nikodym. Lorsque Q n'est pas a.c.p.r.à P , nous posons $D_\varphi(Q, P) := +\infty$. Pour toute m.p. P , l'application $Q \in M \mapsto D_\varphi(Q, P)$ est convexe, positive et vérifie $D_\varphi(P, P) = 0$. Si φ est strictement convexe au voisinage de $x = 1$, alors

$$D_\varphi(Q, P) = 0 \text{ si et seulement si } Q = P. \quad (5.2)$$

Dans toute la suite, nous supposons que φ est strictement convexe sur son domaine (a_φ, b_φ) et deux fois continument différentiable sur $]a_\varphi, b_\varphi[$, l'intérieur de son domaine. Donc, $\varphi'(1) = 0$, $\varphi''(x) > 0$ pour tout $x \in]a_\varphi, b_\varphi[$, et la propriété (5.2) est vérifiée.

Si Ω est une partie de M , nous définissons la φ -divergence, entre Ω est une m.p. P , par

$$D_\varphi(\Omega, P) := \inf_{Q \in \Omega} D_\varphi(Q, P).$$

Une mesure finie $Q^* \in \Omega$, telle que $D_\varphi(Q^*, P) < \infty$ et

$$D_\varphi(Q^*, P) \leq D_\varphi(Q, P) \text{ pour tout } Q \in \Omega,$$

est dite une projection de P sur Ω . Cette projection peut ne pas exister, et peut ne pas être unique.

Par ailleurs, les problèmes d'optimisation avec contraintes peuvent être transformés en des problèmes sans contraintes, et ceci grâce à la notion de la dualité de Lagrange et de la transformée de Fenchel-Legendre. La transformée de Fenchel-Legendre (ou la conjuguée convexe) de φ est définie par

$$\varphi^* : t \in \mathbb{R} \mapsto \varphi^*(t) := \sup_{x \in \mathbb{R}} \{tx - \varphi(x)\}.$$

La fonction φ^* est convexe et fermée, son domaine est un intervalle $(a_\varphi^*, b_\varphi^*)$ dont les bornes

$$a_\varphi^* := \lim_{x \rightarrow -\infty} \frac{\varphi(x)}{x} \quad \text{et} \quad b_\varphi^* := \lim_{x \rightarrow +\infty} \frac{\varphi(x)}{x}$$

vérifient $a_\varphi^* < 0 < b_\varphi^*$ avec $\varphi^*(0) = 0$.

La convexité stricte de φ sur son domaine (a_φ, b_φ) est équivalente au fait que sa conjuguée φ^* soit "essentiellement régulière", i.e., dérivable avec

$$\lim_{t \downarrow a_\varphi^*} \varphi^*(t) = -\infty \text{ si } a_\varphi^* \text{ est fini,}$$

$$\lim_{t \uparrow b_\varphi^*} \varphi^*(t) = +\infty \text{ si } b_\varphi^* \text{ est fini.}$$

Inversement, φ est essentiellement régulière sur son domaine (a_φ, b_φ) si et seulement si φ^* est strictement convexe sur son domaine $(a_\varphi^*, b_\varphi^*)$.

Dans la suite, nous supposons que φ est essentiellement régulière. Donc, φ^* est strictement convexe sur son domaine $(a_\varphi^*, b_\varphi^*)$, et il s'ensuit que

$$a_\varphi^* = \lim_{x \rightarrow -\infty} \frac{\varphi(x)}{x} = \lim_{x \downarrow a_\varphi} \frac{\varphi(x)}{x} = \lim_{x \downarrow a_\varphi} \varphi'(x), \quad b_\varphi^* = \lim_{x \rightarrow +\infty} \frac{\varphi(x)}{x} = \lim_{x \uparrow b_\varphi} \frac{\varphi(x)}{x} = \lim_{x \uparrow b_\varphi} \varphi'(x)$$

et

$$\varphi^*(t) = t\varphi'^{-1}(t) - \varphi(\varphi'^{-1}(t)) \text{ pour tout } t \in]a_\varphi^*, b_\varphi^*[,$$

où φ'^{-1} désigne la fonction inverse de φ' . De plus, φ^* est deux fois continument différentiable sur $]a_\varphi^*, b_\varphi^*[$ avec

$$\varphi^{*'}(t) = \varphi'^{-1}(t) \text{ et } \varphi^{*''}(t) = \frac{1}{\varphi''(\varphi'^{-1}(t))}.$$

En particulier, $\varphi^{*'}(0) = 1$ et $\varphi^{*''}(0) = 1$. Evidemment, puisque φ est supposée fermée, nous avons

$$\varphi(a_\varphi) = \lim_{x \downarrow a_\varphi} \varphi(x) \text{ et } \varphi(b_\varphi) = \lim_{x \uparrow b_\varphi} \varphi(x),$$

qui peuvent être finies ou non. De même, d'après la fermeture de φ^* , nous avons

$$\varphi^*(a_\varphi^*) = \lim_{t \downarrow a_\varphi^*} \varphi^*(t) \text{ et } \varphi^*(b_\varphi^*) = \lim_{t \uparrow b_\varphi^*} \varphi^*(t).$$

Finalement, la première et la deuxième dérivée de φ aux points a_φ et b_φ sont définies comme étant les limites de $\varphi'(x)$ et $\varphi''(x)$ quand $x \downarrow a_\varphi$ et quand $x \uparrow b_\varphi$. La première et la deuxième dérivée de φ^* aux points a_φ^* et b_φ^* sont définies de la même façon.

Le tableau suivant donne quelques exemples usuels de la fonction φ et sa conjuguée convexe φ^* .

La divergence	φ	dom_φ	φ^*	dom_{φ^*}
Kullback-Leibler	$\varphi(x) := x \log x - x + 1$	$[0, +\infty[$	$\varphi^*(t) := e^t - 1$	\mathbb{R}
Kullback-Leibler modifiée	$\varphi(x) := -\log x + x - 1$	$]0, +\infty[$	$\varphi^*(t) := -\log(1 - t)$	$] -\infty, 1[$
χ^2	$\varphi(x) := \frac{1}{2}(x - 1)^2$	\mathbb{R}	$\varphi^*(t) := \frac{1}{2}t^2 + t$	\mathbb{R}
χ^2 modifiée	$\varphi(x) := \frac{1}{2} \frac{(x-1)^2}{x}$	$]0, +\infty[$	$\varphi^*(t) := 1 - \sqrt{1 - 2t}$	$] -\infty, \frac{1}{2}[$
Hellinger	$\varphi(x) := 2(\sqrt{x} - 1)^2$	$[0, +\infty[$	$\varphi^*(t) := \frac{2t}{2-t}$	$] -\infty, 2[$

TABLE 5.1: Les divergences usuelles.

Soit θ une valeur donnée dans Θ , l'estimateur "plug-in" de $D_\varphi(\mathcal{M}_\theta, P_X)$ est

$$\widehat{D}_\varphi(\mathcal{M}_\theta, P_X) := \inf_{Q \in \mathcal{M}_\theta} D_\varphi(Q, \mathbb{P}_n^R).$$

Si la projection $Q_\theta^{(n)}$ de \mathbb{P}_n^R sur \mathcal{M}_θ existe, alors il est clair que $Q_\theta^{(n)}$ est a.c.p.r. à \mathbb{P}_n^R ; ceci veut dire que le support de $Q_\theta^{(n)}$ doit être inclus dans l'ensemble $\{Z_i, \text{ tq } i \in \{1, \dots, n\} \text{ et } \Delta_i = 1\}$. Considérons alors la partie suivante de M

$$\mathcal{M}_\theta^{(n)} := \left\{ Q \in M \text{ tq } Q \text{ a.c.p.r. à } \mathbb{P}_n^R, \sum_{i=1}^n Q(Z_i) = 1 \text{ et } \sum_{i=1}^n Q(Z_i)g(Z_i, \theta) = 0 \right\},$$

donc l'estimateur $\widehat{D}_\varphi(\mathcal{M}_\theta, P_X)$ peut être écrit comme suit

$$\widehat{D}_\varphi(\mathcal{M}_\theta, P_X) = \inf_{Q \in \mathcal{M}_\theta^{(n)}} D_\varphi(Q, \mathbb{P}_n^R). \quad (5.3)$$

Ce problème d'optimisation avec contraintes peut être transformé en un problème sans contraintes, en utilisant la conjuguée convexe φ^* de φ . Ceci est précisé dans la proposition suivante, qui est un cas particulier du Théorème 1 de Broniatowski et Keziou (2006).

Proposition 1. *S'il existe Q_0 dans $\mathcal{M}_\theta^{(n)}$ telle que*

$$a_\varphi < nS_R^{(n)}(Z_i^-)Q_0(Z_i) < b_\varphi \quad \text{pour tout } i = 1, \dots, n,$$

alors

$$\inf_{Q \in \mathcal{M}_\theta^{(n)}} D_\varphi(Q, \mathbb{P}_n^R) = \sup_{t \in \mathbb{R}^{1+l}} \left\{ t_0 - \frac{1}{n} \sum_{i=1}^n \varphi^* \left(t_0 + \sum_{j=1}^l t_j g_j(Z_i, \theta) \right) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} \right\}, \quad (5.4)$$

avec une atteinte duale. Réciproquement, s'il existe une solution optimale duale $\hat{t} := (\hat{t}_0, \hat{t}_1, \dots, \hat{t}_l)^\top \in \mathbb{R}^{1+l}$ telle que

$$a_\varphi^* < \hat{t}_0 + \sum_{j=1}^l \hat{t}_j g_j(Z_i, \theta) < b_\varphi^* \quad \text{pour tout } i = 1, \dots, n \text{ tel que } \Delta_i = 1,$$

alors l'égalité (5.4) est vérifiée, et l'unique projection de \mathbb{P}_n^R sur $\mathcal{M}_\theta^{(n)}$ est donnée par

$$Q_\theta^{(n)}(Z_i) = \frac{\Delta_i}{nS_R^{(n)}(Z_i^-)} \varphi'^{-1} \left(\hat{t}_0 + \sum_{j=1}^l \hat{t}_j g_j(Z_i, \theta) \right), \quad i = 1, \dots, n,$$

où $\hat{t} := (\hat{t}_0, \hat{t}_1, \dots, \hat{t}_l)^\top$ est solution du système d'équations

$$\begin{cases} 1 - \frac{1}{n} \sum_{i=1}^n \varphi'^{-1} \left(\hat{t}_0 + \sum_{j=1}^l \hat{t}_j g_j(Z_i, \theta) \right) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} = 0, \\ -\frac{1}{n} \sum_{i=1}^n g_j(Z_i, \theta) \varphi'^{-1} \left(\hat{t}_0 + \sum_{k=1}^l \hat{t}_k g_k(Z_i, \theta) \right) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} = 0, \quad j = 1, \dots, l. \end{cases}$$

Cette proposition correspond à la proposition 4.1. de Broniatowski et Keziou (2012) dans le cas des données complètes. La proposition 4.2. du même papier, qui est rappelée ci-dessous, concerne les contreparties asymptotiques.

Proposition 2. *Supposons que $\int_{\mathbb{R}} |g_j(x, \theta)| dP_X(x) < \infty$, pour tout $j = 1, \dots, l$. S'il existe Q_0 dans \mathcal{M}_θ avec $D_\varphi(Q_0, P_X) < \infty$ et*

$$a_\varphi < \inf_{x \in \mathbb{R}} \frac{dQ_0}{dP_X}(x) \leq \sup_{x \in \mathbb{R}} \frac{dQ_0}{dP_X}(x) < b_\varphi, \quad P_X - p.s.$$

(les inégalités strictes signifient que $P_X\{x \in \mathbb{R} \text{ tq } (dQ_0/dP_X)(x) \leq a_\varphi\} = P_X\{x \in \mathbb{R} \text{ tq } (dQ_0/dP_X)(x) \geq b_\varphi\} = 0$)

alors

$$\inf_{Q \in \mathcal{M}_\theta} D_\varphi(Q, P_X) = \sup_{t \in \mathbb{R}^{1+l}} \left\{ t_0 - \int_{\mathbb{R}} \varphi^* \left(t_0 + \sum_{j=1}^l t_j g_j(x, \theta) \right) dP_X(x) \right\}, \quad (5.5)$$

avec une atteinte duale. Réciproquement, s'il existe une solution optimale duale t^ qui est un point intérieur de l'ensemble*

$$\left\{ t \in \mathbb{R}^{1+l} \text{ tq } \int_{\mathbb{R}} \left| \varphi^* \left(t_0 + \sum_{j=1}^l t_j g_j(x, \theta) \right) \right| dP_X(x) < \infty \right\}$$

alors l'égalité (5.5) est vérifiée, et l'unique projection de P_X sur \mathcal{M}_θ est donnée par

$$\frac{dQ_\theta^*}{dP_X}(x) = \varphi'^{-1} \left(t_0^* + \sum_{j=1}^l t_j^* g_j(x, \theta) \right),$$

où $t^ := (t_0^*, t_1^*, \dots, t_l^*)^\top$ est solution du système d'équations*

$$\begin{cases} 1 - \int \varphi'^{-1} \left(t_0^* + \sum_{j=1}^l t_j^* g_j(x, \theta) \right) dP_X(x) = 0, \\ - \int g_j(x, \theta) \varphi'^{-1} \left(t_0^* + \sum_{j=1}^l t_j^* g_j(x, \theta) \right) dP_X(x) = 0, \quad j = 1, \dots, l. \end{cases}$$

De plus, t^ est unique si les fonctions $1_{\mathbb{R}}, g_1(\cdot, \theta), \dots, g_l(\cdot, \theta)$ sont linéairement indépendantes dans le sens où $P_X\{x \in \mathbb{R} \text{ tq } t_0 + \sum_{j=1}^l t_j g_j(x, \theta) \neq 0\} > 0$ pour tout $t \in \mathbb{R}^{1+l}$ avec $t \neq 0$.*

En tenant compte de ces propositions, nous allons redéfinir l'estimateur $\widehat{D}_\varphi(\mathcal{M}_\theta, P_X)$ donné dans (5.3). Soit

$$\bar{g} := (1_{\mathbb{R} \times \Theta}, g_1, \dots, g_l),$$

$$t^\top \bar{g}(x, \theta) := t_0 + \sum_{j=1}^l t_j g_j(x, \theta),$$

$$m(x, \theta, t) := t_0 - \varphi^*(t^\top \bar{g}(x, \theta)) \quad \text{pour tout } x \in \mathbb{R}, \theta \in \Theta \subset \mathbb{R}^d, t \in \mathbb{R}^{1+l},$$

$$\Lambda_\theta^{(n)} := \{t \in \mathbb{R}^{1+l} \text{ tq } a_\varphi^* < t^\top \bar{g}(Z_i, \theta) < b_\varphi^* \quad \text{pour tout } i = 1, \dots, n \text{ tq } \Delta_i = 1\}$$

et

$$\Lambda_\theta := \left\{ t \in \mathbb{R}^{1+l} \text{ tq } \int_{\mathbb{R}} \left| \varphi^* \left(t_0 + \sum_{j=1}^l t_j g_j(x, \theta) \right) \right| dP_X(x) < \infty \right\}.$$

Notons, pour toute m.p. P sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et pour toute fonction f intégrable par rapport à P

$$Pf := \int_{\mathbb{R}} f(x) dP(x).$$

Nous redéfinissons $\widehat{D}_\varphi(\mathcal{M}_\theta, P_X)$ comme suit

$$\widehat{D}_\varphi(\mathcal{M}_\theta, P_X) := \sup_{t \in \Lambda_\theta^{(n)}} \frac{1}{n} \sum_{i=1}^n m(Z_i, \theta, t) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} := \sup_{t \in \Lambda_\theta^{(n)}} \mathbb{P}_n^R m(\theta, t), \quad (5.6)$$

et nous estimons $D_\varphi(\mathcal{M}, P_X)$ et θ_V , par analogie au cas des données complètes (voir Broniatowski et Keziou (2012)), par

$$\widehat{D}_\varphi(\mathcal{M}, P_X) := \inf_{\theta \in \Theta} \sup_{t \in \Lambda_\theta^{(n)}} \frac{1}{n} \sum_{i=1}^n m(Z_i, \theta, t) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} := \inf_{\theta \in \Theta} \sup_{t \in \Lambda_\theta^{(n)}} \mathbb{P}_n^R m(\theta, t) \quad (5.7)$$

et

$$\widehat{\theta}_\varphi := \arg \inf_{\theta \in \Theta} \sup_{t \in \Lambda_\theta^{(n)}} \frac{1}{n} \sum_{i=1}^n m(Z_i, \theta, t) \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} := \arg \inf_{\theta \in \Theta} \sup_{t \in \Lambda_\theta^{(n)}} \mathbb{P}_n^R m(\theta, t) \quad (5.8)$$

respectivement.

5.4 Les propriétés asymptotiques des estimateurs étudiés

Dans le but d'étudier quelques propriétés asymptotiques des estimateurs que nous avons définis dans la section précédente, nous avons besoin d'appliquer une LFUGN dans le cas des données censurées à droite. Donc, nous allons d'abord montrer ce résultat. Pour cela, nous allons adapter le Théorème 7, pour obtenir la convergence uniforme en probabilité de $S_R^{(n)}$ sous l'hypothèse suivante.

Hypothèse 0. (a) $F_R(T_Z) < 1$; (b) il existe $p \in]0, 1/2[$ tel que

$$\limsup_{x \rightarrow T_Z} \frac{\left(\int_x^{T_Z} S_X dF_R \right)^{1-p}}{S_X(x)} < \infty.$$

Dans toute la suite, $\|v\| := \sup_i |v_i|$ dénote la norme d'un vecteur $v := (v_1, \dots, v_m)^\top \in \mathbb{R}$ et $\|A\| := \sup_{i,j} |a_{i,j}|$ dénote la norme d'une matrice $A := (a_{i,j})$.

Proposition 3. Soit $h : \mathbb{R}_+ \times \Theta \rightarrow \mathbb{R}^m$ une fonction mesurable. Sous l'Hypothèse 0, et si Θ est compact, $h(X, \theta)$ est continue en tout point $\theta \in \Theta$ avec la probabilité un et $E(\sup_{\theta \in \Theta} \|h(X, \theta)\|) < \infty$, alors on a

$$\sup_{\theta \in \Theta} \|\mathbb{P}_n^R h(\theta) - P_X h(\theta)\| \xrightarrow{P} 0, \text{ quand } n \rightarrow \infty.$$

Démonstration.

$$\text{Puisque } E\left(\frac{\Delta}{S_R(Z^-)} h(Z, \theta)\right) = E\left(\frac{h(X, \theta)}{S_R(X^-)} E(\Delta|X)\right) = E(h(X, \theta)) = P_X h(\theta),$$

nous obtenons

$$\begin{aligned}
\|\mathbb{P}_n^R h(\theta) - P_X h(\theta)\| &= \left\| \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} h(Z_i, \theta) - E \left(\frac{\Delta}{S_R(Z^-)} h(Z, \theta) \right) \right\| \\
&\leq \left\| \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} h(Z_i, \theta) - \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R(Z_i^-)} h(Z_i, \theta) \right\| \\
&\quad + \left\| \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R(Z_i^-)} h(Z_i, \theta) - E \left(\frac{\Delta}{S_R(Z^-)} h(Z, \theta) \right) \right\| \\
&=: \|A_n(\theta)\| + \|B_n(\theta)\|. \tag{5.9}
\end{aligned}$$

Pour traiter le terme $B_n(\theta)$, nous allons appliquer la LFUGN dans le cas des données complètes (voir le Lemme 2.4 de Newey et McFadden (1994)) sur la fonction $v(Z, \Delta, \theta) = \frac{\Delta}{S_R(Z^-)} h(Z, \theta)$, pour cela nous devons vérifier que $E(\sup_{\theta \in \Theta} \|v(Z, \Delta, \theta)\|) < \infty$. En effet,

$$E \left(\sup_{\theta \in \Theta} \|v(X, \Delta, \theta)\| \right) \leq \frac{1}{S_R(T_Z)} E \left(\sup_{\theta \in \Theta} \|h(X, \theta)\| \right) < \infty.$$

D'où

$$\sup_{\theta \in \Theta} \|B_n(\theta)\| \xrightarrow{P} 0. \tag{5.10}$$

Il reste à traiter le terme $A_n(\theta)$.

$$\begin{aligned}
\|A_n(\theta)\| &\leq \sup_{x \in [0, T_Z]} \left| \frac{1}{S_R^{(n)}(x)} - \frac{1}{S_R(x)} \right| \times \left\{ \frac{1}{n} \sum_{i=1}^n \|h(X_i, \theta)\| \right\} \\
&\leq \frac{\sup_{x \in [0, T_Z]} |S_R^{(n)}(x) - S_R(x)|}{S_R^{(n)}(T_Z) S_R(T_Z)} \times \left\{ \frac{1}{n} \sum_{i=1}^n \|h(X_i, \theta)\| \right\}. \tag{5.11}
\end{aligned}$$

En vertu de la LFUGN (appliquée sur la fonction $\|h(X, \theta)\|$), nous avons

$$\sup_{\theta \in \Theta} \left\{ \frac{1}{n} \sum_{i=1}^n \|h(X_i, \theta)\| \right\} = O_P(1) \tag{5.12}$$

et sous l'Hypothèse 0, nous pouvons appliquer le Théorème 7 sur $S_R^{(n)}$, en inversant le temps, pour avoir

$$\sup_{x \in [0, T_Z]} |S_R^{(n)}(x) - S_R(x)| = O_P(n^{-p}).$$

En combinant ceci avec (5.9), (5.10), (5.11), (5.12) et le fait que $P(S_R^{(n)}(T_Z) < S_R(T_Z)/2) \leq P(\sup_{x \in [0, T_Z]} |S_R^{(n)}(x) - S_R(x)| > S_R(T_Z)/2)$, nous obtenons le résultat visé. \square

Nous avons besoin également du résultat suivant, qui représente un TCL multidimensionnel dans le cas de la censure à droite.

Proposition 4. *Soit $h : \mathbb{R}_+ \times \Theta \rightarrow \mathbb{R}$ une fonction mesurable. Si $\theta_V \in \Theta$ vérifie $E(h(X, \theta_V)) = 0$, et si la matrice $\Sigma = (\sigma_{i,j})_{1 \leq i,j \leq m}$, où*

$$\sigma_{i,j} := \int (h_i(x, \theta_V) - \tilde{h}_i(x, \theta_V))(h_j(x, \theta_V) - \tilde{h}_j(x, \theta_V)) \frac{dF(x)}{S_R(x^-)},$$

\tilde{h}_i étant la transformée de Efron et Johnstone (1990) de h_i , définie par

$$\tilde{h}_i(x) := \frac{1}{1 - F(x)} \int_{]x, \infty[} h_i(u) dF(u),$$

existe et elle est inversible, alors $\mathbb{P}_n^R h(\theta_V)$ converge en loi vers un vecteur gaussien centré, de matrice de covariance Σ .

Démonstration.

Voir Zhou (2015), Lemme 21. \square

Dans la suite, nous allons étendre les résultats asymptotiques établis dans Broniatowski et Keziou (2012) au cas des données censurées à droite. Les hypothèses que nous allons considérer correspondent à celles supposées dans ce papier.

Hypothèse 1. (a) $P_X \in \mathcal{M}$ et $\theta_V \in \Theta$ est l'unique solution de $E(g(X, \theta)) = 0$; (b) $\Theta \subset \mathbb{R}^d$ est compact; (c) $g(X, \theta)$ est continue en tout point $\theta \in \Theta$ avec la probabilité un; (d) $E(\sup_{\theta \in \Theta} \|g(X, \theta)\|^\alpha) < \infty$ pour un certain $\alpha > 2$; (e) la matrice $\Omega := E(g(X, \theta_V)g(X, \theta_V)^\top)$ est inversible; (f) la matrice $V_1 := (v_{i,j}^{(1)})_{1 \leq i,j \leq l}$, où $v_{i,j}^{(1)} := \int (g_i(x, \theta_V) - \tilde{g}_i(x, \theta_V))(g_j(x, \theta_V) - \tilde{g}_j(x, \theta_V)) \frac{dF(x)}{S_R(x^-)}$,

est bien définie et inversible.

Les hypothèses 1a-1e sont supposées dans Broniatowski et Keziou (2012), et l'Hypothèse 1f permet d'appliquer la Proposition 4. Une condition suffisante qui garantit l'existence des termes $v_{i,j}^{(1)}$ est précisée dans Akritas (2000)

$$\int \frac{g_i^2(x, \theta_V)}{S_R(x^-)} dF(x) < \infty, \quad \forall i \in \{1, \dots, l\}.$$

Théorème 21. *Sous les Hypothèses 0 et 1, avec une probabilité approchant un (a.p.a.1) lorsque $n \rightarrow \infty$, l'estimateur $\hat{\theta}_\varphi$ existe, et converge vers θ_V en probabilité. $\frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(Z_i^-)} g(Z_i, \hat{\theta}_\varphi) = O_P(1/\sqrt{n})$, $\hat{t}(\hat{\theta}_\varphi) := \arg \sup_{t \in \Lambda_{\hat{\theta}_\varphi}^{(n)}} \mathbb{P}_n^R m(\hat{\theta}_\varphi, t)$ existe et appartient à $\text{int}(\Lambda_{\hat{\theta}_\varphi}^{(n)})$ a.p.a.1 lorsque $n \rightarrow \infty$, et $\hat{t}(\hat{\theta}_\varphi) = O_P(1/\sqrt{n})$.*

Démonstration.

Dans toute la suite, nous utilisons les notations suivantes : $\bar{g}_i(\theta) := \bar{g}(X_i, \theta)$, $b_i := \sup_{\theta \in \Theta} \|\bar{g}_i(\theta)\|$, $\hat{g}(\theta) := \mathbb{P}_n^R g(\theta)$ et $\hat{\Omega}(\theta) := n^{-1} \sum_{i=1}^n \bar{g}_i(\theta) \bar{g}_i(\theta)^\top \Delta_i / S_R^{(n)}(X_i^-)$.

Pour démontrer le Théorème 21, nous nous inspirons largement de la preuve du Théorème 3.1 de Newey et Smith (2004). Pour cela, nous avons besoin des lemmes suivants.

Lemme 2. *Sous l'Hypothèse 1, nous avons pour tous ζ tel que $1/\alpha < \zeta < 1/2$ et $\Lambda_n := \{t \in \mathbb{R}^{1+l} : \|t\| \leq n^{-\zeta}\}$, $\sup_{\theta \in \Theta, t \in \Lambda_n, 1 \leq i \leq n} |t^\top \bar{g}_i(\theta)| \xrightarrow{P} 0$ et a.p.a.1, $\Lambda_n \subseteq \Lambda_\theta^{(n)}$ pour tout $\theta \in \Theta$.*

Preuve.

Sous l'Hypothèse 1, nous avons $\max_{1 \leq i \leq n} b_i = O_P(n^{1/\alpha})$. En effet, soit

$\varepsilon > 0$, les inégalités de Markov et Hölder permettent d'écrire pour tout $\delta > 0$

$$\begin{aligned}
P\left(\max_{1 \leq i \leq n} b_i \geq \delta n^{1/\alpha}\right) &\leq E\left(\frac{\max_{1 \leq i \leq n} b_i}{\delta n^{1/\alpha}}\right) \\
&\leq \frac{1}{\delta} \left(E\left(\frac{(\max_{1 \leq i \leq n} b_i)^\alpha}{n}\right)\right)^{1/\alpha} \\
&= \frac{1}{\delta} \left(E\left(\frac{\max_{1 \leq i \leq n} \sup_{\theta \in \Theta} \|\bar{g}_i(\theta)\|^\alpha}{n}\right)\right)^{1/\alpha} \\
&\leq \frac{1}{\delta} \left(E\left(\frac{1}{n} \sum_{i=1}^n \sup_{\theta \in \Theta} \|\bar{g}(X_i, \theta)\|^\alpha\right)\right)^{1/\alpha} \\
&= \frac{1}{\delta} \left(E\left(\sup_{\theta \in \Theta} \|\bar{g}(X, \theta)\|^\alpha\right)\right)^{1/\alpha}
\end{aligned}$$

il suffit donc de prendre $\delta := \varepsilon^{-1} (E(\sup_{\theta \in \Theta} \|\bar{g}(X, \theta)\|^\alpha))^{1/\alpha}$.

En vertu de l'inégalité de Cauchy-Schwartz, nous avons

$$\sup_{\theta \in \Theta, t \in \Lambda_n, 1 \leq i \leq n} |t^\top \bar{g}_i(\theta)| \leq n^{-\zeta} \max_{1 \leq i \leq n} b_i = O_P(n^{-\zeta+1/\alpha}) \xrightarrow{P} 0,$$

ce qui prouve le premier résultat du lemme. De plus, puisque $0 \in]a_\varphi^*, b_\varphi^*[$, nous avons a.p.a.1 $t^\top \bar{g}_i(\theta) \in]a_\varphi^*, b_\varphi^*[$ pour tous $\beta \in \Theta$ et t tq $\|t\| \leq n^{-\zeta}$. \square

Lemme 3. *Supposons que les Hypothèses 0 et 1 sont vérifiées et considérons une suite $\bar{\theta} = \bar{\theta}_n$ de v.a. à valeurs dans Θ telle que $\bar{\theta} \xrightarrow{P} \theta_V$ et $\hat{g}(\bar{\theta}) = O_P(n^{-1/2})$. Nous avons $\bar{t} := \arg \max_{t \in \Lambda_{\bar{\theta}}^{(n)}} \mathbb{P}_n^R m(\bar{\theta}, t)$ existe a.p.a.1, $\bar{t} = O_P(n^{-1/2})$ et $\sup_{t \in \Lambda_{\bar{\theta}}^{(n)}} \mathbb{P}_n^R m(\bar{\theta}, t) = O_P(n^{-1})$.*

Preuve.

Posons $\bar{\Omega} := \hat{\Omega}(\bar{\theta})$ et $\tilde{\Omega} := E(\bar{g}(X, \theta_V) \bar{g}(X, \theta_V)^\top)$, nous avons

$$\begin{aligned}
\|\bar{\Omega} - \tilde{\Omega}\| &\leq \|\hat{\Omega}(\bar{\theta}) - E(\bar{g}(X, \bar{\theta}) \bar{g}(X, \bar{\theta})^\top)\| \\
&\quad + \|E(\bar{g}(X, \bar{\theta}) \bar{g}(X, \bar{\theta})^\top) - E(\bar{g}(X, \theta_V) \bar{g}(X, \theta_V)^\top)\| \\
&\leq \sup_{\theta \in \Theta} \|\hat{\Omega}(\theta) - E(\bar{g}(X, \theta) \bar{g}(X, \theta)^\top)\| \\
&\quad + \|E(\bar{g}(X, \bar{\theta}) \bar{g}(X, \bar{\theta})^\top) - E(\bar{g}(X, \theta_V) \bar{g}(X, \theta_V)^\top)\|
\end{aligned}$$

or

$$\sup_{\theta \in \Theta} \|\widehat{\Omega}(\theta) - E(\bar{g}(X, \theta)\bar{g}(X, \theta)^\top)\| \xrightarrow{P} 0,$$

d'après la Proposition 3 et

$$\|E(\bar{g}(X, \bar{\theta})\bar{g}(X, \bar{\theta})^\top) - E(\bar{g}(X, \theta_V)\bar{g}(X, \theta_V)^\top)\| \xrightarrow{P} 0,$$

par continuité de la fonction $\theta \mapsto E(\bar{g}(X, \theta)\bar{g}(X, \theta)^\top)$ (voir Newey et McFadden (1994) Lemme 2.4).

Nous déduisons alors que

$$\bar{\Omega} \xrightarrow{P} \widetilde{\Omega}.$$

Puisque $\widetilde{\Omega}$ est une matrice définie positive, la plus petite valeur propre de $\bar{\Omega}$ est minorée par un réel strictement positif a.p.a.1.

Par ailleurs, d'après le Lemme 2 et le fait que φ^* est deux fois continument différentiable sur $]a_\varphi^*, b_\varphi^*[$, nous déduisons que $\mathbb{P}_n^R m(\bar{\theta}, t)$ est deux fois continument différentiable sur Λ_n a.p.a.1. Donc $\tilde{t} := \arg \max_{t \in \Lambda_n} \mathbb{P}_n^R m(\bar{\theta}, t)$ existe a.p.a.1. De plus, le développement de Taylor de la fonction $\mathbb{P}_n^R m(\bar{\theta}, t)$ montre l'existence d'un point \hat{t} à l'intérieur du segment qui relie \tilde{t} et 0 tel que

$$0 = \mathbb{P}_n^R m(\bar{\theta}, 0) \leq \mathbb{P}_n^R m(\bar{\theta}, \tilde{t}) = \tilde{t}^\top \mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, 0) + \frac{1}{2} \tilde{t}^\top \mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\bar{\theta}, \hat{t}) \tilde{t}.$$

Grâce au Lemme 2 et le fait que $\varphi^{*\prime\prime}(0) = 1$, nous avons $\min_{1 \leq i \leq n} \varphi^{*\prime\prime}(\hat{t}^\top \bar{g}_i(\bar{\theta})) > 1/2$ a.p.a.1. Donc

$$\tilde{t}^\top \mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\bar{\theta}, \hat{t}) \tilde{t} \leq -\frac{1}{2} \tilde{t}^\top \bar{\Omega} \tilde{t}$$

et

$$0 \leq \mathbb{P}_n^R m(\bar{\theta}, \tilde{t}) \leq \tilde{t}^\top \mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, 0) - C \|\tilde{t}\|^2 \leq \|\tilde{t}\| \left\| \mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, 0) \right\| - C \|\tilde{t}\|^2, \quad (5.13)$$

où $C > 0$ est une constante générique.

D'où

$$\|\tilde{t}\| \leq \frac{1}{C} \left\| \mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, 0) \right\| = \frac{1}{C} \|\widehat{g}(\bar{\theta})\| = O_P(n^{-1/2}) = o_P(n^{-\zeta}).$$

Donc, a.p.a.1 $\tilde{t} \in \text{int}(\Lambda_n)$ et par conséquent $\mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, \tilde{t}) = 0$. D'après le Lemme 2, a.p.a.1 $\tilde{t} \in \Lambda_{\bar{\theta}}^{(n)}$, donc la concavité stricte de la fonction $t \mapsto \mathbb{P}_n^R m(\bar{\theta}, t)$ et la convexité de $\Lambda_{\bar{\theta}}^{(n)}$ montrent que $\mathbb{P}_n^R m(\bar{\theta}, \tilde{t}) = \sup_{t \in \Lambda_{\bar{\theta}}^{(n)}} \mathbb{P}_n^R m(\bar{\theta}, t)$, ce qui donne le premier et le deuxième résultat visé avec $\bar{t} = \tilde{t}$. Finalement, la relation (5.13) et le fait que $\|\mathbb{P}_n^R \frac{\partial}{\partial t} m(\bar{\theta}, 0)\| = \|\widehat{g}(\bar{\theta})\| = O_P(n^{-1/2})$ et $\|\tilde{t}\| = O_P(n^{-1/2})$ entraînent

$$\mathbb{P}_n^R m(\bar{\theta}, \bar{t}) = O_P(n^{-1}).$$

□

Lemme 4. *Sous les Hypothèses 0 et 1, nous avons $\|\widehat{g}(\widehat{\theta}_\varphi)\| = O_P(n^{-1/2})$.*

Preuve.

Posons $\underline{t} := -n^{-\zeta} \widehat{g}(\widehat{\theta}_\varphi) / \|\widehat{g}(\widehat{\theta}_\varphi)\|$ et $\tilde{t} := (0, \underline{t})^\top$, où ζ est comme dans le Lemme 2. D'après le Lemme 2, $\tilde{t} \in \Lambda_{\widehat{\theta}_\varphi}^{(n)}$ a.p.a.1, donc selon le développement de Taylor, il existe t à l'intérieur du segment qui relie \tilde{t} et 0 tel que

$$\mathbb{P}_n^R m(\widehat{\theta}_\varphi, \tilde{t}) = -\underline{t}^\top \widehat{g}(\widehat{\theta}_\varphi) + \frac{1}{2} \tilde{t}^\top \mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\widehat{\theta}_\varphi, t) \tilde{t}$$

et comme $\varphi^{*''}(0) = 1$, nous avons d'après le Lemme 2, a.p.a.1 $\max_{1 \leq i \leq n} \varphi^{*''}(t^\top \bar{g}_i(\widehat{\theta}_\varphi)) < 2$. Donc

$$\tilde{t}^\top \mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\widehat{\theta}_\varphi, t) \tilde{t} \geq -2 \tilde{t}^\top \widehat{\Omega}(\widehat{\theta}_\varphi) \tilde{t}.$$

De plus, en vertu de l'inégalité de Cauchy-Schwartz et de la proposition 3, nous avons

$$\widehat{\Omega}(\widehat{\theta}_\varphi) \leq \left(\frac{1}{n} \sum_{i=1}^n b_i^2 \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \right) I \xrightarrow{P} CI,$$

où I est la matrice carrée d'ordre $l+1$ dont tous les éléments valent 1. Donc la plus grande valeur propre de $\widehat{\Omega}(\widehat{\theta}_\varphi)$ est bornée et par conséquent

$$\mathbb{P}_n^R m(\widehat{\theta}_\varphi, \tilde{t}) \geq -\underline{t}^\top \widehat{g}(\widehat{\theta}_\varphi) - C \|\tilde{t}\|^2 = n^{-\zeta} \|\widehat{g}(\widehat{\theta}_\varphi)\| - C n^{-2\zeta}.$$

Grâce à la proposition 4, les hypothèses du lemme 3 sont vérifiées pour $\bar{\theta} = \theta_V$ et comme $(\hat{\theta}_\varphi, \hat{t}(\hat{\theta}_\varphi))$ est un point selle de $\mathbb{P}_n^R m(\theta, t)$, cette dernière relation et le lemme 3 donnent

$$n^{-\zeta} \|\hat{g}(\hat{\theta}_\varphi)\| - Cn^{-2\zeta} \leq \mathbb{P}_n^R m(\hat{\theta}_\varphi, \tilde{t}) \leq \mathbb{P}_n^R m(\hat{\theta}_\varphi, \hat{t}(\hat{\theta}_\varphi)) \leq \sup_{t \in \Lambda_{\theta_V}^{(n)}} \mathbb{P}_n^R m(\theta_V, t) = O_P(n^{-1}). \quad (5.14)$$

D'où

$$\|\hat{g}(\hat{\theta}_\varphi)\| \leq O_P(n^{\zeta-1}) + Cn^{-\zeta} = O_P(n^{-\zeta}) \quad (\zeta - 1 < -\zeta \text{ car } \zeta < 1/2).$$

Considérons maintenant une suite (ε_n) qui converge vers 0 et posons $\bar{t} := -\varepsilon_n \hat{g}(\hat{\theta}_\varphi)$ et $\bar{t} := (0, \bar{t})^\top$, nous avons $\bar{t} = o_P(n^{-\zeta})$ donc $\bar{t} \in \Lambda_n$ a.p.a.1. Donc comme dans la relation (5.14)

$$-\bar{t}^\top \hat{g}(\hat{\theta}_\varphi) - C\|\bar{t}\|^2 = (1 - \varepsilon_n C) \varepsilon_n \|\hat{g}(\hat{\theta}_\varphi)\|^2 = O_P(n^{-1})$$

et comme pour n assez grand, $(1 - \varepsilon_n C)$ est minoré par un réel strictement positif, il s'ensuit que $\varepsilon_n \|\hat{g}(\hat{\theta}_\varphi)\|^2 = O_P(n^{-1})$, ceci étant vrai pour toute suite (ε_n) convergente vers 0, nous déduisons que $\|\hat{g}(\hat{\theta}_\varphi)\| = O_P(n^{-1/2})$. \square

Nous sommes maintenant en mesure de démontrer le Théorème 21. Nous avons

$$\begin{aligned} \|P_X g(\hat{\theta}_\varphi)\| &\leq \|P_X g(\hat{\theta}_\varphi) - \hat{g}(\hat{\theta}_\varphi)\| + \|\hat{g}(\hat{\theta}_\varphi)\| \\ &\leq \sup_{\theta \in \Theta} \|\mathbb{P}_n^R g(\theta) - P_X g(\theta)\| + \|\hat{g}(\hat{\theta}_\varphi)\| \xrightarrow{P} 0, \end{aligned}$$

en vertu de la Proposition 3 et du Lemme 4.

Par ailleurs, d'après le Lemme 2.4 de Newey et McFadden (1994), la fonction $P_X g(\theta)$ est continue et comme θ_V est l'unique solution de l'équation $P_X g(\theta) = 0$, $\|P_X g(\theta)\|$ est minorée par un réel strictement positif en dehors de tout voisinage de θ_V . Donc $\hat{\theta}_\varphi$ doit appartenir à tout voisinage de θ_V a.p.a.1, i.e. $\hat{\theta}_\varphi \xrightarrow{P} \theta_V$, ce qui donne le premier résultat du Théorème. Le deuxième

résultat découle du Lemme 4. De plus, notons que les hypothèses du Lemme 3 sont vérifiées pour $\bar{\theta} = \hat{\theta}_\varphi$, donc les deux derniers résultats du Théorème 21 découlent du Lemme 3. \square

Nous avons besoin d'hypothèses supplémentaires pour établir la normalité asymptotique. Considérons les matrices

$$G := E(\partial g(X, \theta_V)/\partial \theta), \Sigma := (G^\top \Omega^{-1} G)^{-1}, H := \Sigma G^\top \Omega^{-1}$$

$$\text{et } P := \Omega^{-1} - \Omega^{-1} G \Sigma G^\top \Omega^{-1}.$$

Notons également $\underline{\hat{t}}(\hat{\theta}_\varphi) := (\hat{t}_1, \dots, \hat{t}_l)^\top$, où $\hat{t}_0, \hat{t}_1, \dots, \hat{t}_l$ sont les composantes du vecteur $\hat{t}(\hat{\theta}_\varphi)$.

Hypothèse 2. (a) $\theta_V \in \text{int}(\Theta)$; (b) avec la probabilité un, $g(X, \theta)$ est continuellement différentiable dans un voisinage N_{θ_V} de θ_V , et $E(\sup_{\theta \in N_{\theta_V}} \|\partial g(X, \theta)/\partial \theta\|) < \infty$; (c) $\text{rang}(G) = d$.

Théorème 22. *Sous les Hypothèses 0, 1 et 2, nous avons $\sqrt{n}(\hat{\theta}_\varphi - \theta_V, \underline{\hat{t}}(\hat{\theta}_\varphi))^\top$ converge en loi vers un vecteur gaussien centré, de matrice de covariance*

$$V := \begin{pmatrix} HV_1 H^\top & HV_1 P \\ PV_1 H^\top & PV_1 P \end{pmatrix}.$$

Démonstration.

Puisque $\hat{t}(\hat{\theta}_\varphi)$ maximise $\mathbb{P}_n^R m(\hat{\theta}_\varphi, t)$ et comme il appartient à $\text{int}(\Lambda_{\hat{\theta}_\varphi}^{(n)})$, nous avons d'après les conditions d'ordre un

$$\frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*'} \left(\hat{t}(\hat{\theta}_\varphi)^\top \bar{g}_i(\hat{\theta}_\varphi) \right) g(X_i, \hat{\theta}_\varphi) = 0,$$

a.p.a.1. De plus, le fait que $\mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\hat{\theta}_\varphi, \hat{t}(\hat{\theta}_\varphi)) \xrightarrow{P} -E(\bar{g}(X, \theta_V) \bar{g}(X, \theta_V)^\top)$ implique que $\mathbb{P}_n^R \frac{\partial^2}{\partial t^2} m(\hat{\theta}_\varphi, \hat{t}(\hat{\theta}_\varphi))$ est inversible a.p.a.1. Donc le théorème des fonctions implicites et le théorème de l'enveloppe donnent

$$\frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*'} \left(\hat{t}(\hat{\theta}_\varphi)^\top \bar{g}_i(\hat{\theta}_\varphi) \right) G_i(\hat{\theta}_\varphi)^\top \underline{\hat{t}}(\hat{\theta}_\varphi) = 0,$$

où $G_i(\theta) := \partial g(X_i, \theta) / \partial \theta$.

Donc

$$\begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*'} \left(\widehat{t}(\widehat{\theta}_\varphi)^\top \bar{g}_i(\widehat{\theta}_\varphi) \right) G_i(\widehat{\theta}_\varphi)^\top \widehat{t}(\widehat{\theta}_\varphi) \\ f(\widehat{\theta}_\varphi, \widehat{t}(\widehat{\theta}_\varphi)) \end{pmatrix} = 0, \quad (5.15)$$

où $f : \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R}^l$ est définie par

$$f(\theta, \underline{t}) := \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*'}(\widehat{t}_0 + \underline{t}^\top g(X_i, \widehat{\theta}_\varphi)) g(X_i, \theta).$$

Par un développement de Taylor, il existe $(\bar{\theta}, \bar{t})^\top$ à l'intérieur du segment qui relie $(\widehat{\theta}_\varphi, \widehat{t}(\widehat{\theta}_\varphi))^\top$ et $(\theta_V, 0)^\top$ tel que

$$\begin{aligned} f(\widehat{\theta}_\varphi, \widehat{t}(\widehat{\theta}_\varphi)) &= \varphi^{*'}(\widehat{t}_0) \widehat{g}(\theta_V) \\ &+ \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*'}(\widehat{t}_0 + \bar{t}^\top g(X_i, \widehat{\theta}_\varphi)) G_i(\bar{\theta})^\top \\ \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i}{S_R^{(n)}(X_i^-)} \varphi^{*''}(\widehat{t}_0 + \bar{t}^\top g(X_i, \widehat{\theta}_\varphi)) g(X_i, \widehat{\theta}_\varphi)^\top g(X_i, \bar{\theta}) \end{pmatrix}^\top \begin{pmatrix} \widehat{\theta}_\varphi - \theta_V \\ \widehat{t}(\widehat{\theta}_\varphi) \end{pmatrix}. \end{aligned}$$

Donc (5.15) implique que

$$0 = \begin{pmatrix} 0 \\ \varphi^{*'}(\widehat{t}_0) \widehat{g}(\theta_V) \end{pmatrix} + \bar{M} \begin{pmatrix} \widehat{\theta}_\varphi - \theta_V \\ \widehat{t}(\widehat{\theta}_\varphi) \end{pmatrix}, \quad (5.16)$$

où

$$\bar{M} := \begin{pmatrix} 0 & \sum_{i=1}^n \frac{\Delta_i}{n S_R^{(n)}(X_i^-)} \varphi^{*'}(\widehat{t}(\widehat{\theta}_\varphi)^\top \bar{g}_i(\widehat{\theta}_\varphi)) G_i(\widehat{\theta}_\varphi)^\top \\ \sum_{i=1}^n \frac{\Delta_i}{n S_R^{(n)}(X_i^-)} \varphi^{*'}(\widehat{t}_0 + \bar{t}^\top g_i(\widehat{\theta}_\varphi)) G_i(\bar{\theta}) & \sum_{i=1}^n \frac{\Delta_i}{n S_R^{(n)}(X_i^-)} \varphi^{*''}(\widehat{t}_0 + \bar{t}^\top g_i(\widehat{\theta}_\varphi)) g_i(\bar{\theta})^\top g_i(\widehat{\theta}_\varphi) \end{pmatrix},$$

avec $g_i(\theta) := g(X_i, \theta)$.

Puisque $\widehat{t}(\widehat{\theta}_\varphi) = O_P(n^{-1/2})$, on peut montrer comme dans le Lemme 2 que $\max_{1 \leq i \leq n} |\widehat{t}(\widehat{\theta}_\varphi)^\top \bar{g}_i(\widehat{\theta}_\varphi)| \xrightarrow{P} 0$. En combinant ceci avec la Proposition 3, il vient que $\bar{M} \xrightarrow{P} M$, où

$$M := \begin{pmatrix} 0 & G^\top \\ G & \Omega \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} -\Sigma & H \\ H^\top & P \end{pmatrix}.$$

En résolvant l'équation (5.16), nous obtenons

$$\begin{aligned}
\sqrt{n} \begin{pmatrix} \widehat{\theta}_\varphi - \theta_V \\ \widehat{t}(\widehat{\theta}_\varphi) \end{pmatrix} &= -\sqrt{n} \overline{M}^{-1} \begin{pmatrix} 0 \\ \varphi^{*'}(\widehat{t}_0) \widehat{g}(\theta_V) \end{pmatrix} \\
&= -\sqrt{n} M^{-1} \begin{pmatrix} 0 \\ \varphi^{*'}(\widehat{t}_0) \widehat{g}(\theta_V) \end{pmatrix} + o_P(1) \\
&= -\sqrt{n} \varphi^{*'}(\widehat{t}_0) \begin{pmatrix} H \\ P \end{pmatrix} \widehat{g}(\theta_V) + o_P(1). \tag{5.17}
\end{aligned}$$

Le résultat visé découle facilement de la Proposition 4 et du fait que $\varphi^{*'}(\widehat{t}_0) \xrightarrow{P} 1$. \square

5.5 Etude de simulation

Dans le but de présenter les performances de l'estimateur $\widehat{\theta}_\varphi$, nous allons mener une étude de simulation. Dans cette étude, nous considérons deux modèles définis par des contraintes linéaires de la forme $g(x, \theta) = (x - \theta, x^2 - h(\theta))^\top$, où h est une fonction connue. Dans ce cas, la vraie valeur du paramètre est $\theta_V = E(X)$. Nous comparons l'estimateur $\widehat{\theta}_\varphi$ avec l'estimateur EL de Wang et Jing (2001) défini par

$$\widehat{\theta}_{EL} := \arg \inf_{\theta \in \Theta} \sup_{(\lambda_1, \lambda_2) \in \mathbb{R}^2} \sum_{i=1}^n \log \{1 + \lambda_1(V_{ni} - \theta_V) + \lambda_2(W_{ni} - h(\theta_V))\},$$

où

$$V_{ni} := \frac{Z_i \Delta_i}{S_R^{(n)}(Z_i)}$$

et

$$W_{ni} := \frac{h(Z_i) \Delta_i}{S_R^{(n)}(Z_i)}$$

et avec l'estimateur "plug-in" de θ_V donné par

$$\tilde{\theta} := \int Z d\mathbb{P}_n^R = \frac{1}{n} \sum_{i=1}^n Z_i \frac{\Delta_i}{S_R^{(n)}(Z_i^-)}.$$

Les divergences utilisées sont celles de Kullback-Leibler (KL), Kullback-Leibler modifiée (KL_m), Hellinger (H) et χ^2 . Ces divergences sont définies au tableau 5.1.

Nous générons des échantillons de taille 100 des variables latentes et nous calculons les estimateurs $\widehat{\theta}_\varphi$, $\widehat{\theta}_{EL}$ et $\widetilde{\theta}$ qui correspondent à chaque échantillon. Nous calculons ensuite, la moyenne des valeurs obtenues pour chaque estimateur, sur la base de 1000 réplifications, afin de la comparer avec la vraie valeur θ_V . Nous calculons également l'erreur quadratique moyenne (MSE) de chacun des estimateurs.

Exemple 2. Considérons le modèle $\mathcal{M} := \bigcup_{\theta \in \Theta} \mathcal{M}_\theta$, où

$$\mathcal{M}_\theta := \left\{ Q \in M \text{ tq } \int_{\mathbb{R}} dQ(x) = 1 \text{ et } \int_{\mathbb{R}} g(x, \theta) dQ(x) = 0 \right\},$$

avec $\Theta :=]0, \infty[$ et $g(x, \theta) := (x - \theta, x^2 - 2\theta^2)^\top$.

La loi exponentielle de paramètre a (notée $\mathcal{E}(a)$) appartient à ce modèle, avec la vraie valeur $\theta_V = 1/a$.

Partant d'une variable aléatoire X de loi $\mathcal{E}(a)$ et censurée à droite par une variable aléatoire R , de loi $\mathcal{E}(b)$, nous avons calculé les estimateurs $\widehat{\theta}_\varphi$, $\widehat{\theta}_{EL}$ et $\widetilde{\theta}$ comme décrit ci-dessus et nous avons obtenu les résultats suivants. Nous avons pris des valeurs différentes des paramètres a et b pour avoir des taux de censures différents.

a	4	7	3	2
b	1	3	2	3
Taux de censure	20%	30%	40%	60%
θ_V	0.25	0.1429	0.3333	0.5
$\widehat{\theta}_{KL}$	0.2408	0.1328	0.3015	0.4444
$\widehat{\theta}_{KL_m}$	0.2417	0.1335	0.3020	0.4379
$\widehat{\theta}_H$	0.2412	0.1335	0.3042	0.4546
$\widehat{\theta}_{\chi^2}$	0.2402	0.1324	0.3006	0.4225
$\widehat{\theta}_{EL}$	0.2436	0.1333	0.2941	0.4018
$\widetilde{\theta}$	0.2431	0.1336	0.2976	0.3332
$MSE(\widehat{\theta}_{KL})$	2.9767×10^{-4}	2.6034×10^{-4}	1.4245×10^{-3}	3.7850×10^{-3}
$MSE(\widehat{\theta}_{KL_m})$	2.4860×10^{-4}	2.5693×10^{-4}	1.4041×10^{-3}	4.5210×10^{-3}
$MSE(\widehat{\theta}_H)$	2.9367×10^{-4}	2.5282×10^{-4}	1.2803×10^{-3}	3.7817×10^{-3}
$MSE(\widehat{\theta}_{\chi^2})$	3.0235×10^{-4}	2.5987×10^{-4}	1.4687×10^{-3}	6.1722×10^{-3}
$MSE(\widehat{\theta}_{EL})$	2.8988×10^{-4}	3.5163×10^{-4}	2.0980×10^{-3}	1.0121×10^{-2}
$MSE(\widetilde{\theta})$	9.4864×10^{-4}	4.8119×10^{-4}	4.3519×10^{-3}	3.9879×10^{-2}

TABLE 5.2: Les résultats obtenus pour le modèle exponentiel.

Pour une taille d'échantillon modérée, nous remarquons que les estimateurs $\widehat{\theta}_\varphi$ ont une bonne performance. L'estimateur $\widehat{\theta}_{KL_m}$ est meilleur que $\widehat{\theta}_{EL}$ et les estimateurs $\widehat{\theta}_\varphi$ et $\widehat{\theta}_{EL}$ sont meilleurs que l'estimateur plug-in $\widetilde{\theta}$.

Exemple 3. Considérons le modèle $\mathcal{M} := \bigcup_{\theta \in \Theta} \mathcal{M}_\theta$, où

$$\mathcal{M}_\theta := \left\{ Q \in M \text{ tq } \int_{\mathbb{R}} dQ(x) = 1 \text{ et } \int_{\mathbb{R}} g(x, \theta) dQ(x) = 0 \right\},$$

avec $\Theta :=]0, \infty[$ et $g(x, \theta) := (x - \theta, x^2 - 4\theta^2/\pi)^\top$.

La loi de Rayleigh de paramètre σ (notée $\mathcal{R}(\sigma)$) appartient à ce modèle, avec la vraie valeur $\theta_V = \sqrt{\pi/2} \times \sigma$.

Dans cet exemple X est distribuée selon la loi $\mathcal{R}(\sigma_1)$ et R est distribuée

selon la loi $\mathcal{R}(\sigma_2)$. Nous avons procédé de la même façon que dans l'exemple précédent et nous avons obtenu les résultats suivants.

σ_1	1	$\sqrt{3}$	$\sqrt{2}$	$\sqrt{3}$
σ_2	2	$\sqrt{7}$	$\sqrt{3}$	$\sqrt{2}$
Taux de censure	20%	30%	40%	60%
θ_V	1.2533	2.1708	1.7725	2.1708
$\widehat{\theta}_{KL}$	1.2463	2.1522	1.7448	2.1097
$\widehat{\theta}_{KL_m}$	1.2465	2.1523	1.7450	2.1077
$\widehat{\theta}_H$	1.2464	2.1527	1.7452	2.1229
$\widehat{\theta}_{\chi^2}$	1.2413	2.1520	1.7448	2.1051
$\widehat{\theta}_{EL}$	1.2304	2.1227	1.7136	2.1005
$\widetilde{\theta}$	1.2385	2.1045	1.6617	1.7113
$MSE(\widehat{\theta}_{KL})$	3.8637×10^{-4}	6.6201×10^{-4}	1.3477×10^{-3}	4.2393×10^{-3}
$MSE(\widehat{\theta}_{KL_m})$	3.8314×10^{-4}	6.5697×10^{-4}	1.3338×10^{-3}	4.4342×10^{-3}
$MSE(\widehat{\theta}_H)$	3.8448×10^{-4}	6.3557×10^{-4}	1.3228×10^{-3}	3.1199×10^{-3}
$MSE(\widehat{\theta}_{\chi^2})$	6.5856×10^{-4}	6.6469×10^{-4}	1.3479×10^{-3}	4.7033×10^{-3}
$MSE(\widehat{\theta}_{EL})$	1.0567×10^{-3}	2.3652×10^{-3}	3.5645×10^{-3}	4.9612×10^{-3}
$MSE(\widetilde{\theta})$	6.6769×10^{-3}	2.9672×10^{-2}	4.2394×10^{-2}	3.5667×10^{-1}

TABLE 5.3: Les résultats obtenus pour le modèle de Rayleigh.

D'après ces résultats, nous pouvons déduire les mêmes conclusions que l'exemple précédent.

Perspectives

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

- Montrer des résultats de convergence en moyenne quadratique similaires à ceux du chapitre 4 dans le cas de la censure double.
- Regarder l'application du résultat principal du chapitre 3 à d'autres types de données incomplètes (données censurées par intervalle, status data, ...).
- Etablir la convergence faible des processus associés aux estimateurs étudiés dans le même chapitre (plus forte que la normalité asymptotique).
- Mettre en œuvre une technique du choix de la fenêtre spécifique aux modèles de la censure double et la censure mixte.
- Etudier les tests de modèle et de la valeur du paramètre liés aux modèles semi-paramétriques traités au chapitre 5.
- Etudier ces modèles avec d'autres types de données censurées (censure double, censure mixte, ...).
- Comparer les estimateurs proposés dans ce cadre en termes d'efficacité au second ordre et de robustesse.
- Etudier les problèmes à deux échantillons relatifs à ce même genre de modèles.

Bibliographie

- Aalen, O. O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics*, 6, 701–726.
- Akritis, M. G. (2000). The central limit theorem under censoring. *Bernoulli*, 1109–1120.
- Berg, A. S. (2007). *Nonparametric function estimation with infinite order kernels and applications*, Phd Thesis, University of California, San Diego.
- Bertholon, H. (2001). *Une modélisation du vieillissement*, Phd Thesis, Université Joseph Fourier, Grenoble.
- Bih-Sheue, S. et Cheun-Der, L. (2004). The jump sizes of self-consistent estimators of the survival function under doubly censored data. *Journal of Statistical Research*, 38(2), 183–196.
- Billingsley, P. (1995). *Probability and measure*. John Wiley & Sons.
- Bitouzé, D., Laurent, B., et Massart, P. (1999). A Dvoretzky-Kiefer-Wolfowitz type inequality for the Kaplan-Meier estimator. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 35(6), 735–763.
- Blum, J. R. et Susarla, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. *In Multivariate Analysis 5* (P.R. Krishnaiah ed.), 213–222.

- Bochner, S. (1955). *Harmonic analysis and the theory of probability*. Univ. of California Press.
- Boukeloua, M. (2015). Rates of mean square convergence of density and failure rate estimators under twice censoring. *Statistics and Probability Letters*, 106, 121–128.
- Boukeloua, M. et Messaci F. (2016). Asymptotic normality of kernel estimators based upon incomplete data. *Journal of Nonparametric Statistics*, 28(3), 469–486.
- Bousquet, N., Bertholon, H., et Celeux, G. (2006). An alternative competing risk model to the Weibull distribution for modelling aging in lifetime data analysis. *Lifetime Data Analysis*, 12, 481–504.
- Breslow, N., et Crowley J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, 2(3), 437–453.
- Broniatowski, M. et Keziou, A. (2006). Minimization of φ -divergences on sets of signed measures. *Studia Scientiarum Mathematicarum Hungarica*, 43(4), 403–442.
- Broniatowski, M. et Keziou, A. (2009). Parametric estimation and tests through divergences and the duality technique. *Journal of Multivariate Analysis*, 100(1), 16–36.
- Broniatowski, M. et Keziou, A. (2012). Divergences and duality for estimation and test under moment condition models. *Journal of Statistical Planning and Inference*, 142(9), 2554–2573.
- Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari*, 4, 421–424.

- Chang, M. N. (1990). Weak convergence of a self-consistent estimator of the survival function with doubly censored data. *The Annals of Statistics*, 18(1), 391–404.
- Chang, J., Chen, S. X., et Chen, X. (2015). High dimensional generalized empirical likelihood for moment restrictions with dependent data. *Journal of Econometrics*, 185(1), 283–304.
- Chang, M. N., et Yang G. L. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *The Annals of Statistics*, 15(4), 1536–1547.
- Chen K., et Lo S;-H. (1997). On the rate of uniform convergence of the product-limit estimator : strong and weak laws. *The Annals of Statistics*, 25(3), 1050–1087.
- Chung, K.-L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Transactions of the American Mathematical Society*, 67(1), 36–50.
- Corcoran, S. (1998). Bartlett adjustment of empirical discrepancy statistics. *Biometrika*, 967–972.
- Deheuvels, P. (1991). Laws of the iterated logarithm for density estimators. *In G. Roussas, éditeur : Nonparametric Functional estimators and Related Topics, NATO adv.Sci.Ins. C*, 19–29. Kluwer Academic.
- Devroye, L. (1992). A note on the usefulness of superkernels in density estimation. *Ann. Stat.*, 20(4), 2037–2056.
- Diehl, S., et Stute, W. (1988). Kernel density and hazard function estimation in the presence of censoring. *Journal of Multivariate Analysis*, 25, 299–310.

- Donsker M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *The Annals of Mathematical Statistics*, 23, 277–281.
- Dvoretzky, A., Kiefer, J., et Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pages 642–669.
- Efron, B. et Johnstone, I. M. (1990). Fisher's information in terms of the hazard rate. *The Annals of Statistics*, pages 38–62.
- Fang, K.-T., Li, G., Lu, X., et Qin, H. (2013). An empirical likelihood method for semiparametric linear regression with right censored data. *Computational and mathematical methods in medicine*, DOI : 10.1155/2013/469373.
- Földes, A., et Rejtő, L. (1981a). A LIL type result for the product limit estimator. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 56, 75–86.
- Földes, A., et Rejtő, L. (1981b). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *The Annals of Statistics*., 9, 122–129.
- Földes, A., Rejtő, L., et Winter, B. B. (1980). Strong consistency properties of nonparametric estimators for randomly censored data, I : The product-limit estimator. *Periodica Mathematica Hungarica*, 11, 233–250.
- Földes, A., Rejtő, L., et Winter, B. B. (1981). Strong consistency properties of nonparametric estimators for randomly censored data, II : Estimation of density and failure rate. *Periodica Mathematica Hungarica*, 12(1), 15–29.
- Gannoun, A. et Saracco, J. (2002). A new proof of strong consistency of kernel estimation of density function and mode under random censorship. *Statist. Probab. Lett.*, 59, 61–66.

- Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. *The Annals of Statistics*, 11(1), 49–58.
- Glivenko, V. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari*, 4, 92–99.
- Gu, M. G. et Zhang, C.-H. (1993). Asymptotic properties of self-consistent estimators based on doubly censored data. *The Annals of Statistics*, 21(2), 611–624.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica : Journal of the Econometric Society*, 1029–1054.
- He, S. Y. et Liang, W. (2014). Empirical likelihood for right censored data with covariables. *Science China Mathematics*, 57(6), 1275–1286.
- He, S., Liang, W., Shen, J. et Yang, G. (2015). Empirical likelihood for right censored lifetime data. *Journal of the American Statistical Association*. DOI : 10.1080/01621459.2015.1024058.
- Huang, J. (1999). Asymptotic properties of nonparametric estimation based on partly intervalcensored data. *Statistica Sinica*, 9, 501–519.
- Jeong, J.-H. (2014). *Statistical inference on residual life*. Springer.
- Kagba, N. S. (2004). *On kernel density estimation for censored data*, Phd Thesis, University of California, San Diego.
- Kaplan, E. L., et Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, 53(282), 457–481.

- Karunamuni, R. J. et Yang, S. (1991). Weak and strong uniform consistency rates of kernel density estimates for randomly censored data. *The Canadian Journal of Statistics*, 19(4), 349–359.
- Keziou, A. (2003). Dual representation of ϕ -divergences and applications. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 336(10), 857–862.
- Keziou, A. et Leoni-Aubin, S. (2008). On empirical likelihood for semiparametric two-sample density ratio models. *Journal of Statistical Planning and Inference*, 138(4), 915–928.
- Kiefer, J. (1961). On large deviations of the empiric df of vector chance variables and a law of the iterated logarithm. *Pacific J. Math*, 11(3), 649–660.
- Kitouni, A., Boukeloua, M., et Messaci, F. (2015). Rate of strong consistency for nonparametric estimators based on twice censored data. *Statistics & Probability Letters*, 96, 255–261.
- Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari*, 4, 83–91.
- Leiderman, P. H., Babu, D., Kagia, J., Kraemer, H. C. et Leiderman, G. F. (1973). African infant precocity and some social influences during the first year. *Nature*, 242, 247–249.
- Louani, D. (1998). On the asymptotic normality of the kernel estimators of the density function and its derivatives under censoring. *Communications in Statistics - Theory and Methods*, 27(12), 2909–2924.
- Marron, J. S. et Padgett, W. J. (1987). Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples. *The Annals of Statistics*, 15(4), 1520–1535.

- Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *The Annals of Probability*, 18(3), 1269–1283.
- Messaci, F., et Nemouchi, N. (2011). A law of the iterated logarithm for the product limit estimator with doubly censored data. *Statistics & Probability Letters*, 81(8), 1241–1244.
- Messaci, F., et Nemouchi, N. (2013). Erratum to “A law of the iterated logarithm for the product limit estimator with doubly censored data” [Statist. Probab. Lett. 81 (2011) 1241–1244]. *Statistics & Probability Letters*, 83(9), 2142.
- Mielniczuk, J. (1986). Some asymptotic properties of kernel estimators of a density function in case of censored data. *The Annals of Statistics*, 14(2), 766–773.
- Morales, D., Pardo, L., Quesada, V. (1991). Bayesian survival estimation for incomplete data when the life distribution is proportionally related to the censoring time distribution. *Comm. Statist. Theory Methods.*, 20, 831–850. MR1131189.
- Muller, H.-G. et Wang, J.-L. (1994). Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics*, 50(1), 61–76.
- Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology*, 1, 27–52.
- Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14, 945–966.
- Newey, W. K. et McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4, 2111–2245.

- Newey, W. K. et Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica*, 72(1), 219–255.
- Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75, 237–249.
- Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18(1), 90–120.
- Owen, A. B. (2001). *Empirical Likelihood*. Chapman & Hall/CRC.
- Pan, X.-R. et Zhou, M. (2002). Empirical likelihood ratio in terms of cumulative hazard function for censored data. *Journal of Multivariate Analysis*, 80(1), 166–188.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33, 1065–1076.
- Patilea, V., et Rolin, J.-M. (2001). Product limit estimators of the survival function for doubly censored data. Rapport technique, Institut de Statistique, Université Catholique de Louvain.
- Patilea, V., et Rolin, J.-M. (2004). Product-limit estimators of the survival function with left or right censored data. Rapport technique, Institut de Statistique, Université Catholique de Louvain.
- Patilea, V., et Rolin, J.-M. (2006). Product limit estimators of the survival function with twice censored data. *The Annals of Statistics*, 34(2), 925–938.
- Peto, R. (1973). Experimental survival curves for interval censored data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 22(1), 86–91.

- Pollard, D. (1984). *Convergence of stochastic processes*. Springer-Verlag.
- Qin, J. et Lawless, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics*, 22(1), 300–325.
- Qin, G. et Zhao, Y. (2007). Empirical likelihood inference for the mean residual life under random censorship. *Statistics & probability letters*, 77(5), 549–557.
- Ren, J-J. (1997). On self-consistent estimators and kernel density estimators with doubly censored data. *Journal of Statistical Planning and Inference*, 64, 27–43.
- Ren, J-J. (2001). Weighted empirical likelihood ratio confidence intervals for the mean with censored data. *Annals of the Institute of Statistical Mathematics*, 53(3), 498–516.
- Rosenblatt, M. (1956). Remarks on some nonparametric estimates of density function. *The Annals of Mathematical Statistics*, 27, 832–837.
- Rosenblatt, M. (1971). Curve estimates. *The Annals of Mathematical Statistics*, 42(6), 1815–1842.
- Samuelson, S. O. (1989). Asymptotic theory for non-parametric estimators from doubly censored data. *Scandinavian Journal of Statistics*, 16, 1–21.
- Schuster, E. F. (1969). Estimation of a probability density and its derivatives. *Ann. Math. Statist.*, 40, 1187–1196.
- Shen, P.-S. (2011). Nonparametric estimators of the survival function with twice censored data. *Ann. Inst. Statist. Math.*, 63, 1207–1219.
- Shen, P.-S. (2012). Modified self-consistent estimators of the survival function with twice censored data. *J. Statist. Plann. Inference*, 142, 1549–1556.

- Shen, J., Yuen, K. C. et Liu, C. (2016). Empirical likelihood confidence regions for one-or two-samples with doubly censored data. *Computational Statistics & Data Analysis*, 93, 285–293.
- Silverman, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, 6(1), 177–184.
- Smirnov, N. (1939). Sur les écarts de la courbe de distribution empirique. *Recueil Mathématique [Matematicheskii Sbornik]*, 6(48)(1), 3–26.
- Smith, R. J. (1997). Alternative Semi-parametric Likelihood Approaches to Generalised Method of Moments Estimation. *The Economic Journal*, 107(441), 503–519.
- Smith, R. J. (2011). GEL criteria for moment condition models. *Econometric Theory*, 27(6), 1192–1235.
- Stute W., et Wang J.-L. (1993). The strong law under random censorship. *The Annals of Statistics*, 21(3), 1591–1607.
- Tanner, M. A. et Wong W. H. (1983). The estimation of the hazard function from randomly censored data by the kernel method. *The annals of statistics*, 11(3), 989–993.
- Tsai, W. Y. et Crowley J. (1985). A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency. *Ann. Statist.*, 13, 1317–1334.
- Turnbull, B. W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association*, 69, 169–173. MR0381120.

- Van Ryzin, J. (1969). On strong consistency of density estimates. *Ann. Math. Statist.*, 40, 1765–1772.
- van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press. Cambridge.
- Wang, Q.-H. et Jing, B.-Y. (2001). Empirical likelihood for a class of functionals of survival distribution with censored data. *Annals of the Institute of Statistical Mathematics*, 53(3), 517–527.
- Wang, Q.-H. et Wang, J.-L. (2001). Inference for the mean difference in the two-sample random censorship model. *Journal of Multivariate Analysis*, 79(2), 295–315.
- Watson, G. S., et Leadbetter, M. R. (1964). Hazard analysis I. *Biometrika*, 51 (1 and 2), 175–184.
- Winter, B. B., Földes, A., et Rejtő, L. (1978). Glivenko-Cantelli theorems for the product limit estimate. *Problems of Control and Information Theory*, 7, 213–225.
- Wu, T. T., Li, G. et Tang, C. (2015). Empirical Likelihood for Censored Linear Regression and Variable Selection. *Scandinavian Journal of Statistics*, 42(3), 798–812.
- Xiang, X. (1994). Law of the logarithm for density and hazard rate estimation for censored data. *J. Multivariate Anal.*, 49, 278–286.
- Zheng, J., Shen, J. et He, S. (2014). Adjusted empirical likelihood for right censored lifetime data. *Statistical Papers*, 55(3), 827–839.
- Zhou, M. (2015). *Empirical Likelihood Method in Survival Analysis*. CRC Press.

ملخص

في هذه الرسالة، نهتم بطريقة النواة لتقدير الكثافة و نسبة المجازفة استنادا إلى معطيات محجوبة. توجد عدة أنواع من الحجب، و دراستنا تختص بنماذج الحجب من اليمين و الحجب المضاعف و الحجب المزدوج. نبدأ بدراسة سياق عام للحجب يشمل كل هذه الأنواع و نبرهن على التقارب الطبيعي لمقدر الكثافة ذو النواة الذي نقوم باقتراحه. هذه النتيجة تسمح لنا باستنتاج التقارب الطبيعي لمقدرات الكثافة و نسبة المجازفة من أجل نماذج الحجب المذكورة أعلاه. نبرهن كذلك على التقارب وفق المعدل التربيعي لنفس المقدرات في حالة الحجب المزدوج، مع تحديد سرعة التقارب. في القسم الثاني من الرسالة، نقوم بدراسة نماذج نصف وسيطية معرفة بواسطة معادلات خطية. بفرض أن المتغير المدروس محجوب من اليمين، نستعمل نظرية التباعدات لانشاء مقدرات للوسيط محل الاهتمام. يتم عرض دراسات محاكات لظهار جودة مختلف المقدرات المدروسة.

الكلمات المفتاحية : معطيات محجوبة، المقدر ذو النواة، الكثافة، نسبة المجازفة، نماذج المعادلات الخطية، التباعدات.

Abstract

In this dissertation, we are interested in kernel estimation of the density and the failure rate for censored data. There exist several kinds of censorship and we focus on right, doubly and twice censored data models. We consider a general framework of censorship, including all these models, and we prove a result on the asymptotic normality of a kernel density estimator that we introduce. This result allows us to deduce the asymptotic normality of the density and failure rate estimates for the above-mentioned censorship models. We also establish the mean square convergence, with rates, of the same estimators in the case of twice censored data. In a second part of the dissertation, we study semiparametric models which verify linear constraints involving an unknown parameter. We assume that the variable of interest is right censored and we use the theory of divergences to construct estimates for the parameter of interest. Simulation studies are presented in order to illustrate the performances of the different studied estimators.

Keywords : censored data, kernel estimators, density, failure rate, linear constraints models, divergences.

Résumé

Dans cette thèse, nous nous intéressons à l'estimation par la méthode du noyau de la densité et du taux de hasard pour des données censurées. Il existe plusieurs types de censure et nous nous concentrons sur les modèles de censure à droite, censure double et censure mixte. Nous considérons un cadre général de censure qui englobe tous ces modèles, et nous prouvons un résultat de normalité asymptotique d'un estimateur à noyau de la densité que nous introduisons. Ce résultat nous permet de déduire la normalité asymptotique des estimateurs de la densité et du taux de hasard pour les modèles de censure cités ci-dessus. Nous établissons également la convergence en moyenne quadratique, avec taux, des mêmes estimateurs dans le cas de la censure mixte. Dans une deuxième partie de la thèse, nous étudions des modèles semi-paramétriques qui vérifient des contraintes linéaires faisant appel à un paramètre inconnu. Nous supposons que la variable d'intérêt est censurée à droite et nous utilisons la théorie des divergences pour construire des estimateurs pour le paramètre d'intérêt. Des études de simulation sont présentées pour illustrer les performances des différents estimateurs étudiés.

Mots-clés : données censurées, estimateurs à noyau, densité, taux de hasard, modèles de contraintes linéaires, divergences.