



HAL
open science

Contributions à l'estimation à noyau de fonctionnelles de la fonction de répartition avec applications en sciences économiques et de gestion

Soffana Madani

► **To cite this version:**

Soffana Madani. Contributions à l'estimation à noyau de fonctionnelles de la fonction de répartition avec applications en sciences économiques et de gestion. Gestion et management. Université de Lyon, 2017. Français. NNT : 2017LYSE1183 . tel-01691084

HAL Id: tel-01691084

<https://theses.hal.science/tel-01691084v1>

Submitted on 23 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2017LYSE1183

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale 486
Sciences Économiques et de Gestion

Spécialité de doctorat : Statistiques/Sciences Économiques

Soutenue publiquement le 29 Septembre 2017, par :
Soffana MADANI

**Contributions à l'estimation à noyau de
fonctionnelles de la fonction de répartition
avec applications en sciences économiques et
de gestion**

Devant le jury composé de :

REY-FOURNIER Béatrice, Professeure, Université Lumière Lyon 2 Présidente

CARDOT Hervé, Professeur, Université de Bourgogne Rapporteur
FLACHAIRE Emmanuel, Professeur, Université d'Aix-Marseille Rapporteur

JIAO Ying, Professeure, Université Claude Bernard Lyon 1 Examinatrice
VANHEMS Anne, Professeure, Toulouse Business School Examinatrice

NORBERG Ragnar, Professeur Émérite, London School of Economics Directeur de thèse
ROBERT Christian, Professeur, Université Claude Bernard Lyon 1 Co-directeur de thèse

Remerciements

J'avais rêvé de ce moment... Qu'il devienne réalité aujourd'hui m'émeut beaucoup. C'est le moment où je mets enfin un point d'orgue à ma thèse, un long chapitre de ma vie de presque trois années et demi. C'est également le moment où les souvenirs les plus éprouvants refont surface, comme les sacrifices qu'une thèse nécessite. Je suis tellement heureuse de pouvoir tourner la page, mais je tiens d'abord à mettre en avant toutes les personnes qui m'ont permis d'en arriver là.

Mes premiers remerciements vont à Monsieur Ragnar NORBERG, mon directeur de thèse. Je n'oublierai jamais ce moment où vous m'avez proposé de m'encadrer, quelle chance ! Vous représentez la gentillesse, la bienveillance, la patience, l'humilité... merci d'avoir cru en moi. Merci aussi de montrer comment la recherche peut être belle. Cette aventure avec vous a été riche et vraiment particulière : les discussions techniques où vous parlez en anglais et moi en français, les déplacements entre Lyon et Londres comme on peut, l'annonce de votre cancer du pancréas il y a presque un an et demi, l'annonce d'un miracle pour votre tumeur, ou encore nos échanges de mails, votre aide et vos encouragements malgré la maladie. Vous êtes un exemple de force et de courage, une inspiration et comme je vous l'ai dit, je vous considère comme mon grand-père.

Je suis très honorée que Messieurs Hervé CARDOT et Emmanuel FLACHAIRE aient accepté de rapporter ma thèse, surtout en saison estivale. Je souhaite leur adresser ma profonde gratitude. Leur point de vue et leurs remarques sont précieux. Je suis très touchée que Madame Béatrice REY-FOURNIER ait accepté de présider le jury de ma soutenance de thèse et je tiens à la remercier chaleureusement. Je suis très heureuse que Mesdames Ying JIAO et Anne VANHEMS fassent également partie de mon jury.

J'aimerais maintenant témoigner ma reconnaissance aux personnes sans qui la soutenance n'aurait pu voir le jour. À Madame Véronique MAUME-DESCHAMPS, pour avoir pris le relais de mon directeur de thèse à des moments cruciaux, notamment avant la soumission de ma thèse. J'ai été très touchée par votre aide. À Monsieur Christophe EVERAERE pour son écoute attentive et sa grande humanité. À Madame Marie-France JOUBERT et Monsieur Jean-Marc LANCELIN pour s'être souciés du bien-être des doctorants, merci du fond du cœur. À Madame Gisèle BARBIER pour son aide et la facilitation des démarches administratives, surtout ces derniers mois, et également à Mesdames Nathalie ROGLIN et Fabienne MACRO pour faciliter l'inscription des doctorants chaque année à l'Université. À Madame RÉGINVAL pour sa compréhension et pour m'avoir permis de me sentir mieux et à Madame BLAISOT qui m'a encouragée dans mon projet professionnel.

J'ai eu l'occasion de rencontrer à l'Université des personnes que je considère au-

jourd'hui comme des proches. Je pense à Pierre RIBEREAU, Esterina MASIELLO et Areski COUSIN. Merci pour votre altruisme, vos conseils et vos encouragements. Merci aussi d'égayer ma vie, que ce soit pour découvrir de nouveaux restaurants ou discuter autour d'un bon verre. J'espère que nous resterons en contact encore longtemps.

Un merci spécial à Madame Patricia BARTOLO pour sa bonne humeur et son aide pour mes impressions, vous m'avez sauvé la vie ! Merci à Monsieur Jean-Daniel HUET d'avoir toujours résolu mes problèmes informatiques avec gentillesse. Merci à Madame Delphine CLEMENS et Monsieur Samy BRIAS pour leur présence agréable à l'ISFA.

Je souhaite bon courage et bonne chance aux doctorants qui vont bientôt terminer leur thèse : Jean de Dieu, Saker, Antoine, Gwladys. Je remercie les anciens de m'avoir montré la voie, ainsi que pour leurs discussions intéressantes et toujours positives, en particulier Matthieu, Romain, Viou, Huy, Khalil, Joachim, la team Montpel avec Rémi, Olivier et bien d'autres. Un message à Manel et Sawssen, je vous remercie pour tous les bons moments passés ensemble, votre optimisme et je vous souhaite beaucoup de joie, ainsi qu'à vos proches.

Je remercie l'univers dans lequel je travaille depuis tellement d'années, l'enseignement. Je veux exprimer ma gratitude à Messieurs MONDÉSERT et MÜLLNER qui m'ont témoigné leur confiance. Un merci tout particulier à Monsieur MÉLINAND qui m'a arrangé mon emploi du temps afin de pouvoir terminer ma thèse dans les meilleures conditions. J'ai rencontré des collègues formidables, je pense à Mado, Danièle, Christine,... et des étudiants et élèves qui m'offrent des moments de bonheur au quotidien. Concilier la thèse et les enseignements parfois à plein temps pouvait être compliqué mais c'est grâce à eux que j'ai pu garder le moral tout au long de la thèse.

J'ai tellement de chance d'avoir rencontré des amis à Lyon et j'aimerais leur dire merci pour tous les moments que l'on partage ensemble, toujours dans la joie. J'ai conscience que nos amitiés sont précieuses et j'espère qu'elles dureront. Abdou et Elmira, je suis vraiment touchée par votre gentillesse et votre écoute, ainsi que votre bienveillance. Merci pour les Tcheps (tellement pimentés!!!), du pur bonheur à chaque fois, et pour les sorties cinéma. Geoffroy, Muriel et Pierre, nous ne parlons jamais du boulot et c'est tellement bien ! Ne changez rien. Merci de m'avoir intégrée bien que nous ne supportons pas la même équipe de football !

Mes remerciements vont désormais à mes parents pour leur mélange rare, pour les valeurs qu'ils m'ont inculquées et aussi parce que je sais qu'ils ont toujours voulu faire de leur mieux. Merci à mes frères de partager les mêmes racines et de pouvoir comprendre ce que je vis.

Un énorme merci à Christine et Bernard pour leurs sourires, leur écoute et leur réconfort, pour les discussions profondes parfois. Merci pour votre soutien, pour m'offrir un lieu de ressourcement où j'ai à chaque fois l'impression que le temps s'arrête, en compagnie de Marie, Davy et Kemi. Merci à vos familles respectives qui m'ont tellement bien accueillie, j'ai beaucoup de chance.

Ma Jacky, 12 années que tu es à mes côtés, malgré la distance qui nous sépare. Merci d'être ce que tu es, de me soutenir, tu es une vraie amie.

Enfin, je souhaiterais conclure avec une personne qui a joué un immense rôle dans l'aboutissement de mon travail. Une personne qui m'a permis de libérer du temps libre quand il m'en manquait, une personne qui a été d'un très grand réconfort lorsque les doutes s'emparaient de moi... Une personne généreuse, bienveillante, attentionnée, et qui me fait tant rire. La vie semble tellement simple à tes côtés et je te remercie Samy pour le bonheur que tu m'apportes tous les jours.

"Nul ne peut atteindre l'aube sans passer par la nuit."

Khalil Gibran

"La bonté d'un seul homme est plus puissante que la méchanceté de mille."

Panaït Istrati

À Ragnar.

Résumé

La répartition des revenus d'une population, la distribution des instants de défaillance d'un matériel et l'évolution des bénéficiaires des contrats d'assurance vie - étudiées en sciences économiques et de gestion - sont liées à des fonctions continues appartenant à la classe des fonctionnelles de la fonction de répartition.

Notre thèse porte sur l'estimation à noyau de fonctionnelles de la fonction de répartition avec applications en sciences économiques et de gestion.

Dans le premier chapitre, nous proposons des estimateurs polynomiaux locaux dans le cadre i.i.d. de deux fonctionnelles de la fonction de répartition, notées L_F et T_F , utiles pour produire des estimateurs lisses de la courbe de Lorenz et du temps total de test normalisé (*scaled total time on test transform*). La méthode d'estimation est décrite dans Abdous, Berlinet et Hengartner (2003) et nous prouvons le bon comportement asymptotique des estimateurs polynomiaux locaux. Jusqu'alors, Gastwirth (1972) et Barlow et Campo (1975) avaient défini des estimateurs continus par morceaux de la courbe de Lorenz et du temps total de test normalisé, ce qui ne respectait pas la propriété de continuité des courbes initiales. Des illustrations sur données simulées et réelles sont proposées.

Le second chapitre a pour but de fournir des estimateurs polynomiaux locaux dans le cadre i.i.d. des dérivées successives des fonctionnelles de la fonction de répartition explorées dans le chapitre précédent. À part l'estimation de la dérivée première de la fonction T_F qui se traite à l'aide de l'estimation lisse de la fonction de répartition, la méthode d'estimation employée est l'approximation polynomiale locale des fonctionnelles de la fonction de répartition détaillée dans Berlinet et Thomas-Agnan (2004). Divers types de convergence ainsi que la normalité asymptotique sont obtenus, y compris pour la densité et ses dérivées successives. Des simulations apparaissent et sont commentées.

Le point de départ du troisième chapitre est l'estimateur de Parzen-Rosenblatt (Rosenblatt (1956), Parzen (1964)) de la densité. Nous améliorons dans un premier temps le biais de l'estimateur de Parzen-Rosenblatt et de ses dérivées successives à l'aide de noyaux d'ordre supérieur (Berlinet (1993)). Nous démontrons ensuite les nouvelles conditions de normalité asymptotique de ces estimateurs. Enfin, nous construisons une méthode de correction des effets de bord pour les estimateurs des dérivées de la densité, grâce aux dérivées d'ordre supérieur.

Le dernier chapitre s'intéresse au taux de hasard, qui contrairement aux deux fonctionnelles de la fonction de répartition traitées dans le premier chapitre, n'est pas un rapport de deux fonctionnelles linéaires de la fonction de répartition. Dans le cadre i.i.d., les estimateurs à noyau du taux de hasard et de ses dérivées successives sont construits à partir des estimateurs à noyau de la densité et ses dérivées successives. La normalité asymptotique des premiers estimateurs est logiquement obtenue à partir de celle des seconds. Nous nous plaçons ensuite dans le modèle à intensité multiplicative, un cadre plus général englobant des données censurées et dépendantes. Nous menons la procédure à terme de Ramlau-Hansen (1983) afin d'obtenir les bonnes propriétés asymptotiques des estimateurs du taux de hasard et de ses dérivées successives puis

nous tentons d'appliquer l'approximation polynomiale locale dans ce contexte. Le taux d'accumulation du surplus dans le domaine de la participation aux bénéfiques pourra alors être estimé non paramétriquement puisqu'il dépend des taux de transition (taux de hasard d'un état vers un autre) d'une chaîne de Markov (Ramlau-Hansen (1991), Norberg (1999)).

Mots Clés : estimation à noyau non paramétrique, espaces de Hilbert à noyau reproduisant, approximation polynomiale locale, fonctionnelles de la fonction de répartition, dérivées successives, densité, courbe de Lorenz, temps total de test normalisé, modèle à intensité multiplicative, taux de hasard, participation aux bénéfiques, taux d'accumulation du surplus, effets de bord.

Abstract

The income distribution of a population, the distribution of failure times of a system and the evolution of the surplus in with-profit policies - studied in economics and management - are related to continuous functions belonging to the class of functionals of the distribution function.

Our thesis covers the kernel estimation of some functionals of the distribution function with applications in economics and management.

In the first chapter, we offer local polynomial estimators in the i.i.d. case of two functionals of the distribution function, written L_F and T_F , which are useful to produce the smooth estimators of the Lorenz curve and the scaled total time on test transform. The estimation method is described in Abdous, Berline and Hengartner (2003) and we prove the good asymptotic behavior of the local polynomial estimators. Until now, Gastwirth (1972) and Barlow and Campo (1975) have defined continuous piecewise estimators of the Lorenz curve and the scaled total time on test transform, which do not respect the continuity of the original curves. Illustrations on simulated and real data are given.

The second chapter is intended to provide smooth estimators in the i.i.d. case of the derivatives of the two functionals of the distribution function presented in the last chapter. Apart from the estimation of the first derivative of the function T_F with a smooth estimation of the distribution function, the estimation method is the local polynomial approximation of functionals of the distribution function detailed in Berline and Thomas-Agnan (2004). Various types of convergence and asymptotic normality are obtained, including the probability density function and its derivatives. Simulations appear and are discussed.

The starting point of the third chapter is the Parzen-Rosenblatt estimator (Rosenblatt (1956), Parzen (1964)) of the probability density function. We first improve the bias of this estimator and its derivatives by using higher order kernels (Berline (1993)). Then we find the modified conditions for the asymptotic normality of these estimators. Finally, we build a method to remove boundary effects of the estimators of the probability density function and its derivatives, thanks to higher order derivatives.

We are interested, in this final chapter, in the hazard rate function which, unlike the two functionals of the distribution function explored in the first chapter, is not a fraction of two linear functionals of the distribution function. In the i.i.d. case, kernel estimators of the hazard rate and its derivatives are produced from the kernel estimators of the probability density function and its derivatives. The asymptotic normality of the first estimators is logically obtained from the second ones. Then, we are placed in the multiplicative intensity model, a more general framework including censored and dependent data. We complete the described method in Ramlau-Hansen (1983) to obtain good asymptotic properties of the estimators of the hazard rate and its derivatives and we try to adopt the local polynomial approximation in this context. The surplus rate in with-profit policies will be nonparametrically estimated as its mathematical expression depends on transition rates (hazard rates from one state to another) in a Markov chain (Ramlau-Hansen (1991), Norberg (1999)).

Key words : nonparametric kernel estimation, reproducing kernel Hilbert spaces, local polynomial estimation, functionals of the distribution function, derivatives, probability density function, Lorenz curve, scaled total time on test transform curve, multiplicative intensity point process model, hazard rate function, with-profit policies, surplus, boundary effects.

Table des matières

Remerciements	3
Résumé	9
Abstract	11
Introduction générale	17
1 Estimation lisse de la courbe de Lorenz et du temps total de test normalisé	33
1.1 Introduction	33
1.2 Cadre d'étude	35
1.3 Hypothèses et notations	37
1.4 Estimation lisse de la courbe de Lorenz et application à l'indice de Gini	38
1.4.1 Estimation lisse de la courbe de Lorenz	38
1.4.2 Application à l'indice de Gini	42
1.5 Estimation lisse du temps total de test normalisé et application au remplacement de composants non réparables	42
1.5.1 Estimation lisse du temps total de test normalisé	42
1.5.2 Détermination du moment optimal de remplacement des composants non réparables	45
1.6 Preuves	47
1.6.1 Preuve du Lemme 1.4.1	48
1.6.2 Preuve du Théorème 1.4.1	48
1.6.3 Preuve du Théorème 1.4.2	50
1.6.4 Preuve du Théorème 1.4.3	50
1.6.5 Preuve du Théorème 1.4.4	54
1.6.6 Preuve du Lemme 1.5.1	55
1.6.7 Preuve du Théorème 1.5.1	55
1.6.8 Preuve du Théorème 1.5.2	58
1.6.9 Preuve du Théorème 1.5.3	58
1.7 Simulations	61
1.7.1 Cadre	61
1.7.2 Données simulées	62
1.7.3 Données réelles	66

2	Estimation à noyau de fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives	69
2.1	Introduction	69
2.2	Cadre d'étude	70
2.2.1	Hypothèses et notations	70
2.3	Estimation polynomiale locale des fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives	72
2.4	Estimation polynomiale locale de la densité et de ses dérivées	75
2.5	Estimation à noyau des dérivées de la fonction T_F	77
2.5.1	Dérivées d'ordre supérieur ou égal à 2 à l'aide de l'approximation polynomiale locale	78
2.5.2	Dérivée première par la méthode à noyau classique	79
2.6	Estimation polynomiale locale des dérivées de la fonction L_F	80
2.7	Choix de fenêtre	83
2.8	Preuves	83
2.8.1	Preuve du Lemme 2.3.1	83
2.8.2	Preuve du Lemme 2.4.1	84
2.8.3	Preuve du Théorème 2.4.1	85
2.8.4	Preuve du Théorème 2.4.2	86
2.8.5	Preuve du Théorème 2.5.1	88
2.8.6	Preuve du Théorème 2.5.2	88
2.8.7	Preuve du Théorème 2.5.3	90
2.8.8	Preuve du Théorème 2.5.4	91
2.8.9	Preuve du Lemme 2.6.1	94
2.8.10	Preuve du Théorème 2.6.1	95
2.8.11	Preuve du Théorème 2.6.2	96
2.9	Simulations	100
2.9.1	Cadre	100
2.9.2	Données simulées	101
2.9.3	Données réelles	104
3	Estimation à noyau de la densité et de ses dérivées successives et correction des effets de bord	107
3.1	Introduction	107
3.2	Estimation à noyau classique de la densité et de ses dérivées	107
3.3	Une amélioration à l'aide de noyaux réducteurs de biais	111
3.4	Application à une méthode de correction des effets de bord	112
3.5	Preuves	114
3.5.1	Preuve du Lemme 3.2.1	114
3.5.2	Preuve du Lemme 3.2.2	115
3.5.3	Preuve du Théorème 3.2.1	116
3.5.4	Preuve du Théorème 3.2.2	116
3.5.5	Preuve du Lemme 3.3.1	117
3.5.6	Preuve du Théorème 3.3.1	117
3.5.7	Preuve du Théorème 3.4.1	118
3.5.8	Preuve du Théorème 3.4.2	118

4	Estimation à noyau du taux de hasard et de ses dérivées successives et application au taux d'accumulation du surplus en actuariat	121
4.1	Introduction	121
4.2	Résultat préliminaire	122
4.3	Estimation du taux de hasard et de ses dérivées successives pour des variables aléatoires indépendantes et identiquement distribuées	123
4.4	Estimation du taux de hasard et de ses dérivées successives dans le modèle à intensité multiplicative	126
4.4.1	Cadre d'étude	126
4.4.2	Hypothèses et notations	128
4.4.3	Estimation à noyau classique du taux de hasard et de ses dérivées successives	131
4.4.4	Estimation polynomiale locale du taux de hasard et de ses dérivées successives	133
4.4.5	Choix de fenêtre	137
4.5	Un mot sur les effets de bord	137
4.6	Preuves	138
4.6.1	Preuve du Lemme 4.4.1	139
4.6.2	Preuve du Lemme 4.4.2	140
4.6.3	Preuve du Théorème 4.4.1	143
4.6.4	Preuve du Théorème 4.4.2	143
4.6.5	Preuve du Lemme 4.4.3	147
4.6.6	Preuve du Lemme 4.4.4	148
4.6.7	Preuve du Théorème 4.4.3	151
4.6.8	Preuve du Théorème 4.4.4	151
4.7	Application au taux d'accumulation du surplus en actuariat	154
4.7.1	Cadre	154
4.7.2	Estimation du taux d'accumulation du surplus	158
4.7.3	Preuve du Théorème 4.7.1	160
4.7.4	Preuve du Théorème 4.7.2	160
	Conclusion	163
A	Espaces de Hilbert à noyaux reproduisants et approximation polynomiale locale	165
A.1	Projection	165
A.2	Sous-espaces de dimension finie	165
A.3	Espaces de fonctions	166
A.4	Sous-espaces de polynômes	167
A.5	Fonctionnelles de la fonction de répartition	168
B	Calcul de noyaux d'ordre supérieur d'une hiérarchie	169
C	Convolution	173
C.1	Convolutions	173
C.2	Estimation à noyau	174
D	Martingales	177

E	Estimation de probabilités de transition d'un processus de Markov non homogène	179
E.1	Définition de l'estimateur de Aalen-Johansen	179
E.2	Propriétés asymptotiques de l'estimateur de Aalen-Johansen	180
F	Théorèmes et résultats utiles	183
F.1	Loi faible des grands nombres	183
F.2	Théorème de la limite centrale multidimensionnel	183
F.3	Méthode Delta	183
F.4	Théorème de convergence dominée de Lebesgue	183
F.5	Théorème de Glivenko-Cantelli	184
F.6	Loi forte des grands nombres	184
F.7	Théorème des fonctions continues	184
F.8	Convergence en probabilité et convergence \mathbb{L}^p , $p \in \mathbb{N}^*$	184
F.9	Théorème de Lindeberg-Feller	185
F.10	Règles de Cramér	186
F.11	Inégalité de Markov	186
F.12	Inégalité de Jensen	186
F.13	Inégalité de Cauchy-Schwartz	186
F.14	Théorème de Rebolledo	186
	Table des figures	189
	Bibliographie	191

Introduction générale

La Courbe de Lorenz et la mesure des inégalités

INÉGALITÉ n.f. Caractère, état de choses ou de personnes inégales entre elles.

INÉGAL adj. Qui n'est pas égal à quelque chose ou quelqu'un d'autre.

ÉGAL adj. Semblable en nature, en quantité, en qualité, en valeur.

(Définitions extraites du *Grand Larousse Illustré 2016*)

Depuis le milieu du XIX^e siècle, avec la révolution industrielle, les inégalités économiques entre individus dans le Monde ont fortement augmenté. Les pays du Nord ont en effet connu une forte croissance, creusant ainsi l'écart avec les populations du Sud. Ce processus s'est ralenti au cours des années 1980-1990, puis renversé depuis une vingtaine d'années. Le ralentissement des taux de croissance des pays riches, dépassés par ceux des pays émergents, a conduit mécaniquement à une réduction des inégalités à l'échelle mondiale.

Mais dans le même temps, après une tendance à la baisse au cours du XX^e siècle, les inégalités au sein des pays tendent à augmenter. Les pays riches connaissent depuis les années 1980, et plus récemment en France depuis les années 2000, une réouverture des inégalités internes. La tendance est même plus marquée dans les pays émergents, ceci impliquant une répartition inéquitable des produits de la croissance.

L'exemple de la Chine, devenue rapidement la seconde puissance mondiale, illustre parfaitement ces deux processus engagés depuis la fin du XX^e siècle. Du fait de son poids très important dans la population mondiale, elle a fortement contribué par son enrichissement rapide à la réduction des inégalités entre les habitants de la planète, mais elle a en même temps connu une forte hausse des inégalités à l'intérieur du pays, qui perdure aujourd'hui.

Pour plus de détails sur l'évolution de l'inégalité mondiale et son interprétation, nous renvoyons le lecteur à Bourguignon ([36]).

Comment a-t-on pu affirmer que les inégalités augmentaient ou baissaient ? Impossible sans avoir préalablement réussi à mesurer les inégalités, qui est un enjeu essentiel en économie politique. Savoir décrire les inégalités pour espérer les comprendre, comparer les inégalités au sein d'une population sur plusieurs périodes afin d'avoir une idée de leur évolution dans le temps ou d'examiner l'impact de politiques économiques ou sociales mises en place par un gouvernement, comparer les inégalités sur des populations diverses dans le but d'avoir une vision plus globale... voici des objectifs tant recherchés, et en particulier en ce qui concerne les inégalités économiques.

Pour mesurer celles-ci, introduisons le *concept du revenu* présenté par Hicks. Il le définit comme la consommation maximale d'un individu sur une période donnée sans appauvrissement, c'est-à-dire vérifiant l'égalité de son patrimoine en début et fin de période. En partant de ce concept, plusieurs types de revenus sont définis pour permettre des mesures utilisables statistiquement.

L'Institut National de la Statistique et des Etudes Economiques (INSEE) définit les *revenus primaires* comme les "revenus directement liés à une participation des ménages au processus de production. La majeure partie des revenus primaires des ménages est constituée de la rémunération des salariés, laquelle comprend les salaires et les cotisations sociales. Ces revenus comprennent aussi des revenus de la propriété résultant du prêt ou de la location d'actifs financiers ou de terrains (intérêts, dividendes, revenus fonciers...)".

Navarro décrit dans [123] le *revenu disponible* défini par l'INSEE comme la "somme du revenu primaire, des pensions de retraite, du chômage, des indemnités journalières (en cas d'arrêt maladie), des prestations familiales, des allocations logement, des minima sociaux de laquelle on déduit les cotisations sociales, les contributions sociales (CSG, CRDS, prélèvements sociaux), les impôts directs (impôts sur le revenu, prélèvement forfaitaire et taxe d'habitation)". En fait, la notion de revenu disponible suppose l'intervention de l'État. Son rôle théorique est de redistribuer les revenus de telle manière à en fournir une répartition plus égalitaire.

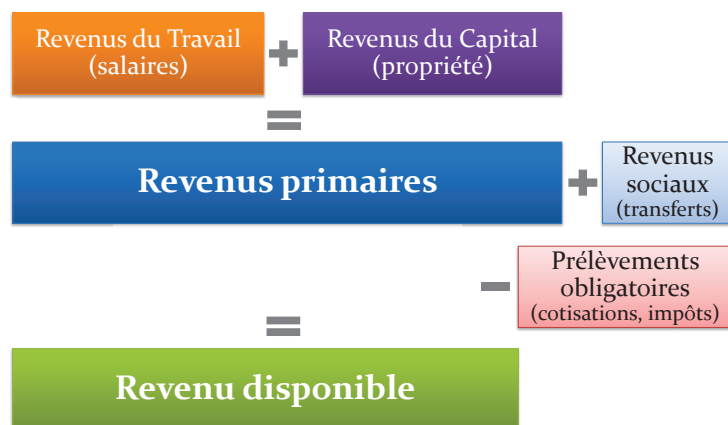


FIGURE 1 – Schéma reliant les revenus primaires et le revenu disponible.

Le revenu disponible ne tient pas compte du nombre de personnes composant les ménages qu'on appelle les *unités de consommation* (UC). Pour l'INSEE, le premier adulte du ménage compte pour 1 UC, les autres personnes âgées d'au moins 14 ans comptent pour 0,5 UC et les enfants âgés de moins de 14 ans comptent pour 0,3 UC. Pour effectuer des comparaisons entre ménages, il peut être judicieux de considérer le revenu disponible par unité de consommation. Le résultat de ce calcul correspond au *niveau de vie*.

Nous sommes malheureusement confrontés à plusieurs difficultés pour obtenir des données exploitables. Tout d'abord, comment choisir entre les revenus primaires, le revenu disponible et le niveau de vie ? Les revenus non monétaires (en nature) ne sont pris en compte dans aucune des trois définitions, ne devrait-on pas pourtant les inclure pour avoir une idée plus juste des revenus ?

Ensuite, les données proviennent uniquement des revenus déclarés. Les hauts revenus ont tendance à dissimuler une partie de leurs revenus et l'économie dite "souterraine" (travail clandestin, fraude, trafics) est mal connue.

Enfin, les données fournies par les organismes statistiques sont peu nombreuses en raison du coût élevé pour leur gestion et leur accès est long. Dans [9], D'Agostino et Trombert citent les principales sources permettant de recueillir régulièrement les revenus du travail, les revenus du capital et les revenus sociaux, indispensables au calcul des revenus primaires et du revenu disponible mais indique que l'INSEE les exploite uniquement tous les 5 ans. Des enquêtes supplémentaires sont parfois lancées mais à basse fréquence comme l'enquête sur les revenus fiscaux des ménages réalisée par l'INSEE qui a seulement lieu tous les 5 ans. Les résumés d'enquêtes se font aussi attendre, souvent durant plusieurs années, à cause du traitement de l'information fastidieux.

Pour un état des lieux récent de la France, Navarro relève dans [123] que les revenus primaires brut de l'ensemble des ménages est estimé en 2013 à 1522 milliards d'euros, le revenu disponible brut de l'ensemble des ménages est estimé à 1326,3 milliards d'euros et le niveau de vie annuel médian est de 20 603 euros en 2012. Elle note la diminution des inégalités de revenus primaires au cours du XX^e siècle suivie d'un accroissement à partir de l'année 2000. La même tendance et la même année de transition sont observées pour les inégalités de niveaux de vie. Entre 2002 et 2012, le niveau de vie des 10% aux plus hauts revenus s'est élevé tandis que celui des 10% les plus pauvres a baissé, ce qui a causé indéniablement l'augmentation des inégalités de niveaux de vie au XXI^e siècle.

Plusieurs instruments de mesure ont été proposés pour analyser la distribution des revenus d'une population. Ils se classent en trois catégories : mesure des disparités des revenus (étendue, moyenne, médiane, comparaison de la moyenne à la médiane...), mesure de dispersion des revenus (rapports interdéciles, rapports interquartiles...), mesure de concentration des revenus (courbe de Lorenz, indice de Gini). Nous renvoyons le lecteur aux ouvrages [9], [114] et [123] pour la description des deux premières catégories et nous présentons la dernière.

Max Otto Lorenz ([109]) a eu l'idée de relier la part cumulée des ménages les moins riches à la part cumulée correspondante du revenu total (somme des revenus de l'ensemble des ménages) pour étudier la concentration des revenus. Le graphique obtenu se situant dans le carré unitaire et passant par les points (0,0) et (1,1) se nomme la *courbe de Lorenz* en hommage à son créateur. Une répartition des revenus égalitaire implique que 1% (respectivement 2%, 3%...) des ménages les moins riches détient 1% (respectivement 2%, 3%...) du revenu total et se représente par la diagonale dans le carré unitaire appelée *ligne d'égalité parfaite*. Une répartition totalement inégalitaire a lieu lorsque tous les ménages sauf un possède un revenu nul. L'illustration est une ligne brisée dans le carré unitaire qu'on nomme *ligne de parfaite inégalité*. Finalement, plus la courbe de Lorenz s'approche de la diagonale et plus les inégalités de revenus sont faibles.

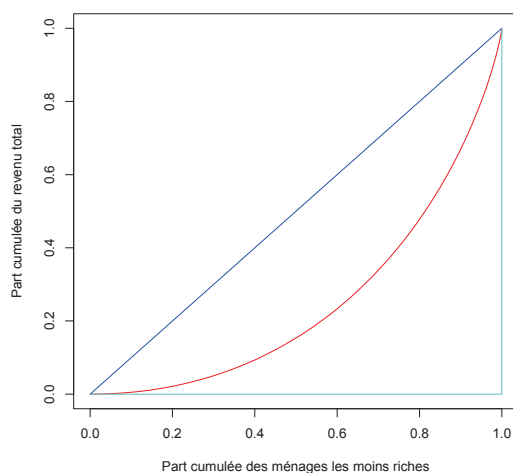


FIGURE 2 – Un exemple de courbe de Lorenz (en rouge) avec la ligne d'égalité parfaite (en bleu) et la ligne de parfaite inégalité (en turquoise).

Considérons deux courbes de Lorenz illustrant deux (ou plusieurs) distributions de revenus ou une même distribution de revenu sur deux (ou plusieurs) périodes différentes. Alors la courbe de Lorenz la plus proche de la diagonale correspond à la distribution la plus égalitaire, à condition qu'elles ne se croisent pas.

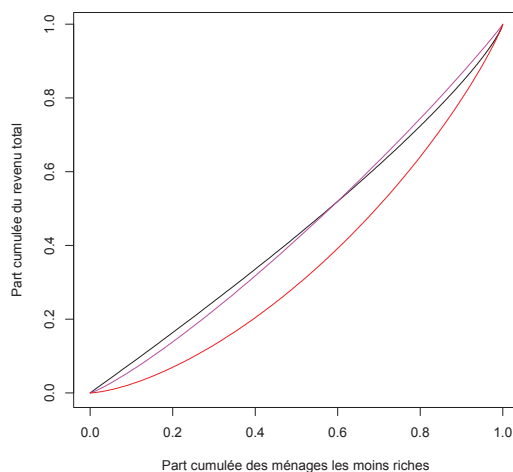


FIGURE 3 – 3 courbes de Lorenz : loi de Pareto (3,5) (en noir), loi de Weibull (3,5.5) (en magenta) et loi de Maxwell (en rouge).

Sur cet exemple, la courbe rouge est la plus éloignée de la diagonale et ne croise aucune des deux autres courbes donc la loi de Maxwell est la plus inégalitaire. Impossible de déterminer la loi la plus égalitaire, nous pouvons uniquement affirmer que la loi de Pareto (3,5) est plus égalitaire pour les 60% les moins riches tandis qu'il s'agit de la loi de Weibull (3,5.5) pour les 40% les plus riches.

L'*indice de Gini*, introduit par le statisticien Corrado Gini, est une valeur numérique associée à la courbe de Lorenz. Il s'écrit comme le quotient de l'aire comprise entre la courbe de Lorenz et la diagonale et de l'aire sous la diagonale. On comprend que l'indice de Gini est nul si l'aire entre la courbe de Lorenz et la diagonale est égale à 0, c'est-à-dire si la courbe de Lorenz et la ligne d'égalité parfaite sont confondues. Par suite, l'indice de Gini est égal à 1 si la courbe de Lorenz est confondue avec la ligne de parfaite inégalité. La réalité se situe entre ces deux cas extrêmes, c'est-à-dire que l'indice de Gini est toujours compris entre 0 et 1, et les inégalités varient dans le même sens que l'indice de Gini.

Lorsque nous souhaitons comparer plusieurs distributions de revenus, celle ayant l'indice de Gini le plus faible est la plus égalitaire, ce qui permet de surmonter la difficulté rencontrée en cas de croisement des courbes de Lorenz. En cas d'égalité des indices de Gini, il est impossible de conclure. Si l'on reprend l'exemple illustré par la Figure 3, l'indice de Gini associé à la courbe noire est plus faible que celui associé à la courbe magenta donc la loi de Pareto (3,5) est la plus égalitaire.

Vue la diversité des outils de mesure des trois catégories, nous pouvons nous questionner sur le choix le plus pertinent. Il n'y en a pas, la diversité est complémentaire. Nous pouvons cependant émettre certaines réserves : les instruments de mesure des disparités ne sont pas des indices synthétiques, ceux de mesure de dispersion s'expriment en fonction des extrêmes (premier et dernier quartiles pour le rapport interquartiles, premier et dernier déciles pour le rapport interdéciles...) d'où une représentativité incomplète de la répartition des revenus, l'utilisation limitée de la courbe de Lorenz pour comparer plusieurs courbes de Lorenz qui se croisent, l'indice de Gini, passant au-delà de toutes ces limites, mais incapable de décrire la distribution de revenus étudiée.

L'objectif de notre thèse est de rendre plus juste la mesure des inégalités de revenus. Pour cela, nous nous concentrons sur les outils statistiques les plus employés de nos jours : la courbe de Lorenz et l'indice de Gini. La courbe de Lorenz, comme son nom l'indique, est une courbe, c'est pourquoi nous tenterons de construire un estimateur lisse pour cette fonction, d'où découle une estimation de l'indice de Gini. L'approche est non paramétrique, c'est-à-dire que nous nous plaçons dans un cadre libre, sans aucune hypothèse et sans imposer de distribution sur les revenus. Gastwirth ([75]) avait jusque-là défini un estimateur mais continu par morceaux.

Le Temps Total de Test normalisé, indicateur de fiabilité

Le concept de *fiabilité* a émergé dans l'entre-deux-guerres pour étudier les fréquences des pannes dans l'aéronautique, menaçant directement la sécurité des personnes. La discipline se développe rapidement jusque dans les années 1950 avec son application aux systèmes électroniques, introduits là encore dans l'aéronautique mais également dans la défense, l'aérospatial ou le nucléaire. Un *système* étant un ensemble de *composants* en interaction destiné à accomplir une tâche donnée, la fiabilité d'un système s'explique par la fiabilité de ses composants.

Les décennies suivantes connaissent la généralisation de son utilisation (en mécanique, hydraulique, électricité, informatique...) et aujourd'hui la fiabilité est devenue un outil clef dans l'étude de la plupart des composants, produits et processus "grand public" : transport, énergie, bâtiments, composants électroniques, composants mécaniques, logi-

ciels, etc.

Le terme de "fiabilité" est parfois utilisé à tort pour parler de la *sûreté de fonctionnement* : c'est un abus de langage car la fiabilité n'en est qu'une composante. La sûreté de fonctionnement - en anglais *dependability* - est une notion plus générale qui mesure la qualité de service délivrée par un système, de manière à ce que l'utilisateur ait en lui une confiance justifiée. Elle s'est développée dans les années 1980, à partir de la fiabilité et la sécurité, avec l'ajout d'autres composantes au service de la maîtrise des risques comme la *maintenabilité* et la *disponibilité* : voir ci-dessous une illustration simplifiée des interactions.

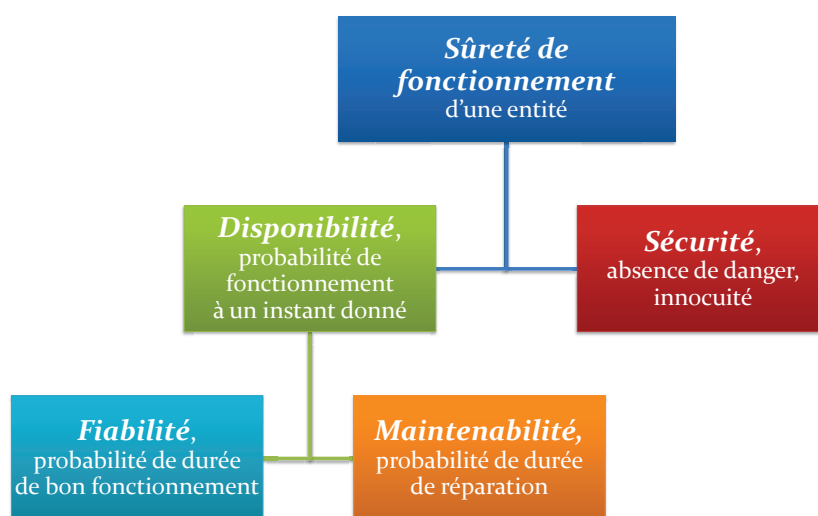


FIGURE 4 – Composantes de la sûreté de fonctionnement.

Après cette clarification, définissons plus précisément la notion de *fiabilité* (*reliability* en anglais). Selon la norme AFNOR X 60-500, la fiabilité est "l'aptitude d'une entité à accomplir une fonction requise, dans des conditions données, pendant un intervalle de temps donné". La valeur de la fiabilité est donc la probabilité de n'avoir aucune *défaillance* - état dans lequel l'entité ne peut plus accomplir sa mission - pendant une durée de fonctionnement donnée et dans des conditions données.

Prévoir la fiabilité d'un système est un enjeu essentiel pour des raisons de sécurité (éviter les défaillances d'un système de freinage, système nucléaire, système informatique, système électrique, système de transports, etc.) et pour des raisons économiques : coûts de maintenance (réparation, gestion de personnel et des stocks de pièces de rechange), de non production, d'immobilisation, de non qualité et parfois quasi impossibilité de réparer certains matériels (un satellite par exemple).

Pour la suite nous nous placerons dans le domaine industriel en étudiant comme objet le composant, mais l'approche est applicable aussi bien à un matériel, un système, une unité de production, un service, etc.

Les composants des systèmes industriels sont classés en deux catégories distinctes : les composants non réparables et réparables.

Un *composant non réparable* est un composant qui ne peut pas être remis en état de fonctionnement lorsque sa première défaillance survient. Les raisons sont essentiellement économiques (faible coût du composant défectueux, réparation onéreuse) et/ou techniques (complexité de réparation du composant défectueux). Le remplacement du composant défaillant est donc privilégié à sa réparation. Un fusible, une batterie, une dalle d'écran en sont des exemples.

Un *composant réparable* est un composant qui, au contraire, peut être remis en état de fonctionnement plusieurs fois au cours de sa vie. En général, il est maintenu préventivement, par anticipation d'une anomalie, mais en cas de défaillance, des opérations de maintenance corrective (ou réparation) lui permettent de continuer à accomplir sa mission. Une pompe, un robinet, un moteur font partie de cette catégorie.

La modélisation utilisée pour évaluer la fiabilité d'un composant dépend de sa catégorie.

Un composant non réparable est caractérisé par un instant de défaillance (ou une durée de vie ou une durée de fonctionnement). Pour réaliser une étude statistique, il faut disposer d'un échantillon de composants semblables dans les mêmes conditions et noter leurs instants de défaillances. Dans ces conditions, les instants de défaillances sont indépendants et identiquement distribués.

Un composant réparable passe par divers états suite à ses différentes dégradations et est donc caractérisé par les durées entre ses défaillances successives qui ne sont généralement ni indépendantes ni identiquement distribuées. En effet, l'historique du processus des défaillances et de leur réparation conditionne les défaillances suivantes. Les instants de défaillances sont alors modélisés par un processus ponctuel d'intensité aléatoire.

Nous traitons dans cette thèse uniquement de la fiabilité des composants non réparables. Afin de présenter une méthode graphique performante pour évaluer la fiabilité d'un composant, nous avons d'abord besoin d'introduire le *concept du temps total de test* de Barlow et Campo ([16]).

Pour comprendre de quoi il s'agit, supposons que les instants de défaillance des n composants d'un système sont décrits par une variable aléatoire discrète. Alors les n composants fonctionnent jusqu'à la première défaillance, $n - 1$ composants fonctionnent après la première défaillance et jusqu'à la seconde défaillance, ..., et 1 composant (le dernier en vie) fonctionne après l'avant-dernière défaillance et jusqu'à la dernière défaillance (sa propre défaillance).

Le *temps total de test jusqu'à la k^e défaillance* est le temps de fonctionnement de l'ensemble des composants jusqu'à la défaillance du k^e composant, où k est un entier compris entre 0 et n (il est nul si $k = 0$).

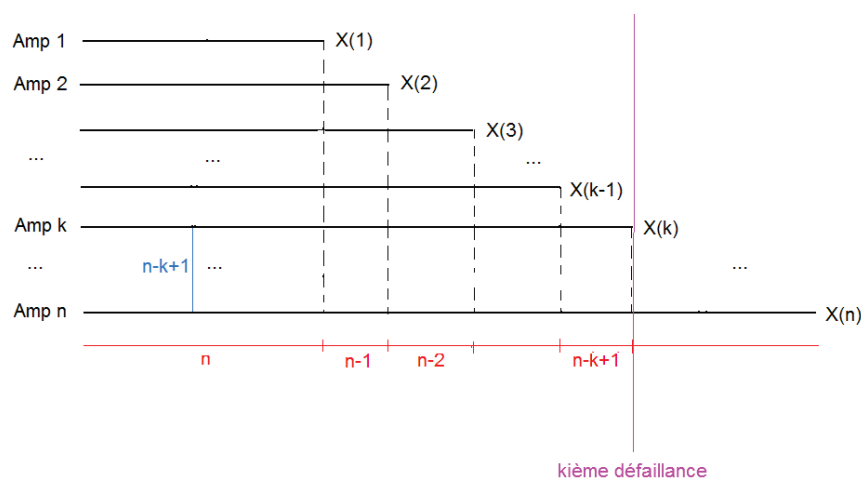


FIGURE 5 – Illustration du temps total de test de n ampoules jusqu'à la k^{e} défaillance, où k est un entier compris entre 1 et n . Les durées de vie ordonnées des n ampoules sont notées $X(1), \dots, X(n)$.

La durée moyenne de fonctionnement de l'ensemble des composants jusqu'à la k^{e} défaillance est donc obtenue en divisant le temps total de test jusqu'à la k^{e} défaillance par le nombre total de composants n , où k est un entier compris entre 0 et n . Barlow et Campo ([16]) ont eu l'idée de relier la durée moyenne de fonctionnement de l'ensemble des composants jusqu'à la k^{e} défaillance à la proportion $\frac{k}{n}$ de composants défaillants, pour tout k entier compris entre 0 et n , dans le but de détecter une tendance au niveau du composant.

Les durées moyennes de fonctionnement de l'ensemble des composants entre défaillances successives sont obtenues par différence de la durée moyenne de fonctionnement de l'ensemble des composants jusqu'à la k^{e} défaillance et de la durée moyenne de fonctionnement de l'ensemble des composants jusqu'à la $(k-1)^{\text{e}}$ défaillance, pour tout k compris entre 1 et n . Si elles augmentent alors il y a un accroissement de fiabilité du composant. Si elles diminuent alors il y a un vieillissement du composant. Si elles sont constantes alors il y a une absence de tendance. Sur la Figure 6, le premier composant a tendance à se détériorer tandis que le second gagne en fiabilité. Les durées moyennes de fonctionnement entre défaillances successives sont respectivement 0, 4, 6.25, 8.25, 10 et 0, 1.5, 2.5, 5, 9. Le troisième composant débute avec une phase d'accroissement de fiabilité puis se termine par une phase de vieillissement. Aucune tendance n'est constatée au niveau du dernier système.

Barlow et Campo ([16]) et Bergman ([29]) ont remarqué que la durée moyenne de fonctionnement de l'ensemble des composants était invariante par changement d'échelle donc qu'il était possible de lui appliquer un coefficient de normalisation sans modifier la distribution des instants de défaillance. Les nombres sont désormais compris entre 0 et 1 sur l'axe des ordonnées et le graphique obtenu dans le carré unitaire, plus simple à lire, est appelé *TTT-plot*. Les nombres sur l'axe des ordonnées n'ont alors plus d'unité, ce qui permet de comparer les tendances de plusieurs systèmes en représentant leurs TTT-plots. De plus, on peut tracer la diagonale qui caractérise l'absence de tendance.

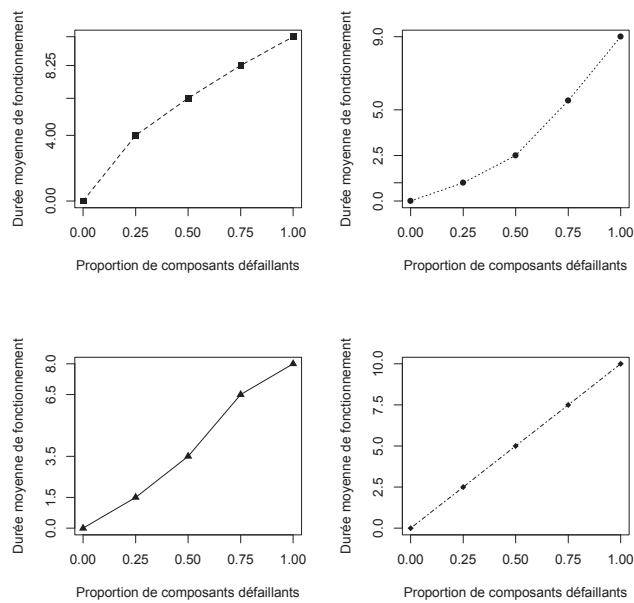


FIGURE 6 – Durée moyenne de fonctionnement en fonction de la proportion de composants défectueux de plusieurs systèmes de 4 composants.

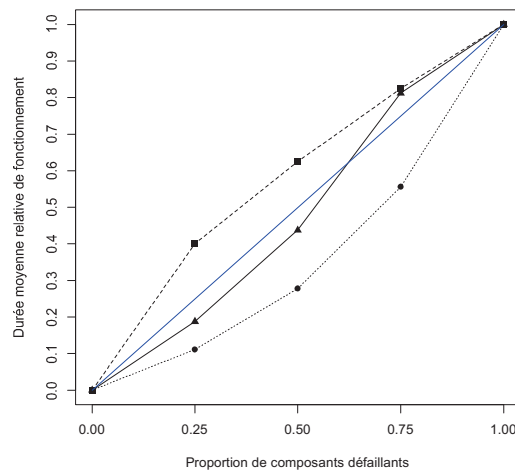


FIGURE 7 – TTT-plots associés aux 3 premiers systèmes de la figure précédente. La courbe bleue représente l'absence de tendance.

Sur la Figure 7, les allures des TTT-plots sont identiques à celle des 3 premiers graphiques de la Figure 6 donc les conclusions sont identiques. Nous pouvons ajouter que l'accroissement de fiabilité du troisième composant est moins élevé que celui du second.

La variable aléatoire correspondant aux instants de défaillance est en général continue puisqu'elle peut prendre toutes les valeurs possibles réelles. Le TTT-plot devient

alors une courbe dans le carré unitaire nommée le *temps total de test normalisé*. Cette courbe est convexe si le système gagne en fiabilité, et concave si le système se détériore.

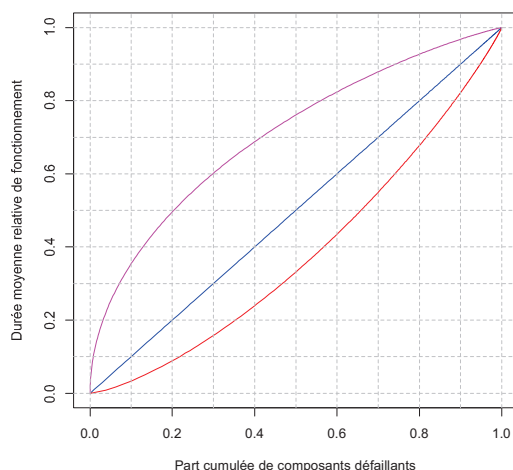


FIGURE 8 – Temps total de test normalisé traduisant un accroissement de fiabilité (en rouge), une tendance au vieillissement (en magenta) et ne traduisant aucune tendance (en bleu).

Le but dans notre thèse est d'analyser la fiabilité des composants non réparables afin de mieux maîtriser le risque industriel et la sécurité. Le temps total de test normalisé est une méthode graphique non paramétrique utile pour détecter une tendance au niveau du composant (tendance au vieillissement et à la détérioration ou accroissement de fiabilité), mais également pour identifier la loi des défaillances. Par exemple, l'absence de tendance illustrée par la diagonale dans le carré unitaire caractérise une loi exponentielle. Bergman ([29]) utilise graphiquement le temps total de test normalisé afin de déterminer le moment optimal de remplacement du composant défaillant. Rechercher un estimateur du temps total de test normalisé est ainsi essentiel pour rendre une vision plus conforme de la fiabilité des composants non réparables. Le temps total de test normalisé est une fonction continue, c'est pourquoi nous emploierons une méthode à noyau. Barlow et Campo ([16]) ont déjà proposé un estimateur non paramétrique mais il ne vérifie pas la continuité.

La Participation aux Bénéfices pour gérer le risque non diversifiable

"La marge de profit - dit M. Ch. Robert, dans la conclusion de son intéressant rapport sur la Participation au jury de l'Exposition d'Économie sociale - qui appartient aujourd'hui presque exclusivement à l'entrepreneur, peut se partager entre lui et son personnel". Cet extrait de D'Eichtal ([65]) illustre les prémices de la participation aux bénéfices en France. Le député Gaston Guillemet a souhaité qu'on la définisse par la suite comme "un contrat par lequel les entrepreneurs s'engagent à réserver à leurs ouvriers ou employés une portion déterminée de leurs bénéfices avec droit de contrôle des

co-partageants et sans que ceux-ci aident à concourir aux pertes". La participation des bénéficiaires était donc à l'origine en direction des salariés. D'Eichtal a mentionné dans [65] qu'elle était surtout pratiquée par des banquiers et des compagnies d'assurances alors que peu d'établissements l'avaient mise en place. Aujourd'hui, elle est obligatoire pour les entreprises d'au moins 50 salariés et se nomme toujours participation aux bénéficiaires ou *participation des salariés aux bénéficiaires* ou *participation des salariés*.

Une variante du système décrit plus haut a vu le jour et se nomme également *participation aux bénéficiaires*. Elle concerne uniquement les compagnies d'assurances et les bénéficiaires sont versés non pas à leurs salariés mais aux souscripteurs de contrats d'assurance. La loi française du 4 Janvier 1994 relative au droit général de la collectivité des souscripteurs exige que : "Les entreprises d'assurance sur la vie ou de capitalisation doivent faire participer les assurés aux bénéficiaires techniques et financiers qu'elles réalisent" (Article L331-3 du Code des Assurances). Dans la suite, la participation aux bénéficiaires désigne cette variante dans le cadre de l'assurance vie.

"Un *contrat d'assurance vie* est un contrat d'assurance dont le risque, objet du contrat, dépend de la durée de la vie humaine.", d'après l'Institut National de la Consommation (INC). L'*assureur* ou la compagnie d'assurances s'engage à verser à un *bénéficiaire* un capital ou une rente (*prestation*) en contrepartie de *primes* payées par le *souscripteur* pendant une période déterminée lorsque le risque survient pour la personne désignée au contrat (l'*assuré*, souvent le souscripteur). La prestation est versée à une date déterminée dans le contrat si l'assuré est en vie. En cas de décès, elle est versée au décès de l'assuré.

Le taux de rémunération garanti par un contrat d'assurance vie, appelé aussi *taux technique*, est fixé par le Code des Assurances. Il est déterminé de manière prudente, c'est-à-dire de telle sorte que le taux d'intérêt observé en réalité soit supérieur. Les placements d'une partie de l'épargne des assurés sur les marchés financiers sont donc sous-évalués et des *bénéfices financiers* sont réalisés par la compagnie d'assurances.

Afin de prédire le risque de mortalité d'une population, celle-ci emploie des tables de mortalité. On y lit en général un nombre de décès plus élevé que la réalité, ce qui conduit à un *bénéfice de mortalité*.

Le montant des primes inclut des frais (frais de dossier, frais de gestion, frais de versement, frais d'arbitrage...) surestimés par la compagnie d'assurances. Le *bénéfice de gestion* est la différence entre les frais prélevés et les charges réellement supportées par la compagnie d'assurances.

On désigne par *bénéfices techniques*, la somme du bénéfice de mortalité et du bénéfice de gestion.

La participation aux bénéficiaires, imposée par la loi, consiste à associer les assurés aux bénéficiaires techniques et financiers. Les taux de redistribution des bénéfices envers les assurés sont fixés de la manière suivante : 90% des bénéfices techniques et 85% des bénéfices financiers.

L'unique liberté de la compagnie d'assurances concerne le délai de la redistribution. Soit cette dernière est immédiate - alors une réduction de prime, un versement en espèces... peuvent être proposés -, soit elle est différée, alors les bénéfices sont intégrés à une réserve, appelée *provision pour participation aux bénéfices* ou *provision pour participation aux excédents*, mais devront être reversés aux assurés sous 8 ans maximum, limite imposée par la loi. La provision pour participation aux bénéfices présente l'avantage pour la compagnie d'assurances de pouvoir lisser ses rendements (années moins

favorables que d'autres) et d'avoir une sécurité financière en cas de coup dur. On note une augmentation du provisionnement des bénéfices par les compagnies d'assurances ces dernières années, sans doute à cause de la baisse continue des taux d'intérêt.

Finalement, la participation aux bénéfices en France est presque totalement encadrée par la loi. On dit que la participation aux bénéfices est *réglementée*. De nombreux pays, notamment en Afrique, sont basés sur ce modèle français. Nous pouvons citer l'Algérie, le Maroc, la Conférence Interafricaine des Marchés d'Assurances (zone CIMA) entre autres.

Il existe un modèle de participation aux bénéfices très répandu et très différent, appliqué dans les pays Anglo-saxons mais également dans certains pays Africains comme l'Afrique du Sud, le Ghana, le Kenya, le Nigéria...

Ce modèle d'inspiration anglo-saxonne est fondé sur la liberté des compagnies d'assurances. Les tarifs des contrats d'assurance ne requièrent aucune validation par une quelconque autorité de supervision, aucun taux d'actualisation ni aucune table de mortalité ne sont imposés, aucun montant minimal annuel de participation aux bénéfices n'est fixé.

Les règles de partage des bénéfices entre les assurés et la compagnie d'assurances sont libres, mais la compagnie d'assurances n'en conserve en général que 10% à cause de la concurrence impitoyable. De plus, une autorité de supervision pourrait intervenir en cas d'abus de la compagnie d'assurances.

La participation aux bénéfices est dite *contractuelle* ou *sous liberté contractuelle*.

Les contrats d'assurance vie sont habituellement à long terme avec des horizons temporels suffisamment lointains pour apercevoir les changements significatifs des taux d'intérêt, taux de mortalité, taux d'incapacité... L'évolution incertaine de tels taux soumet les compagnies d'assurances à un risque qui est non diversifiable, c'est-à-dire qu'il ne peut être éliminé en augmentant la taille du portefeuille. Traditionnellement, ce risque est géré par la participation aux bénéfices.

L'objet de notre thèse est de gérer le risque non diversifiable des compagnies d'assurance, non pas en proposant un nouveau modèle de participation aux bénéfices, mais en tentant de décrire précisément l'évolution des bénéfices dans ce cadre. Norberg délivre une approche moderne de la participation aux bénéfices et générale puisqu'elle n'exige pas de connaître le modèle (réglementaire, liberté contractuelle...) dans [127], où le processus des bénéfices s'exprime mathématiquement en fonction des taux instantanés de transition ou intensités de transition des assurés à chaque instant (la formule n'est pas nécessaire à ce stade pour la compréhension du lecteur qui attendra le Chapitre 4). Nous avons par conséquent besoin de définir le comportement des intensités de transition, ce qui nécessite leur estimation ainsi que celle de leurs dérivées successives. Nous serons capables de produire un estimateur du bénéfice instantané de la compagnie d'assurances lorsqu'un assuré se situe dans un certain état et par suite du taux d'accumulation des bénéfices. Comme les hypothèses paramétriques ne sont pas légitimes, nous appliquerons des méthodes à noyau.

Modélisation

Soit X la variable aléatoire positive correspondant au revenu d'un individu, de fonction de répartition continue F , de densité $f = F'$ et d'espérance μ finie. La fonction inverse de F est définie pour tout $t \in [0, 1]$ par

$$F^{-1}(t) = \inf_{x>0} \{x : F(x) \geq t\}. \quad (1)$$

Kendall et Stuart dans [94] ont donné la définition de la courbe de Lorenz sous la forme de deux équations :

$$p = F(x) \quad (2)$$

et

$$L(p) = \frac{1}{\mu} \int_0^x t f(t) dt. \quad (3)$$

En injectant la solution $x = F^{-1}(p)$ de l'équation (2) dans l'équation (3) puis en effectuant le changement de variable $s = F(t)$, Gastwirth ([74]) a proposé la définition de la courbe de Lorenz suivante :

$$L(p) = \frac{1}{\mu} \int_0^p F^{-1}(s) ds, \quad 0 \leq p \leq 1.$$

Tous les articles évoquant la courbe de Lorenz ont utilisé cette définition, à quelques exceptions près. En vertu de l'équation (2), Hart ([82]) a remplacé p par $F(x)$ dans l'équation (3), d'où

$$L(F(x)) = \frac{1}{\mu} \int_0^x s f(s) ds, \quad x > 0.$$

Abdous, Berlinet et Hengartner ([4]) et Berlinet et Thomas-Agnan ([34]) ont fait apparaître la formule de l'espérance μ et introduit une nouvelle notation :

$$L_F(x) := \frac{\int_0^x s dF(s)}{\int_0^{+\infty} s dF(s)}, \quad x > 0. \quad (4)$$

Le graphique qui pour tout $x > 0$, associe à $F(x)$ la quantité $L_F(x)$ s'intitule la courbe de Lorenz.

Soit X une variable aléatoire positive correspondant à l'instant de défaillance d'un composant arbitraire, de fonction de répartition continue F et d'espérance μ finie. La fonction inverse de F est donnée par (1). Barlow et Doksum ([18]), Barlow et al. ([15]) et Barlow et Campo ([16]) ont étudié le temps total de test (*Total Time on Test-transform* ou *TTT-transform* en anglais) noté

$$H_F^{-1}(p) = \int_0^{F^{-1}(p)} (1 - F(s)) ds, \quad 0 \leq p \leq 1. \quad (5)$$

Barlow et Campo ([16]) ont pensé à normaliser le temps total de test en raison de son invariance par changement d'échelle puis Bergman ([29]) a énoncé la formule du temps total de test normalisé (*scaled TTT-transform* en anglais) :

$$\frac{H_F^{-1}(p)}{H_F^{-1}(1)} = \frac{1}{\mu} \int_0^{F^{-1}(p)} (1 - F(s)) ds, \quad 0 \leq p \leq 1. \quad (6)$$

Grâce au changement de variable $p = F(x)$ et en écrivant l'expression de μ à l'aide d'une intégration par parties, Abdous, Berlinet et Hengartner ([4]) et Berlinet et Thomas-Agnan ([34]) ont proposé la notation suivante :

$$T_F(x) := \frac{\int_0^x (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds}, \quad x > 0. \quad (7)$$

Le temps total de test normalisé est le graphique qui pour tout $x > 0$, associe à $F(x)$ la quantité $T_F(x)$.

Soit X une variable aléatoire réelle de fonction de répartition continue F et de densité $f = F'$. Alors le taux de hasard est défini par

$$\alpha_F(x) := \frac{f(x)}{1 - F(x)}, \quad x \in \mathbb{R}. \quad (8)$$

En raison de l'intérêt porté à la courbe de Lorenz, au temps total de test normalisé et au taux hasard respectivement en sciences économiques, fiabilité et science actuarielle, obtenir des estimateurs avec de bonnes propriétés asymptotiques semble essentiel. Nous avons insisté précédemment sur la recherche d'une estimation lisse plus judicieuse, ce qui motive l'utilisation des méthodes à noyau. L'approche non paramétrique permet de se placer dans un cadre libre, sans fixer d'hypothèses ni de distribution sur les données. En observant les formules (4), (7) et (8), on constate que les fonctions L_F , T_F et α_F appartiennent à la famille des fonctionnelles de la fonction de répartition F , c'est-à-dire les fonctions Φ telles que :

$$\begin{aligned} \Phi(\cdot, F) : I \subset \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \Phi(x, F), \end{aligned}$$

Ce point commun justifie notre volonté de les regrouper dans un même document. Cette classe de fonctions a été explorée tout d'abord par Abdous, Berlinet et Hengartner ([4]) qui ont développé une formule pour les estimateurs des fonctionnelles lisses de la fonction de répartition et de leurs dérivées successives pour des variables aléatoires indépendantes et identiquement distribuées dans l'espace des polynômes de degré inférieur ou égal à r , $r > 0$. Berlinet et Thomas-Agnan ([34]) ont ensuite généralisé cette formule aux espaces de Hilbert à noyau reproduisant. La spécificité de leurs estimateurs vient de l'emploi de noyaux d'ordre supérieur réduisant considérablement le biais (Berlinet ([32])). Nous renvoyons le lecteur à l'Annexe A pour les détails de leurs travaux.

Abdous, Berlinet et Hengartner ont prouvé dans [4] la convergence des estimateurs des fonctionnelles lisses de la fonction de répartition et de leurs dérivées successives sous des conditions très générales. Mais est-il possible d'identifier des conditions plus précises pour les fonctions L_F et T_F ainsi que leurs dérivées successives ? Sous quelles hypothèses obtenir la normalité asymptotique ? Sont-elles identiques quelle que soit la fonctionnelle de la fonction de répartition étudiée ? Peut-on par exemple traiter les fonctions L_F et T_F , qui sont des rapports de fonctionnelles linéaires de la fonction de répartition, de la même manière que la fonctionnelle de la fonction de répartition α_F qui n'en est pas un ? Les résultats dans [4] se généralisent-ils aux variables censurées et dépendantes ? Si oui, qu'apportent-ils par rapport à ceux démontrés par Ramlau-Hansen ([134]) ?

Diverses applications découlent de l'estimation des fonctions L_F , T_F et α_F . Hart ([82]) a écrit la formule de l'indice de Gini en fonction de la courbe de Lorenz :

$$G_F := 1 - 2 \int_0^{+\infty} L_F(x) dF(x). \quad (9)$$

Norberg ([127], [128], [129]) a exprimé le taux d'accumulation des bénéfices (appelé également *taux d'accumulation du surplus*) à chaque instant $x \geq 0$ en fonction des intensités de transition, une autre manière d'appeler le taux de hasard dans la situation particulière d'une chaîne de Markov.

Organisation de la thèse et contributions

L'atout principal de notre thèse est sans aucun doute son emplacement au carrefour des probabilités, de la statistique, des sciences économiques et de gestion (fiabilité et science actuarielle).

Quatre chapitres seront consacrés à l'estimation lisse non paramétrique de fonctionnelles de la fonction de répartition en sciences économiques et de gestion.

Dans le premier chapitre, nous proposons une estimation non paramétrique de la courbe de Lorenz et de l'indice de Gini en sciences économiques puis du temps total de test normalisé en fiabilité. La détermination du moment optimal de remplacement des composants non réparables à partir de l'estimation du graphique du temps total de test normalisé est évoquée. Construire des estimateurs lisses comme les courbes initiales est nouveau. Le cadre d'étude suppose uniquement des variables aléatoires indépendantes et identiquement distribuées de fonction de répartition commune continue. La méthode d'estimation utilisée est l'approximation polynomiale locale découlant de la théorie des espaces de Hilbert à noyau reproduisant (Berlinet et Thomas-Agnan ([34])). Divers types de convergence ainsi que la normalité asymptotique sont obtenus, puis démontrés. Les illustrations des estimateurs sur données simulées et données réelles sont présentées en fin de chapitre.

L'objectif du second chapitre est de compléter l'étude des fonctions L_F et T_F en se concentrant sur leurs dérivées. Pour cela, nous considérons plus largement, l'estimation non paramétrique des fonctionnelles linéaires de la fonction de répartition continûment différentiables un certain nombre de fois et de leurs dérivées successives. Le cadre d'étude ainsi que la méthode d'estimation employée sont identiques à ceux du chapitre précédent. Après avoir construit les estimateurs à noyau des fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives, le biais est étudié. Dans l'impossibilité de fournir une vitesse de convergence générale, nous examinons tout d'abord le cas incontournable de la densité et de ses dérivées successives. N'étant pourtant pas notre but initial, les résultats limites de leurs estimateurs sont riches. Nous constatons ensuite qu'ils servent d'intermédiaire à l'obtention des propriétés asymptotiques des estimateurs des dérivées de la fonction T_F , bien que la dérivée première doive être traitée séparément de manière classique. Enfin, le comportement asymptotique des estimateurs des dérivées de la fonction L_F est exploré. Les preuves des résultats statistiques énumérés sont détaillées puis des simulations sur données réelles sont notamment posées en fin de chapitre.

Dans le troisième chapitre, nous nous focalisons sur l'estimation à noyau de la densité et de ses dérivées successives à l'aide de la convolution classique pour des variables aléatoires indépendantes et identiquement distribuées. Le but de ce chapitre est de détecter les différences et similitudes par rapport à l'approximation polynomiale locale, méthode d'estimation utilisée dans les deux chapitres précédents. De nombreux résultats asymptotiques peuvent être obtenus dans un premier temps. Nous proposons dans un second temps l'emploi de noyaux d'ordre supérieur et analysons leur effet bénéfique sur le biais de ces estimateurs non paramétriques. Dans un dernier temps, nous développons une méthode de correction des effets de bord des estimateurs à noyau de la densité et de ses dérivées successives, grâce aux estimateurs à noyau des dérivées d'ordre supérieur de la densité. Cette méthode, qui pourrait se généraliser à de nombreuses autres fonctions, marque encore une fois l'utilité de savoir estimer des dérivées. Les démonstrations des propriétés asymptotiques des estimateurs corrigés ou non sont rédigées en fin de chapitre.

Le quatrième chapitre se focalise sur l'estimation non paramétrique du taux de hasard, ainsi que de ses dérivées successives. Dans le cadre privilégié des variables aléatoires indépendantes et identiquement distribuées, la normalité asymptotique des estimateurs à noyau du taux de hasard et de ses dérivées successives est prouvée aisément car ces derniers sont construits à l'aide des estimateurs à noyau de la densité et de ses dérivées successives, dont nous possédons des résultats limites. Dans le modèle à intensité multiplicative, les variables aléatoires peuvent être dépendantes et censurées. Quelques rappels sur les martingales dans ce cadre d'étude plus général et quelques hypothèses sont nécessaires pour démarrer le travail d'estimation. Ramlau-Hansen ([134]) a démontré des résultats de convergence et de normalité asymptotique pour l'estimateur à noyau de l'intensité à l'aide de la convolution classique. Les résultats sont résumés dans Andersen et al. ([11]). Il a indiqué la procédure pour obtenir les propriétés asymptotiques des dérivées successives, que nous menons à terme. Nous examinons ensuite la possibilité d'adopter l'approximation polynomiale locale dans ce contexte. Après avoir mentionné les différences et similitudes entre les deux méthodes d'estimation, nous évoquons le choix de fenêtre. Les preuves des propriétés asymptotiques des estimateurs à noyau obtenus sont ensuite détaillées. Enfin, nous nous intéressons au taux d'accumulation du surplus en actuariat. Nous fournissons un estimateur à noyau ayant un bon comportement asymptotique, qui sera démontré en fin de chapitre.

Chapitre 1

Estimation lisse de la courbe de Lorenz et du temps total de test normalisé

1.1 Introduction

Dans ce premier chapitre, nous poursuivons deux objectifs. Le premier consiste à mesurer plus justement les inégalités économiques et pour cela, nous tentons de construire un estimateur lisse de la courbe de Lorenz ayant de bonnes propriétés asymptotiques, d'où peut découler une estimation de l'indice de Gini. Le second concerne la visualisation plus précise de la fiabilité des composants non réparables. Nous nous focalisons sur le temps total de test normalisé qui représente un outil pertinent, et nous essayons également de proposer un estimateur lisse possédant un bon comportement asymptotique. Nous supposons uniquement que les revenus et les instants de défaillance sont indépendants et identiquement distribués et nous nous plaçons dans un cadre non paramétrique, c'est-à-dire qu'aucune hypothèse sur la distribution des revenus ou celle des instants de défaillance n'est imposée.

La toute première définition mathématique de la courbe de Lorenz revient à Kendall et Stuart ([94]), qui l'ont exprimée sous la forme de deux équations en supposant une distribution de revenus absolument continue. Deux ans plus tard, Gastwirth ([74]) a fourni une définition générale de la courbe de Lorenz, s'appliquant aux lois continues comme aux lois discrètes, sous la forme d'une unique formule :

$$L(p) = \frac{1}{\mu} \int_0^p F^{-1}(u) du, \quad 0 \leq p \leq 1, \quad (1.1)$$

où la distribution de revenus et sa fonction inverse sont respectivement notées F et F^{-1} , et μ désigne l'espérance. En raison de ses applications nombreuses dans des domaines variés tels que les sciences économiques (Gastwirth ([75]) et l'indice de Gini, Hart ([82]), Gail et Gastwirth ([72], [73]) et les tests, la pêche (Thompson ([155])) ou encore la bibliométrie (Burrell ([37], [38], [39]), Sarabia ([141])), la courbe de Lorenz a suscité de nombreux travaux en estimation non paramétrique. Gastwirth ([75]) a proposé un

estimateur naturel de la courbe de Lorenz, défini par

$$L_n(p) = \frac{1}{\bar{X}_n} \int_0^p F_n^{-1}(u) du, \quad 0 \leq p \leq 1, \quad (1.2)$$

où \bar{X}_n représente la moyenne empirique d'un échantillon de n observations indépendantes X_1, \dots, X_n et de même distribution F , F_n la fonction de répartition empirique de cet échantillon, c'est-à-dire

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[X_i \leq x]}, \quad x > 0,$$

et

$$F_n^{-1}(t) = \inf_{x>0} \{x : F_n(x) \geq t\}, \quad 0 \leq t \leq 1,$$

l'inverse de la fonction de répartition empirique. À partir de cet estimateur, il a pu construire un estimateur de l'indice de Gini. Les propriétés asymptotiques de L_n ont été examinées, indépendamment par Gail et Gastwirth ([73]) qui ont démontré sa convergence presque sûre et sa normalité asymptotique à condition que le premier moment soit fini et non nul, et par Goldie ([78]) qui a prouvé sa convergence presque sûre uniforme sous la même hypothèse. Csörgő et al. ([50]) ont publié une étude unifiée sur la convergence en loi et la convergence presque sûre uniforme du processus de Lorenz $L_n - L$ ainsi que d'autres processus reliés à celui-ci. Rao et Zhao ([137]) ont ensuite démontré la loi du logarithme itéré du processus de Lorenz qu'ont amélioré Csörgő et Zitikis ([56], [58]) en assouplissant leurs hypothèses.

La situation de censure a également été étudiée. Csörgő et Horváth ([51]) ont tout d'abord construit un estimateur pour des données censurées à droite et prouvé sa convergence en loi, ainsi que sa convergence presque sûre uniforme. Tse ([156]) a ensuite considéré des observations tronquées à gauche et censurées à droite puis démontré la loi du logarithme itéré du processus de Lorenz et la convergence en loi de l'estimateur de l'indice de Gini qui en découle.

La limite que l'on peut trouver à tous ces travaux est de proposer des estimateurs continus par morceaux de la courbe de Lorenz alors que cette dernière est une fonction continue. Ces dernières années, Luo ([111]) s'est intéressé à l'estimation à noyau de la courbe de Lorenz et a obtenu la normalité asymptotique. Quant à Zhang, Wu et Li ([169]), ils ont produit un estimateur grâce à une méthode de lissage par splines.

Nous construisons également un estimateur lisse de la courbe de Lorenz et nous établissons sa convergence presque sûre, puis sa normalité asymptotique sous des conditions plus souples que celles posées par Luo ([111]).

Le temps total de test d'une distribution F s'écrit

$$H_F^{-1}(p) = \int_0^{F^{-1}(p)} (1 - F(s)) ds, \quad 0 \leq p \leq 1,$$

où F^{-1} correspond à la fonction inverse de la distribution F . Barlow et Doksum ([18]), Barlow et al. ([15]) et Barlow et Campo ([16]) ont analysé ses propriétés. Ayant constaté une invariance d'échelle, Barlow et Campo ([16]) ont introduit le temps total de test normalisé

$$T(p) = \frac{1}{\mu} \int_0^{F^{-1}(p)} (1 - F(s)) ds, \quad 0 \leq p \leq 1, \quad (1.3)$$

où μ représente l'espérance, puis ont démontré que sa version empirique

$$T_n(p) = \frac{1}{\bar{X}_n} \int_0^{F_n^{-1}(p)} (1 - F_n(s)) ds, \quad 0 \leq p \leq 1, \quad (1.4)$$

appelée TTT-plot, converge vers cette fonction lorsque que le nombre d'observations est suffisamment élevé. Ce résultat fort permet au TTT-plot d'être un outil graphique intéressant pour identifier la loi suivie par les observations (à condition qu'elles soient suffisamment nombreuses), en le comparant au temps total de test normalisé de lois connues. Ils ont de plus, grâce au temps total de test normalisé, caractérisé les distributions à taux de défaillance croissant (IFR) ainsi que celles à taux de défaillance décroissant (DFR), tout comme Barlow ([14]). Klefsjö ([95]) a illustré d'autres propriétés possibles de vieillissement (DMRL, NBUE, HNBUE) d'une distribution à l'aide du temps total de test normalisé et nous renvoyons le lecteur à son article pour un récapitulatif de toutes les définitions des lois IFR, DFR, DMRL, NBUE et HNBUE. Bergman ([28], [29]) a par ailleurs prouvé l'utilité du temps total de test normalisé pour déterminer le moment optimal de remplacement de matériel dû au vieillissement. Chandra Singpurwalla ([41]) ont remarqué le lien existant entre le temps total de test (normalisé) et la courbe de Lorenz (et l'indice de Gini), ainsi que Klefsjö ([97]), d'où l'application possible de la courbe de Lorenz en fiabilité et l'application du temps total de test normalisé en sciences économiques. Plusieurs chercheurs ont étudié ces deux fonctions. Csörgő et al. ([50]) ont développé les propriétés statistiques des processus de Lorenz et du temps total de test normalisé $T_n - T$, Csörgő et Horváth ([51]) ont construit un estimateur du temps total de test normalisé en situation de censure de la même manière que pour la courbe de Lorenz puis Csörgő et Zitikis ([59]) ont prouvé la loi du logarithme itéré pour le processus du temps total de test normalisé $T_n - T$, un an après avoir publié un article sur la loi du logarithme itéré du processus de Lorenz. Le TTT-plot, dont la procédure de construction est décrite dans Klefsjö ([98]), ne possède malheureusement pas la propriété de continuité du temps total de test normalisé. Il s'agit par la suite de proposer un estimateur lisse du temps total de test normalisé et de démontrer son bon comportement asymptotique à travers des théorèmes sur sa convergence presque sûre et sa normalité asymptotique.

1.2 Cadre d'étude

Soient X_1, \dots, X_n des variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F . Leurs espérances et variance communes sont considérées finies. La fonction de répartition empirique notée F_n associée à ce n -échantillon s'écrit

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

Nous rappelons que F_n converge vers F uniformément presque sûrement. De plus, la distribution asymptotique de F_n est donnée par

$$\sqrt{n}(F_n - F) \xrightarrow{\mathcal{L}} B, \quad (1.5)$$

où B est un processus Gaussien d'espérance $\mathbb{E} B(x) = 0$ et de fonction de covariance $\text{Cov}[B(x), B(y)] = F(\min(x, y))[1 - F(\max(x, y))]$. Cela découle de

$$\mathbb{E} \mathbf{1}_{[X_i \leq x]} = F(x),$$

$$\mathbb{E} [\mathbf{1}_{[X_i \leq x]} \mathbf{1}_{[X_i \leq y]}] = \mathbb{E} [\mathbf{1}_{[X_i \leq \min(x, y)]}] = F(\min(x, y)),$$

et

$$\text{Cov} [\mathbf{1}_{[X_i \leq x]} \mathbf{1}_{[X_i \leq y]}] = F(\min(x, y)) - F(x)F(y) = F(\min(x, y)) [1 - F(\max(x, y))].$$

Soit $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^d$ une fonction mesurable de carré intégrable. Alors nous pouvons déterminer les distributions asymptotiques des fonctions

$$\mathbf{g}_n := \int \mathbf{g} dF_n,$$

c'est-à-dire

$$\mathbf{g}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{g}(X_i).$$

Les vecteurs aléatoires $\mathbf{g}(X_1), \dots, \mathbf{g}(X_n)$ sont i.i.d. et la fonction \mathbf{g} mesurable nous assure que $\mathbb{E} \mathbf{g}(X_1) < +\infty$ donc d'après la loi faible des grands nombres F.1, nous avons :

$$\mathbf{g}_n \xrightarrow{\mathbb{P}} \mathbb{E} \mathbf{g},$$

où $\mathbb{E} \mathbf{g} := \mathbb{E} \mathbf{g}(X_1) = \int \mathbf{g} dF$.

La fonction \mathbf{g} est de carré intégrable donc les conditions sont réunies pour pouvoir appliquer le théorème de la limite centrale F.2 aux vecteurs aléatoires i.i.d. $\mathbf{g}(X_1), \dots, \mathbf{g}(X_n)$. Nous en déduisons :

$$\sqrt{n} (\mathbf{g}_n - \mathbb{E} \mathbf{g}) \xrightarrow{\mathcal{L}} \int \mathbf{g} dB \sim N_d(\mathbf{0}, \Sigma_{\mathbf{g}}) \quad (1.6)$$

avec

$$\Sigma_{\mathbf{g}} = (\text{Cov}[g_i(X_1), g_j(X_1)])_{i=1, \dots, d}^{j=1, \dots, d}.$$

Grâce à (1.6), nous présentons un résultat sur les fonctions d'intégrales par rapport à la distribution empirique F_n .

Lemme 1.2.1. *Soit $\mathbf{h} : \mathbb{R}^d \mapsto \mathbb{R}^e$ une fonction différentiable en $\mathbb{E} \mathbf{g}$. Alors*

$$\sqrt{n} (\mathbf{h}(\mathbf{g}_n) - \mathbf{h}(\mathbb{E} \mathbf{g})) \xrightarrow{\mathcal{L}} N_e(\mathbf{0}, D\mathbf{h}(\mathbb{E} \mathbf{g}) \Sigma_{\mathbf{g}} D\mathbf{h}^T(\mathbb{E} \mathbf{g})), \quad (1.7)$$

où $D\mathbf{h}$ est la matrice jacobienne de \mathbf{h} .

Ce résultat est obtenu à l'aide de la méthode Delta F.3 avec $k = d$, $m = e$, $T_n = \mathbf{g}_n$, $\theta = \mathbb{E} \mathbf{g}$ et $\phi = \mathbf{h}$.

Le Lemme 1.2.1 sera d'une aide précieuse pour démontrer les normalités asymptotiques des fonctions L_F et T_F définies respectivement par :

$$L_F(x) = \frac{\int_0^x s dF(s)}{\int_0^{+\infty} s dF(s)}, \quad x > 0, \quad (1.8)$$

et

$$T_F(x) = \frac{\int_0^x (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds}, \quad x > 0. \quad (1.9)$$

obtenus respectivement à partir de (1.1) et (1.3) en posant $p = F(x)$.

Pour la suite du chapitre, nous devons ajouter des hypothèses supplémentaires sur les variables aléatoires X_1, \dots, X_n et introduire quelques notations.

1.3 Hypothèses et notations

En plus d'être équidistribuées et indépendantes, X_1, \dots, X_n désignent des variables aléatoires positives pour respecter les définitions de la courbe de Lorenz et du temps total de test normalisé. Nous avons mentionné dans la partie Modélisation de l'Introduction que ces deux fonctions sont uniquement définies pour une variable aléatoire positive. De plus, la fonction de répartition commune aux X_1, \dots, X_n est supposée continue et la densité de probabilité $F' = f$ est de support $[0, +\infty[$. Les espérance et variance communes sont supposées finies, strictement positives et notées respectivement μ_1 et σ_1^2 .

La fonction K_0 représente une densité de probabilité de support $[-1, 1]$. Autrement dit,

$$K_0(u) \geq 0 \text{ pour tout } u \in [-1, 1]$$

et

$$\int_{-1}^1 K_0(u) du = 1.$$

Des illustrations possibles sont les noyaux Uniforme, d'Epanechnikov, Biweight, Triweight, et bien d'autres. Désignons par $K_r^{(0)}$ un noyau satisfaisant (A.16) pour $m = 0$ et borné sur le support $[-1, 1]$, ce qui signifie :

Hypothèse (H1)

(i)

$$\int_{-1}^1 u^i K_r^{(0)}(u) du = \begin{cases} 1, & i = 0, \\ 0, & i \in \{1, \dots, r\}, \\ C_{r+1}, & i > r, \end{cases}$$

avec C_{r+1} une constante finie ,

(ii)

$$\sup_{u \in [-1, 1]} |K_r^{(0)}(u)| < +\infty,$$

(iii) Le support de $K_r^{(0)}$ est $[-1, 1]$.

Pour une idée plus précise des noyaux $K_r^{(0)}$, nous renvoyons le lecteur à l'Annexe B où il trouvera le calcul détaillé des 5 premiers termes de la hiérarchie de noyaux associée à la densité d'Epanechnikov (c'est-à-dire les noyaux $K_r^{(0)}$ pour r variant de 0 à 4).

Afin de définir un estimateur à noyau des fonctions L_F et T_F , nous aurons besoin d'un paramètre de lissage qu'on appelle *fenêtre*. Nous adoptons la notation h_n puisque la fenêtre dépend du nombre n d'observations puis fixons les deux conditions suivantes :

Hypothèse (H2)

(i) (h_n) est une suite décroissante de nombres réels strictement positifs,

(ii) $h_n \rightarrow 0$ lorsque $n \rightarrow +\infty$.

Les hypothèses précédentes sont tout-à-fait classiques en estimation non paramétrique. Nous pouvons maintenant aborder l'estimation lisse de la courbe de Lorenz, de l'indice de Gini et du temps total de test normalisé.

1.4 Estimation lisse de la courbe de Lorenz et application à l'indice de Gini

Dans cette section, x désigne un réel strictement positif.

Il s'agit de produire un estimateur lisse de la courbe de Lorenz $(F(x), L_F(x))$, $x > 0$ puis de proposer ensuite une estimation de l'indice de Gini. L'idée est de construire des estimateurs lisses de la fonction de répartition F et de la fonction L_F au point x .

Introduisons \hat{F}_n l'estimateur à noyau de la fonction de répartition F défini par

$$\hat{F}_n(x) = \frac{1}{h_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad (1.10)$$

où K est une fonction de répartition telle que

$$K(u) = \int_{-\infty}^u K_0(t) dt. \quad (1.11)$$

Yamato ([165]) a prouvé les bonnes propriétés asymptotiques de cet estimateur et notamment sa convergence uniforme.

Nous tentons maintenant de construire un estimateur à noyau de la fonction L_F au point x .

1.4.1 Estimation lisse de la courbe de Lorenz

La fonction L_F est une fonctionnelle de la fonction de répartition. Afin d'approximer $L_F(x)$ en tout point x , nous proposons un estimateur dont la construction est donnée

par (A.18) que nous notons $\hat{L}_n^{(0)}(x)$. Il s'écrit de la manière suivante :

$$\hat{L}_n^{(0)}(x) = \int_0^{+\infty} L_{F_n}(z) \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz, \quad (1.12)$$

où L_{F_n} est l'estimateur empirique de la fonction L_F décrit dans [146], obtenu en substituant F_n à F dans (1.8), c'est-à-dire

$$L_{F_n}(x) = \frac{\int_0^x s dF_n(s)}{\int_0^{+\infty} s dF_n(s)}, \quad x > 0. \quad (1.13)$$

La formule de L_{F_n} résulte également du changement de variable $p = F_n(x)$ dans (1.2). Ainsi, nous avons :

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} [\int_0^z s dF_n(s)] \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz}{\int_0^{+\infty} s dF_n(s)},$$

donc

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} \left[\int_0^{+\infty} l(s, z) dF_n(s) \right] dz}{\int_0^{+\infty} s dF_n(s)},$$

où la fonction $l : \mathbb{R}^2 \rightarrow \mathbb{R}$ est telle que $l(s, z) := s \mathbf{1}_{[s \leq z]} \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right)$. Remarquons que :

$$\begin{aligned} s &\mapsto \int_0^{+\infty} l(s, z) dz \text{ est } F_n\text{-intégrable sur } [0, +\infty[, \\ z &\mapsto \int_0^{+\infty} l(s, z) dF_n(s) \text{ est intégrable sur } [0, +\infty[. \end{aligned}$$

En effet, le changement de variable $u = \frac{z-x}{h_n}$ et l'hypothèse (H1(ii)-(iii)) impliquent que :

$$\begin{aligned} \int_0^{+\infty} \left[\int_0^{+\infty} l(s, z) dz \right] dF_n(s) &= \int_0^{x+h_n} s \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right] dF_n(s) \\ &\leq \int_0^{x+h_n} s \left[\int_{-1}^1 |K_r^{(0)}(u)| du \right] dF_n(s) \\ &= 2 \sup_{u \in [-1, 1]} |K_r^{(0)}(u)| \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{[X_i \leq x+h_n]} \\ &\leq 2(x+h_n) \sup_{u \in [-1, 1]} |K_r^{(0)}(u)| \\ &< +\infty, \end{aligned}$$

et

$$\begin{aligned}
 \int_0^{+\infty} \left[\int_0^{+\infty} l(s, z) dF_n(s) \right] dz &= \int_{x-h_n}^{x+h_n} \frac{1}{h_n} K_r^{(0)} \left(\frac{z-x}{h_n} \right) \left[\frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{[X_i \leq z]} \right] dz \\
 &= \frac{1}{n} \sum_{i=1}^n X_i \int_{-1}^1 K_r^{(0)}(u) \mathbf{1}_{[X_i \leq x+h_n u]} du \\
 &\leq 2(x+h_n) \sup_{u \in [-1,1]} |K_r^{(0)}(u)| \\
 &< +\infty.
 \end{aligned}$$

Nous pouvons donc appliquer le théorème de Fubini :

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} \left[\int_0^{+\infty} s \mathbf{1}_{[s \leq z]} \frac{1}{h_n} K_r^{(0)} \left(\frac{z-x}{h_n} \right) dz \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)},$$

d'où

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} s \left[\int_s^{+\infty} \frac{1}{h_n} K_r^{(0)} \left(\frac{z-x}{h_n} \right) dz \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)}.$$

À l'aide du changement de variable $u = \frac{z-x}{h_n}$ et par (H1)-(iii), nous arrivons à :

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} s \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)}.$$

En posant pour tout $s > 0$,

$$\Psi_{n,0}(s; x) := \mathbf{1}_{[s < x+h_n]} \int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du, \tag{1.14}$$

il vient que :

$$\hat{L}_n^{(0)}(x) = \frac{\int_0^{+\infty} s \Psi_{n,0}(s; x) dF_n(s)}{\int_0^{+\infty} s dF_n(s)}. \tag{1.15}$$

Grâce à l'expression obtenue (1.15), nous présentons une propriété de $\hat{L}_n^{(0)}$.

Lemme 1.4.1. *Sous les hypothèses (H1) et (H2), on a pour tout $x > 0$,*

$$\left| \hat{L}_n^{(0)}(x) \right| \leq 1.$$

Le résultat du Lemme 1.4.1 montre une similitude entre l'estimateur $\hat{L}_n^{(0)}$ et la fonction L_F . En effet, pour tout $x > 0$,

$$|L_F(x)| = \frac{\int_0^x s dF(s)}{\int_0^{+\infty} s dF(s)} \leq 1.$$

Remarque :

Nous pouvons également écrire $L_F(x) \geq 0$ pour tout $x > 0$. Cependant, il nous est

impossible d'obtenir $\hat{L}_n^{(0)}(x) \geq 0$ pour tout $x > 0$ car pour tout $r \geq 0$, $K_r^{(0)}$ n'est pas une fonction positive.

Pour compléter l'étude de notre estimateur $\hat{L}_n^{(0)}(x)$, nous nous focalisons désormais sur son comportement asymptotique. Nous établissons plusieurs résultats dont la convergence presque sûre.

Théorème 1.4.1. *Nous supposons que les hypothèses (H1) et (H2) sont vérifiées. Alors pour tout $x > 0$, l'estimateur $\hat{L}_n^{(0)}$ est convergent presque sûrement,*

$$\hat{L}_n^{(0)}(x) \xrightarrow{p.s} L_F(x). \quad (1.16)$$

Les convergences en probabilité et en moyenne quadratique de l'estimateur $\hat{L}_n^{(0)}(x)$ découlent du Théorème 1.4.1, d'où le théorème suivant :

Théorème 1.4.2. *Soit $x > 0$ fixé. Alors sous les hypothèses (H1) et (H2) :*

(i) *l'estimateur $\hat{L}_n^{(0)}$ est convergent en probabilité,*

$$\hat{L}_n^{(0)}(x) \xrightarrow{\mathbb{P}} L_F(x). \quad (1.17)$$

De plus,

(ii) *l'estimateur $\hat{L}_n^{(0)}$ est convergent en moyenne quadratique,*

$$\mathbb{E} \left[\hat{L}_n^{(0)}(x) - L_F(x) \right]^2 \rightarrow 0. \quad (1.18)$$

Nous terminons par un résultat sur la convergence en loi de l'estimateur $\hat{L}_n^{(0)}(x)$.

Théorème 1.4.3. *Soit $x > 0$ fixé. Si les hypothèses (H1) et (H2) sont satisfaites, alors nous avons*

$$\sqrt{n} \left(\hat{L}_n^{(0)}(x) - L_F(x) \right) \xrightarrow{\mathcal{L}} N(0, \sigma_L^2(x)), \quad (1.19)$$

avec

$$\sigma_L^2(x) = \frac{\mu_2(x) \mu_1^2 - 2\mu_1(x) \mu_2(x) \mu_1 + \mu_1^2(x) \mu_2}{\mu_1^4}, \quad (1.20)$$

où $\mu_k(x) = \int_0^x s^k dF(s)$ et $\mu_k = \mu_k(+\infty)$ pour $k = 1, 2$.

Remarque :

Le terme $\sigma_L^2(x)$ est une généralisation du second paramètre de la loi normale asymptotique obtenue dans le cas discret avec L_{F_n} dans [73].

Corollaire au Théorème 1.4.3. *Un intervalle de confiance asymptotique pour $L_F(x)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\hat{L}_n^{(0)}(x) \pm z_{1-\epsilon/2} \sqrt{\frac{\sigma_L^2(x)}{n}}, \quad (1.21)$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite.

Ainsi, la courbe de Lorenz $(F(x), L_F(x))$, $x > 0$ peut être estimée par $(\hat{F}_n(x), \hat{L}_n^{(0)}(x))$, $x > 0$, où \hat{F}_n et $\hat{L}_n^{(0)}$ sont des estimateurs à noyau respectivement donnés par (1.10) et (1.15).

1.4.2 Application à l'indice de Gini

Rappelons la formule dans [81] de l'indice de Gini noté G découlant de la fonction L_F :

$$G = \int_0^{+\infty} L_F(z) dF(z). \quad (1.22)$$

Nous considérons naturellement son estimation en injectant l'estimateur polynomial local $\hat{L}_n^{(0)}$ à la place de la fonction L_F , c'est-à-dire par

$$\hat{G}_n = \int_0^{+\infty} \hat{L}_n^{(0)}(z) d\hat{F}_n(z), \quad (1.23)$$

où \hat{F}_n est donné par (1.10).

Le théorème suivant expose la convergence de l'estimateur \hat{G}_n vers l'indice de Gini G .

Théorème 1.4.4. *Sous les conditions (H1) et (H2), on a*

$$\hat{G}_n \rightarrow G \quad (1.24)$$

lorsque $n \rightarrow +\infty$.

1.5 Estimation lisse du temps total de test normalisé et application au remplacement de composants non réparables

1.5.1 Estimation lisse du temps total de test normalisé

Soit x un réel strictement positif.

Il s'agit de produire un estimateur lisse du temps total de test normalisé $(F(x), T_F(x))$, $x > 0$. L'idée est de construire des estimateurs lisses de la fonction de répartition F et de la fonction T_F au point x .

Nous utiliserons \hat{F}_n donné par (1.10) comme estimateur lisse de la fonction de répartition. Il nous reste à déterminer un estimateur lisse de la fonction T_F , ce qui constitue l'objectif de cette section. Nous reproduisons les étapes de la section précédente pour approximer $T_F(x)$ en tout point x car la fonction T_F est une fonctionnelle de la fonction de répartition. Commençons par donner l'écriture de l'estimateur $\hat{T}_n^{(0)}(x)$ dont la construction est décrite par (A.18) :

$$\hat{T}_n^{(0)}(x) = \int_0^{+\infty} T_{F_n}(z) \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz, \quad (1.25)$$

où T_{F_n} est donné par

$$T_{F_n}(x) = \frac{\int_0^x (1 - F_n(s)) ds}{\int_0^{+\infty} (1 - F_n(s)) ds}, \quad x > 0. \quad (1.26)$$

Cette formule découle du changement de variable $p = F_n(x)$ dans (1.4). T_{F_n} est également l'estimateur empirique de la fonction T_F présenté dans [146], obtenu en substituant F_n à F dans (1.9). En remplaçant T_{F_n} par son expression (1.26) dans (1.25), nous avons :

$$\hat{T}_n^{(0)}(x) = \frac{\int_0^{+\infty} [\int_0^z (1 - F_n(s)) ds] \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz}{\int_0^{+\infty} (1 - F_n(s)) ds}$$

donc

$$\hat{T}_n^{(0)}(x) = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} [\int_0^{+\infty} t(s, z) ds] dz}{\int_0^{+\infty} (1 - F_n(s)) ds},$$

où la fonction $t : \mathbb{R}^2 \rightarrow \mathbb{R}$ est telle que $t(s, z) := \mathbf{1}_{[s \leq z]} \mathbf{1}_{[s < X_i]} \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right)$. Nous obtenons que les deux fonctions :

$$s \mapsto \int_0^{+\infty} t(s, z) dz \quad \text{et} \quad z \mapsto \int_0^{+\infty} t(s, z) ds$$

sont dans $\mathbb{L}^1([0, +\infty[)$ d'après le changement de variable $u = \frac{z-x}{h_n}$ et l'hypothèse (H1(ii)-(iii)). Voici les détails :

$$\begin{aligned} \int_0^{+\infty} \left[\int_0^{+\infty} t(s, z) dz \right] ds &= \int_0^{x+h_n} \mathbf{1}_{[s < X_i]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right] ds \\ &\leq \int_0^{\min(X_i, x+h_n)} \left[\int_{-1}^1 |K_r^{(0)}(u)| du \right] ds \\ &\leq 2(x+h_n) \sup_{u \in [-1, 1]} |K_r^{(0)}(u)| \\ &< +\infty, \end{aligned}$$

et

$$\begin{aligned} \int_0^{+\infty} \left[\int_0^{+\infty} t(s, z) ds \right] dz &= \int_{x-h_n}^{x+h_n} \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) \left[\int_0^{\min(X_i, z)} ds \right] dz \\ &= \int_{-1}^1 K_r^{(0)}(u) \min(X_i, x+h_n u) du \\ &\leq 2(x+h_n) \sup_{u \in [-1, 1]} |K_r^{(0)}(u)| \\ &< +\infty. \end{aligned}$$

Ainsi, l'utilisation légitime du théorème de Fubini entraîne que :

$$\hat{T}_n^{(0)}(x) = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s \leq z]} \mathbf{1}_{[s < X_i]} \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz \right] ds}{\int_0^{+\infty} (1 - F_n(s)) ds},$$

c'est-à-dire

$$\hat{T}_n^{(0)}(x) = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) dz \right] ds}{\int_0^{+\infty} (1 - F_n(s)) ds}.$$

En appliquant le changement de variable $u = \frac{z-x}{h_n}$ et utilisant l'hypothèse (H1(iii)), nous obtenons :

$$\hat{T}_n^{(0)}(x) = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) dz \right] ds}{\int_0^{+\infty} (1 - F_n(s)) ds},$$

et grâce à (1.14), il vient que :

$$\hat{T}_n^{(0)}(x) = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds}{\int_0^{+\infty} (1 - F_n(s)) ds}. \quad (1.27)$$

Un premier résultat concernant l'estimateur $\hat{T}_n^{(0)}$ peut être proposé grâce à l'analyse de l'écriture (1.27).

Lemme 1.5.1. *Sous les hypothèses (H1) et (H2), on a pour tout $x > 0$,*

$$|\hat{T}_n^{(0)}(x)| \leq 1.$$

Le Lemme 1.5.1 met en évidence une inégalité commune à l'estimateur $\hat{T}_n^{(0)}$ et à la fonction T_F . En effet, pour tout $x > 0$,

$$|T_F(x)| = \frac{\int_0^x (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds} \leq 1.$$

Remarque :

Pour la même raison que $\hat{L}_n^{(0)}(x)$, il nous est impossible d'obtenir $\hat{T}_n^{(0)}(x) \geq 0$ pour tout $x > 0$. Pourtant, nous pouvons écrire $T_F(x) \geq 0$ pour tout $x > 0$.

Plusieurs propriétés asymptotiques de l'estimateur $\hat{T}_n^{(0)}$ ont été établies. La convergence presque sûre est le sujet du théorème suivant.

Théorème 1.5.1. *Nous supposons que les hypothèses (H1) et (H2) sont vérifiées. Alors pour tout $x > 0$, l'estimateur $\hat{T}_n^{(0)}$ est convergent presque sûrement,*

$$\hat{T}_n^{(0)}(x) \xrightarrow{p.s.} T_F(x). \quad (1.28)$$

Par suite, nous énonçons un théorème qui traite des convergences en probabilité et en moyenne quadratique de l'estimateur $\hat{T}_n^{(0)}$.

Théorème 1.5.2. Soit $x > 0$ fixé. Alors sous les hypothèses (H1) et (H2) :

(i) l'estimateur $\hat{T}_n^{(0)}$ est convergent en probabilité,

$$\hat{T}_n^{(0)}(x) \xrightarrow{\mathbb{P}} T_F(x). \quad (1.29)$$

De plus,

(ii) l'estimateur $\hat{T}_n^{(0)}$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{T}_n^{(0)}(x) - T_F(x) \right]^2 \rightarrow 0. \quad (1.30)$$

L'exploration de l'estimateur $\hat{T}_n^{(0)}$ s'achève avec l'obtention de sa convergence en loi.

Théorème 1.5.3. Soit $x > 0$ fixé. Si les hypothèses (H1) et (H2) sont satisfaites, alors nous avons

$$\sqrt{n} \left(\hat{T}_n^{(0)}(x) - T_F(x) \right) \xrightarrow{\mathcal{L}} N(0, \sigma_T^2(x)), \quad (1.31)$$

avec

$$\sigma_T^2(x) = \frac{\nu_2(x) \mu_1^2 - 2\nu_1(x) \nu_3(x) \mu_1 + \nu_1^2(x) \mu_2}{\mu_1^4}, \quad (1.32)$$

où $\mu_k(x) = \int_0^x s^k dF(s)$, $\mu_k = \mu_k(+\infty)$, $\nu_1(x) = \int_0^x s dF(s) + x \int_x^{+\infty} dF(s)$, $\nu_2(x) = \int_0^x s^2 dF(s) + x^2 \int_x^{+\infty} dF(s)$ et $\nu_3(x) = \int_0^x s^2 dF(s) + x \int_x^{+\infty} s dF(s)$.

Corollaire au Théorème 1.5.3. Un intervalle de confiance asymptotique pour $T_F(x)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes

$$\hat{T}_n^{(0)}(x) \pm z_{1-\epsilon/2} \sqrt{\frac{\sigma_T^2(x)}{n}}, \quad (1.33)$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite.

Ainsi, le temps total de test normalisé $(F(x), T_F(x))$, $x > 0$ peut être estimé par $(\hat{F}_n(x), \hat{T}_n^{(0)}(x))$, $x > 0$, où \hat{F}_n et $\hat{T}_n^{(0)}$ sont des estimateurs à noyau respectivement donnés par (1.10) et (1.27).

1.5.2 Détermination du moment optimal de remplacement des composants non réparables

Bergman ([29]) puis Bergman et Klefsjö ([30]) ont expliqué l'utilité du temps total de test normalisé dans les stratégies de remplacement des composants non réparables.

Supposons que des composants d'un même type sont indépendants et que leurs instants de défaillance sont décrits par la distribution continue et strictement croissante F .

À tout moment, un composant peut être remplacé à condition de payer un *coût de remplacement* noté C_1 . S'il y a défaillance de ce composant, un coût supplémentaire C_2 , que l'on appelle *coût de défaillance*, doit être réglé.

Le moment optimal de remplacement des composants non réparables est déterminé

par un problème d'optimisation. Il s'agit de trouver l'instant $T = T^*$ qui minimise la fonction de coût donnée par

$$C(T) = \frac{C_1 + C_2 F(T)}{\int_0^T (1 - F(s)) ds}, \quad (1.34)$$

correspondant aux coûts générés par les remplacements et défaillances successifs des composants.

La fonction F est continue et strictement croissante donc la fonction inverse F^{-1} de la fonction F existe. Il résulte que pour tout t appartenant à $[0, 1]$, nous pouvons écrire $t = F(T)$ et

$$\begin{aligned} C(T) &= \frac{C_1 + C_2 t}{\int_0^{F^{-1}(t)} (1 - F(s)) ds} \\ &= \frac{C_1 + C_2 t}{\mu T(t)} \end{aligned}$$

où $T(p) := \frac{H_F^{-1}(p)}{H_F^{-1}(1)}$, $0 \leq p \leq 1$, est le temps total de test normalisé présenté dans l'Introduction. Nous remarquons que minimiser la fonction de coût C en T revient à :

$$\text{Maximiser}_t g(t) = \frac{\mu T(t)}{C_1 + C_2 t}. \quad (1.35)$$

La fonction g est dérivable sur $[0, 1]$. D'une part, la dérivée de la fonction g s'annule lorsque

$$T(t) = \frac{T'(t)(C_1 + C_2 t)}{C_2} := l(t). \quad (1.36)$$

D'autre part, la dérivée de la fonction g s'annule lorsque

$$g(t) = \frac{\mu T'(t)}{C_2}. \quad (1.37)$$

La définition (1.36) de la fonction l permet d'écrire que $l\left(-\frac{C_1}{C_2}\right) = 0$. En utilisant (1.37), il vient que le problème d'optimisation (1.35) est équivalent au problème d'optimisation suivant :

$$\text{Maximiser}_t T'(t). \quad (1.38)$$

Avec (1.36) et (1.38), le maximum t^* du problème d'optimisation (1.35) est atteint lorsque la courbe du temps total de test normalisé coupe la droite passant par le point $\left(-\frac{C_1}{C_2}, 0\right)$ ayant la pente la plus élevée. Une représentation pour déterminer graphiquement t^* se situe ci-après :

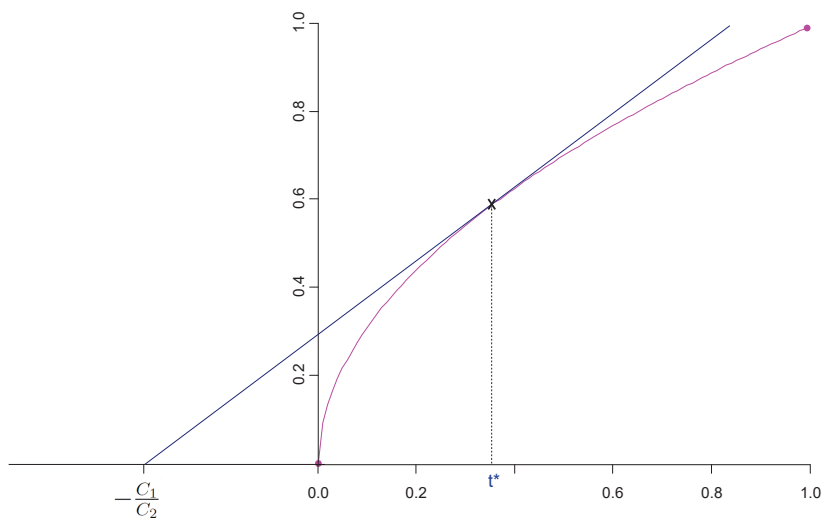


FIGURE 1.1 – Méthode graphique pour déterminer le moment optimal de remplacement des composants non réparables, à partir du graphique du temps total de test normalisé (en magenta).

Une dernière étape est nécessaire pour définir le moment optimal de remplacement des composants (qui minimise la fonction de coût C). Il faut calculer $T^* = F^{-1}(t^*)$, où F^{-1} désigne la fonction inverse de la fonction de répartition F , c'est-à-dire

$$F^{-1}(t) = \inf_{x>0} \{x : F(x) \geq t\}. \quad (1.39)$$

Lorsque la fonction de répartition F est inconnue mais que nous disposons d'un échantillon de données issu de F , nous pouvons construire un estimateur lisse ($\hat{F}_n(x), T_n^{(0)}(x)$), $x > 0$, du temps total de test normalisé ($F(x), T_F(x)$), $x > 0$, ou encore $(t, T(t))$, $0 \leq t \leq 1$, où $\hat{F}_n(x)$ est l'estimateur à noyau (1.10) de la fonction de répartition F et $T_n^{(0)}(x)$ est l'estimateur polynomial local de la fonction T_F au point x . Alors la méthode graphique permettra de déterminer une estimation $\hat{F}_n(x^*)$ de t^* puis $\hat{F}_n^{-1}(\hat{F}_n(x^*)) = x^*$ estimera le moment optimal de remplacement des composants T^* , avec

$$\hat{F}_n^{-1}(t) = \inf_{x>0} \{x : \hat{F}_n(x) \geq t\}. \quad (1.40)$$

1.6 Preuves

Dans toutes les preuves, x représente un réel strictement positif.

1.6.1 Preuve du Lemme 1.4.1

D'après (1.15), l'estimateur $\hat{L}_n^{(0)}$ au point x a pour expression :

$$\left| \hat{L}_n^{(0)}(x) \right| = \frac{\left| \int_0^{+\infty} s \Psi_{n,0}(s; x) dF_n(s) \right|}{\left| \int_0^{+\infty} s dF_n(s) \right|}$$

ce qui entraîne l'inégalité suivante :

$$\left| \hat{L}_n^{(0)}(x) \right| \leq \frac{\int_0^{+\infty} s \cdot \sup_{s>0} |\Psi_{n,0}(s; x)| dF_n(s)}{\int_0^{+\infty} s dF_n(s)}. \quad (1.41)$$

Or, grâce à la définition (1.14), nous écrivons :

$$\sup_{s>0} |\Psi_{n,0}(s; x)| = \sup_{s>0} \mathbf{1}_{[s < x+h_n]} \left| \int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right|.$$

Puisque pour tout $s > 0$,

$$\mathbf{1}_{[s < x+h_n]} \leq 1$$

et

$$\left| \int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right| \leq \left| \int_{-1}^1 K_r^{(0)}(u) du \right| = 1 \quad (1.42)$$

par (H1(i)), nous déduisons que :

$$\sup_{s>0} |\Psi_{n,0}(s; x)| \leq 1, \quad (1.43)$$

d'où en injectant dans (1.41) :

$$\left| \hat{L}_n^{(0)}(x) \right| \leq \frac{\int_0^{+\infty} s dF_n(s)}{\int_0^{+\infty} s dF_n(s)} = 1. \blacksquare$$

1.6.2 Preuve du Théorème 1.4.1

D'après (1.8), nous écrivons

$$L_F(x) = \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\int_0^{+\infty} s dF(s)} = \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\mu_1}. \quad (1.44)$$

La définition (1.15) entraîne donc que :

$$\begin{aligned}
& \left| \hat{L}_n^{(0)}(x) - L_F(x) \right| \\
&= \left| \frac{\int_0^{+\infty} s \Psi_{n,0}(s; x) dF_n(s) - \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s)}{\int_0^{+\infty} s dF_n(s)} \right| \\
&\quad + \left| \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s) - \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\int_0^{+\infty} s dF_n(s)} \right| \\
&\quad + \left| \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\int_0^{+\infty} s dF_n(s)} - L_F(x) \right| \\
&:= \frac{|B_{1,n}|}{\bar{X}_n} + \left| \frac{(\mu_1 - \bar{X}_n) \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\mu_1 \bar{X}_n} \right|. \tag{1.45}
\end{aligned}$$

Analysons tout d'abord le terme $|B_{1,n}|$.

$$\begin{aligned}
|B_{1,n}| &= \left| \int_0^{+\infty} s \Psi_{n,0}(s; x) dF_n(s) - \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s) \right| \\
&= \left| \int_0^{+\infty} s \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right] dF_n(s) - \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s) \right| \\
&= \left| \int_0^{+\infty} s \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right] dF_n(s) + \int_x^{x+h_n} s dF_n(s) \right|.
\end{aligned}$$

En utilisant l'inégalité triangulaire, il résulte que

$$|B_{1,n}| \leq \left| \int_0^{+\infty} s \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right] dF_n(s) \right| + \left| \int_x^{x+h_n} s dF_n(s) \right| \tag{1.46}$$

Par le résultat (1.42), la quantité $\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du$ est bornée. L'utilisation du théorème de convergence dominée et la convergence de $\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du$ vers 0 impliquent que le premier terme de (1.46) converge vers 0. Quant au second terme de (1.46), constatons que

$$\int_x^{x+h_n} s dF_n(s) \leq (x+h_n) (F_n(x+h_n) - F_n(x)).$$

Or :

$$\begin{aligned}
& |F_n(x+h_n) - F_n(x)| \\
&\leq |F_n(x+h_n) - F(x+h_n)| + |F(x) - F_n(x)| + |F(x+h_n) - F(x)|. \tag{1.47}
\end{aligned}$$

Les deux premières valeurs absolues convergent vers 0 presque sûrement grâce au théorème de Glivenko-Cantelli F.5. Quant à la dernière valeur absolue, elle converge également vers 0 presque sûrement en raison de la continuité de F et du fait que $x+h_n \rightarrow x$ quand $n \rightarrow +\infty$ grâce à (H2(ii)). Ainsi,

$$|B_{1,n}| \xrightarrow{p.s.} 0. \tag{1.48}$$

Étudions ensuite le terme $|B_{2,n}|$. Comme les variables aléatoires $X_1 \mathbf{1}_{[X_1 \leq x]}, \dots,$

$X_n \mathbf{1}_{[X_n \leq x]}$ sont i.i.d. et $\mathbb{E} X_1^2 \mathbf{1}_{[X_1 \leq x]} < +\infty$, la loi forte des grands nombres F.6 entraîne que :

$$|B_{2,n}| \xrightarrow{p.s.} 0. \quad (1.49)$$

La loi forte des grands nombres appliquées aux variables aléatoires i.i.d. X_1, \dots, X_n , nous permet par ailleurs d'obtenir :

$$\overline{X_n} \xrightarrow{p.s.} \mu_1 < +\infty. \quad (1.50)$$

De (1.48), (1.49), (1.50) et d'après le théorème des fonctions continues F.7 (car μ_1 n'est pas nul) , nous déduisons :

$$\frac{|B_{1,n}| + |B_{2,n}|}{\overline{X_n}} \xrightarrow{p.s.} 0. \quad (1.51)$$

Enfin, l'intégrale $\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)$ est égale à μ_1 supposée finie. En utilisant (1.50) et en appliquant le théorème des fonctions continues,

$$\left| \frac{(\mu_1 - \overline{X_n}) \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\mu_1 \overline{X_n}} \right| \xrightarrow{p.s.} 0. \quad (1.52)$$

Afin d'obtenir la convergence presque sûre de l'expression (1.45), nous devons appliquer une dernière fois le théorème des fonctions continues (à (1.51) et à (1.52)). Nous pouvons alors conclure :

$$\left| \hat{L}_n^{(0)}(x) - L_F(x) \right| \xrightarrow{p.s.} 0. \blacksquare$$

1.6.3 Preuve du Théorème 1.4.2

La convergence presque sûre entraîne la convergence en probabilité donc le Théorème 1.4.1 entraîne (i).

Le résultat (i) couplé avec celui du Lemme 1.4.1 nous place dans les conditions de l'Annexe F.8, d'où (ii). \blacksquare

1.6.4 Preuve du Théorème 1.4.3

Reprenons (1.44) :

$$L_F(x) = \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s)}{\int_0^{+\infty} s dF(s)},$$

et introduisons

$$\tilde{L}_n(x) := \frac{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s)}{\int_0^{+\infty} s dF_n(s)}, \quad (1.53)$$

d'où

$$\begin{aligned}\hat{L}_n^{(0)}(x) &= \tilde{L}_n(x) \frac{\int_0^{+\infty} s \Psi_{n,0}(s; x) dF_n(s)}{\int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF_n(s)} \\ &= \tilde{L}_n(x) \frac{\frac{1}{n} \sum_{i=1}^n X_i \Psi_{n,0}(X_i; x)}{\frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{[X_i \leq x]}}.\end{aligned}\quad (1.54)$$

Remarque :

Le support de la densité f étant $[0, +\infty]$, on suppose pour tout $x > 0$ fixé, qu'il existe à la fois des réalisations de X_i inférieures ou égales à x ainsi que supérieures à x . Le dénominateur ne peut donc être nul.

Lemme 1.6.1. *Soit $x > 0$ fixé. Alors pour tout $s > 0$, nous avons*

$$\Psi_{n,0}(s; x) \rightarrow \mathbf{1}_{[s \leq x]}, \quad (1.55)$$

lorsque $n \rightarrow +\infty$.

Grâce à (1.55), nous pouvons écrire lorsque $n \rightarrow +\infty$,

$$\Psi_{n,0}(X_i; x) \rightarrow X_i \mathbf{1}_{[X_i \leq x]} \text{ pour tout } 1 \leq i \leq n,$$

donc

$$\frac{1}{n} \sum_{i=1}^n X_i \Psi_{n,0}(X_i; x) \sim \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{[X_i \leq x]}$$

donc le second facteur de (1.54) converge en probabilité vers 1. Par suite, $\hat{L}_n^{(0)}(x)$ et $\tilde{L}_n(x)$ ont la même loi asymptotique puis la distribution limite de

$$\sqrt{n} \left(\hat{L}_n^{(0)}(x) - L_F(x) \right)$$

est identique à celle de

$$\sqrt{n} \left(\tilde{L}_n(x) - L_F(x) \right).$$

Soit $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ la fonction telle que pour tout $s > 0$, $\mathbf{g}(s) = (g_1(s), g_2(s))$ avec

$$g_1(s) = s \mathbf{1}_{[s \leq x]} \quad \text{et} \quad g_2(s) = s.$$

Alors pour la variable aléatoire positive X_1 suivant la fonction de répartition F , on a :

$$\begin{aligned}\mathbb{E} g_1(X_1) &= \mathbb{E} \left[X_1 \mathbf{1}_{[X_1 \leq x]} \right] \\ &= \int_0^{+\infty} s \mathbf{1}_{[s \leq x]} dF(s) \\ &= \int_0^x s dF(s) \\ &:= \mu_1(x),\end{aligned}$$

$$\begin{aligned}\mathbb{E} g_2(X_1) &= \mathbb{E} X_1 \\ &= \int_0^{+\infty} s dF(s) \\ &:= \mu_1,\end{aligned}$$

$$\begin{aligned}
 \mathbb{E} [g_1(X_1)]^2 &= \mathbb{E} [X_1^2 \mathbf{1}_{[X_1 \leq x]}] \\
 &= \mathbb{E} [X_1^2 \mathbf{1}_{[X_1 \leq x]}] \\
 &= \int_0^{+\infty} s^2 \mathbf{1}_{[s \leq x]} dF(s) \\
 &= \int_0^x s^2 dF(s) \\
 &:= \mu_2(x),
 \end{aligned}$$

$$\begin{aligned}
 \mathbb{E} [g_2(X_1)]^2 &= \mathbb{E} X_1^2 \\
 &= \int_0^{+\infty} s^2 dF(s) \\
 &:= \mu_2
 \end{aligned}$$

et

$$\begin{aligned}
 \mathbb{E} [g_1(X_1) g_2(X_1)] &= \mathbb{E} [X_1^2 \mathbf{1}_{[X_1 \leq x]}] \\
 &= \int_0^{+\infty} s^2 \mathbf{1}_{[s \leq x]} dF(s) \\
 &= \int_0^x s^2 dF(s) \\
 &:= \mu_2(x).
 \end{aligned}$$

D'où

$$\begin{aligned}
 \text{Var } g_1(X_1) &= \mathbb{E} [g_1(X_1)]^2 - \mathbb{E}^2 g_1(X_1) \\
 &= \mu_2(x) - \mu_1^2(x),
 \end{aligned}$$

$$\begin{aligned}
 \text{Var } g_2(X_1) &= \mathbb{E} [g_2(X_1)]^2 - \mathbb{E}^2 g_2(X_1) \\
 &= \mu_2 - \mu_1^2
 \end{aligned}$$

et

$$\begin{aligned}
 \text{Cov} [g_1(X_1), g_2(X_1)] &= \mathbb{E} [g_1(X_1) g_2(X_1)] - \mathbb{E} g_1(X_1) \mathbb{E} g_2(X_1) \\
 &= \mu_2(x) - \mu_1(x) \mu_1.
 \end{aligned}$$

D'après (1.6), nous avons

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} g_1(s) dF_n(s) \\ \int_0^{+\infty} g_2(s) dF_n(s) \end{pmatrix} - \begin{pmatrix} \mathbb{E} g_1(X_1) \\ \mathbb{E} g_2(X_1) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} N_2(0, \Sigma_{\mathbf{g}})$$

c'est-à-dire

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} s \mathbf{1}_{(s \leq x)} dF_n(s) \\ \int_0^{+\infty} s dF_n(s) \end{pmatrix} - \begin{pmatrix} \mu_1(x) \\ \mu_1 \end{pmatrix} \right) \xrightarrow{\mathcal{L}} N_2(0, \Sigma_{\mathbf{g}})$$

avec

$$\Sigma_{\mathbf{g}} = \begin{pmatrix} \mu_2(x) - \mu_1^2(x) & \mu_2(x) - \mu_1(x)\mu_1 \\ \mu_2(x) - \mu_1(x)\mu_1 & \mu_2 - \mu_1^2 \end{pmatrix}.$$

Remarque :

On applique le théorème de Lindeberg-Feller F.9 aux vecteurs aléatoires i.i.d.

$$Y_{n,i} = \begin{pmatrix} \frac{1}{\sqrt{n}} X_i \mathbf{1}_{[X_i \leq x]} \\ \frac{1}{\sqrt{n}} X_i \end{pmatrix}.$$

Considérons $\mathbf{h} : \mathbb{R}^2 \mapsto \mathbb{R}$ la fonction telle que pour tout couple $(u, v) \in \mathbb{R}^2$,

$$\mathbf{h}(u, v) = \frac{u}{v},$$

alors la matrice des dérivées premières partielles de h s'écrit

$$D\mathbf{h} = \begin{pmatrix} \frac{\partial}{\partial u} h(u, v) & \frac{\partial}{\partial v} h(u, v) \end{pmatrix} = \begin{pmatrix} 1 & -\frac{u}{v^2} \end{pmatrix} \quad (1.56)$$

Par (1.7), nous obtenons

$$\sqrt{n} \left(\frac{\int_0^{+\infty} s \mathbf{1}_{(s \leq x)} dF_n(s)}{\int_0^{+\infty} s dF_n(s)} - \frac{\mu_1(x)}{\mu_1} \right) \xrightarrow{\mathcal{L}} \mathbf{N}(0, \sigma_L^2),$$

avec

$$\begin{aligned} \sigma_L^2 &= D\mathbf{h}(\mathbb{E} \mathbf{g}) \Sigma_{\mathbf{g}} D\mathbf{h}^T(\mathbb{E} \mathbf{g}) \\ &= \left(\frac{1}{\mu_1} \right)^2 (\mu_2(x) - \mu_1^2(x)) + 2 \frac{1}{\mu_1} \left(-\frac{\mu_1(x)}{\mu_1^2} \right) (\mu_2(x) - \mu_1(x)\mu_1) + \left(-\frac{\mu_1(x)}{\mu_1^2} \right)^2 (\mu_2 - \mu_1^2) \\ &= \frac{\mu_2(x)\mu_1^2 - 2\mu_1(x)\mu_2(x)\mu_1 + \mu_1^2(x)\mu_2}{\mu_1^4}. \end{aligned}$$

Il en résulte que :

$$\sqrt{n} (\tilde{L}_n(x) - L_F(x)) \xrightarrow{\mathcal{L}} \mathbf{N}(0, \sigma_L^2),$$

avec σ_L^2 exprimé ci-dessus.

Pour terminer la preuve, il reste à démontrer le Lemme 1.6.1.

1er cas : si $s > x$

Tout d'abord,

$$s > x \implies \mathbf{1}_{[s \leq x]} = 0.$$

Par ailleurs,

$$\begin{aligned} s > x &\implies s - x > 0 \\ &\implies \frac{s - x}{h_n} \rightarrow +\infty \text{ grâce à (H2)} \\ &\implies \int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \rightarrow 0 \text{ par (H1(iii))} \\ &\implies \Psi_{n,0}(s; x) \rightarrow 0 \text{ car } \mathbf{1}_{[s < x + h_n]} = 0 \text{ ou } 1. \end{aligned}$$

Ainsi,

$$s > x \implies \Psi_{n,0}(s; x) \rightarrow \mathbf{1}_{(s \leq x)} \text{ lorsque } n \rightarrow +\infty.$$

2nd cas : si $s \leq x$

Tout d'abord,

$$s \leq x \implies \mathbf{1}_{[s \leq x]} = 1.$$

Par ailleurs,

$$\begin{aligned} s \leq x &\implies s < x + h_n \text{ grâce à (H1(iii)) et (H2(i))} \\ &\implies \mathbf{1}_{[s < x + h_n]} = 1 \\ &\implies \Psi_{n,0}(s; x) = \int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \\ &\implies \Psi_{n,0}(s; x) \rightarrow 1 \end{aligned}$$

car $s - x \leq 0$ donc $\frac{s-x}{h_n} \rightarrow -\infty$ grâce à (H2(ii)-(iii)).

Ainsi,

$$s \leq x \implies \Psi_{n,0}(s; x) \rightarrow \mathbf{1}_{[s \leq x]} \text{ lorsque } n \rightarrow +\infty. \blacksquare$$

1.6.5 Preuve du Théorème 1.4.4

Grâce aux définitions (1.22) et (1.23), nous écrivons

$$\begin{aligned} |\hat{G}_n - G| &= \left| \int_0^{+\infty} \hat{L}_n^{(0)}(x) d\hat{F}_n(x) - \int_0^{+\infty} L_F(x) d\hat{F}_n(x) \right. \\ &\quad \left. + \int_0^{+\infty} L_F(x) d\hat{F}_n(x) - \int_0^{+\infty} L_F(x) dF(x) \right| \end{aligned}$$

En appliquant l'inégalité triangulaire, nous obtenons

$$\begin{aligned} |\hat{G}_n - G| &\leq \left| \int_0^{+\infty} (\hat{L}_n^{(0)}(x) - L_F(x)) d\hat{F}_n(x) \right| + \left| \int_0^{+\infty} L_F(x) (d\hat{F}_n(x) - dF(x)) \right| \\ &\leq \int_0^{+\infty} |\hat{L}_n^{(0)}(x) - L_F(x)| d\hat{F}_n(x) + \left| \int_0^{+\infty} L_F(x) d(\hat{F}_n(x) - F(x)) \right| \end{aligned}$$

Or on a

$$|L_F(x)| \leq 1$$

et

$$|\hat{L}_n^{(0)}(x) - L_F(x)| \leq |\hat{L}_n^{(0)}(x)| + |L_F(x)| \leq 2$$

d'après le Lemme 1.4.1, donc nous pouvons utiliser le théorème de convergence dominée F.4. En vertu de (1.18) et des bonnes propriétés de l'estimateur à noyau \hat{F}_n (voir [165]) de la fonction de répartition F , nous écrivons

$$|\hat{L}_n^{(0)}(x) - L_F(x)| \rightarrow 0$$

et

$$d(\hat{F}_n(x) - F(x)) \rightarrow 0.$$

On en déduit que \hat{G}_n converge vers G lorsque $n \rightarrow +\infty$. \blacksquare

1.6.6 Preuve du Lemme 1.5.1

D'après (1.27), l'estimateur $\hat{T}_n^{(0)}$ au point x a pour expression :

$$|\hat{T}_n^{(0)}(x)| = \frac{\left| \frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds \right|}{\left| \int_0^{+\infty} (1 - F_n(s)) ds \right|},$$

ce qui entraîne l'inégalité suivante :

$$|\hat{T}_n^{(0)}(x)| \leq \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{X_i} \sup_{s>0} |\Psi_{n,0}(s; x)| ds}{\int_0^{+\infty} (1 - F_n(s)) ds}. \quad (1.57)$$

En injectant (1.43) dans (1.57), nous obtenons :

$$|\hat{T}_n^{(0)}(x)| \leq \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{X_i} ds}{\int_0^{+\infty} (1 - F_n(s)) ds},$$

c'est-à-dire

$$|\hat{T}_n^{(0)}(x)| \leq \frac{\frac{1}{n} \sum_{i=1}^n X_i}{\frac{1}{n} \sum_{i=1}^n X_i} = 1. \blacksquare$$

1.6.7 Preuve du Théorème 1.5.1

Constatons que (1.9) peut se réécrire comme

$$T_F(x) = \frac{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds} = \frac{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\mu_1}. \quad (1.58)$$

et

$$\begin{aligned} \int_0^{+\infty} (1 - F_n(s)) ds &= \int_0^{+\infty} \left[1 - \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(X_i \leq s)} \right] ds \\ &= \frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{(X_i > s)} ds \\ &= \frac{1}{n} \sum_{i=1}^n \int_0^{X_i} ds \\ &= \frac{1}{n} \sum_{i=1}^n X_i \\ &= \int_0^{+\infty} s dF_n(s). \end{aligned} \quad (1.59)$$

Alors la définition (1.27) entraîne que :

$$\begin{aligned}
 & \left| \hat{T}_n^{(0)}(x) - T_F(x) \right| \\
 &= \left| \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds - \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F_n(s)) ds}{\int_0^{+\infty} s dF_n(s)} \right. \\
 & \quad \left. + \frac{\left| \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F_n(s)) ds - \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds \right|}{\int_0^{+\infty} s dF_n(s)} \right. \\
 & \quad \left. + \left| \frac{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\int_0^{+\infty} s dF_n(s)} - T_F(x) \right| \right| \\
 &:= \frac{|D_{1,n}| + |D_{2,n}|}{\bar{X}_n} + \left| \frac{(\mu_1 - \bar{X}_n) \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\mu_1 \bar{X}_n} \right|. \tag{1.60}
 \end{aligned}$$

Analysons tout d'abord le terme $|D_{1,n}|$.

$$\begin{aligned}
 |D_{1,n}| &= \left| \frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds - \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F_n(s)) ds \right| \\
 &= \left| \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \Psi_{n,0}(s; x) ds \right] dF_n(t) - \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s \leq x]} ds \right] dF_n(t) \right| \\
 &= \left| \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s < x+h_n]} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du \right] ds \right] dF_n(t) \right. \\
 & \quad \left. - \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s \leq x]} ds \right] dF_n(t) \right| \\
 &= \left| \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s < x+h_n]} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right. \\
 & \quad \left. + \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s < x+h_n]} ds \right] dF_n(t) - \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[s < t]} \mathbf{1}_{[s \leq x]} ds \right] dF_n(t) \right|.
 \end{aligned}$$

Grâce à l'inégalité triangulaire, on a

$$\begin{aligned}
 |D_{1,n}| &\leq \left| \int_0^{+\infty} \left[\int_0^{\min(t, x+h_n)} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right| \\
 & \quad + \left| \int_0^{+\infty} \left[\int_0^{\min(t, x+h_n)} ds \right] dF_n(t) - \int_0^{+\infty} \left[\int_0^{\min(t, x)} ds \right] dF_n(t) \right|. \tag{1.61}
 \end{aligned}$$

D'où

$$\begin{aligned}
|D_{1,n}| &\leq \left| \int_0^{x+h_n} \left[\int_0^t \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right| \\
&\quad + \left| \int_{x+h_n}^{+\infty} \left[\int_0^{x+h_n} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right| \\
&\quad + \left| \int_0^{x+h_n} \left[\int_0^t ds \right] dF_n(t) + \int_{x+h_n}^{+\infty} \left[\int_0^{x+h_n} ds \right] dF_n(t) \right| \\
&\quad + \left| \int_0^x \left[\int_0^t ds \right] dF_n(t) + \int_x^{+\infty} \left[\int_0^x ds \right] dF_n(t) \right| \\
&= \left| \int_0^{x+h_n} \left[\int_0^t \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right| \tag{1.62}
\end{aligned}$$

$$+ \left| \int_{x+h_n}^{+\infty} \left[\int_0^{x+h_n} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du - 1 \right) ds \right] dF_n(t) \right| \tag{1.63}$$

$$+ \left| \int_x^{x+h_n} \left[\int_0^t ds - \int_0^x ds \right] dF_n(t) - \int_{x+h_n}^{+\infty} \left[\int_x^{x+h_n} ds \right] dF_n(t) \right| \tag{1.64}$$

La quantité $\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du$ est bornée d'après (1.42). Le théorème de convergence dominée, la convergence de $\int_{\frac{s-x}{h_n}}^1 K_r^{(0)}(u) du$ vers 0 et $\lim_{u \rightarrow +\infty} F(u) = 1$ impliquent que les termes (1.62) et (1.63) convergent vers 0. Quant au terme (1.64), nous remarquons qu'il est majoré par

$$(t+x)|F_n(x+h_n) - F_n(x)| + h_n|1 - F_n(x+h_n)|$$

Le théorème de Glivenko-Cantelli F.5, la continuité de F et le fait que $x+h_n \rightarrow x$ quand $n \rightarrow +\infty$ appliqués à la décomposition (1.47) de $|F(x+h_n) - F(x)|$ puis la majoration de F_n par 1 permettent d'obtenir

$$|D_{1,n}| \xrightarrow{p.s.} 0. \tag{1.65}$$

Étudions ensuite le terme $|D_{2,n}|$. Nous pouvons écrire :

$$|D_{2,n}| = \left| \int_0^x (F(s) - F_n(s)) ds \right| \leq x \cdot \sup_{s>0} |F(s) - F_n(s)|$$

En appliquant le théorème de Glivenko-Cantelli, on a :

$$|D_{2,n}| \xrightarrow{p.s.} 0. \tag{1.66}$$

De (1.65), (1.66), (1.50) et d'après le théorème des fonctions continues F.7 (car μ_1 est non nul), nous déduisons :

$$\frac{|D_{1,n}| + |D_{2,n}|}{X_n} \xrightarrow{p.s.} 0. \tag{1.67}$$

Enfin, l'intégrale $\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds$ est égale à μ_1 supposée finie. En utilisant

(1.50) et en appliquant le théorème des fonctions continues,

$$\left| \frac{(\mu_1 - \bar{X}_n) \int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\mu_1 \bar{X}_n} \right| \xrightarrow{p.s.} 0. \quad (1.68)$$

La convergence presque sûre de l'expression (1.60) :

$$\left| \hat{T}_n^{(0)}(x) - T_F(x) \right| \xrightarrow{p.s.} 0.$$

découle du théorème des fonctions continues appliqué à (1.67) et à (1.68). ■

1.6.8 Preuve du Théorème 1.5.2

La démonstration est identique à celle du Théorème 1.4.2.

La convergence en probabilité (i) est obtenue à l'aide de la convergence presque sûre du Théorème 1.5.1.

Les conditions de l'Annexe F.8 sont satisfaites grâce à la convergence en probabilité (i) utilisée conjointement avec Lemme 1.5.1. Par conséquent, nous avons (ii). ■

1.6.9 Preuve du Théorème 1.5.3

En utilisant (1.58), nous avons :

$$T_F(x) = \frac{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds}, \quad (1.69)$$

et introduisons

$$\tilde{T}_n(x) := \frac{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F_n(s)) ds}{\int_0^{+\infty} (1 - F_n(s)) ds}, \quad (1.70)$$

d'où

$$\begin{aligned} \hat{T}_n^{(0)}(x) &= \tilde{T}_n(x) \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds}{\int_0^{+\infty} \mathbf{1}_{[s \leq x]} (1 - F_n(s)) ds} \\ &= \tilde{T}_n(x) \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds}{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s \leq x]} ds}. \end{aligned} \quad (1.71)$$

Grâce à (1.55) et par le théorème de convergence dominée, nous pouvons écrire

$$\int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds \rightarrow \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s \leq x]} ds \text{ pour tout } 1 \leq i \leq n,$$

donc

$$\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \Psi_{n,0}(s; x) ds \sim \frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s \leq x]} ds$$

donc le second facteur de (1.71) converge en probabilité vers 1. Par suite, les lois

asymptotiques de $\hat{T}_n^{(0)}(x)$ et de $\tilde{T}_n(x)$ sont identiques, d'où la distribution limite de

$$\sqrt{n}(\hat{T}_n^{(0)}(x) - T_F(x))$$

est la même que celle de

$$\sqrt{n}(\tilde{T}_n(x) - T_F(x)).$$

L'égalité (1.59) conduit à :

$$\begin{aligned} \tilde{T}_n(x) &= \frac{\frac{1}{n} \sum_{i=1}^n \int_0^{+\infty} \mathbf{1}_{(s < X_i)} \mathbf{1}_{(s \leq x)} ds}{\int_0^{+\infty} s dF_n(s)} \\ &= \frac{\int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{(t < s)} \mathbf{1}_{(t \leq x)} dt \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)}. \end{aligned}$$

Nous introduisons $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ la fonction telle que pour tout $s > 0$, $\mathbf{g}(s) = (g_3(s), g_2(s))$ avec

$$g_3(s) = \int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \quad \text{et} \quad g_2(s) = s.$$

Alors pour la variable aléatoire positive X_1 suivant la fonction de répartition F , on a :

$$\begin{aligned} \mathbb{E} g_3(X_1) &= \mathbb{E} \int_0^{+\infty} \mathbf{1}_{[t < X_1]} \mathbf{1}_{[t \leq x]} dt \\ &= \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \right] dF(s) \\ &= \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) \\ &= \int_0^x \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) + \int_x^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) \\ &= \int_0^x \left[\int_0^s dt \right] dF(s) + \int_x^{+\infty} \left[\int_0^x dt \right] dF(s) \\ &= \int_0^x s dF(s) + x \int_x^{+\infty} dF(s) \\ &:= \nu_1(x), \end{aligned}$$

$$\mathbb{E} g_2(X_1) = \mu_1,$$

$$\begin{aligned}
 \mathbb{E} [g_3(X_1)]^2 &= \mathbb{E} \left[\int_0^{+\infty} \mathbf{1}_{[t < X_1]} \mathbf{1}_{[t \leq x]} dt \right]^2 \\
 &= \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \right]^2 dF(s) \\
 &= \int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right]^2 dF(s) \\
 &= \int_0^x \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right]^2 dF(s) + \int_x^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right]^2 dF(s) \\
 &= \int_0^x \left[\int_0^s dt \right]^2 dF(s) + \int_x^{+\infty} \left[\int_0^x dt \right]^2 dF(s) \\
 &= \int_0^x s^2 dF(s) + x^2 \int_x^{+\infty} dF(s) \\
 &:= \nu_2(x),
 \end{aligned}$$

$$\mathbb{E} [g_2(X_1)]^2 = \mu_2$$

et

$$\begin{aligned}
 \mathbb{E} [g_3(X_1) g_2(X_1)] &= \mathbb{E} \left[X_1 \int_0^{+\infty} \mathbf{1}_{[t < X_1]} \mathbf{1}_{[t \leq x]} dt \right] \\
 &= \int_0^{+\infty} s \left[\int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \right] dF(s) \\
 &= \int_0^{+\infty} s \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) \\
 &= \int_0^x s \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) + \int_x^{+\infty} s \left[\int_0^{+\infty} \mathbf{1}_{[t < \min(s, x)]} dt \right] dF(s) \\
 &= \int_0^x s \left[\int_0^s dt \right] dF(s) + \int_x^{+\infty} s \left[\int_0^x dt \right] dF(s) \\
 &= \int_0^x s^2 dF(s) + x \int_x^{+\infty} s dF(s) \\
 &:= \nu_3(x).
 \end{aligned}$$

D'où

$$\begin{aligned}
 \text{Var } g_3(X_1) &= \mathbb{E} [g_3(X_1)]^2 - \mathbb{E}^2 g_3(X_1) \\
 &= \nu_2(x) - \nu_1^2(x),
 \end{aligned}$$

$$\text{Var } g_2(X_1) = \mu_2 - \mu_1^2$$

et

$$\begin{aligned}
 \text{Cov} [g_3(X_1), g_2(X_1)] &= \mathbb{E} [g_3(X_1) g_2(X_1)] - \mathbb{E} g_3(X_1) \mathbb{E} g_2(X_1) \\
 &= \nu_3(x) - \nu_1(x) \mu_1.
 \end{aligned}$$

D'après (1.6), nous avons :

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} g_3(s) \mathbf{1}_{(s \leq x)} dF_n(s) \\ \int_0^{+\infty} g_2(s) dF_n(s) \end{pmatrix} - \begin{pmatrix} \mathbb{E} g_3(X_1) \\ \mathbb{E} g_2(X_1) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} N_2(0, \Sigma_{\mathbf{g}})$$

c'est-à-dire

$$\sqrt{n} \left(\left(\frac{\int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)} \right) - \begin{pmatrix} \nu_1(x) \\ \mu_1 \end{pmatrix} \right) \xrightarrow{\mathcal{L}} N_2(0, \Sigma_g)$$

avec

$$\Sigma_g = \begin{pmatrix} \nu_2(x) - \nu_1^2(x) & \nu_3(x) - \nu_1(x)\mu_1 \\ \nu_3(x) - \nu_1(x)\mu_1 & \mu_2 - \mu_1^2 \end{pmatrix}.$$

Remarque :

On applique le théorème de Lindeberg-Feller F.9 aux vecteurs aléatoires i.i.d.

$$Y_{n,i} = \begin{pmatrix} \frac{1}{\sqrt{n}} \int_0^{+\infty} \mathbf{1}_{[s < X_i]} \mathbf{1}_{[s \leq x]} ds \\ \frac{1}{\sqrt{n}} X_i \end{pmatrix}.$$

Nous nous reportons à la Preuve du Théorème 1.4.3 pour le choix de la fonction \mathbf{h} . Nous obtenons d'après (1.7) :

$$\sqrt{n} \left(\frac{\int_0^{+\infty} \left[\int_0^{+\infty} \mathbf{1}_{[t < s]} \mathbf{1}_{[t \leq x]} dt \right] dF_n(s)}{\int_0^{+\infty} s dF_n(s)} - \frac{\nu_1(x)}{\mu_1} \right) \xrightarrow{\mathcal{L}} N(0, \sigma_T^2),$$

où

$$\sigma_T^2 = D\mathbf{h}(\mathbb{E} \mathbf{g}) \Sigma_g D\mathbf{h}^T(\mathbb{E} \mathbf{g}),$$

avec $D\mathbf{h}$ défini en (1.56). Ainsi,

$$\begin{aligned} \sigma_T^2 &= \left(\frac{1}{\mu_1} \right)^2 (\nu_2(x) - \nu_1^2(x)) + 2 \frac{1}{\mu_1} \left(-\frac{\nu_1(x)}{\mu_1^2} \right) (\nu_3(x) - \nu_1(x)\mu_1) + \left(-\frac{\nu_1(x)}{\mu_1^2} \right)^2 (\mu_2 - \mu_1^2) \\ &= \frac{\nu_2(x)\mu_1^2 - 2\nu_1(x)\nu_3(x)\mu_1 + \nu_1^2(x)\mu_2}{\mu_1^4}. \end{aligned}$$

Nous en déduisons que :

$$\sqrt{n} (\tilde{T}_n(x) - T_F(x)) \xrightarrow{\mathcal{L}} N(0, \sigma_T^2),$$

avec σ_T^2 exprimé ci-dessus. ■

1.7 Simulations

1.7.1 Cadre

On suppose que les n données X_1, \dots, X_n d'un échantillon sont positives, indépendantes et identiquement distribuées. La fonction de répartition commune est notée F .

Les fonctions L_F et T_F sont définies respectivement par

$$L_F(x) = \frac{\int_0^x s dF(s)}{\int_0^{+\infty} s dF(s)}, \quad x > 0, \quad (1.72)$$

et

$$T_F(x) = \frac{\int_0^x (1 - F(s)) ds}{\int_0^{+\infty} (1 - F(s)) ds}, \quad x > 0. \quad (1.73)$$

Deux types d'estimateurs ont été décrits dans ce chapitre.

Tout d'abord, les estimateurs construits naturellement à partir des formules (1.72) et (1.73), en substituant la fonction de répartition empirique F_n du n -échantillon à la fonction de répartition F dans (1.72) et dans (1.73) :

$$L_{F_n}(x) = \frac{\int_0^x s dF_n(s)}{\int_0^{+\infty} s dF_n(s)}, \quad x > 0, \quad (1.74)$$

et

$$T_{F_n}(x) = \frac{\int_0^x (1 - F_n(s)) ds}{\int_0^{+\infty} (1 - F_n(s)) ds}, \quad x > 0. \quad (1.75)$$

Ensuite, les estimateurs lisses des fonctions L_F et T_F , obtenus par la méthode d'approximation polynomiale locale :

$$\hat{L}_n^{(0)}(x) = \int_0^{+\infty} L_{F_n}(z) \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz, \quad x > 0, \quad (1.76)$$

et

$$\hat{T}_n^{(0)}(x) = \int_0^{+\infty} T_{F_n}(z) \frac{1}{h_n} K_r^{(0)}\left(\frac{z-x}{h_n}\right) dz, \quad x > 0, \quad (1.77)$$

avec $r = 1$. Ces deux estimateurs à noyau s'écrivent à l'aide d'un noyau $K_1^{(0)}$ et d'un paramètre de lissage (ou d'une fenêtre) h_n .

Désignons par K_0 la densité d'Epanechnikov, c'est-à-dire une fonction positive, d'intégrale égale à 1, donnée par

$$K_0(x) = \frac{3}{4} (1 - x^2) \mathbf{1}_{[-1;1]}.$$

D'après (B.1), le noyau $K_1^{(0)}$ coïncide avec la densité K_0 . Le paramètre de lissage est considéré fixe.

1.7.2 Données simulées

Pour les simulations suivantes, la taille d'échantillon est fixée à $n = 100$. De plus, on choisit $h = 0.1$ pour la loi Exponentielle de paramètre 3 et pour la loi de Maxwell puis $h = 0.8$ pour la loi de Weibull de paramètres 2 et 1.2.

Nous traçons dans un premier temps, la fonction qui à tout $x > 0$ associe $L_F(x)$, $L_{F_n}(x)$ et $\hat{L}_n^{(0)}(x)$, sur un même graphique, afin d'illustrer le comportement asymptotique (les théorèmes obtenus) de l'estimateur lisse $\hat{L}_n^{(0)}(x)$ de la fonction L_F au point x . Dans un second temps, nous tentons d'approcher la courbe de Lorenz $(F(x), L_F(x))$, $x > 0$, permettant d'identifier à sa visualisation le degré d'inégalités au sein de la population, de laquelle est extrait un échantillon de n données. Pour arriver à cette fin, nous utilisons l'estimateur empirique $(F_n(x), L_{F_n}(x))$, $x > 0$, et l'estimateur lisse

$(\hat{F}_n(x), \hat{L}_n^{(0)}(x))$, $x > 0$, où $\hat{F}_n(x)$ est un estimateur à noyau (lisse) de la fonction de répartition F défini par (C.7), soit

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad (1.78)$$

avec K une fonction de répartition telle que

$$K(u) = \int_{-1}^u K_0(t) dt.$$

La figure suivante illustre l'estimation de la fonction L_F en chaque point x (à gauche) et de la courbe de Lorenz $(F(x), L_F(x))$, $x > 0$, (à droite) pour la loi Exponentielle de paramètre 3 et la loi de Maxwell.

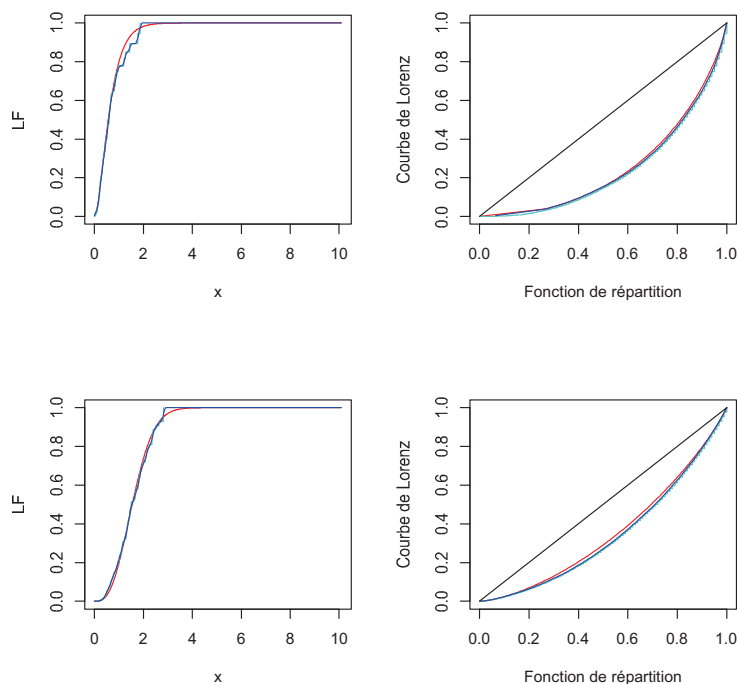


FIGURE 1.2 – Courbe de Lorenz théorique (en rouge) avec son estimateur empirique (en turquoise), son estimateur lisse (en bleu) et la ligne d'égalité parfaite (en noir).

On constate que $\hat{L}_n^{(0)}$ et $(\hat{F}_n(x), \hat{L}_n^{(0)}(x))$, $x > 0$, fournissent une estimation lisse, contrairement aux estimateurs existants L_{F_n} et $(F_n(x), L_{F_n}(x))$, $x > 0$. La distribution de Maxwell est moins inégalitaire que la distribution Exponentielle de paramètre 3, puisque sa courbe est plus proche de la ligne d'égalité parfaite. Les estimations sont cohérentes.

Le tableau suivant présente les erreurs d'estimation des deux estimateurs de la courbe de Lorenz $(F(x), L_F(x))$, $x > 0$. La colonne nommée "moyenne" indique l'erreur en moyenne quadratique intégrée MISE (Mean Integrated Square Error) de ces estimateurs. Les résultats sont le résumé statistique de 500 échantillons.

Taille de l'échantillon	écart-type		max		moyenne		min	
	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.
30	0.0077646	0.006806134	0.0757238	0.0403561	0.0072834	0.0066536	0.0002763	0.0002286
100	0.001989103	0.002100703	0.019263	0.02013	0.0020319	0.001882	0.00012	0.00005285
500	0.000387622	0.000352281	0.003702	0.002163	0.0003914	0.0003446	0.00002943	0.00002134

Taille de l'échantillon	écart-type		max		moyenne		min	
	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.
30	0.006130775	0.005678702	0.0639543	0.0525646	0.0062464	0.0057703	0.0002885	0.0002914
100	0.001640137	0.001368311	0.0160085	0.0088978	0.0017767	0.0015312	0.000133	0.0001326
500	0.000312997	0.000299917	0.003696	0.002768	0.0003425	0.0003231	0.00002863	0.00002064

FIGURE 1.3 – Erreur quadratique intégrée de l'estimateur empirique (Emp.) et de l'estimateur lisse (Poly.) de la courbe de Lorenz pour la loi Exponentielle de paramètre 3 (en haut) puis pour la loi de Maxwell (en bas), en fonction de 3 tailles d'échantillon.

Nous remarquons que les MISE des deux estimateurs diminuent et leur écart se réduit lorsque la taille de l'échantillon augmente. L'estimateur lisse semble plus performant puisque par sa MISE est plus faible que celle de l'estimateur empirique.

Intéressons-nous maintenant au temps total de test normalisé. Dans un premier temps, nous traçons sur un même graphique, la fonction qui à tout $x > 0$ associe $T_F(x)$, $T_{F_n}(x)$ et $\hat{T}_n^{(0)}(x)$ pour illustrer le comportement asymptotique (les théorèmes obtenus) de l'estimateur lisse $\hat{T}_n^{(0)}(x)$ de la fonction T_F au point x . Dans un second temps, nous tentons d'approcher le temps total de test normalisé $(F(x), T_F(x))$, $x > 0$, utile pour détecter une tendance sur la fiabilité d'une population de matériels non réparables, de laquelle est extrait un échantillon de n données, mais également pour déterminer l'instant optimal de remplacement des matériels défectueux. Afin de réaliser cet objectif, nous disposons de l'estimateur empirique $(F_n(x), T_{F_n}(x))$, $x > 0$, et de l'estimateur lisse $(\hat{F}_n(x), \hat{T}_n^{(0)}(x))$, $x > 0$, où $\hat{F}_n(x)$ est donné par (1.78).

La Figure 1.4 permet de visualiser l'estimation de la fonction T_F (à gauche) et du temps total de test normalisé $(F(x), T_F(x))$, $x > 0$, (à droite) pour la loi Exponentielle de paramètre 3 et la loi de Weibull de paramètres 0.7 et 1.2. Le contraste est visible entre les estimateurs lisses $\hat{T}_n^{(0)}$ et $(\hat{F}_n(x), \hat{T}_n^{(0)}(x))$, $x > 0$, et les estimateurs existants T_{F_n} et $(F_n(x), T_{F_n}(x))$, $x > 0$, seulement continus par morceaux. La distribution Exponentielle de paramètre 3 correspond à une absence de tendance car sa courbe coïncide avec la ligne noire, tandis que la distribution de Weibull de paramètres 2 et 1.2 traduit une tendance au vieillissement. Les estimations vont dans ce sens.

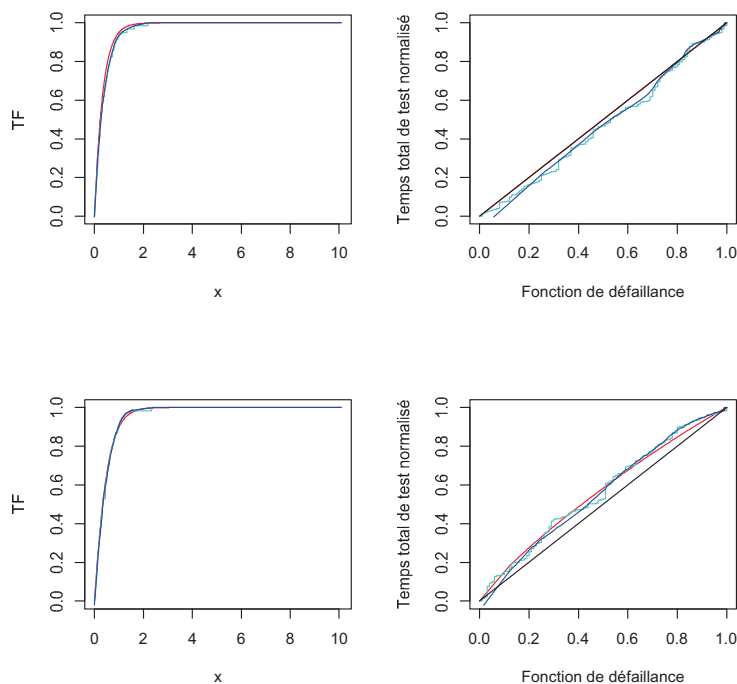


FIGURE 1.4 – Temps total de test normalisé théorique (en rouge) avec son estimateur empirique (en turquoise), son estimateur lisse (en bleu) et la droite d’absence de tendance (en noir).

Le tableau suivant décrit les erreurs d’estimation des deux estimateurs du temps total de test normalisé $(F(x), T_F(x))$, $x > 0$. La colonne nommée "moyenne" indique l’erreur en moyenne quadratique intégrée MISE (Mean Integrated Square Error) de ces estimateurs.

Taille de l'échantillon	écart-type		max		moyenne		min	
	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.
30	0.005625935	0.003458306	0.07424	0.02634	0.003926	0.00242	0.00001813	0.00003843
100	0.001639935	0.001030242	0.01976	0.007238	0.001198	0.0007761	0.000005291	0.000005693
500	0.000290162	0.000185665	0.002781	0.00161	0.0002266	0.0001399	0.00000061	0.000001753

Taille de l'échantillon	écart-type		max		moyenne		min	
	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.	Emp.	Poly.
30	0.00485692	0.00384356	0.08094	0.02955	0.003229	0.002772	0.00001015	0.00002311
100	0.001183899	0.000993763	0.01238	0.006494	0.0008727	0.0008258	0.00000357	0.000006359
500	0.000222012	0.000198933	0.004034	0.001296	0.0001681	0.0001648	0.000001145	0.000001407

FIGURE 1.5 – Erreur quadratique intégrée de l’estimateur empirique (Emp.) et de l’estimateur lisse (Poly.) du temps total de test normalisé pour la loi Exponentielle de paramètre 3 (en haut) puis pour la loi de Weibull de paramètres 2 et 1.2 (en bas), en fonction de 3 tailles d’échantillon.

En observant le résumé statistique de 500 échantillons, le constat est identique à celui de la Figure 1.3 : lorsque la taille de l'échantillon augmente, les MISE des deux estimateurs diminuent et leur écart se réduit. De plus, l'estimateur lisse se distingue par sa MISE inférieure à celle de l'estimateur empirique.

Finalement, la possession du critère de continuité (propriété des courbes théoriques) et d'une MISE plus compétitive justifie la préférence des estimateurs lisses $(\hat{F}_n(x), \hat{L}_n^{(0)}(x))$, $x > 0$, et $(\hat{F}_n(x), \hat{T}_n^{(0)}(x))$, $x > 0$, aux estimateurs existants $(F_n(x), L_{F_n}(x))$, $x > 0$ et $(F_n(x), T_{F_n}(x))$, $x > 0$.

1.7.3 Données réelles

Ampoules chez Norris Electronics

L'entreprise Norris Electronics fabrique des ampoules à forte intensité intervenant dans de nombreux produits électriques. Souhaitant augmenter leur performance, le groupe de recherche de l'entreprise a créé un nouveau filament dans le but d'allonger la durée de vie des ampoules. L'entreprise s'est ensuite intéressée à évaluer les avantages de ce nouveau filament en menant une étude à partir d'un échantillon de 200 ampoules. Les données indiquent le nombre d'heures d'éclairage avant que le filament ne grille.

La liste des données est fournie dans [12].

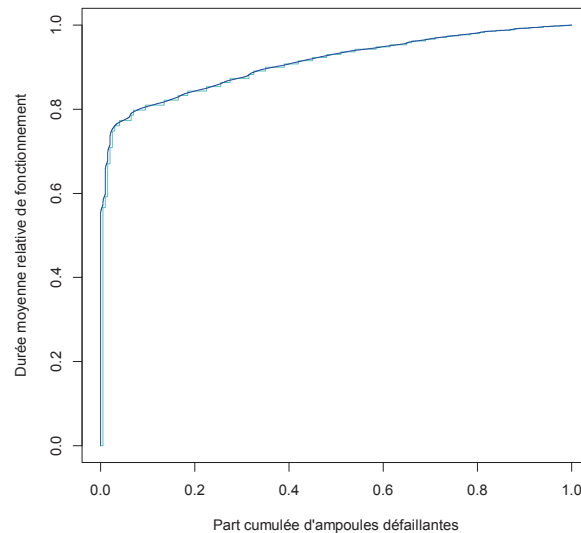


FIGURE 1.6 – Estimateur empirique (en turquoise) et estimateur lisse (en bleu) du temps total de test normalisé des ampoules, avec $h = 0.8$.

Nous remarquons que les deux estimateurs sont très proches. Cependant, l'estimateur lisse est continu comme la courbe initiale alors que l'estimateur empirique ne l'est pas. Une tendance au fort vieillissement est détectée puisque les estimateurs sont au-dessus

et très éloignés de la diagonale dans le carré unitaire. Si l'on connaissait la fonction de coût à minimiser, nous pourrions déterminer le moment optimal de remplacement des ampoules. Il semble que l'on doive considérer un remplacement assez rapide des ampoules.

Chapitre 2

Estimation à noyau de fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives

2.1 Introduction

Le but de ce second chapitre est de compléter l'étude des fonctions L_F et T_F . En s'intéressant à leurs dérivées successives, nous tentons de détecter les particularités de ces deux courbes. Nous proposons d'estimer les dérivées successives à l'aide de méthodes à noyau afin de conserver la continuité. Nous supposons que les revenus et les instants de défaillance sont indépendants et identiquement distribués.

L'estimation polynomiale locale a été introduite par Stone ([150], [151]) et Cleveland ([46]). Fan et Gijbels ([68]), Loader ([108]) puis Hjort et Jones ([84]) l'ont appliquée pour estimer la densité, Lejeune et Sarda ([107]) pour obtenir un estimateur lisse de la fonction de répartition ou encore Abdous et Berred ([5]) pour déterminer les estimateurs de la durée de vie résiduelle et de sa dérivée première. Abdous, Berlinet et Hengartner ([4]) ont ensuite utilisé l'approximation polynomiale locale pour estimer la classe des fonctionnelles de la fonction de répartition (toutes les fonctions dont l'expression dépend de la fonction de répartition) ayant une certaine régularité, ainsi que leurs dérivées successives. Berlinet ([32]) avait motivé l'emploi de noyaux d'ordre supérieur permettant de réduire considérablement le biais des estimateurs ainsi que d'une hiérarchie de noyaux afin de décrire tous les estimateurs des dérivées d'une fonction d'intérêt.

Abdous, Berlinet et Hengartner ([4]) ont obtenu la consistance des estimateurs sans donner de conditions précises. Nous essayons d'identifier ces conditions en considérant toutes les fonctionnelles linéaires de la fonction de répartition. La courbe de Lorenz et le temps total de test normalisé, étant deux illustrations de fonctionnelles de la fonction de répartition, nous employons la méthode décrite par Abdous, Berlinet et Hengartner ([4]) pour construire les estimateurs de leurs dérivées successives. Nous examinons ensuite leurs propriétés asymptotiques telles que divers types de convergences et la normalité asymptotique.

2.2 Cadre d'étude

Soient X_1, \dots, X_n des variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F .

Nous considérons la famille des fonctionnelles de la fonction de répartition F , c'est-à-dire l'ensemble des fonctions :

$$\begin{aligned} \Phi(\cdot, F) : I \subset \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \Phi(x, F) \end{aligned}$$

et notre défi est de proposer des estimateurs (possédant de bonnes propriétés asymptotiques) des dérivées, si elles existent, de ces fonctions.

Afin de décrire brièvement la construction des estimateurs, nous introduisons plusieurs hypothèses et notations.

2.2.1 Hypothèses et notations

Soit x un réel fixé.

Hypothèse (H1)

F_n est un estimateur consistant de F .

Hypothèse (H2)

$\Phi(x, \cdot)$ est continue par rapport à la norme sup.

Nous donnons plus de détails sur l'hypothèse (H2). Elle signifie que pour toute suite de fonctions de répartition F_n vérifiant $\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \rightarrow 0$, nous obtenons $\sup_{x \in \mathbb{R}} |\Phi(x, F_n) - \Phi(x, F)| \rightarrow 0$.

Désormais, F_n désignera la fonction de répartition empirique classique basée sur les observations X_1, \dots, X_n :

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

Alors l'hypothèse (H1) est satisfaite et puisque nous supposons (H2), le théorème de Glivenko-Cantelli (F.5) entraîne la convergence de $\Phi(x, F_n)$ vers $\Phi(x, F)$ en tout point x réel.

Soit r un entier non nul. On a besoin de poser l'hypothèse suivante :

Hypothèse (H3)

$\Phi(\cdot, F)$ est $(r + 1)$ fois continûment différentiable en tout point x .

Les dérivées de $\Phi(x, F)$ jusqu'à l'ordre r sont notées $\Phi^{(m)}(x, F)$, $m = 0, \dots, r$. Ce sont les coefficients dans le développement de Taylor au point x comme montré ci-dessous :

$$\Phi(x + hv, F) = \sum_{j=0}^r \frac{\Phi^{(j)}(x, F)}{j!} (hv)^j + O((hv)^{r+1}), \quad (2.1)$$

avec h et v deux réels fixés.

Comme $\Phi(x + hv, F_n)$ estime $\Phi(x + hv, F)$, une idée naturelle est d'approcher $\Phi(x + hv, F_n)$ localement par un polynôme et d'utiliser la valeur du polynôme et ses dérivées comme estimateurs de $\Phi^{(m)}(x, F)$. Plus clairement, soit K_0 une densité de probabilité ayant pour support $[-1, 1]$ et formons la projection de $\Phi(x + h\cdot, F_n)$ sur l'espace \mathbb{V} des polynômes d'ordre inférieur ou égal à r . Cela revient à minimiser la fonction objectif :

$$J(a_0(x), \dots, a_r(x)) = \int \left[\Phi(x + hv, F_n) - \sum_{j=0}^r \frac{a_j(x)}{j!} (hv)^j \right]^2 K_0(v) dv.$$

Nommons le minimiseur $\hat{a}_j(x)$, $j = 0, \dots, r$. Comme $\Phi(x + hv, F_n)$ estime $\Phi(x + hv, F)$ et ce dernier a la représentation (2.1), la m^e dérivée de $\Phi(\cdot, F)$ au point x , notée $\Phi^{(m)}(x, F)$, est estimée par

$$\hat{\Phi}_n^{(m)}(x) = \hat{a}_m(x). \quad (2.2)$$

Puisque le polynôme minimiseur est la projection de $\Phi(x + h\cdot, F_n)$ sur l'espace \mathbb{V} des polynômes d'ordre inférieur ou égal à r , nous avons également la représentation (A.11), d'où

$$\sum_{j=0}^r \frac{\hat{a}_j(x)}{j!} (hv)^j = \sum_{j=0}^r \int \Phi(x + hu, F_n) Q_j(u) d\mu(u) Q_j(v),$$

avec $(Q_j)_{j=0, \dots, r}$ une base orthonormale de l'espace \mathbb{V} des polynômes d'ordre inférieur ou égal à r . Or

$$\hat{a}_m(x) = \frac{1}{h^m} \frac{\partial^m}{\partial v^m} \sum_{j=0}^r \frac{\hat{a}_j(x)}{j!} (hv)^j \Big|_{v=0}$$

donc grâce à (2.2), l'estimateur de $\Phi^{(m)}(x, F)$ peut être réécrit comme

$$\hat{\Phi}_n^{(m)}(x) = \frac{1}{h^m} \int \Phi(z, F_n) \frac{1}{h} K_r^{(m)} \left(\frac{z-x}{h} \right) dz \quad (2.3)$$

avec $K_r^{(m)}$ donné par (A.15) et après avoir effectué le changement de variable $z = x + hu$. On retrouve la formule (A.18).

Nous imposons au noyau $K_r^{(m)}(x)$ de vérifier (A.16) puis d'être borné et de support $[-1, 1]$, ce qui se traduit par :

Hypothèse (H4)

(i)

$$\int_{-1}^1 u^i K_r^{(m)}(u) du = \begin{cases} m!, & i = 0, \\ 0, & i \neq m \text{ et } i \in \{1, \dots, r\}, \\ C_{r+1}, & i > r, \end{cases}$$

avec C_{r+1} une constante finie,

(ii)

$$C := \sup_{u \in [-1, 1]} |K_r^{(m)}(u)| < +\infty,$$

(iii) Le support de $K_r^{(m)}$ est $[-1, 1]$.

Pour une idée plus distincte des noyaux $K_r^{(m)}$, nous renvoyons le lecteur à l'Annexe B où il y consultera les calculs des premiers termes $K_r^{(m)}$ de la hiérarchie de noyaux associée à la densité d'Epanechnikov pour $r \leq 5$ et $m \leq r$.

Nous supposons que la fenêtre notée h_n dépend du nombre n d'observations et vérifie l'hypothèse suivante :

Hypothèse (H5)

(h_n) est une suite décroissante de nombres réels strictement positifs telle que $h_n \rightarrow 0$ lorsque $n \rightarrow +\infty$.

Ainsi, l'expression (2.3) de l'estimateur de $\Phi^{(m)}(x, F)$ est légèrement modifiée et devient :

$$\hat{\Phi}_n^{(m)}(x) = \int \Phi(z, F_n) \frac{1}{h_n^{m+1}} K_r^{(m)}(z - x) dz. \quad (2.4)$$

Dans [4], il est énoncé que l'estimateur est consistant, mais sans fournir de conditions précises. Notre travail consistera à les identifier. Nous nous intéresserons également à la normalité asymptotique. Nous essayons tout d'abord dans le cas des fonctionnelles linéaires de la fonction de répartition.

2.3 Estimation polynomiale locale des fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives

Soit $x \in \mathbb{R}$. Nous nous plaçons maintenant dans la situation où la fonctionnelle $\Phi(x, \cdot)$ est linéaire, c'est-à-dire de la forme

$$\Phi(x, F) = \int \phi(x, s) dF(s). \quad (2.5)$$

Alors l'estimateur de $\Phi^{(m)}(x, F)$ s'écrit :

$$\begin{aligned}\hat{\Phi}_n^{(m)}(x) &= \int \int \phi(z, s) dF_n(s) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz \\ &= \int \int \phi(z, s) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz dF_n(s) \\ &= \frac{1}{n} \sum_{i=1}^n \int \phi(z, X_i) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz.\end{aligned}\quad (2.6)$$

On reconnaît une moyenne empirique de copies i.i.d. d'une variable aléatoire

$$\int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz,$$

où X_1 a pour distribution F .

Le lemme suivant fournit le biais de l'estimateur $\hat{\Phi}_n^{(m)}$ au point x .

Lemme 2.3.1. *Soient x un réel fixé et m un entier compris entre 0 et r . Si les hypothèses (H2) à (H4) sont vérifiées, alors*

$$\mathbb{E} \hat{\Phi}_n^{(m)}(x) = \Phi^{(m)}(x, F) + O\left(h_n^{r+1-m}\right). \quad (2.7)$$

On dit que $\hat{\Phi}_n^{(m)}(x)$ est un estimateur asymptotiquement sans biais de $\Phi^{(m)}(x, F)$ sous l'hypothèse (H5).

L'égalité (2.6) combinée à l'indépendance et à l'équidistribution des variables aléatoires X_1, \dots, X_n nous amène à :

$$\text{Var} \hat{\Phi}_n^{(m)}(x) = \frac{1}{n} \text{Var} \int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz, \quad (2.8)$$

où

$$\begin{aligned}\text{Var} \int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz \\ = \mathbb{E} \left[\int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz \right]^2\end{aligned}\quad (2.9)$$

$$- \left(\mathbb{E} \int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)}\left(\frac{z-x}{h_n}\right) dz \right)^2. \quad (2.10)$$

Afin de déterminer la variance exacte asymptotique de notre estimateur $\hat{\Phi}_n^{(m)}(x)$, nous

avons besoin de calculer le premier terme du membre de droite (2.9). Donc :

$$\begin{aligned}
 & \mathbb{E} \left[\int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \right]^2 \\
 &= \int \left(\int \phi(z, s) \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \right)^2 dF(s) \\
 &= \frac{1}{h_n^{2m}} \int \left(\int \phi(z, s) \frac{1}{h_n} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \right)^2 f(s) ds \\
 &= \frac{1}{h_n^{2m}} \int \left(\int \phi(x + h_n u, s) K_r^{(m)}(u) du \right)^2 f(s) ds. \tag{2.11}
 \end{aligned}$$

À partir d'ici, il est intéressant de remarquer que le prolongement du calcul est impossible sans posséder d'information supplémentaire sur la fonction ϕ . De plus, il semble compliqué d'affirmer que l'ordre de convergence est identique pour toute fonction ϕ . Rappelons qu'à une fonction ϕ donnée est associée une fonction $\Phi(x, F)$ ainsi que les estimateurs $\hat{\Phi}_n^{(m)}(x)$, $0 \leq m \leq r$, permettant d'approximer respectivement les fonctions $\Phi^{(m)}(x, F)$, $0 \leq m \leq r$. Donc si nous souhaitons estimer la m^e dérivée d'une fonction $\Phi(x, F)$, ce qui consiste notamment à déterminer la variance de son estimateur $\hat{\Phi}_n^{(m)}(x)$, alors il devient nécessaire de spécifier la fonction ϕ .

Nous comprenons donc qu'il est impossible de fournir les conditions de consistance des fonctionnelles linéaires de la fonction de répartition globalement. Au contraire, nous analyserons au cas par cas et tenterons d'obtenir la consistance et la normalité asymptotique pour chacune des fonctionnelles de la fonction de répartition étudiées dans ce manuscrit.

Grâce à l'expression (2.11), nous constatons également que les cas $m = 0$ et $m \geq 1$ doivent être distingués, en raison des propriétés des noyaux $K_r^{(m)}$. La différence essentielle découle de (H4)-(i) et s'écrit :

$$\int K_r^{(m)}(u) du = \begin{cases} 1, & \text{si } m = 0, \\ 0, & \text{si } m \geq 1. \end{cases}$$

Voici la raison pour laquelle nous avons traité séparément l'estimation de la courbe de Lorenz et du temps total de test normalisé (Chapitre 1) et celle de leurs dérivées successives (dans la Section 2.6).

Pour la suite du chapitre, nous ajoutons que la fonction de répartition commune aux X_1, \dots, X_n est supposée continue et la densité de probabilité $F' = f$ est de support $[0, +\infty[$.

2.4 Estimation polynomiale locale de la densité et de ses dérivées

Nous nous intéressons dans cette section à l'estimation de la densité f et de ses dérivées successives au point $x \in \mathbb{R}$. Nous optons pour

$$\phi(x, s) = \mathbf{1}_{(-\infty, x]}(s) \quad (2.12)$$

afin d'obtenir à l'aide de (2.5) :

$$\Phi(x, F) = F(x). \quad (2.13)$$

Soit $r > 0$. Notre but est d'étudier les estimateurs $\hat{\Phi}_n^{(m)}(x)$ pour $1 \leq m \leq r$ ou de manière équivalente les estimateurs $\hat{f}_n^{(m)}(x)$ des dérivées $f^{(m)}(x)$ pour $0 \leq m \leq r - 1$.

La fonction de répartition F supposée continue implique que (H2) est satisfaite et l'hypothèse (H3) se transforme en :

Hypothèse (H3')

f est r fois continûment différentiable en tout point x .

Nous injectons (2.13) dans (2.6) mais nous veillons à prendre l'ordre $(m + 1)$ de $\hat{\Phi}_n(x)$. En effet, pour tout m supérieur ou égal à 0,

$$\hat{f}_n^{(m)}(x) = \hat{\Phi}_n^{(m+1)}(x, F) \quad (2.14)$$

pour le choix (2.13). L'estimateur de la dérivée $f^{(m)}(x)$ s'exprime donc comme suit :

$$\hat{f}_n^{(m)}(x) = \frac{1}{n} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n^{m+2}} K_r^{(m+1)}\left(\frac{z-x}{h_n}\right) dz \quad (2.15)$$

$$= \frac{1}{nh_n^{m+1}} \sum_{i=1}^n \int_{\frac{X_i-x}{h_n}}^1 K_r^{(m+1)}(u) du \cdot \mathbf{1}_{[x-h_n \leq X_i \leq x+h_n]}, \quad (2.16)$$

la dernière ligne résultant du changement de variable $u = \frac{z-x}{h_n}$.

Nous commençons par montrer un résultat similaire à (3.4), sur la variance exacte asymptotique de l'estimateur $\hat{f}_n^{(m)}(x)$.

Lemme 2.4.1. *Soient x un réel fixé et m un entier compris entre 0 et $r - 1$. Alors sous les hypothèses (H3')-(H4),*

$$\mathbb{V}ar \hat{f}_n^{(m)}(x) = \frac{1}{nh_n^{2m+1}} (b_{m+1}(x) + O(h_n)) - \frac{\left(f^{(m)}(x) + O(h_n^{r-m})\right)^2}{n}, \quad (2.17)$$

avec

$$b_{m+1}(x) = f(x) \int_{-1}^1 \left(\int_v^1 K_r^{(m+1)}(u) du \right) dv. \quad (2.18)$$

Nous continuons avec un théorème donnant les conditions de convergence \mathbb{L}^2 et en probabilité de l'estimateur $\hat{f}_n^{(m)}(x)$.

Théorème 2.4.1. *Soient x un réel fixé et m un entier compris entre 0 et $r - 1$. Sous $(H3')$ - $(H4)$ et si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (2.19)$$

suffisamment lentement tel que

$$n h_n^{2m+1} \rightarrow +\infty, \quad (2.20)$$

alors l'estimateur $\hat{f}_n^{(m)}(x)$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{f}_n^{(m)}(x) - f^{(m)}(x) \right]^2 \rightarrow 0, \quad (2.21)$$

donc également convergent en probabilité,

$$\hat{f}_n^{(m)}(x) \xrightarrow{\mathbb{P}} f^{(m)}(x). \quad (2.22)$$

Nous établissons maintenant la normalité asymptotique de l'estimateur $\hat{f}_n^{(m)}(x)$, ce qui conduit à la construction possible d'intervalles de confiance.

Théorème 2.4.2. *Soient x un réel fixé et m un entier compris entre 0 et $r - 1$. Alors sous les hypothèses $(H3')$ - $(H4)$ et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$n h_n \rightarrow +\infty \quad (2.23)$$

et suffisamment vite tel que

$$n h_n^{2r+1} \rightarrow 0, \quad (2.24)$$

alors

$$\sqrt{\frac{n h_n^{2m+1}}{b_{m+1}(x)}} \left(\hat{f}_n^{(m)}(x) - f^{(m)}(x) \right) \xrightarrow{\mathcal{L}} N(0, 1), \quad (2.25)$$

avec $b_{m+1}(x)$ donné par (2.18).

La condition (2.24) assure le biais négligeable de l'estimateur $\hat{f}_n^{(m)}(x)$, ce qui permet d'obtenir la formule (2.25).

Corollaire au Théorème 2.4.2. *Un intervalle de confiance asymptotique pour $f^{(m)}(x)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\hat{f}_n^{(m)}(x) \pm z_{1-\epsilon/2} \sqrt{\frac{b_{m+1}(x)}{n h_n^{2m+1}}}, \quad (2.26)$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite. Sa longueur tend vers 0 si h_n tend vers 0 suffisamment lentement tel que la condition de consistance (2.20) est satisfaite :

$$n h_n^{2m+1} \rightarrow +\infty.$$

Nous constatons que les conditions de consistance sont plutôt convenables : la première (2.19) est identique pour tout m tandis que la seconde (2.20) devient plus contraignante lorsque m augmente. Les conditions (2.23) et (2.24) pour la normalité asymptotique sont très différentes : la première est la même pour tout m alors que la seconde, indépendante de m , dépend de r . Puisque r est fixé avant d'effectuer l'estimation, nous aurons intérêt à le choisir assez grand afin de réaliser plus facilement la condition (2.24).

Si $m < r - 1$, alors nous verrons que la condition (2.24) dans le Théorème 2.4.2 sera plus faible que celle à fixer dans le contexte de l'estimation à noyau classique (Chapitre 3). Cela est dû à la réduction de biais, qui est l'avantage de la méthode d'approximation locale polynomiale par rapport à la méthode de convolution.

À partir de maintenant, les variables aléatoires X_1, \dots, X_n , en plus d'être équidistribuées et indépendantes, seront positives afin de respecter les définitions de la courbe de Lorenz et du temps total en test normalisé et leur espérance commune sera notée $\mu_1 > 0$. Nous avons vu dans l'Introduction que ces deux fonctions sont uniquement définies pour une variable aléatoire positive. La moyenne empirique associée à ce n -échantillon est nommée \bar{X}_n .

Nous fixons également x dans \mathbb{R}_*^+ .

2.5 Estimation à noyau des dérivées de la fonction T_F

Nous rappelons tout d'abord la formule de la fonction T_F pour tout $x > 0$:

$$T_F(x) = \frac{\int_0^x (1 - F(s)) ds}{\mu_1}.$$

Soit $r > 0$. L'objectif de cette section est de produire les estimateurs de ses dérivées $T_F^{(m)}$, pour $1 \leq m \leq r$, au point x . Par une investigation rapide, nous percevons une régularité dans l'expression de $T_F^{(m)}(x)$:

$$T_F^{(m)}(x) = \begin{cases} \frac{1-F(x)}{\mu_1} & \text{si } m = 1, \\ -\frac{f^{(m-2)}(x)}{\mu_1} & \text{si } m \geq 2. \end{cases} \quad (2.27)$$

Effectivement, un lien existe entre les dérivées $T_F^{(m)}$ et les dérivées de f si $m \geq 2$. D'où l'hypothèse (H3) qui devient :

Hypothèse (H3'')

f est $(r - 1)$ fois continûment différentiable en tout point x .

Quant à la dérivée première, elle dépend de la fonction de répartition F .

Nous désignons par $\hat{T}_n^{(m)}(x)$ l'estimateur de la dérivée m^e de la fonction T_F au point x . Et nous nous concentrons dans un premier temps sur les dérivées d'ordre supérieur

à 2 de de la fonction T_F .

2.5.1 Dérivées d'ordre supérieur ou égal à 2 à l'aide de l'approximation polynomiale locale

Soient r un entier strictement supérieur à 0 et m un entier compris entre 2 et r . Alors d'après (2.27), il est naturel d'exprimer les estimateurs $\hat{T}_n^{(m)}(x)$ de la manière suivante :

$$\hat{T}_n^{(m)}(x) = \frac{\hat{f}_n^{(m-2)}(x)}{\bar{X}_n} \quad (2.28)$$

c'est-à-dire

$$\hat{T}_n^{(m)}(x) = \frac{1}{nh_n^{m-1}\bar{X}_n} \sum_{i=1}^n \int_{\frac{x_i-x}{h_n}}^1 K_r^{(m-1)}(u) du \cdot \mathbf{1}_{[x-h_n \leq X_i \leq x+h_n]}, \quad (2.29)$$

en utilisant (2.16).

Grâce au travail effectué dans la section précédente, nous déduisons le théorème suivant sur la convergence en probabilité de l'estimateur $\hat{T}_n^{(m)}(x)$ pour $m \geq 2$.

Théorème 2.5.1. *Soient $x > 0$ et m un entier supérieur ou égal à 2. Sous (H3'')-(H4) et si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (2.30)$$

suffisamment lentement tel que

$$nh_n^{2m-3} \rightarrow +\infty, \quad (2.31)$$

alors l'estimateur $\hat{T}_n^{(m)}(x)$ est convergent en probabilité,

$$\hat{T}_n^{(m)}(x) \xrightarrow{\mathbb{P}} T_F^{(m)}(x). \quad (2.32)$$

Nous énonçons enfin deux résultats concernant le comportement asymptotique de l'estimateur $\hat{T}_n^{(m)}(x)$ pour $m \geq 2$.

Théorème 2.5.2. *Soient $x > 0$ et m un entier supérieur ou égal à 2. Sous (H3'')-(H4) et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$nh_n \rightarrow +\infty \quad (2.33)$$

et suffisamment vite tel que

$$nh_n^{2r+1} \rightarrow 0, \quad (2.34)$$

alors

$$\sqrt{\frac{nh_n^{2m-3}\mu_1^2}{b_{m-1}(x)}} \left(\hat{T}_n^{(m)}(x) - T_F^{(m)}(x) \right) \xrightarrow{\mathcal{L}} N(0, 1), \quad (2.35)$$

où

$$b_{m-1}(x) = f(x) \int_{-1}^1 \left(\int_v^1 K_r^{(m-1)}(u) du \right) dv. \quad (2.36)$$

Corollaire au Théorème 2.5.2. *Un intervalle de confiance asymptotique pour $T_F^{(m)}(x)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\hat{T}_n^{(m)}(x) \pm z_{1-\epsilon/2} \sqrt{\frac{b_{m-1}(x)}{n h_n^{2m-3} \mu_1^2}}, \quad (2.37)$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite. Sa longueur tend vers 0 si h_n tend vers 0 suffisamment lentement tel que la condition de consistance (2.31) est satisfaite :

$$n h_n^{2m-3} \rightarrow +\infty.$$

La première condition de consistance (2.30) coïncide avec (2.19). Quant à la seconde, nous nous attendions à obtenir (2.31), qui correspond à (2.20) avec $m - 2$ à la place de m . Cela vient de la relation (2.28) entre les estimateurs des dérivées de la fonction T_F et de celles de la densité f . On constate le même phénomène entre les expressions (2.36) et (2.18). Les conditions pour la normalité asymptotique (2.33) et (2.34) sont analogues à celles de $\hat{f}_n^{(m)}(x)$.

Il nous reste à présent à examiner le cas $m = 1$, c'est le sujet de la prochaine sous-section.

2.5.2 Dérivée première par la méthode à noyau classique

Grâce à (2.27), un choix naturel d'estimateur vient pour la dérivée première T_F' de la fonction T_F au point $x > 0$:

$$\hat{T}_n^{(1)}(x) = \frac{1 - \hat{F}_n(x)}{\bar{X}_n}$$

où \hat{F}_n est un estimateur à noyau de F . Il n'est pas judicieux d'utiliser la fonction de répartition empirique F_n car l'intérêt de notre thèse est de construire des estimateurs lisses. Ainsi, nous proposons :

$$\hat{T}_n^{(1)}(x) = \frac{1 - \frac{1}{n} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}{\bar{X}_n}$$

avec K une fonction de répartition telle que

$$K(u) = \int_{-1}^u K_0(t) dt. \quad (2.38)$$

En admettant uniquement l'hypothèse classique sur la fenêtre (h_n), la convergence presque sûre de l'estimateur ainsi que la normalité asymptotique de $\hat{T}_n^{(1)}(x)$ sont établies.

Théorème 2.5.3. *Soit $x > 0$ fixé. Sous (H5), l'estimateur $\hat{T}_n^{(1)}(x)$ converge presque sûrement vers $T_F'(x)$,*

$$\hat{T}_n^{(1)}(x) \xrightarrow{p.s.} T_F'(x). \quad (2.39)$$

Théorème 2.5.4. *Soit $x > 0$ fixé. Si l'hypothèse (H5) est satisfaite alors*

$$\sqrt{n} \left(\hat{T}_n^{(1)}(x) - T_F'(x) \right) \xrightarrow{\mathcal{L}} N(0, \sigma_{T'}^2), \quad (2.40)$$

avec

$$\sigma_{T'}^2 = \frac{\mu_2(x) \mu_1^2 - 2\mu_1(x) \mu_2(x) \mu_1 + \mu_1^2(x) \mu_2}{\mu_1^4}, \quad (2.41)$$

où $\mu_k(x) = \int_0^x s^k dF(s)$ et $\mu_k = \mu_k(+\infty)$ pour $k = 1, 2$.

2.6 Estimation polynomiale locale des dérivées de la fonction L_F

Soit $r > 0$. Dans cette section, il est question d'analyser le comportement asymptotique des estimateurs, notés $\hat{L}_n^{(m)}(x)$ pour $1 \leq m \leq r$, des dérivées m^e pour $1 \leq m \leq r$ de la fonction L_F au point $x > 0$. Pour cela, nous choisissons de travailler avec la fonction ϕ telle que

$$\phi(x, s) = s \mathbf{1}_{]-\infty, x]}(s) \quad (2.42)$$

ce qui entraîne :

$$\Phi(x, F) = \int s \mathbf{1}_{[s \leq x]} F(s), \quad (2.43)$$

à l'aide de (2.5).

Posons :

$$\hat{G}_n^{(m)}(x) := \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{[X_i \leq x]} \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \quad (2.44)$$

$$= \frac{1}{nh_n^m} \sum_{i=1}^n X_i \int_{\frac{X_i-x}{h_n}}^1 K_r^{(m)}(u) du \cdot \mathbf{1}_{[x-h_n \leq X_i \leq x+h_n]}, \quad (2.45)$$

obtenu en injectant (2.42) dans (2.6) puis en effectuant le changement de variable $u = \frac{z-x}{h_n}$.

Avec le choix (2.43), on constate que (H2) est satisfaite car la fonction de répartition F est supposée continue. De plus, on remarque que

$$L_F(x) = \frac{\Phi(x, F)}{\mu_1}.$$

On en déduit l'estimateur naturel de $L_F^{(m)}(x)$, la m^e -dérivée de la fonction L_F , au point x :

$$\hat{L}_n^{(m)}(x) = \frac{\hat{G}_n^{(m)}(x)}{\bar{X}_n}$$

c'est-à-dire

$$\hat{L}_n^{(m)}(x) = \frac{1}{nh_n^m \bar{X}_n} \sum_{i=1}^n X_i \int_{\frac{X_i-x}{h_n}}^1 K_r^{(m)}(u) du \cdot \mathbf{1}_{[x-h_n \leq X_i \leq x+h_n]}, \quad (2.46)$$

avec (2.45).

Nous donnons ci-dessous la formule des dérivées $L_F^{(m)}$ au point $x > 0$, pour $1 \leq m \leq r$:

$$L_F^{(m)}(x) = \begin{cases} \frac{xf(x)}{\mu_1} & \text{si } m = 1, \\ \frac{(m-1)f^{(m-2)}(x) + xf^{(m-1)}(x)}{\mu_1} & \text{si } m \geq 2. \end{cases} \quad (2.47)$$

Comme l'ordre le plus élevé de la dérivée de f dont dépend la dérivée $L_F^{(m)}$ est $(m-1)$ alors l'hypothèse (H3) doit être remplacée par l'hypothèse (H3'), définie dans la Section 2.4.

En remplaçant (2.42) dans (2.11), nous aboutissons à l'ordre de variance de l'estimateur $\hat{G}_n^{(m)}(x)$ qui est le sujet du lemme suivant.

Lemme 2.6.1. *Soient $x > 0$ et m un entier compris entre 1 et r . Alors sous (H3')-(H4),*

$$\text{Var} \hat{G}_n^{(m)}(x) = \frac{1}{nh_n^{2m-1}} (b_m(x) + O(h_n)) - \frac{(\mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}))^2}{n}, \quad (2.48)$$

avec

$$b_m(x) = x^2 f(x) \int_{-1}^1 \left(\int_v^1 K_r^{(m)}(u) du \right) dv. \quad (2.49)$$

Nous découvrons que les ordres de variance de $\hat{f}_n^{(m)}(x)$ et $\hat{G}_n^{(m)}(x)$ ne sont pas égaux (respectivement $\frac{1}{nh_n^{2m+1}}$ dans le Lemme 2.4.1 et $\frac{1}{nh_n^{2m-1}}$ dans le Lemme 2.6.1). De plus, les expressions (2.49) et (2.18) ont une forme semblable mais sont tout de même différentes. D'où notre motivation de dissocier l'estimation des dérivées de la densité et l'estimation des dérivées m^e de la fonction L_F pour $1 \leq m \leq r$.

Le prochain théorème explicite les conditions de convergence en probabilité de l'estimateur $\hat{L}_n^{(m)}(x)$ pour $m \geq 1$.

Théorème 2.6.1. *Soient $x > 0$ et m un entier compris entre 1 et r . Sous (H3')-(H4) et si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (2.50)$$

suffisamment lentement tel que

$$nh_n^{2m-1} \rightarrow +\infty, \quad (2.51)$$

alors l'estimateur $\hat{L}_n^{(m)}$ est convergent en probabilité,

$$\hat{L}_n^{(m)}(x) \xrightarrow{\mathbb{P}} L_F^{(m)}(x). \quad (2.52)$$

Nous obtenons les résultats suivants sur le comportement asymptotique de l'estimateur $\hat{L}_n^{(m)}(x)$ pour $m \geq 1$.

Théorème 2.6.2. *Soient $x > 0$ et m un entier compris entre 1 et r . Sous les hypothèses $(H3')$ - $(H4)$ et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$n h_n \rightarrow +\infty \tag{2.53}$$

et suffisamment vite tel que

$$n h_n^{2r+1} \rightarrow 0, \tag{2.54}$$

alors

$$\sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \left(\hat{L}_n^{(m)}(x) - L_F^{(m)}(x) \right) \xrightarrow{\mathcal{L}} N(0, 1), \tag{2.55}$$

avec $b_m(x)$ donné par (2.49).

Corollaire au Théorème 2.6.2. *Un intervalle de confiance asymptotique pour $L_F^{(m)}(x)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\hat{L}_n^{(m)}(x) \pm z_{1-\epsilon/2} \sqrt{\frac{b_m(x)}{n h_n^{2m-1} \mu_1^2}}, \tag{2.56}$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite. Sa longueur tend vers 0 si h_n tend vers 0 suffisamment lentement tel que la condition de consistance (2.51) est satisfaite :

$$n h_n^{2m-1} \rightarrow +\infty.$$

Les conditions de consistance sont intuitivement raisonnables : la première (2.50) est équivalente à (2.19) et indépendante de m tandis que la deuxième (2.51) dépend de m et devient plus stricte quand m augmente, comme (2.20), mais son écriture n'est pas la même. Les conditions pour la normalité asymptotique (2.53) et (2.54) concordent avec (2.23) et (2.24). Remarquons que la seconde (2.54) dépend de l'entier r fixé au départ et coïncide également avec (2.34). Peut-on pour autant en déduire qu'elle est identique pour toutes les fonctionnelles linéaires de la fonction de répartition ? Et que dire de la classe des fonctionnelles de la fonction de répartition ? Le Chapitre 4 nous éclairera sur cette dernière interrogation.

Les estimateurs présentés dans ce chapitre dépendent d'une fenêtre (h_n). Pour obtenir la consistance et la normalité asymptotique ainsi que pour construire des intervalles de confiance, des hypothèses sur ce paramètre ont été fixées. Dès lors, nous sommes en mesure de nous demander s'il est possible de trouver une fenêtre (h_n) satisfaisant toutes ces conditions. C'est l'objet de la section suivante.

2.7 Choix de fenêtre

Soient r un entier non nul fixé.

Les hypothèses (H5), (2.23) et (2.24) sont communes pour les estimateurs polynomiaux locaux de toutes les fonctions étudiées dans ce chapitre, sauf de la dérivée première de la fonction T_F , et permettent d'aboutir à leur normalité asymptotique. Il faudra vérifier en plus, l'hypothèse (2.20), pour tout m compris entre 0 et $r - 1$, pour la densité f et toutes ses dérivées, l'hypothèse (2.31), pour tout m compris entre 2 et r , pour les dérivées de la fonction T_F d'ordre supérieur à 2 et l'hypothèse (2.51), pour tout m compris entre 1 et r , pour les dérivées successives de L_F afin d'obtenir les intervalles de confiance (et la consistance).

Nous pouvons constater que (2.20) pour tout m compris entre 0 et $r - 1$, (2.31) pour tout m compris entre 2 et r , et (2.51) pour tout m compris entre 1 et r , impliquent (2.23). Donc cette dernière hypothèse est à ignorer. De plus, (2.20), (2.31) et (2.51) sont de plus en plus restrictives lorsque l'ordre de dérivation m augmente. Il suffit donc de poser :

- $nh_n^{2r-1} \rightarrow +\infty$ pour la densité f et ses dérivées successives, ainsi que pour les dérivées de la fonction L_F ,
- $nh_n^{2r-3} \rightarrow +\infty$ pour les dérivées d'ordre supérieur à 2 de la fonction L_F .

Ainsi, les conditions permettant de construire des estimateurs polynomiaux locaux consistants et asymptotiquement normaux avec intervalles de confiance sont indépendantes de l'ordre de dérivation m .

2.8 Preuves

2.8.1 Preuve du Lemme 2.3.1

Pour justifier que la méthode d'approximation polynomiale locale est intéressante en matière de réduction de biais, nous calculons l'espérance des estimateurs des fonctionnelles linéaires de la fonction de répartition et de leurs dérivées successives. Nous avons par indépendance et équidistribution des variables aléatoires X_1, \dots, X_n :

$$\begin{aligned}
 \mathbb{E} \hat{\Phi}_n^{(m)}(x) &= \mathbb{E} \int \phi(z, X_1) \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \\
 &= \int \left[\int \phi(z, s) dF(s) \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) \right] dz \\
 &= \frac{1}{h_n^m} \int \Phi(z, F) \frac{1}{h_n} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz
 \end{aligned} \tag{2.57}$$

grâce à la linéarité de l'espérance et à la linéarité de $\Phi(x, \cdot)$. Nous appliquons le changement de variable $u = \frac{z-x}{h_n}$, d'où :

$$\begin{aligned} \mathbb{E} \hat{\Phi}_n^{(m)}(x) &= \frac{1}{h_n^m} \int_{-1}^1 \Phi(x + h_n u, F) K_r^{(m)}(u) du \\ &= \frac{1}{h_n^m} \int_{-1}^1 \sum_{i=0}^r \Phi^{(i)}(x, F) \frac{(h_n u)^i}{i!} K_r^{(m)}(u) du \\ &\quad + \frac{1}{h_n^m} \int_{-1}^1 \Phi^{(r+1)}(x + u', F) \frac{(h_n u)^{r+1}}{(r+1)!} K_r^{(m)}(u) du \end{aligned} \tag{2.58}$$

d'après la formule de Taylor-Lagrange pour u' compris entre 0 et $h_n u$, avec $u \in [-1, 1]$. En utilisant (H4(i)) sur la première ligne puis (H3), (H4(i)) sur la seconde ligne, il résulte que :

$$\mathbb{E} \hat{\Phi}_n^{(m)}(x) = \Phi^{(m)}(x, F) + O\left(h_n^{r+1-m}\right). \blacksquare \tag{2.59}$$

2.8.2 Preuve du Lemme 2.4.1

Nous reprenons l'intégrale $\int \phi(x + h_n u, s) K_r^{(m)}(u) du$ dans l'expression (2.11) et la simplifions dans le cas particulier :

$$\phi(x, s) = \mathbf{1}_{[s \leq x]}. \tag{2.60}$$

Elle devient :

$$\int_{-1}^1 \mathbf{1}_{[s \leq x + h_n u]} K_r^{(m)}(u) du = \begin{cases} 0, & s > x + h_n, \\ \int_{\frac{s-x}{h_n}}^1 K_r^{(m)}(u) du, & |s - x| \leq h_n, \\ \int_{-1}^1 K_r^{(m)}(u) du = \delta_{m,0}, & s < x - h_n, \end{cases}$$

où $\delta_{m,0}$ égal à 1 si $m = 0$ et 0 sinon, d'après (H4(i)).

Ainsi, en reprenant (2.11), on a pour $m \geq 1$:

$$\begin{aligned} &\mathbb{E} \left[\int \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n^{m+1}} K_r^{(m)}(z - x) dz \right]^2 \\ &= \frac{1}{h_n^{2m}} \int_{x-h_n}^{x+h_n} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(m)}(u) du \right)^2 f(s) ds \\ &= \frac{1}{h_n^{2m}} \int_{x-h_n}^{x+h_n} \left(\int_{\frac{s-x}{h_n}}^1 K_r^{(m)}(u) du \right)^2 (f(x) + f'(s')(s-x)) ds \\ &= \frac{1}{h_n^{2m}} \int_{-1}^1 \left(\int_v^1 K_r^{(m)}(u) du \right)^2 (f(x) h_n + f'(s') v h_n^2) dv, \end{aligned}$$

en utilisant la formule de Taylor-Lagrange pour s' compris entre x et s , avec $s \in [x - h_n, x + h_n]$, puis le changement de variable $v = \frac{s-x}{h_n}$. On obtient donc :

$$\begin{aligned} & \mathbb{E} \left[\int \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n^{m+1}} K_r^{(m)}(z-x) dz \right]^2 \\ &= \frac{1}{h_n^{2m}} \left(\int_{-1}^1 \left(\int_v^1 K_r^{(m)}(u) du \right)^2 dv f(x) h_n + O(h_n^2) \right) \\ &= \frac{1}{h_n^{2m-1}} (b_m(x) + O(h_n)) , \end{aligned} \quad (2.61)$$

avec

$$b_m(x) = f(x) \int_{-1}^1 \left(\int_v^1 K_r^{(m)}(u) du \right)^2 dv , \quad (2.62)$$

et grâce à (H3'), (H4(ii)-(iii)), $|v| \leq 1$,

$$|O(h_n)| \leq \sup\{|f'(s')| ; s' \in [-h_n, h_n]\} 8C^2 h_n.$$

À partir de maintenant, m désigne un entier compris entre 0 et $r-1$. Puisque

$$\hat{f}_n^{(m)}(x) = \Phi_n^{(m+1)}(x, F) \quad (2.63)$$

pour le choix (2.60), alors grâce à (2.59), nous obtenons

$$\mathbb{E} \hat{f}_n^{(m)}(x) = f^{(m)}(x) + O(h_n^{r-m}) . \quad (2.64)$$

Nous utilisons d'une part (2.63) conjointement à (2.61) puis nous substituons à (2.9). D'autre part, nous injectons (2.64) dans (2.10). Ainsi, grâce à (2.8), nous déduisons, pour tout m compris entre 0 et $r-1$, que :

$$\text{Var} \hat{f}_n^{(m)}(x) = \frac{1}{n h_n^{2m+1}} (b_{m+1}(x) + O(h_n)) - \frac{(f^{(m)}(x) + O(h_n^{r-m}))^2}{n} . \quad \blacksquare \quad (2.65)$$

2.8.3 Preuve du Théorème 2.4.1

Pour $m \geq 0$, l'erreur en moyenne quadratique se décompose en un terme de variance et un terme de biais :

$$\begin{aligned} \mathbb{E} \left[\hat{f}_n^{(m)}(x) - f^{(m)}(x) \right]^2 &= \text{Var} \hat{f}_n^{(m)}(x) + \left(\mathbb{E} \hat{f}_n^{(m)}(x) - f^{(m)}(x) \right)^2 \\ &= \frac{1}{n h_n^{2m+1}} (b_{m+1}(x) + O(h_n)) - \frac{(f^{(m)}(x) + O(h_n^{r-m}))^2}{n} \\ &\quad + O(h_n^{2(r-m)}) , \end{aligned}$$

grâce aux résultats (2.65) et (2.64) de la preuve précédente. Sous les hypothèses (2.20) et (H5), les termes de variance et de biais convergent respectivement vers 0.

La convergence en probabilité de $\hat{f}_n^{(m)}(x)$ vers $f^{(m)}(x)$ est la conséquence de la convergence \mathbb{L}^2 de $\hat{f}_n^{(m)}(x)$ vers $f^{(m)}(x)$. \blacksquare

2.8.4 Preuve du Théorème 2.4.2

Soit $m \geq 0$. Nous pouvons écrire :

$$\sqrt{\frac{n h_n^{2m+1}}{b_{m+1}(x)}} \left(\hat{f}_n^{(m)}(x) - f^{(m)}(x) \right) = \sqrt{\frac{\text{Var} \hat{f}_n^{(m)}(x)}{\frac{b_{m+1}(x)}{n h_n^{2m+1}}}} \frac{\hat{f}_n^{(m)}(x) - f^{(m)}(x)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(x)}}. \quad (2.66)$$

On considère dans un premier temps la différence normalisée entre l'estimateur et la valeur à estimer :

$$\frac{\hat{f}_n^{(m)}(x) - f^{(m)}(x)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(x)}} = Z_n + \delta_n,$$

où

$$Z_n = \frac{\hat{f}_n^{(m)}(x) - \mathbb{E} \hat{f}_n^{(m)}(x)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(x)}}$$

et

$$\delta_n = \frac{\mathbb{E} \hat{f}_n^{(m)}(x) - f^{(m)}(x)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(x)}}.$$

On prouve tout d'abord que $Z_n \xrightarrow{\mathcal{L}} N(0, 1)$. Nous obtenons les résultats suivants pour $m = 0, \dots, r-1$:

$$Z_n = \frac{\frac{1}{n h_n^{m+2}} \sum_{i=1}^n \left(\int \mathbf{1}_{[X_i \leq z]} K_r^{(m+1)} \left(\frac{z-x}{h_n} \right) dz - \mathbb{E} \int \mathbf{1}_{[X_i \leq z]} K_r^{(m+1)} \left(\frac{z-x}{h_n} dz \right) \right)}{\sqrt{\frac{1}{n h_n^{2m+1}} (b_{m+1}(x) + O(h_n)) - \frac{(f^{(m)}(x) + O(h_n^{r-m}))^2}{n}}}$$

est équivalent à la variable aléatoire

$$\tilde{Z}_n = \frac{\sum_{i=1}^n \left(\int_{-1}^1 \mathbf{1}_{[X_i \leq x+h_n u]} K_r^{(m+1)}(u) du - \mathbb{E} \int_{-1}^1 \mathbf{1}_{[X_i \leq x+h_n u]} K_r^{(m+1)}(u) du \right)}{\sqrt{n h_n (b_{m+1}(x) + O(h_n))}}$$

donc Z_n a la même loi asymptotique que \tilde{Z}_n . Nous sommes dans le cadre des théorèmes de la limite centrale pour tableaux triangulaires avec $k_n = n$ et Y_{ni} des copies i.i.d. de

$$Y_n = \frac{\int_{-1}^1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m+1)}(u) du}{\sqrt{n h_n (b_{m+1}(x) + O(h_n))}}.$$

Puisque $\text{Var} Z_n = 1$ pour tout n , alors $\lim_{n \rightarrow \infty} \text{Var} \tilde{Z}_n = 1$ donc la condition F.9(iv) est satisfaite. Afin de vérifier la condition de Lindeberg-Feller F.9(iii), la fenêtre doit être choisie telle que

$$\forall \epsilon > 0, n \mathbb{E} \left[Y_n^2 \mathbf{1}_{|Y_n| > \epsilon} \right] \rightarrow 0.$$

Par le Lemme F.9.2, il suffit d'avoir

$$\forall \epsilon > 0, \mathbb{P}[|Y_n| > \epsilon] = 0.$$

Soit $\epsilon > 0$. Alors :

$$\mathbb{P}[|Y_n| > \epsilon] = 0 \iff \mathbb{P}\left[\left|\frac{\int_{-1}^1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m+1)}(u) du}{\sqrt{nh_n (b_{m+1}(x) + O(h_n))}}\right| > \epsilon\right] = 0,$$

ce qui est équivalent à :

$$\mathbb{P}\left[\left|\int_{-1}^1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m+1)}(u) du\right| > \epsilon \sqrt{nh_n (b_{m+1}(x) + O(h_n))}\right] = 0.$$

Il suffira d'exiger $nh_n \rightarrow +\infty$ car le numérateur de Y_n est borné.

En effet,

$$\begin{aligned} \left|\int_{-1}^1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m+1)}(u) du\right| &\leq \int_{-1}^1 |K_r^{(m+1)}(u)| du \\ &\leq 2 \sup_{u \in [-1,1]} |K_r^{(m+1)}(u)| \\ &< +\infty. \end{aligned}$$

Ensuite, l'ordre du terme de biais δ_n est obtenu en utilisant (2.64) et (2.65) :

$$\delta_n = \frac{O(h_n^{r+1-m})}{\sqrt{\frac{1}{nh_n^{2m+1}} (b_{m+1}(x) + O(h_n)) - \frac{(f^{(m)}(x) + O(h_n^{r-m}))^2}{n}}} = O\left(\sqrt{nh_n^{2r+1}}\right).$$

Donc sous les conditions (2.23) et (2.24), nous avons :

$$\frac{\hat{f}_n^{(m)}(x) - f^{(m)}(x)}{\sqrt{\text{Var } \hat{f}_n^{(m)}(x)}} \xrightarrow{\mathcal{L}} N(0, 1), \quad (2.67)$$

d'après la règle de Cramér F.10-1(i).

Nous remarquons dans un second temps, grâce à (2.65), que :

$$\sqrt{\frac{\text{Var } \hat{f}_n^{(m)}(x)}{\frac{b_{m+1}(x)}{nh_n^{2m+1}}}} \xrightarrow{\mathbb{P}} 1. \quad (2.68)$$

Comme nous avons la décomposition (2.66), la règle de Cramér F.10-1(ii) appliquée à (2.67) et à (2.68) nous permet de conclure que :

$$\sqrt{\frac{nh_n^{2m+1}}{b_{m+1}(x)}} (\hat{f}_n^{(m)}(x) - f^{(m)}(x)) \xrightarrow{\mathcal{L}} N(0, 1). \blacksquare$$

2.8.5 Preuve du Théorème 2.5.1

Soit m un entier compris entre 2 et r . Alors la formule de l'estimateur de la dérivée m^e de T_F au point x est la suivante :

$$\hat{T}_n^{(m)}(x) = \frac{\hat{f}_n^{(m-2)}(x)}{\bar{X}_n}. \quad (2.69)$$

La loi faible des grands nombres F.1 appliquée aux variables aléatoires X_1, \dots, X_n de même espérance μ_1 entraîne la convergence en probabilité de \bar{X}_n vers μ_1 . Avec l'hypothèse (2.97), puis en observant que $\mu_1 \neq 0$, la règle de Cramér F.10-1(iii) implique que

$$\hat{T}_n^{(m)}(x) = \frac{\hat{f}_n^{(m-2)}(x)}{\bar{X}_n} \xrightarrow{\mathbb{P}} \frac{f^{(m-2)}(x)}{\mu_1} = \hat{T}_F^{(m-2)}(x). \blacksquare$$

2.8.6 Preuve du Théorème 2.5.2

Soient $x > 0$ et m un entier compris entre 2 et r . Nous allons prouver la convergence en loi de l'estimateur $\hat{T}_n^{(m)}(x)$. Pour notre aider dans notre quête, constatons grâce à (2.69) que l'expression :

$$\sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \left(\hat{T}_n^{(m)}(x) - T_F^{(m)}(x) \right)$$

est égale à :

$$\sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \left(\frac{\hat{f}_n^{(m-2)}(x)}{\bar{X}_n} - \frac{f^{(m-2)}(x)}{\mu_1} \right),$$

ou encore à :

$$\sqrt{\frac{n h_n^{2m-3}}{b_{m-1}(x)}} \left(\hat{f}_n^{(m-2)}(x) - f^{(m-2)}(x) \right) \quad (2.70)$$

$$+ \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \hat{f}_n^{(m-2)}(x). \quad (2.71)$$

Concentrons-nous dans un premier temps sur le terme (4.79). Le Théorème 2.4.2 expose la normalité asymptotique de $\hat{f}_n^{(m)}(x)$ pour tout entier m compris entre 0 et $r-1$ sous plusieurs conditions : (H3')-(H4) doivent être vérifiées et lorsque $n \rightarrow +\infty$, la fenêtre h_n doit tendre vers 0 suffisamment lentement telle que $nh_n \rightarrow +\infty$ et suffisamment vite telle que $nh_n^{2r+1} \rightarrow 0$. Nous en déduisons que :

$$\sqrt{\frac{n h_n^{2m-3}}{b_{m-1}(x)}} \left(\hat{f}_n^{(m-2)}(x) - f^{(m-2)}(x) \right) \xrightarrow{\mathcal{L}} \text{N}(0, 1),$$

sous les mêmes hypothèses.

Nous nous focalisons dans un second temps sur le terme (4.80). À l'aide de l'écriture (2.29), nous avons :

$$\begin{aligned}
& \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \hat{f}_n^{(m-2)}(x) \\
&= \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu} \right) \sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \frac{1}{n} \sum_{i=1}^n \int \mathbf{1}_{(X_i \leq z)} \frac{1}{h_n^{m-1}} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \\
&= \sqrt{n}(\mu_1 - \bar{X}_n) \cdot \frac{1}{\sqrt{b_{m-1}(x) \bar{X}_n}} \cdot \frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{(X_i \leq z)} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz. \quad (2.72)
\end{aligned}$$

Remarquons tout d'abord que $\sqrt{n}(\mu_1 - \bar{X}_n)$ est borné dans \mathbb{L}^2 donc également en probabilité, d'où :

$$\sqrt{n}(\mu_1 - \bar{X}_n) = O_{\mathbb{P}}(1). \quad (2.73)$$

Ensuite, nous prouvons que $\frac{1}{\sqrt{b_{m-1}(x) \mu_1 \bar{X}_n}}$ converge en probabilité vers $\sqrt{b_{m-1}(x) \mu_1^2}$, en vertu de la loi faible des grands nombres F.1 et des règles de Cramér F.10-1(ii)-(iii). Cela entraîne que :

$$\frac{1}{\sqrt{b_{m-1}(x) \bar{X}_n}} = O_{\mathbb{P}}(1). \quad (2.74)$$

Enfin, nous appliquons l'inégalité de Markov F.11 pour $p = 1$ au terme de droite de (2.72). Pour tout $\epsilon > 0$, nous avons :

$$\begin{aligned}
& \epsilon \mathbb{P} \left[\left| \frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \right| > \epsilon \right] \\
& \leq \mathbb{E} \left[\left| \frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \right| \right] \\
& \leq \mathbb{E} \left[\frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} \left| K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) \right| dz \right] \\
& = \mathbb{E} \left[\frac{1}{\sqrt{h_n}} \int \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n} \left| K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) \right| dz \right].
\end{aligned}$$

L'espérance devient :

$$\frac{1}{\sqrt{h_n}} \int_{x-h_n}^{x+h_n} \left[\int_{\frac{s-x}{h_n}}^1 \left| K_r^{(m)}(u) \right| du \right] f(s) ds$$

par le changement de variable $u = \frac{z-x}{h_n}$, la définition de l'espérance et l'hypothèse (H4(iii)). En utilisant successivement la formule de Taylor-Lagrange, le changement de variable $v = \frac{s-x}{h_n}$ et l'hypothèse (H4(ii)), l'expression précédente est majorée par :

$$\sqrt{h_n} \left(4 f(x) \sup_{u \in [-1,1]} \left| K_r^{(m-1)}(u) \right| + O(h_n) \right),$$

qui converge vers 0 sous (H5). Il vient alors, par le théorème des gendarmes, que pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left[\left| \frac{1}{n\sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \right| > \epsilon \right] = 0,$$

c'est-à-dire

$$\frac{1}{n\sqrt{h_n}} \sum_{i=1}^n \int \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz = o_{\mathbb{P}}(1). \quad (2.75)$$

En remplaçant (2.73), (2.74) et (2.75) dans (2.72), nous déduisons que

$$\left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \hat{f}_n^{(m-2)}(x) = O_{\mathbb{P}}(1) \cdot O_{\mathbb{P}}(1) \cdot o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1). \quad (2.76)$$

La règle de Cramér (F.10)-1(i) appliquée à (2.8.6) et à (2.76) nous permet de conclure que :

$$\sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \left(\hat{T}_n^{(m)}(x) - T_F^{(m)}(x) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad \blacksquare$$

2.8.7 Preuve du Théorème 2.5.3

Rappelons que

$$\hat{T}'_F(x) = \frac{1 - F(x)}{\mu_1} \quad (2.77)$$

et

$$\hat{T}_n^{(1)}(x) = \frac{1 - \hat{F}_n(x)}{\bar{X}_n}. \quad (2.78)$$

La convergence presque sûre de $\hat{T}_n^{(1)}(x)$ vers $T'_F(x)$ découle de la convergence presque sûre obtenue par Yamato ([165]) de l'estimateur à noyau \hat{F}_n de la fonction de répartition F , défini par :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right), \quad (2.79)$$

où K est donné par (2.38).

D'après (2.77) et (2.78), il vient que

$$\begin{aligned} \hat{T}_n^{(1)}(x) - T'_F(x) &= \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) (1 - \hat{F}_n) + \frac{F - \hat{F}_n}{\mu_1} \\ &= \frac{\mu_1 - \bar{X}_n}{\bar{X}_n \mu_1} (1 - \hat{F}_n) + \frac{F - \hat{F}_n}{\mu_1}. \end{aligned} \quad (2.80)$$

Par la loi forte des grands nombres F.6, nous écrivons

$$\bar{X}_n \xrightarrow{p.s.} \mu_1. \quad (2.81)$$

Par ailleurs, Yamato ([165]) a prouvé la convergence presque sûre de \hat{F}_n vers F . Grâce à ce dernier résultat, à (2.80) et aux 3 règles de Cramér F.10 puisque $\bar{X}_n \neq 0$ et $\mu \neq 0$, nous déduisons la convergence presque sûre de $\hat{T}_n^{(1)}(x) - T_F'(x)$ vers 0, ou de manière équivalente

$$\hat{T}_n^{(1)}(x) \xrightarrow{p.s.} T_F'(x). \blacksquare \quad (2.82)$$

2.8.8 Preuve du Théorème 2.5.4

Par définition, on a

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[X_i \leq x]}. \quad (2.83)$$

Alors l'estimateur (2.79) peut être réécrit comme

$$\hat{F}_n(x) = \int K\left(\frac{x-s}{h_n}\right) dF_n(s). \quad (2.84)$$

Lemme 2.8.1. *Soit $x > 0$ fixé. Alors pour tout $s > 0$, nous avons*

$$1 - K\left(\frac{x-s}{h_n}\right) \rightarrow \mathbf{1}_{[s > x]}, \quad (2.85)$$

lorsque $n \rightarrow +\infty$.

Grâce au Lemme 2.8.1, nous pouvons écrire lorsque $n \rightarrow +\infty$,

$$1 - K\left(\frac{x - X_i}{h_n}\right) \rightarrow \mathbf{1}_{[X_i > x]} \text{ pour tout } 1 \leq i \leq n,$$

donc

$$\frac{1}{n} \sum_{i=1}^n \left(1 - K\left(\frac{x - X_i}{h_n}\right)\right) \sim \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[X_i > x]}.$$

Or (2.78) et (2.83) conduisent à

$$\begin{aligned} \hat{T}_n^{(1)}(x) &= \frac{1 - F_n(x)}{\bar{X}_n} \frac{1 - \hat{F}_n(x)}{1 - F_n(x)} \\ &= \frac{1 - F_n(x)}{\bar{X}_n} \frac{\frac{1}{n} \sum_{i=1}^n \left(1 - K\left(\frac{x - X_i}{h_n}\right)\right)}{\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[X_i > x]}} \end{aligned} \quad (2.86)$$

donc le second facteur de (2.86) converge en probabilité vers 1. Par suite, la loi asymptotique de $\hat{T}_n^{(1)}(x)$ et $\frac{1 - F_n(x)}{\bar{X}_n}$ est identique puis la distribution limite de

$$\sqrt{n} \left(\hat{T}_n^{(1)}(x) - T_F'(x) \right)$$

est la même que celle de

$$\sqrt{n} \left(\frac{1 - F_n(x)}{\bar{X}_n} - T_F'(x) \right),$$

où T_F' désigne la dérivée première du temps total de test normalisé T_F .

Soit $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ la fonction telle que pour tout $s > 0$, $\mathbf{g}(s) = (g_4(s), g_2(s))$ avec

$$g_4(s) = \mathbf{1}_{[s>x]} \quad \text{et} \quad g_2(s) = s.$$

Alors pour la variable aléatoire positive X_1 suivant la fonction de répartition F , on a :

$$\begin{aligned} \mathbb{E} g_4(X_1) &= \mathbb{E} [\mathbf{1}_{[X_1>x]}] \\ &= \int_0^{+\infty} \mathbf{1}_{[s>x]} dF(s) \\ &:= 1 - F(x), \end{aligned}$$

$$\begin{aligned} \mathbb{E} g_2(X_1) &= \mathbb{E} X_1 \\ &= \int_0^{+\infty} s dF(s) \\ &:= \mu_1, \end{aligned}$$

$$\begin{aligned} \mathbb{E} [g_4(X_1)]^2 &= \mathbb{E} [\mathbf{1}_{[X_1>x]}^2] \\ &= \mathbb{E} [\mathbf{1}_{[X_1>x]}] \\ &= \int_0^{+\infty} \mathbf{1}_{[s>x]} dF(s) \\ &:= 1 - F(x), \end{aligned}$$

$$\begin{aligned} \mathbb{E} [g_2(X_1)]^2 &= \mathbb{E} X_1^2 \\ &= \int_0^{+\infty} s^2 dF(s) \\ &:= \mu_2 \end{aligned}$$

et

$$\begin{aligned} \mathbb{E} [g_4(X_1) g_2(X_1)] &= \mathbb{E} [X_1 \mathbf{1}_{[X_1>x]}] \\ &= \int_0^{+\infty} s \mathbf{1}_{[s>x]} dF(s) \\ &= \int_x^{+\infty} s dF(s) \\ &:= \rho(x). \end{aligned}$$

D'où

$$\begin{aligned} \text{Var } g_4(X_1) &= \mathbb{E} [g_4(X_1)]^2 - \mathbb{E}^2 g_4(X_1) \\ &= F(x) (1 - F(x)), \end{aligned}$$

$$\begin{aligned} \text{Var } g_2(X_1) &= \mathbb{E} [g_2(X_1)]^2 - \mathbb{E}^2 g_2(X_1) \\ &= \mu_2 - \mu_1^2 \end{aligned}$$

et

$$\begin{aligned} \text{Cov}[g_4(X_1), g_2(X_1)] &= \mathbb{E}[g_4(X_1)g_2(X_1)] - \mathbb{E}g_4(X_1)\mathbb{E}g_2(X_1) \\ &= \rho(x) - (1 - F(x))\mu_1. \end{aligned}$$

D'après (1.6), nous avons

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} g_4(s) dF_n(s) \\ \int_0^{+\infty} g_2(s) dF_n(s) \end{pmatrix} - \begin{pmatrix} \mathbb{E}g_4(X_1) \\ \mathbb{E}g_2(X_1) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \text{N}_2(0, \Sigma_{\mathbf{g}})$$

c'est-à-dire

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} \mathbf{1}_{(s>x)} dF_n(s) \\ \int_0^{+\infty} s dF_n(s) \end{pmatrix} - \begin{pmatrix} 1 - F(x) \\ \mu_1 \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \text{N}_2(0, \Sigma_{\mathbf{g}})$$

avec

$$\Sigma_{\mathbf{g}} = \begin{pmatrix} F(x)(1 - F(x)) & \rho(x) - (1 - F(x))\mu_1 \\ \rho(x) - (1 - F(x))\mu_1 & \mu_2 - \mu_1^2 \end{pmatrix}.$$

Remarque :

On applique le théorème de Lindeberg-Feller F.9 aux vecteurs aléatoires i.i.d.

$$Y_{n,i} = \begin{pmatrix} \frac{1}{\sqrt{n}} \mathbf{1}_{[X_i > x]} \\ \frac{1}{\sqrt{n}} X_i \end{pmatrix}.$$

En choisissant la fonction \mathbf{h} indiquée dans la Preuve du Théorème 1.4.3, nous obtenons à l'aide de (1.7),

$$\sqrt{n} \left(\begin{pmatrix} \int_0^{+\infty} \mathbf{1}_{(s>x)} dF_n(s) \\ \int_0^{+\infty} s dF_n(s) \end{pmatrix} - \begin{pmatrix} 1 - F(x) \\ \mu_1 \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \text{N}(0, \sigma_{T'}^2),$$

où

$$\sigma_{T'}^2 = D\mathbf{h}(\mathbb{E}\mathbf{g})\Sigma_{\mathbf{g}}D\mathbf{h}^T(\mathbb{E}\mathbf{g}), \quad (2.87)$$

avec $D\mathbf{h}$ défini en (1.56). Il résulte que

$$\begin{aligned} \sigma_{T'}^2 &= \left(\frac{1}{\mu_1} \right)^2 F(x)(1 - F(x)) + 2 \frac{1}{\mu_1} \left(-\frac{1 - F(x)}{\mu_1^2} \right) (\rho(x) - (1 - F(x))\mu_1) \\ &\quad + \left(-\frac{1 - F(x)}{\mu_1^2} \right)^2 (\mu_2 - \mu_1^2) \\ &= \frac{(1 - F(x))\mu_1^2 - 2(1 - F(x))\rho(x)\mu_1 + (1 - F(x))^2\mu_2}{\mu_1^4}. \end{aligned}$$

On en déduit que :

$$\sqrt{n} \left(\hat{T}_n^{(1)}(x) - T_F'(x) \right) \xrightarrow{\mathcal{L}} \text{N}(0, \sigma_{T'}^2),$$

avec $\sigma_{T'}^2$, exprimé ci-dessus.

Pour terminer la preuve, il reste à démontrer le Lemme 2.8.1 et pour cela, introduisons la notation suivante

$$\bar{K} \left(\frac{x-s}{h_n} \right) := 1 - K \left(\frac{x-s}{h_n} \right), \quad (2.88)$$

avec K donné par (2.38). Comme K_0 est une densité de probabilité sur le support $[-1, 1]$ alors $\int_{-1}^1 K_0(u) du = 1$, d'où

$$\bar{K}\left(\frac{x-s}{h_n}\right) = \int_{\frac{x-s}{h_n}}^1 K_0(u) du. \quad (2.89)$$

1er cas : si $s \leq x$
 Tout d'abord,

$$s \leq x \implies \mathbf{1}_{[s > x]} = 0.$$

Par ailleurs,

$$\begin{aligned} s \leq x &\implies x - s \geq 0 \\ &\implies \frac{x-s}{h_n} \rightarrow +\infty \text{ grâce à (H5)} \\ &\implies \bar{K}\left(\frac{x-s}{h_n}\right) \rightarrow 0. \end{aligned}$$

Ainsi,

$$s \leq x \implies \bar{K}\left(\frac{x-s}{h_n}\right) \rightarrow \mathbf{1}_{(s > x)} \text{ lorsque } n \rightarrow +\infty.$$

2nd cas : si $s > x$
 Tout d'abord,

$$s > x \implies \mathbf{1}_{[s > x]} = 1.$$

Par ailleurs,

$$\begin{aligned} s > x &\implies x - s < 0 \\ &\implies \frac{x-s}{h_n} \rightarrow -\infty \text{ grâce à (H5)} \\ &\implies \bar{K}\left(\frac{x-s}{h_n}\right) \rightarrow 1. \end{aligned}$$

Ainsi, grâce à (2.88),

$$s > x \implies 1 - K\left(\frac{x-s}{h_n}\right) \rightarrow \mathbf{1}_{[s > x]} \text{ lorsque } n \rightarrow +\infty. \blacksquare$$

2.8.9 Preuve du Lemme 2.6.1

Nous reprenons l'intégrale $\int \phi(x + h_n u, s) K_r^{(m)}(u) du$ dans l'expression (2.11) et avec le choix de la fonction ϕ telle que :

$$\phi(x, s) = s \mathbf{1}_{[s \leq x]}, \quad (2.90)$$

nous obtenons

$$\int_{-1}^1 s \mathbf{1}_{[s \leq x + h_n u]} K_r^{(m)}(u) du = \begin{cases} 0, & s > x + h_n, \\ \int_{\frac{s-x}{h_n}}^1 s K_r^{(m)}(u) du, & |s-x| \leq h_n, \\ \int_{-1}^1 s K_r^{(m)}(u) du = \eta_{m,0}, & s < x - h_n, \end{cases}$$

où $\eta_{m,0}$ égal à s si $m = 0$ et 0 sinon, d'après (H4(i)).

Ainsi, pour $m \geq 1$, (2.11) devient

$$\begin{aligned}
& \mathbb{E} \left[\int X_1 \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n^{m+1}} K_r^{(m)}(z-x) dz \right]^2 \\
&= \frac{1}{h_n^{2m}} \int_{x-h_n}^{x+h_n} \left[\int_{\frac{s-x}{h_n}}^1 s K_r^{(m)}(u) du \right]^2 f(s) ds \\
&= \frac{1}{h_n^{2m}} \int_{x-h_n}^{x+h_n} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(m)}(u) du \right]^2 s^2 f(s) ds \\
&= \frac{1}{h_n^{2m}} \int_{x-h_n}^{x+h_n} \left[\int_{\frac{s-x}{h_n}}^1 K_r^{(m)}(u) du \right]^2 \left(x^2 f(x) + (2s' f(s') + s'^2 f'(s')) (s-x) \right) ds \\
&= \frac{1}{h_n^{2m}} \int_{-1}^1 \left[\int_v^1 K_r^{(m)}(u) du \right]^2 \left(x^2 f(x) h_n + (2s' f(s') + s'^2 f'(s')) v h_n^2 \right) dv,
\end{aligned}$$

les deux dernières lignes résultant respectivement de la formule de Taylor-Lagrange pour s' compris entre x et s , avec $s \in [x-h_n, x+h_n]$, puis du changement de variable $v = \frac{s-x}{h_n}$. Puisque les hypothèses (H3') et (H4(ii)-(iii)) sont admises et $|v| \leq 1$, il vient que :

$$\begin{aligned}
& \mathbb{E} \left[\int X_1 \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n^{m+1}} K_r^{(m)}(z-x) dz \right]^2 \\
&= \frac{1}{h_n^{2m}} \left(\int_{-1}^1 \left[\int_v^1 K_r^{(m)}(u) du \right]^2 dv x^2 f(x) h_n + O(h_n^2) \right) \\
&= \frac{1}{h_n^{2m-1}} (b_m(x) + O(h_n))
\end{aligned} \tag{2.91}$$

avec

$$b_m(x) = x^2 f(x) \int_{-1}^1 \left(\int_v^1 K_r^{(m)}(u) du \right)^2 dv. \tag{2.92}$$

De (2.59), nous déduisons

$$\mathbb{E} \hat{G}_n^{(m)}(x) = \mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}). \tag{2.93}$$

En substituant (2.91) à (2.9), en injectant (2.93) dans (2.10) et en utilisant (2.8), nous avons :

$$\text{Var} \hat{G}_n^{(m)}(x) = \frac{1}{n h_n^{2m-1}} (b_m(x) + O(h_n)) - \frac{(\mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}))^2}{n}. \blacksquare \tag{2.94}$$

2.8.10 Preuve du Théorème 2.6.1

Soient $r > 0$ et m un entier compris entre 1 et r . Alors l'estimateur $\hat{L}_n^{(m)}(x)$ de $L_F^{(m)}(x)$, la dérivée m^e de $L_F(x)$ au point $x > 0$, converge en probabilité. Pour prouver un tel résultat, la méthode consiste à passer par l'objet intermédiaire $\hat{G}_n^{(m)}(x)$ qui possède de bonnes propriétés en matière de convergence. Nous énonçons le lemme suivant.

Lemme 2.8.2. Soient x un réel fixé et m un entier compris entre 1 et r . Sous (H3'')-(H4) et si lorsque $n \rightarrow +\infty$,

$$h_n \rightarrow 0$$

suffisamment lentement tel que

$$n h_n^{2m-1} \rightarrow +\infty, \quad (2.95)$$

alors l'estimateur $\hat{G}_n^{(m)}(x)$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x) \right]^2 \rightarrow 0, \quad (2.96)$$

donc également convergent en probabilité,

$$\hat{G}_n^{(m)}(x) \xrightarrow{\mathbb{P}} \mu_1 L_F^{(m)}(x). \quad (2.97)$$

Grâce au Lemme 2.8.2, l'objet $\hat{G}_n^{(m)}(x)$ converge en probabilité vers $\mu_1 L_F^{(m)}(x)$. Ensuite on a la convergence de \bar{X}_n vers μ_1 en probabilité, comme cela est expliqué dans la preuve du Théorème 2.5.1. Puisque

$$\hat{L}_n^{(m)}(x) = \frac{\hat{G}_n^{(m)}(x)}{\bar{X}_n}, \quad (2.98)$$

il résulte que l'estimateur $\hat{L}_n^{(m)}(x)$ converge en probabilité vers $L_F^{(m)}(x)$, en vertu de la règle de Cramér F.10-1(iii).

Pour achever notre preuve, il reste à démontrer le Lemme 2.8.2. Pour $m \geq 1$, l'erreur en moyenne quadratique est égale à :

$$\begin{aligned} \mathbb{E} \left[\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x) \right]^2 &= \text{Var} \hat{G}_n^{(m)}(x) + \left(\mathbb{E} \hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x) \right)^2 \\ &= \frac{1}{n h_n^{2m-1}} (b_m(x) + O(h_n)) - \frac{\left(\mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}) \right)^2}{n} \\ &\quad + O\left(h_n^{2(r+1-m)}\right), \end{aligned}$$

d'après les formules (2.93) et (2.94). Les termes de variance et de biais convergent vers 0 respectivement sous (2.95) et (H5), qui sont supposées dans le Lemme 2.8.2. Enfin, la convergence \mathbb{L}^2 entraîne la convergence en probabilité. ■

2.8.11 Preuve du Théorème 2.6.2

Soit m un entier supérieur ou égal à 1. Nous prouvons la convergence en loi de $\hat{L}_n^{(m)}(x)$ en deux grandes étapes.

La première étape consiste à obtenir la normalité asymptotique de $\hat{G}_n^{(m)}(x)$. Afin d'atteindre cet objectif, nous adoptons une procédure similaire à celle employée dans la preuve du Théorème 2.4.2. Commençons par noter que :

$$\sqrt{\frac{n h_n^{2m-1}}{b_m(x)}} \left(\hat{G}_n^{(m)}(x) - G^{(m)}(x) \right) = \sqrt{\frac{\text{Var} \hat{G}_n^{(m)}(x)}{\frac{b_m(x)}{n h_n^{2m-1}}}} \frac{\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x)}{\sqrt{\text{Var} \hat{G}_n^{(m)}(x)}}. \quad (2.99)$$

Nous considérons tout d'abord la différence normalisée entre l'estimateur et la valeur à estimer,

$$\frac{\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x)}{\sqrt{\text{Var } \hat{G}_n^{(m)}(x)}} = Z_n + \delta_n,$$

où

$$Z_n = \frac{\hat{G}_n^{(m)}(x) - \mathbb{E} \hat{G}_n^{(m)}(x)}{\sqrt{\text{Var } \hat{G}_n^{(m)}(x)}}$$

et

$$\delta_n = \frac{\mathbb{E} \hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x)}{\sqrt{\text{Var } \hat{G}_n^{(m)}(x)}}.$$

Alors :

$$Z_n = \frac{\frac{1}{nh_n^{m+1}} \sum_{i=1}^n \left(\int X_i \mathbf{1}_{[X_i \leq z]} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz - \mathbb{E} \int X_i \mathbf{1}_{[X_i \leq z]} K_r^{(m)} \left(\frac{z-x}{h_n} dz \right) \right)}{\sqrt{\frac{1}{nh_n^{2m-1}} (b_m(x) + O(h_n)) - \frac{(\mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}))^2}{n}}}$$

a la même loi asymptotique que

$$\tilde{Z}_n = \frac{\sum_{i=1}^n \left(\int_{-1}^1 X_i \mathbf{1}_{[X_i \leq x+h_n u]} K_r^{(m)}(u) du - \mathbb{E} \int_{-1}^1 X_i \mathbf{1}_{[X_i \leq x+h_n u]} K_r^{(m)}(u) du \right)}{\sqrt{nh_n (b_m(x) + O(h_n))}}$$

Nous envisageons donc d'appliquer un théorème de la limite centrale pour tableaux triangulaires avec $k_n = n$ et Y_{ni} des copies i.i.d. de

$$Y_n = \frac{\int_{-1}^1 X_1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m)}(u) du}{\sqrt{nh_n (b_m(x) + O(h_n))}}.$$

La condition F.9(iv) est vérifiée : $\lim_{n \rightarrow +\infty} \text{Var } \tilde{Z}_n = 1$ dû au fait que $\text{Var } Z_n = 1$ pour tout n . En choisissant la fenêtre telle que

$$n \mathbb{E} \left[Y_n^2 \mathbf{1}_{[|Y_n| > \epsilon]} \right] \rightarrow 0.$$

pour tout $\epsilon > 0$, la condition de Lindeberg-Feller F.9(iii) est satisfaite. Pour cela, il suffit d'avoir

$$\forall \epsilon > 0, \mathbb{P}[|Y_n| > \epsilon] = 0.$$

d'après le Lemme F.9.2. Pour $\epsilon > 0$:

$$\mathbb{P}[|Y_n| > \epsilon] = 0 \iff \mathbb{P} \left[\left| \frac{\int_{-1}^1 X_1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m)}(u) du}{\sqrt{nh_n (b_m(x) + O(h_n))}} \right| > \epsilon \right] = 0$$

ce qui équivaut à :

$$\mathbb{P} \left[\left| \int_{-1}^1 X_1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m)}(u) du \right| > \epsilon \sqrt{nh_n (b_m(x) + O(h_n))} \right] = 0.$$

Il suffira d'exiger $nh_n \rightarrow +\infty$ car le numérateur de Y_n est borné.

Nous avons en effet

$$\begin{aligned} \left| \int_{-1}^1 X_1 \mathbf{1}_{[X_1 \leq x+h_n u]} K_r^{(m)}(u) du \right| &\leq \int_{-1}^1 |x + h_n u| |K_r^{(m)}(u)| du \\ &\leq 2(x + h_n) \sup_{u \in [-1,1]} |K_r^{(m)}(u)| \\ &< +\infty. \end{aligned}$$

Le théorème de Lindeberg-Feller (F.9) implique donc que $\tilde{Z}_n \xrightarrow{\mathcal{L}} N(0, 1)$, d'où

$$Z_n \xrightarrow{\mathcal{L}} N(0, 1). \quad (2.100)$$

Afin de déterminer l'ordre du terme de biais δ_n , nous utilisons (2.93) et (2.94). Nous trouvons :

$$\delta_n = \frac{O(h_n^{r+1-m})}{\sqrt{\frac{1}{nh_n^{2m-1}} (b_m(x) + O(h_n)) - \frac{(\mu_1 L_F^{(m)}(x) + O(h_n^{r+1-m}))^2}{n}}} = O\left(\sqrt{nh_n^{2r+1}}\right). \quad (2.101)$$

Il résulte des conditions (2.53) et (2.54) que :

$$\frac{\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x)}{\sqrt{\text{Var } \hat{G}_n^{(m)}(x)}} \xrightarrow{\mathcal{L}} N(0, 1), \quad (2.102)$$

d'après la règle de Cramér F.10-1(i) appliquée à (2.100) et à (2.101).

Ensuite, nous déduisons de (2.94) que :

$$\sqrt{\frac{\text{Var } \hat{G}_n^{(m)}(x)}{\frac{b_m(x)}{nh_n^{2m-1}}}} \xrightarrow{\mathbb{P}} 1. \quad (2.103)$$

Enfin, puisque nous avons la décomposition (2.99), la règle de Cramér F.10-1(ii) appliquée à (2.102) et à (2.103) nous permet de conclure que :

$$\sqrt{\frac{nh_n^{2m-1}}{b_m(x)}} \left(\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x) \right) \xrightarrow{\mathcal{L}} N(0, 1). \quad (2.104)$$

Dans la deuxième et dernière étape, nous utilisons le résultat obtenu à l'étape précédente et en déduisons la normalité asymptotique de $\hat{L}_n^{(m)}(x)$. Nous nous inspirons de la preuve du Théorème 2.5.2. Le lien (2.98) entre $\hat{G}_n^{(m)}(x)$ et $\hat{L}_n^{(m)}(x)$ permet d'écrire que :

$$\sqrt{\frac{nh_n^{2m-1}\mu_1^2}{b_m(x)}} \left(\hat{L}_n^{(m)}(x) - L_F^{(m)}(x) \right)$$

est égal à :

$$\sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \left(\frac{\hat{G}_n^{(m)}(x)}{\bar{X}_n} - \frac{\mu_1 L_F^{(m)}(x)}{\mu_1} \right),$$

ou encore à :

$$\sqrt{\frac{n h_n^{2m-1}}{b_m(x)}} \left(\hat{G}_n^{(m)}(x) - \mu_1 L_F^{(m)}(x) \right) \quad (2.105)$$

$$+ \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \hat{G}_n^{(m)}(x). \quad (2.106)$$

Constatons tout d'abord que le terme (2.105) converge vers la loi normale centrée réduite grâce à (2.104).

Ensuite, intéressons-nous au terme (2.106). Comme dans la preuve du Théorème 2.5.2, il s'agit de le contrôler par le théorème des gendarmes pour montrer qu'il converge en probabilité vers 0. À l'aide de (2.44), nous avons :

$$\begin{aligned} & \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \hat{G}_n^{(m)}(x) \\ &= \left(\frac{1}{\bar{X}_n} - \frac{1}{\mu} \right) \sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \frac{1}{n} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n^{m+1}} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz \\ &= \sqrt{n} (\mu_1 - \bar{X}_n) \cdot \frac{1}{\sqrt{b_m(x) \bar{X}_n}} \cdot \frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz. \end{aligned} \quad (2.107)$$

Par l'inégalité de Markov F.11 pour $p = 1$, nous écrivons que pour tout $\epsilon > 0$, on a :

$$\begin{aligned} & \epsilon \mathbb{P} \left[\left| \frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \right| > \epsilon \right] \\ & \leq \mathbb{E} \left[\frac{1}{n \sqrt{h_n}} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} \left| K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) \right| dz \right] \\ & = \mathbb{E} \left[\frac{1}{\sqrt{h_n}} \int X_1 \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n} \left| K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) \right| dz \right]. \end{aligned}$$

Par le changement de variable $u = \frac{z-x}{h_n}$, la définition de l'espérance et l'hypothèse (H4(iii)) puis en utilisant successivement la formule de Taylor-Lagrange, le changement de variable $v = \frac{s-x}{h_n}$ et l'hypothèse (H4(ii)), il vient que :

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{\sqrt{h_n}} \int X_1 \mathbf{1}_{[X_1 \leq z]} \frac{1}{h_n} \left| K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) \right| dz \right] \\ & \leq \sqrt{h_n} \left(4 x f(x) \sup_{u \in [-1,1]} \left| K_r^{(m-1)}(u) \right| + O(h_n) \right). \end{aligned}$$

Le majorant converge vers 0 sous (H5), d'où pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left[\left| \frac{1}{n\sqrt{h_n}} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz \right| > \epsilon \right] = 0,$$

c'est-à-dire

$$\frac{1}{n\sqrt{h_n}} \sum_{i=1}^n \int X_i \mathbf{1}_{[X_i \leq z]} \frac{1}{h_n} K_r^{(m-1)} \left(\frac{z-x}{h_n} \right) dz = o_{\mathbb{P}}(1). \quad (2.108)$$

En injectant (2.73), (2.74) et (2.108) dans (2.107), il résulte que :

$$\left(\frac{1}{\bar{X}_n} - \frac{1}{\mu_1} \right) \sqrt{\frac{n h_n^{2m-3} \mu_1^2}{b_{m-1}(x)}} \hat{f}_n^{(m-2)}(x) = O_{\mathbb{P}}(1) \cdot O_{\mathbb{P}}(1) \cdot o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1). \quad (2.109)$$

La normalité asymptotique de $\hat{T}_n^{(m)}(x)$ provient de la règle de Cramér F.10-1(i) appliquée à (2.104) et à (2.109) :

$$\sqrt{\frac{n h_n^{2m-1} \mu_1^2}{b_m(x)}} \left(\hat{L}_n^{(m)}(x) - L_F^{(m)}(x) \right) \xrightarrow{\mathcal{L}} N(0, 1). \quad \blacksquare$$

2.9 Simulations

2.9.1 Cadre

On suppose que les n données X_1, \dots, X_n d'un échantillon sont positives, indépendantes et identiquement distribuées. La fonction de répartition commune est notée F .

Les dérivées des fonctions L_F et T_F sont notées respectivement $L_F^{(m)}$ et $T_F^{(m)}$ pour un entier m compris entre 0 et r , $r > 0$.

Soit m un entier compris entre 0 et r , $r > 0$.

Pour $m = 0$, nous renvoyons le lecteur au Chapitre 1 qui se concentre sur l'estimation polynomiale locale des fonctions L_F et T_F .

Les estimateurs lisses des dérivées m^e des fonctions L_F pour $m \geq 1$ et T_F pour $m \geq 2$ sont obtenus par la méthode d'approximation polynomiale locale et s'écrivent :

$$\hat{L}_n^{(m)}(x) = \frac{1}{h_n^m} \int_0^{+\infty} L_{F_n}(z) \frac{1}{h_n} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz, \quad m \geq 1, \quad x > 0, \quad (2.110)$$

et

$$\hat{T}_n^{(m)}(x) = \frac{1}{h_n^m} \int_0^{+\infty} T_{F_n}(z) \frac{1}{h_n} K_r^{(m)} \left(\frac{z-x}{h_n} \right) dz, \quad m \geq 2, \quad x > 0. \quad (2.111)$$

Les deux estimateurs à noyau (2.110) et (2.111) s'expriment en fonction d'un noyau $K_r^{(m)}$ et d'un paramètre de lissage (ou d'une fenêtre) h_n . Désignons par K_0 la densité

d'Epanechnikov, c'est-à-dire une fonction positive, d'intégrale égale à 1, donnée par

$$K_0(x) = \frac{3}{4} (1 - x^2) \mathbf{1}_{[-1;1]}.$$

Alors le noyau $K_r^{(m)}$ est donné par (B.1) et s'écrit à l'aide de la densité K_0 . La densité K_0 étant symétrique, nous fixerons $r = m + 1$ (les détails de ce choix se trouvent dans l'Annexe B). Le paramètre de lissage est considéré fixe.

Quant à la dérivée première T'_F de la fonction F , elle sera estimée par

$$\hat{T}_n^{(1)}(x) = \frac{1 - \frac{1}{n} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}{\bar{X}_n} \quad (2.112)$$

avec K une fonction de répartition telle que

$$K(u) = \int_{-1}^u K_0(t) dt$$

où K_0 représentera la densité d'Epanechnikov, \bar{X}_n la moyenne empirique du n -échantillon X_1, \dots, X_n et h_n un paramètre de lissage que l'on fixera.

2.9.2 Données simulées

Pour les simulations suivantes, la taille d'échantillon est fixée à 100 sauf pour l'estimation de la dérivée seconde de la fonction L_F (1000). De plus, on choisit respectivement $h = 0.7$, $h = 0.8$, $h = 0.1$ et $h = 0.8$ pour les estimateurs de la dérivée première de la fonction L_F , de la dérivée seconde de la fonction L_F , de la dérivée première de la fonction T_F et de la dérivée seconde de la fonction T_F .

Nous traçons dans un premier temps, la fonction qui à tout $x > 0$ associe la dérivée première $L_F^{(1)}(x)$ et son estimateur à noyau $\hat{L}_n^{(1)}(x)$, sur un même graphique.

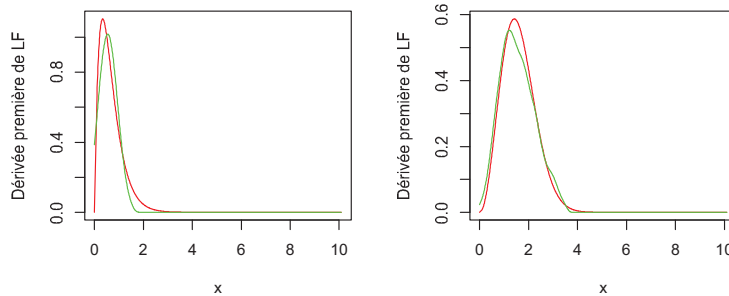


FIGURE 2.1 – Dérivée première de la courbe de la fonction L_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Maxwell (à droite).

Dans un second temps, nous représentons sur le même graphique la fonction qui à tout $x > 0$ associe la dérivée seconde $L_F^{(2)}(x)$ et son estimateur à noyau $\hat{L}_n^{(2)}(x)$.

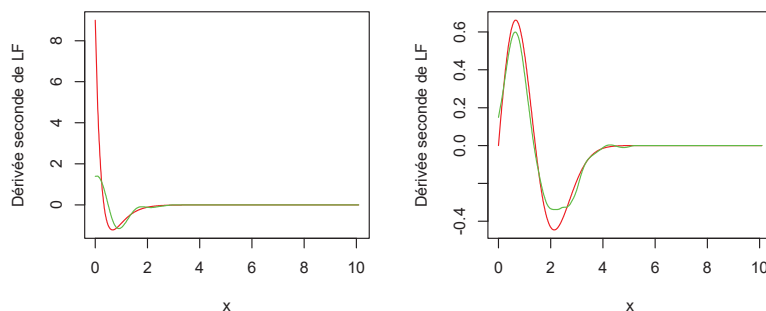


FIGURE 2.2 – Dérivée seconde de la courbe de la fonction L_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Maxwell (à droite).

On constate que les estimateurs $\hat{L}_n^{(1)}(x)$ et $\hat{L}_n^{(2)}(x)$ fournissent une estimation lisse des dérivées première et seconde de la fonction L_F .

Le tableau suivant expose les erreurs d'estimation des estimateurs polynomiaux locaux des deux premières dérivées de la fonction L_F . La colonne nommée "moyenne" indique l'erreur en moyenne quadratique MISE (Mean Integrated Square Error) de ces estimateurs. Les résultats sont le résumé statistique de 500 échantillons.

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.125973	0.7844	0.4017	0.165
100	0.044267	0.3095	0.13119	0.06287
500	0.009604	0.06383	0.02836	0.01096

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.086375	0.7155	0.2915	0.128
100	0.025525	0.20621	0.08676	0.03194
500	0.023576	0.18762	0.08545	0.03589

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.649792	5.692	3.641	2.549
100	0.306712	4.578	3.503	2.804
500	0.176043	3.922	3.455	2.968

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.170307	0.80813	0.25055	0.01247
100	0.062373	0.42687	0.09079	0.01124
500	0.012658	0.076989	0.023961	0.006816

FIGURE 2.3 – Erreur quadratique intégrée de l'estimateur à noyau de la dérivée première (en haut) et de la dérivée seconde (en bas) de la fonction L_F pour la loi Exponentielle de paramètre 3 (à gauche) puis pour la loi de Maxwell (à droite), en fonction de 3 tailles d'échantillon.

Nous remarquons que les MISE des estimateurs à noyau des deux dérivées de la fonction L_F , pour les lois Exponentielle de paramètre 3 et de Maxwell, diminuent lorsque la taille de l'échantillon augmente. La MISE de l'estimateur à noyau de la dérivée seconde pour la loi Exponentielle de paramètre 3 est assez élevée. Le remède employé pour réduire la MISE est d'augmenter la taille de l'échantillon, d'où notre choix initial de prendre 1000 données.

Il s'agit désormais d'illustrer le comportement asymptotique (les théorèmes obtenus) des estimateurs polynomiaux locaux des dérivées première et seconde de la fonction T_F .

Les deux prochaines figures permettent de visualiser la continuité et le bon ajustement des estimateurs construits.

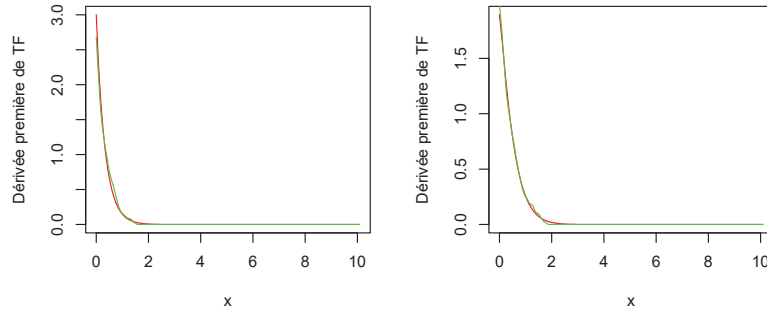


FIGURE 2.4 – Dérivée première de la fonction T_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Weibull de paramètres 2 et 1.2 (à droite).

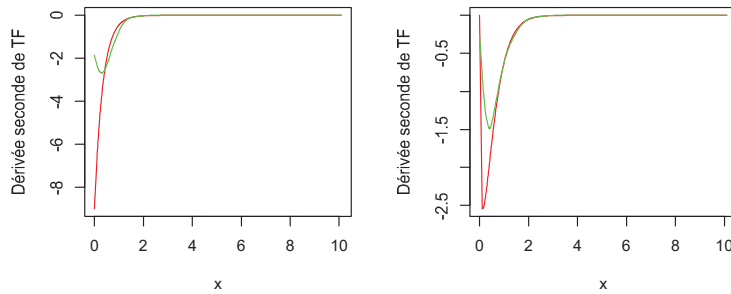


FIGURE 2.5 – Dérivée seconde de la fonction T_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Weibull de paramètres 2 et 1.2 (à droite).

Le tableau suivant décrit les erreurs d'estimation des estimateurs des dérivées première et seconde des fonctions L_F et T_F . La colonne "moyenne" fournit les erreurs en moyenne quadratique MISE (Mean Integrated Square Error) de ces estimateurs. En observant le résumé statistique de 500 échantillons, il vient que les MISE des estimateurs diminuent lorsque la taille de l'échantillon augmente, quelle que soit la loi.

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.042585	0.298997	0.043478	0.002654
100	0.012465	0.078125	0.013406	0.000901
500	0.0027	0.015049	0.002662	0.000347

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.017787	0.10022	0.022439	0.002645
100	0.006807	0.049259	0.006841	0.000799
500	0.001424	0.008294	0.001457	7.32E-05

Taille de l'échantillon	écart-type	max	moyenne	min
30	1.495723	8.418	4.046	1.665
100	0.91377	5.861	3.771	2.068
500	0.387331	4.977	3.731	2.627

Taille de l'échantillon	écart-type	max	moyenne	min
30	0.460801	2.22861	0.69269	0.08398
100	0.245462	1.2704	0.6213	0.1873
500	0.103569	0.8543	0.594	0.3727

FIGURE 2.6 – Erreur quadratique intégrée de l'estimateur à noyau de la dérivée première (en haut) et la dérivée seconde (en bas) de la fonction T_F pour la loi Exponentielle de paramètre 3 (à gauche) puis pour la loi de Weibull de paramètres 2 et 1.2 (à droite), en fonction de 3 tailles d'échantillon.

Finalement, les estimateurs construits des dérivées première et seconde des fonctions L_F et T_F sont performants de part leur continuité et d'autre part, grâce à la valeur basse de leur MISE (proche de 0 et diminuant avec la taille de l'échantillon).

2.9.3 Données réelles

Ampoules chez Norris Electronics

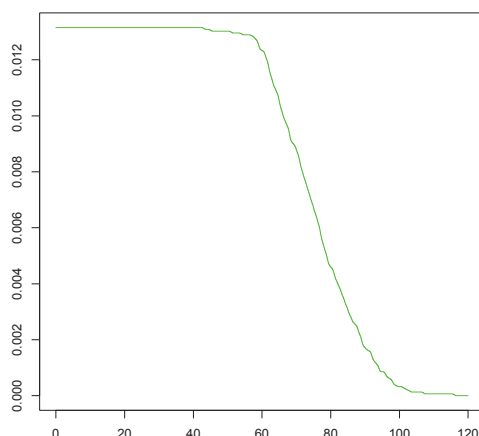


FIGURE 2.7 – Estimateur à noyau de la dérivée première de la fonction T_F , avec $h = 0.8$.

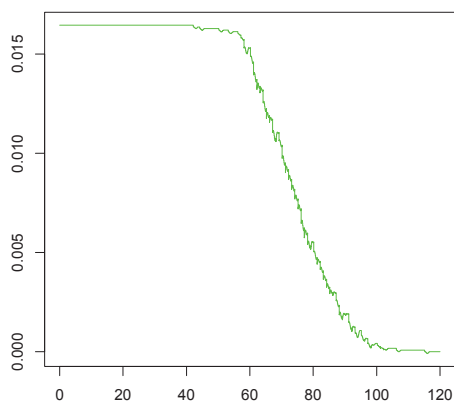


FIGURE 2.8 – Estimateur à noyau de la dérivée seconde de la fonction T_F , avec $h = 0.8$.

La Figure 2.7 montre que la dérivée première de la fonction T_F est estimée par une fonction positive. La courbe T_F semble donc croissante. Par ailleurs, la Figure 2.8 propose une estimation de la dérivée seconde de la fonction T_F . L'estimateur est une fonction positive donc la fonction T_F serait convexe.

Chapitre 3

Estimation à noyau de la densité et de ses dérivées successives et correction des effets de bord

3.1 Introduction

L'estimateur à noyau de la densité a été introduit et étudié par Rosenblatt ([140]), Parzen ([130]) et Watson et Leadbetter ([158]). Plus tard, Schuster ([142]) a prouvé la convergence uniforme des dérivées de la densité.

Lorsque le support d'estimation est compact, les estimateurs à noyau sont sujets à des effets de bord. De nombreuses méthodes ont vu le jour afin d'y remédier. Gasser et Müller ([70]) ont modifié le noyau aux bords du support, tout comme Gasser, Müller et Mammitzsch ([71]) puis Müller ([116]). Schuster ([143]) et Silverman ([147]) ont adopté une méthode de réflexion (et même négative pour le second) autour des bords du support alors que Hall et Wehrly ([80]) ont employé une méthode géométrique alliant la réflexion et la modification du noyau aux bords du support. Cowling et Hall ([48]) ont eu l'idée de générer des pseudodata et de les intégrer à l'estimateur à noyau.

Dans ce chapitre, nous nous concentrons sur l'estimation à noyau classique de la densité et de ses dérivées successives. Nous améliorons tout d'abord le biais des estimateurs à l'aide de noyaux d'ordre supérieur. Nous développons ensuite une méthode de correction des effets de bord à partir des dérivées successives de la fonction estimée, ce qui justifie une fois de plus l'utilité d'estimer les dérivées successives d'une fonction.

3.2 Estimation à noyau classique de la densité et de ses dérivées

Soient $r \in \mathbb{N}^*$ et m un entier tels que $1 \leq m \leq r$. Notons X_1, \dots, X_n des variables aléatoires indépendantes et identiquement distribuées de même fonction de répartition F .

Soit $t \in \mathbb{R}$. Nous faisons suite au cadre C.2 et nous nous tournons maintenant vers le

but principal de l'estimation à noyau, qui est d'estimer les dérivées de la fonction de répartition, $F^{(m)}(t)$ pour $m = 1, \dots, r$. Pour cela, nous avons besoin de poser l'hypothèse suivante :

Hypothèse (H1)

F est $(m + 2)$ fois continûment différentiable en tout point t .

Introduisons un noyau k tel que :

Hypothèse (H2)

(i) k est m fois continûment différentiable ,

(ii) Le support de $k^{(m)}$ est $[-1, 1]$,

(iii)

$$\sup_{u \in [-1, 1]} |k^{(m)}(u)| < +\infty ,$$

puis nommons (h_n) une suite de fenêtres réelles strictement positives vérifiant $h_n \rightarrow 0$ lorsque $n \rightarrow +\infty$.

Par commodité de notation, nous travaillons avec la densité $F' = f$ comme fonction initiale et au lieu de (C.8), considérons l'estimateur de $f^{(m)}(t)$, $m = 0, \dots, r - 1$:

$$\hat{f}_n^{(m)}(t) = \frac{1}{n} \sum_{i=1}^n g_n^{(m)}(t - X_i) , \quad (3.1)$$

où $g_n^{(m)}$ est donné par (C.6). Les propriétés statistiques de l'estimateur $\hat{f}_n^{(m)}(t)$ sont simplement dérivées car il correspond à une moyenne de copies i.i.d. d'une variable aléatoire

$$g_n^{(m)}(t - X) = \frac{1}{h_n^{m+1}} k^{(m)}\left(\frac{t - X}{h_n}\right) . \quad (3.2)$$

Nous énonçons deux lemmes concernant respectivement le biais et la variance de l'estimateur $\hat{f}_n^{(m)}(t)$.

Lemme 3.2.1. *Si t est un réel et m est un entier compris entre 0 et $r - 1$, alors sous les hypothèses (H1)-(H2), nous avons*

$$\mathbb{E} \hat{f}_n^{(m)}(t) = f^{(m)}(t) + O(h_n) . \quad (3.3)$$

Lemme 3.2.2. *Soient $t \in \mathbb{R}$ et m un entier compris entre 0 et $r - 1$. Si (H1) et (H2) sont satisfaites, alors*

$$\text{Var} \hat{f}_n^{(m)}(t) = \frac{1}{n h_n^{2m+1}} (b_m(t) + O(h_n)) - \frac{(f^{(m)}(t) + O(h_n))^2}{n} , \quad (3.4)$$

avec

$$b_m(t) = f(t) \int_{-1}^1 k^{(m)}(u)^2 du. \quad (3.5)$$

Si $m < r - 1$, nous distinguons à quel point l'avantage de l'estimation polynomiale locale est bénéfique. En effet, l'ordre du biais dans (3.3) est supérieur à celui dans (2.64). Les expressions (3.5) et (2.18) intervenant dans la variance des estimateurs ne sont pas tout-à-fait similaires. La première dépend d'un noyau $k^{(m)}(\cdot)$ et dans la seconde, le terme $\int_{-1}^1 K_r^{(m+1)}(u) du$ (qui représente un lissage du noyau $K_r^{(m+1)}$) joue son rôle. Rencontrer $K_r^{(m+1)}$ pour estimer la dérivée m^e de la densité ne doit pas paraître déroutant puisque le noyau $K_r^{(m)}$ est utilisé pour l'estimation de la dérivée m^e de $\Phi(\cdot, F) = F(\cdot)$.

Nous pouvons maintenant prouver deux résultats sur les propriétés asymptotiques de l'estimateur (3.1) lorsque $n \rightarrow +\infty$. Le premier concerne la consistance.

Théorème 3.2.1. *Soient un réel t et un entier m compris entre 0 et $r - 1$ et supposons les hypothèses (H1)-(H2) satisfaites. Si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (3.6)$$

suffisamment lentement tel que

$$n h_n^{2m+1} \rightarrow +\infty, \quad (3.7)$$

alors l'estimateur $\hat{f}_n^{(m)}(t)$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right]^2 \rightarrow 0, \quad (3.8)$$

donc également convergent en probabilité,

$$\hat{f}_n^{(m)}(t) \xrightarrow{\mathbb{P}} f^{(m)}(t). \quad (3.9)$$

Remarque :

Nous devrions chercher des conditions plus faibles que (3.7) pour la convergence en probabilité, et pour cela il faudrait consulter [4].

En quête d'un résultat sur la normalité asymptotique, le point de départ naturel est la différence normalisée entre l'estimateur et la valeur à estimer,

$$\frac{\hat{f}_n^{(m)}(t) - f^{(m)}(t)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(t)}} = Z_n + \delta_n, \quad (3.10)$$

où

$$Z_n = \frac{\hat{f}_n^{(m)}(t) - \mathbb{E} \hat{f}_n^{(m)}(t)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(t)}} \quad (3.11)$$

est une variable aléatoire normalisée se prêtant aux théorèmes de la limite centrale, et

$$\delta_n = \frac{\mathbb{E} \hat{f}_n^{(m)}(t) - f^{(m)}(t)}{\sqrt{\text{Var} \hat{f}_n^{(m)}(t)}} \quad (3.12)$$

est un terme de biais qui peut tendre vers 0 avec un choix de fenêtre approprié.

Théorème 3.2.2. *Posons $t \in \mathbb{R}$ et m un entier tel que $0 \leq m \leq r - 1$ et admettons les hypothèses (H1)-(H2). Si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$n h_n \rightarrow +\infty \tag{3.13}$$

et suffisamment vite tel que

$$n h_n^{2m+3} \rightarrow 0, \tag{3.14}$$

alors

$$\sqrt{\frac{n h_n^{2m+1}}{b_m(t)}} \left(\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right) \xrightarrow{\mathcal{L}} N(0, 1), \tag{3.15}$$

avec $b_m(t)$ donné par (3.5).

Corollaire au Théorème 3.2.2. *Un intervalle de confiance asymptotique pour $f^{(m)}(t)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\hat{f}_n^{(m)}(t) \pm z_{1-\epsilon/2} \sqrt{\frac{b_m(t)}{n h_n^{2m+1}}}, \tag{3.16}$$

où $z_{1-\epsilon/2}$ est le quantile d'ordre $1 - \epsilon/2$ de la loi normale centrée réduite. Sa longueur tend vers 0 si h_n tend vers 0 suffisamment lentement tel que la condition de consistance (3.7) est satisfaite :

$$n h_n^{2m+1} \rightarrow +\infty.$$

Les conditions de consistance (3.6) et (3.7) sont intuitivement acceptables : la première est la même pour tout m , alors que la seconde est plus restrictive quand m augmente. Elles sont analogues à (2.19) et (2.20). Les conditions (3.13) et (3.14) pour la normalité asymptotique sont moins intuitives : la première est identique pour tout m , alors que la seconde devient moins stricte quand m augmente. Ces observations servent à illustrer que les deux modes de convergence – en moyenne quadratique et en loi – sont très différents. Le Corollaire résout ce puzzle en indiquant que la normalité asymptotique est utile seulement si la condition de consistance est satisfaite.

Notons que les conditions de consistance et de normalité asymptotique dépendent de m donc si l'on s'intéresse à l'estimation conjointe des dérivées, on ne peut appliquer la même fenêtre pour tous les ordres de dérivation.

Si $m < r - 1$, alors la condition (2.24) est plus souple que la condition (3.14). Cela est dû à la réduction de biais, qui est l'avantage de la méthode d'approximation locale polynomiale par rapport à la méthode de convolution. Mais nous expliquons dans la section suivante comment modifier le cadre afin d'améliorer la condition (3.14).

3.3 Une amélioration à l'aide de noyaux réducteurs de biais

Soit ν un entier non nul et posons à présent deux nouvelles hypothèses :

Hypothèse (H1')

F est $(m + \nu + 1)$ fois continûment différentiable en tout point t .

Hypothèse (H3)

k est un noyau d'ordre ν , c'est-à-dire tel que

(i)

$$\int u^i k(u) du = 0 \text{ pour tout } i = 1, \dots, \nu - 1$$

et

(ii)

$$\int u^\nu k(u) du = 1.$$

L'ordre d'un noyau est toujours pair, donc d'ordre 2 ou plus, c'est pourquoi ν désigne par la suite un nombre pair. Notons que les noyaux positifs symétriques sont d'ordre 2. Quant aux noyaux d'ordre supérieur à 2, ils possèdent des parties négatives et ne sont pas des densités de probabilité.

Le lemme suivant montre que l'ordre d'un noyau détermine sa capacité à réduire le biais de l'estimateur.

Lemme 3.3.1. *Si t est un réel et m est un entier compris entre 0 et $r - 1$, alors sous les hypothèses (H1')-(H2)-(H3), nous avons*

$$\mathbb{E} \hat{f}_n^{(m)}(t) = f^{(m)}(t) + O(h_n^\nu). \quad (3.17)$$

L'estimateur est asymptotiquement sans biais et le biais est d'ordre h_n^ν , ce qui améliore le résultat du Lemme 3.2.1 puisque $\nu \geq 2$. Le Théorème 3.2.2 peut être ainsi renforcé comme suit.

Théorème 3.3.1. *Soient t un réel et m un entier puis supposons que les hypothèses (H1')-(H2)-(H3) sont satisfaites. Si $h_n \rightarrow 0$ suffisamment lentement tel que*

$$n h_n \rightarrow +\infty \quad (3.18)$$

et suffisamment vite tel que

$$n h_n^{2(m+\nu)+1} \rightarrow 0, \quad (3.19)$$

alors

$$\sqrt{\frac{n h_n^{2m+1}}{b_m(t)}} \left(\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right) \xrightarrow{\mathcal{L}} N(0, 1), \quad (3.20)$$

avec $b_m(t)$ donné par (3.5).

Si $\nu \geq 2$, la condition (3.19) est moins stricte que (3.14) tandis que les conditions (3.13) et (3.18) sont similaires.

3.4 Application à une méthode de correction des effets de bord

Soient m , p et r des entiers, a , b , c et h des réels tels que $a \leq b$ et $h > 0$.

Dans le cadre de l'estimation de la densité, les effets de bord surviennent lorsque f a un support compact $[a, b]$. Les fonctions $f^{(m)}(b)$ représentent des dérivées à gauche et on suppose qu'elles existent. Avec une fenêtre donnée $h < \frac{b-a}{2}$, les estimateurs $\hat{f}_n^{(m)}(t)$ sont uniquement définis sur l'intervalle $[a+h, b-h]$. Pour $t \in [a, a+h] \cup [b-h, b]$, ils doivent donc être modifiés. La méthode standard de correction des effets de bord pour l'estimation de la densité est l'application d'un noyau ayant un support tronqué, par exemple $[-1, \frac{b-t}{h}]$ pour h fixé et $t \in [b-h, b]$.

Cependant, une méthode alternative basée sur un développement de Taylor peut être proposée, à condition que des estimateurs soient connus pour les dérivées successives de la densité f . Considérons l'estimation de $f^{(m)}(b)$, pour $m < r-1$. Grâce à la formule de Taylor-Lagrange de $f^{(m)}$ à l'ordre $p \leq r-m$ en un point $b-c$ tel que $0 < c < b-a$, nous écrivons :

$$f^{(m)}(b) = \sum_{j=0}^p \frac{1}{j!} f^{(m+j)}(b-c) c^j + O(c^{p+1}), \quad (3.21)$$

où

$$O(c^{p+1}) = \frac{1}{(p+1)!} f^{(m+p+1)}(b-c') c^{p+1}$$

avec c' compris entre 0 et c . Un estimateur naturel de $f^{(m)}(b)$ est obtenu en injectant les estimateurs à noyau standards (3.1) des dérivées $f^{(m+j)}(b-c)$ apparaissant dans la somme à droite de (3.21) :

$$\tilde{f}_n^{(m)}(b) = \sum_{j=0}^p \frac{1}{j!} \hat{f}_n^{(m+j)}(b-c) c^j, \quad (3.22)$$

pourvu que les deux hypothèses suivantes soient satisfaites :

Hypothèse (H1'')

F est $(m+p+2)$ fois continûment différentiable en tout point t .

Hypothèse (H2')

- (i) k est $(m + p)$ fois continûment différentiable ,
- (ii) Le support de $k^{(m+j)}$ est $[-1, 1]$ pour $j = 0, \dots, p$,
- (iii)

$$\sup_{u \in [-1, 1]} |k^{(m+j)}(u)| < +\infty \text{ pour } j = 0, \dots, p.$$

Alors l'estimateur $\tilde{f}_n^{(m)}(b)$ est proche de $f^{(m)}(b)$ avec un terme d'erreur d'ordre $O(c_n^{p+1})$ pour n assez grand, d'après la consistance des estimateurs $\hat{f}_n^{(m+j)}(b - c)$.

En quête de résultats asymptotiques plus raffinés, nous remplaçons la constante c par une suite (c_n) convergeant vers 0. Il est désormais nécessaire d'expliquer comment les fenêtres à l'origine des estimateurs $\hat{f}_n^{(m+j)}$ sont déterminées. Rappelant le Théorème 3.2.1, le Théorème 3.2.2 ainsi que son Corollaire, un choix naturel de fenêtre intervenant dans l'expression des estimateurs $\hat{f}_n^{(m+j)}$ est

$$h_{m+j,n} = n^{-\frac{1}{2(m+j+1)}} h, \quad (3.23)$$

où $h > 0$ est une constante choisie convenablement. En effet, nous avons $h_{m+j,n} \rightarrow 0$ et $n h_{m+j,n}^{2(m+j)+1} \rightarrow +\infty$, assurant la consistance, et $n h_{m+j,n}^{2(m+j)+3} \rightarrow 0$, garantissant la normalité asymptotique et le rétrécissement des intervalles de confiance lorsque la taille de l'échantillon n croît.

Insérons $c = c_n$ dans (3.21) et (3.22), alors :

$$\begin{aligned} \tilde{f}_n^{(m)}(b) - f^{(m)}(b) &= \sum_{j=0}^p \frac{1}{j!} \left(\hat{f}_n^{(m+j)}(b - c_n) - f^{(m+j)}(b - c_n) \right) c_n^j + O(c_n^{p+1}) \\ &= \sum_{j=0}^p \frac{\sqrt{b_{m+j}(b - c_n)}}{j!} a_{j,n} Z_{m+j,n} + O(c_n^{p+1}), \end{aligned} \quad (3.24)$$

où

$$a_{j,n} = \left(n h_{m+j,n}^{2(m+j)+1} \right)^{-\frac{1}{2}} c_n^j \quad (3.25)$$

et

$$Z_{m+j,n} = \sqrt{\frac{n h_{m+j,n}^{2(m+j)+1}}{b_{m+j}(b - c_n)}} \left(\hat{f}_n^{(m+j)}(b - c_n) - f^{(m+j)}(b - c_n) \right). \quad (3.26)$$

La décomposition (3.24) est primordiale pour démontrer la consistance de $\tilde{f}_n^{(m)}(b)$, l'estimateur corrigé des effets de bord des dérivées d'ordre supérieur de la densité.

Théorème 3.4.1. *Si les hypothèses (H1'') et (H2') sont vérifiées alors l'estimateur $\tilde{f}_n^{(m)}(b)$ est convergent en probabilité,*

$$\tilde{f}_n^{(m)}(b) \xrightarrow{\mathbb{P}} f^{(m)}(b), \quad (3.27)$$

avec le choix de fenêtre (3.23).

Or la suite (c_n) doit être telle que c_n soit le plus petit possible, et puisque $h_{m+j,n}$ augmente avec j , nous devrions prendre

$$c_n = h_{m,n} = n^{-\frac{1}{2(m+1)}} h \quad (3.28)$$

lorsque nous effectuons le développement de Taylor jusqu'à l'ordre p . Avec un tel choix pour c_n , nous établissons la normalité asymptotique de l'estimateur $\tilde{f}_n^{(m)}(b)$.

Théorème 3.4.2. *Sous les hypothèses (H1'') et (H2'), si la suite (c_n) vérifie (3.28), le choix de fenêtre (3.23) conduit à :*

$$\frac{p!}{\sqrt{b_{m+p}(b)}} n^{\frac{1}{4}(\frac{1}{m+p+1} + \frac{2p}{m+1})} h^{m+\frac{1}{2}} \left(\tilde{f}_n^{(m)}(b) - f^{(m)}(b) \right) \xrightarrow{\mathcal{L}} N(0, 1), \quad (3.29)$$

avec

$$b_{m+p}(b) = f(b) \int_{-1}^1 k^{(m+p)}(u)^2 du. \quad (3.30)$$

Pour $p = 0$, le résultat est conforme à celui obtenu dans le Théorème 3.2.2, en fixant $h_n = n^{-\frac{1}{2(m+1)}} h$.

Corollaire au Théorème 3.4.2. *Un intervalle de confiance pour $f^{(m)}(t)$ au niveau de confiance $1 - \epsilon$ est donné par les bornes*

$$\tilde{f}_n^{(m)}(t) \pm z_{1-\epsilon/2} \frac{\sqrt{b_{m+p}(b)}}{p! h^{m+\frac{1}{2}} n^{\frac{1}{4}(\frac{1}{m+p+1} + \frac{2p}{m+1})}}. \quad (3.31)$$

Le résultat est valable seulement si $b_{m+p}(b) > 0$. La longueur de l'intervalle de confiance est égale à $2\sqrt{b_{m+p}(b)}$ fois un facteur qui décroît avec p : l'exposant $\frac{1}{4}(\frac{1}{m+p+1} + \frac{2p}{m+1})$ augmente avec p , il est supérieur à $\frac{1}{4(m+1)}$ (pour $p = 0$) et converge vers $+\infty$ (lorsque $p \rightarrow +\infty$).

3.5 Preuves

Dans toutes les preuves, t représente un réel et m correspond à un entier compris entre 0 et $r - 1$. ν est un nombre pair supérieur ou égal à 2.

3.5.1 Preuve du Lemme 3.2.1

Nous avons mis en lumière précédemment le fait que l'estimateur $\hat{f}_n^{(m)}(t)$ est une moyenne de copies i.i.d. d'une variable aléatoire

$$g_n^{(m)}(t - X) = \frac{1}{h_n^{m+1}} k^{(m)}\left(\frac{t - X}{h_n}\right).$$

Ainsi, l'espérance de l'estimateur $\hat{f}_n^{(m)}(t)$ est égale au premier moment de $g_n^{(m)}(t - X)$, que nous calculons plus bas.

$$\begin{aligned}\mathbb{E} g_n^{(m)}(t - X) &= \int g_n^{(m)}(t - s) f(s) ds \\ &= \int f^{(m)}(t - s) g_n(s) ds,\end{aligned}$$

d'après l'identité entre (C.4) et (C.5). En utilisant un développement de Taylor-Lagrange avec s' compris entre 0 et s , nous obtenons

$$\mathbb{E} g_n^{(m)}(t - X) = \int_{-h_n}^{h_n} \left(f^{(m)}(t) - f^{(m+1)}(t - s') s \right) g_n(s) ds \quad (3.32)$$

$$= f^{(m)}(t) + O(h_n), \quad (3.33)$$

avec

$$|O(h_n)| \leq \sup\{|f^{(m+1)}(t - s')| ; s' \in [-h_n, h_n]\} 2h_n.$$

D'où

$$\mathbb{E} \hat{f}_n^{(m)}(t) = f^{(m)}(t) + O(h_n). \blacksquare \quad (3.34)$$

3.5.2 Preuve du Lemme 3.2.2

La formule (3.1) utilisée conjointement avec l'indépendance et l'équidistribution des variables $g_n^{(m)}(t - X_i)$ entraîne que

$$\mathbb{V}ar \hat{f}_n^{(m)}(t) = \frac{\mathbb{V}ar g_n^{(m)}(t - X)}{n}. \quad (3.35)$$

Déterminer le second moment de $g_n^{(m)}(t - X)$ s'avère donc crucial pour disposer d'une expression de la variance exacte asymptotique de l'estimateur $\hat{f}_n^{(m)}(t)$. En appliquant successivement la définition de l'espérance, le changement de variable $u = \frac{s-t}{h_n}$ puis la formule de Taylor-Lagrange pour u' compris entre 0 et $h_n u$, avec $u \in [-1, 1]$, on a :

$$\begin{aligned}\mathbb{E} \left[g_n^{(m)}(t - X) \right]^2 &= \mathbb{E} \left[\frac{1}{h_n^{m+1}} k^{(m)} \left(\frac{t - X}{h_n} \right) \right]^2 \\ &= \frac{1}{h_n^{2m+1}} \int \frac{1}{h_n} k^{(m)} \left(\frac{t - s}{h_n} \right)^2 f(s) ds \\ &= \frac{1}{h_n^{2m+1}} \int_{-1}^1 k^{(m)}(-u)^2 f(t + h_n u) du \\ &= \frac{1}{h_n^{2m+1}} \left(f(t) \int_{-1}^1 k^{(m)}(u)^2 du + O(h_n) \right).\end{aligned}$$

Grâce à (3.32), il vient que :

$$\mathbb{V}ar g_n^{(m)}(t - X) = \frac{1}{h_n^{2m+1}} (b_m(t) + O(h_n)) - \left(f^{(m)}(t) + O(h_n) \right)^2, \quad (3.36)$$

avec

$$b_m(t) = f(t) \int_{-1}^1 k^{(m)}(u)^2 du. \quad (3.37)$$

En injectant (3.36) dans (3.35), nous obtenons

$$\text{Var } \hat{f}_n^{(m)}(t) = \frac{1}{n h_n^{2m+1}} (b_m(t) + O(h_n)) - \frac{(f^{(m)}(t) + O(h_n))^2}{n}. \blacksquare \quad (3.38)$$

3.5.3 Preuve du Théorème 3.2.1

L'erreur quadratique moyenne de $\hat{f}_n^{(m)}(t)$ est

$$\begin{aligned} \mathbb{E} \left[\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right]^2 &= \text{Var } \hat{f}_n^{(m)}(t) + [\mathbb{E} \hat{f}_n^{(m)}(t) - f^{(m)}(t)]^2 \\ &= \frac{1}{n h_n^{2m+1}} (b_m(t) + O(h_n)) - \frac{(f^{(m)}(t) + O(h_n))^2}{n} + O(h_n^2) \end{aligned}$$

Par (3.34), le second terme est $O(h_n^2)$, qui tend vers 0 sous la condition (3.6). Le premier terme est donné par (3.38), qui tend vers 0 si la condition (3.7) est satisfaite. Cela prouve (3.8), et ensuite (3.9) découle de l'inégalité de Markov F.11 pour $p = 2$:

$$\mathbb{P} \left[|\hat{f}_n^{(m)}(t) - f^{(m)}(t)| > \epsilon \right] \leq \frac{\mathbb{E} \left[\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right]^2}{\epsilon^2}. \blacksquare$$

3.5.4 Preuve du Théorème 3.2.2

On remarque la décomposition suivante :

$$\sqrt{\frac{n h_n^{2m+1}}{b_m(t)}} \left(\hat{f}_n^{(m)}(t) - f^{(m)}(t) \right) = \sqrt{\frac{\text{Var } \hat{f}_n^{(m)}(t)}{\frac{b_m(t)}{n h_n^{2m+1}}}} (Z_n + \delta_n), \quad (3.39)$$

avec Z_n et δ_n donnés respectivement par (3.11) et (3.12).

Analysons d'abord le terme Z_n , candidat pour être asymptotiquement normal. En insérant les expressions (3.1), (3.2), et (3.4), on a

$$Z_n = \frac{\frac{1}{n h_n^{m+1}} \sum_{i=1}^n \left(k^{(m)} \left(\frac{t-X_i}{h_n} \right) - \mathbb{E} k^{(m)} \left(\frac{t-X_i}{h_n} \right) \right)}{\sqrt{\frac{1}{n h_n^{2m+1}} (b_m(t) + O(h_n)) - \frac{(f^{(m)}(t) + O(h_n))^2}{n}}}.$$

Or la variable Z_n a la même distribution limite que la variable aléatoire \tilde{Z}_n dont l'expression est la suivante :

$$\tilde{Z}_n = \sum_{i=1}^n \frac{\left(k^{(m)} \left(\frac{t-X_i}{h_n} \right) - \mathbb{E} k^{(m)} \left(\frac{t-X_i}{h_n} \right) \right)}{\sqrt{n h_n (b_m(t) + O(h_n))}}.$$

Nous pensons immédiatement à utiliser le théorème de la limite centrale de Lindeberg-Feller F.9 avec $k_n = n$ et $Y_{n,i}$ des copies i.i.d. de

$$Y_n = \frac{k^{(m)}\left(\frac{t-X}{h_n}\right)}{\sqrt{n h_n (b_m(t) + O(h_n))}}.$$

Comme $\text{Var } Z_n = 1$ pour tout n , on en déduit que $\lim_{n \rightarrow +\infty} \text{Var } \tilde{Z}_n = 1$ donc la condition F.9(iv) est satisfaite. Le noyau $k^{(m)}$ est borné selon (H2(iii)) donc par le Lemme F.9.2, si on suppose (3.13), alors la condition de Lindeberg-Feller F.9(iii) :

$$\forall \epsilon > 0, n \mathbb{E} \left[Y_n^2 \mathbf{1}[|Y_n| > \epsilon] \right] \rightarrow 0$$

est vérifiée. Ceci entraîne que \tilde{Z}_n suit une loi normale centrée réduite, donc y compris Z_n .

Concentrons-nous ensuite sur le terme de biais (3.12). En injectant (3.3) et (3.4), nous obtenons

$$\delta_n = \frac{O(h_n)}{\sqrt{\frac{1}{n h_n^{2m+1}} (b_m(t) + O(h_n)) - \frac{(f^{(m)}(t) + O(h_n))^2}{n}}} = O\left(\sqrt{n h_n^{2m+3}}\right),$$

qui converge vers 0 sous la condition (3.14).

Finalement, observons que

$$\frac{\text{Var } \hat{f}_n^{(m)}(t)}{\frac{b_m(t)}{n h_n^{2m+1}}} \xrightarrow{\mathbb{P}} 1.$$

Le résultat énoncé provient alors des règles de Cramér F.10-1(i)-(ii). ■

3.5.5 Preuve du Lemme 3.3.1

L'hypothèse (H1') nous permet d'effectuer un développement de Taylor de $f^{(m)}$ à l'ordre $\nu - 1$ au point t et par suite, nous pouvons remplacer (3.32) par

$$\int_{-h_n}^{h_n} \left(\sum_{j=0}^{\nu-1} f^{(m+j)}(t) \frac{(-s)^j}{j!} + f^{(m+\nu)}(t-s') \frac{s^\nu}{\nu!} \right) g_n(s) ds$$

pour s' compris entre 0 et s et obtenir

$$\mathbb{E} \hat{f}_n^{(m)}(t) = f^{(m)}(t) + O(h_n^\nu). \quad \blacksquare \tag{3.40}$$

3.5.6 Preuve du Théorème 3.3.1

Pour démontrer le Théorème 3.3.1, on s'inspire de la preuve du Théorème 3.2.2. Toutes les étapes doivent être reproduites à l'identique, exceptée celle concernant le calcul du

terme de biais dans (3.12). Grâce à (3.40), celui-ci a pour expression

$$\delta_n = \frac{O(h_n^\nu)}{\sqrt{\frac{1}{nh_n^{2m+1}}(b_m(t) + O(h_n))}} = O\left(\sqrt{nh_n^{2(m+\nu)+1}}\right),$$

et converge vers 0 lorsque $n \rightarrow +\infty$ sous la condition (3.19). ■

3.5.7 Preuve du Théorème 3.4.1

Reprenons l'expression (3.24) :

$$\tilde{f}_n^{(m)}(b) - f^{(m)}(b) = \sum_{j=0}^p \frac{\sqrt{b_{m+j}(b - c_n)}}{j!} a_{j,n} Z_{m+j,n} + O(c_n^{p+1}), \quad (3.41)$$

avec $a_{j,n}$ et $Z_{m+j,n}$ donnés respectivement par (3.25) et (3.26).

Par inspection de la preuve du Théorème 3.2.2, avec $b - c_n$ dans le rôle de t , il vient que

$$Z_{m+j,n} \xrightarrow{\mathcal{L}} N(0, 1) \quad (3.42)$$

pour tout $j = 0, \dots, p$. Puisque $nh_{m+j,n}^{2(m+j)+1} \rightarrow +\infty$, $c_n \rightarrow 0$, et $b_{m+j}(b - c_n) \rightarrow b_{m+j}(b)$, les règles de Cramér F.10-1(i)-(ii) entraînent que $\tilde{f}_n^{(m)}(b)$ est un estimateur consistant de $f^{(m)}(b)$. ■

3.5.8 Preuve du Théorème 3.4.2

Dans le but d'obtenir un résultat sur la distribution asymptotique de $\tilde{f}_n^{(m)}(b)$, nous avons besoin d'identifier le terme $a_{j,n}$ dominant dans l'expression à droite de (3.24). Pour cela, développons le coefficient $a_{j,n}$ donné par (3.25) à l'aide de (3.23) et (3.28), d'où :

$$\begin{aligned} a_{j,n} &= \left(n \left(n^{-\frac{1}{2(m+j+1)}} h \right)^{2(m+j)+1} \right)^{-\frac{1}{2}} \left(n^{-\frac{1}{2(m+1)}} h \right)^j \\ &= n^{-\frac{1}{4} \left(\frac{1}{m+j+1} + \frac{2j}{m+1} \right)} h^{-(m+\frac{1}{2})}. \end{aligned} \quad (3.43)$$

Le terme $a_{j,n}$ dominant étant celui de plus petit ordre, nous sommes amenés à déterminer

$$\operatorname{argmax}_{j=0, \dots, p} g(j),$$

où $g : \mathbb{R}_+ \mapsto \mathbb{R}$ est la fonction définie par

$$g(x) = \frac{1}{m+x+1} + \frac{2x}{m+1}.$$

Puisque

$$g''(x) = \frac{2}{(m+x+1)^3} > 0,$$

la fonction g est convexe, donc elle atteint son maximum en $x = 0$ ou en $x = p$. Le dernier cas se produit si $g(0) < g(p)$, c'est-à-dire si

$$0 < \frac{1}{m+p+1} + \frac{2p-1}{m+1},$$

ce qui équivaut à $0 < (2p+2m+1)p$. Cette inégalité est toujours vérifiée donc le terme $a_{j,n}$ dominant dans la somme à droite de (3.24) est le p^e terme $a_{p,n}$. Une conséquence directe est :

$$\frac{a_{j,n}}{a_{p,n}} \rightarrow 0 \text{ pour } j = 1, \dots, p-1. \quad (3.44)$$

De plus, avec (3.28) et (3.43), on a :

$$\begin{aligned} \frac{c_n^{p+1}}{a_{p,n}} &= \left(n^{-\frac{1}{2(m+1)}} h \right)^{p+1} n^{\frac{1}{4} \left(\frac{1}{m+p+1} + \frac{2p}{m+1} \right)} h^{m+\frac{1}{2}} \\ &= n^{-\frac{1}{4} \left(\frac{2}{m+1} - \frac{1}{m+p+1} \right)} h^{m+p+\frac{3}{2}}, \end{aligned}$$

d'où

$$\frac{c_n^{p+1}}{a_{p,n}} \rightarrow 0. \quad (3.45)$$

En effet,

$$\frac{2}{m+1} - \frac{1}{m+p+1} = \frac{2p+m+1}{(m+1)(m+p+1)} > 0.$$

En multipliant par $\frac{p!}{\sqrt{b_{m+p}(b)} a_{p,n}}$ dans (3.24), il vient que

$$\frac{p!}{\sqrt{b_{m+p}(b)} a_{p,n}} \left(\tilde{f}_n^{(m)}(b) - f^{(m)}(b) \right) = \sum_{j=0}^p \frac{p!}{\sqrt{b_{m+p}(b)}} \frac{\sqrt{b_{m+j}(b-c_n)}}{j!} \frac{a_{j,n}}{a_{p,n}} Z_{m+j,n} + O\left(\frac{c_n^{p+1}}{a_{p,n}}\right)$$

donc d'après (3.44) et (3.45) et la règle de Cramér F.10-2, $\frac{p!}{\sqrt{b_{m+p}(b)} a_{p,n}} \left(\tilde{f}_n^{(m)}(b) - f^{(m)}(b) \right)$

a la même loi asymptotique que $\frac{\sqrt{b_{m+p}(b-c_n)}}{\sqrt{b_{m+p}(b)}} Z_{m+p,n}$. Faisons tendre n vers $+\infty$ dans cette dernière expression puis utilisons le fait que $c_n \rightarrow 0$, $b_{m+p}(b-c_n) \rightarrow b_{m+p}(b)$ et (3.42). La règle de Cramér F.3-1(ii) puis la forme de $a_{p,n}$ obtenue à l'aide de (3.43) pour $j = p$ nous permet d'arriver à

$$\frac{p!}{\sqrt{b_{m+p}(b)}} n^{\frac{1}{4} \left(\frac{1}{m+p+1} + \frac{2p}{m+1} \right)} h^{m+\frac{1}{2}} \left(\tilde{f}_n^{(m)}(b) - f^{(m)}(b) \right) \xrightarrow{\mathcal{L}} N(0, 1), \quad (3.46)$$

avec

$$b_{m+p}(b) = f(b) \int_{-1}^1 k^{(m+p)}(u)^2 du. \blacksquare \quad (3.47)$$

Chapitre 4

Estimation à noyau du taux de hasard et de ses dérivées successives et application au taux d'accumulation du surplus en actuariat

4.1 Introduction

Dans ce chapitre, nous décidons d'explorer l'estimation non paramétrique du taux de hasard, jouant un rôle particulièrement conséquent en fiabilité (connu sous le nom de taux de défaillance) et en actuariat. Il est défini par

$$\alpha(t) = \frac{f(t)}{1 - F(t)}, t \in \mathbb{R}. \quad (4.1)$$

Ce n'est ni une fonctionnelle linéaire de la fonction de répartition, ni un rapport de deux fonctionnelles linéaires de la fonction de répartition (contrairement aux fonctions L_F et T_F).

Watson et Leadbetter ([159], [160]) se sont intéressés à l'estimation de cette fonction pour des données indépendantes et identiquement distribuées et ont introduit trois estimateurs. Rice et Rosenblatt ([139]) ont ensuite étudié les propriétés (espérance, variance) de ces estimateurs.

Tanner et Wong ([154]), Yandell ([166]), Müller et Wang ([119]) puis Jiang et Doksum ([87]) ont exclusivement analysé l'estimation à noyau du taux de hasard pour des durées de vie censurées à droite indépendantes. Les premiers ont prouvé la normalité asymptotique de l'estimateur à noyau tandis que le second a fourni des intervalles de confiance du taux de hasard à l'aide de cet estimateur. Müller et Wang ([119]) ont construit des estimateurs lisses pour les dérivées du taux de hasard, à l'aide de polynômes locaux et de noyaux d'ordre supérieur, puis Jiang et Doksum ([87]) ont démontré la consistance et la normalité de ces estimateurs.

Sans le savoir, Ramlau-Hansen ([134]) a retrouvé les résultats de Tanner et Wong ([154]) en se focalisant sur le modèle à intensité multiplicative pour des processus de comptage, un cadre qui englobe entre autres la situation de censure et les processus markoviens.

Son estimateur est obtenu grâce au lissage de l'estimateur de Nelson-Aalen (Nelson ([124])) et possède de bonnes propriétés asymptotiques telles que la consistance et la normalité asymptotique. En dérivant un certain nombre de fois l'estimateur à noyau, Ramlau-Hansen ([134]) a défini des estimateurs pour les dérivées du taux de hasard. Nielsen et Tangaard ([125]) ont indiqué que les estimateurs (du taux de hasard et de ses dérivées) définis par Ramlau-Hansen ([134]) étaient sujets aux effets de bord. Fan et Gijbels ([68]) ont souligné que l'approximation polynomiale locale pouvait constituer un remède, c'est pourquoi nous tenterons de construire des estimateurs du taux de hasard et ses dérivées successives dans le modèle à intensité multiplicative à l'aide de l'approximation polynomiale et de démontrer son bon comportement asymptotique. Ramlau-Hansen ([134]) avait brièvement décrit comment obtenir la consistance et la normalité asymptotique des estimateurs du taux de hasard et de ses dérivées successives à l'aide de la méthode de convolution. Nous mènerons cette procédure à terme de manière rigoureuse.

Savoir estimer (4.1) présente un intérêt en science actuarielle. (Hoem ([85])) fait appel à des processus de Markov pour décrire la trajectoire des assurés au cours de leur vie et dans le contexte de la participation aux bénéfices dont une approche moderne se trouve dans (Norberg ([127])), le taux d'accumulation des bénéfices (Ramlau-Hansen ([136]), Norberg ([126], [127], [128], [129])) s'exprime en fonction des taux de transition d'une chaîne de Markov. Pourtant indicés par un état de départ et un état d'arrivée, les taux de transition ont la même formule que (4.1) donc leur estimation est similaire à celle de (4.1) dans le modèle à intensité multiplicative, ce qui nous permettra de proposer un estimateur du taux d'accumulation des bénéfices et d'examiner ses propriétés statistiques.

4.2 Résultat préliminaire

Nous énonçons un résultat découlant de la méthode Delta F.3 qui sera utile tout au long du chapitre.

Lemme 4.2.1. *Fixons $k \in \mathbb{N}^*$. Soit $(a_n) = (a_{1,n}, \dots, a_{k-1,n}, a_{k,n})$ une suite telle que*

$$a_{k,n} \rightarrow +\infty, \text{ pour } j = 1, \dots, k,$$

et

$$a_{k,n} = o(a_{j,n}), \text{ pour } j = 1, \dots, k - 1. \quad (4.2)$$

Si $\mathbf{T}_n = (T_{1,n}, \dots, T_{k,n})$, $\theta = (\theta_1, \dots, \theta_k)$ et $\mathbf{T} = (T_1, \dots, T_k)$ sont des vecteurs aléatoires tels que

$$(a_{1,n}(T_{1,n} - \theta_1), \dots, a_{k-1,n}(T_{k-1,n} - \theta_{k-1}), a_{k,n}(T_{k,n} - \theta_k)) \xrightarrow{\mathcal{L}} \mathbf{T},$$

alors :

$$a_{k,n}(T_{1,n} - \theta_1, \dots, T_{k-1,n} - \theta_{k-1}, T_{k,n} - \theta_k) \xrightarrow{\mathcal{L}} (0, \dots, 0, T_k). \quad (4.3)$$

Le Lemme 4.2.1 indique que la loi asymptotique obtenue dans (4.3) est déterminée par la composante de la suite (a_n) ayant la vitesse de convergence la plus faible.

Preuve :

Nous supposons $a_{k,n} = o(a_{j,n})$, pour $j = 1, \dots, k-1$, c'est-à-dire

$$\frac{a_{k,n}}{a_{j,n}} \xrightarrow{\mathbb{P}} 0, \text{ pour } j = 1, \dots, k-1.$$

Donc en appliquant la règle de Cramér F.10-1(ii), nous obtenons

$$\begin{aligned} & \left(\frac{a_{k,n}}{a_{1,n}} \cdot a_{1,n}(T_{1,n} - \theta_1), \dots, \frac{a_{k,n}}{a_{k-1,n}} \cdot a_{k-1,n}(T_{k-1,n} - \theta_{k-1}), a_{k,n}(T_{k,n} - \theta_k) \right) \\ & \xrightarrow{\mathcal{L}} (0.T_1, \dots, 0.T_{k-1}, T_k), \end{aligned}$$

d'où le résultat. ■

4.3 Estimation du taux de hasard et de ses dérivées successives pour des variables aléatoires indépendantes et identiquement distribuées

Soient X_1, \dots, X_n des variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F et de densité de probabilité f . La fonction de répartition empirique notée F_n associée à ce n -échantillon s'écrit

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[X_i \leq t]}, t \in \mathbb{R}.$$

Considérons l'estimation du taux de hasard par sa version empirique

$$\hat{\alpha}_n(t) = \frac{\hat{f}_n(t)}{1 - F_n(t)}, \quad (4.4)$$

où \hat{f}_n représente l'estimateur à noyau de la densité f donné par (3.1) pour $m = 0$.

Rappelant le Théorème 3.2.2, soit une suite (h_n) telle que $nh_n \rightarrow +\infty$ et $nh_n^3 \rightarrow 0$ puis utilisons (1.5). Alors nous pouvons écrire

$$\left(\sqrt{nh_n}(\hat{f}_n(t) - f(t)), \sqrt{n}((1 - F_n(t)) - (1 - F(t))) \right) \xrightarrow{\mathcal{L}} \mathbf{T},$$

pour un certain vecteur aléatoire \mathbf{T} . Grâce au Lemme 4.2.1, nous déduisons que

$$\sqrt{nh_n} \left(\begin{pmatrix} \hat{f}_n(t) \\ 1 - F_n(t) \end{pmatrix} - \begin{pmatrix} f(t) \\ 1 - F(t) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \begin{pmatrix} T_1 \\ 0 \end{pmatrix},$$

où $T_1 \stackrel{\mathcal{L}}{=} N(0, b_0(t))$ avec $b_0(t)$ donné par (3.5) :

$$b_0(t) = f(t) \int_{-1}^1 k(u)^2 du.$$

La méthode Delta F.3 avec le choix $\phi(u, v) = \frac{u}{v}$ implique alors que

$$\sqrt{n h_n} (\hat{\alpha}_n(t) - \alpha(t)) \xrightarrow{\mathcal{L}} N\left(0, \frac{b_0(t)}{(1 - F(t))^2}\right),$$

c'est-à-dire

$$\sqrt{n h_n} (\hat{\alpha}_n(t) - \alpha(t)) \xrightarrow{\mathcal{L}} N\left(0, \frac{\alpha(t) \int_{-1}^1 k(u)^2 du}{1 - F(t)}\right). \quad (4.5)$$

Prendre (a_n) d'ordre $\sqrt{n h_n}$ est l'unique choix utile. Si (a_n) converge plus vite, alors aucune convergence en loi ne serait obtenue, et si (a_n) converge plus lentement, alors on obtiendrait la convergence triviale vers 0 en loi et donc en probabilité.

La variance asymptotique dans (4.5) est la même que celle du Théorème IV.2.4 dans [11] avec $a_n = \sqrt{n}$ et $Y^{(n)} = n[1 - F_n(t)]$ (le nombre de survivants parmi n individus avec des durées de vie T_1, \dots, T_n , i.i.d. non censurées). Le résultat du Théorème IV.2.4 concerne

$$\sqrt{n h_n} (\hat{\alpha}^{(n)}(t) - \tilde{\alpha}(t)),$$

où $\hat{\alpha}^{(n)}(t)$ est la dérivée première de l'estimateur de Nelson-Aalen lissé et $\tilde{\alpha}(t) = \alpha(t) J^{(n)}(t)$. En fait, l'estimateur $\hat{\alpha}^{(n)}(t)$ n'est pas bien défini pour t tel que $J^{(n)}(t) = \mathbf{1}_{[Y^{(n)}(t) > 0]} = 0$. Le fixer égal à 0 pour de tels t signifie que nous utilisons l'estimateur $\hat{\alpha}^{(n)}(t) J^{(n)}(t)$, et devrions chercher la loi asymptotique de

$$\sqrt{n h_n} (\hat{\alpha}^{(n)}(t) J^{(n)}(t) - \alpha(t))$$

ou peut-être

$$\sqrt{n h_n} (\hat{\alpha}^{(n)}(t) J^{(n)}(t) - \tilde{\alpha}(t)).$$

Cependant, ces ajustements sont sans conséquence pour la loi asymptotique puisque nous montrons facilement que

$$\sqrt{n h_n} (1 - J^{(n)}(t)) \xrightarrow{\mathbb{P}} 0.$$

En fait, par l'inégalité de Markov F.11 pour $p = 2$ et puisque $1 - J^{(n)}(t)$ est égal à 0 ou 1 donc est égal à son carré, on a :

$$\forall \epsilon > 0, \mathbb{P}\left[\sqrt{n h_n} (1 - J^{(n)}(t)) > \epsilon\right] \leq \frac{n h_n \mathbb{E}[1 - J^{(n)}(t)]}{\epsilon^2},$$

où

$$n h_n \mathbb{E}[1 - J^{(n)}(t)] = n h_n F(t)^n = h_n e^{\ln(n) + n \ln(F(t))} \rightarrow 0$$

pour tout t tel que $F(t) < 1$.

Supposons désormais que nous sommes intéressés par la dérivée première du taux de hasard

$$\alpha'(t) = \frac{f'(t)}{1 - F(t)} + \frac{f(t)^2}{(1 - F(t))^2}.$$

Nous proposons naturellement l'estimateur $\hat{\alpha}'_n(t)$ défini par

$$\hat{\alpha}'_n(t) = \frac{\hat{f}'_n(t)}{1 - F_n(t)} + \frac{\hat{f}_n(t)^2}{(1 - F_n(t))^2},$$

avec \hat{f}'_n donné par (3.1) pour $m = 1$.

D'après le Théorème 3.2.2, posons la suite (h_n) telle que $nh_n^3 \rightarrow +\infty$ et $nh_n^5 \rightarrow 0$. Le Lemme 4.2.1 entraîne que

$$\sqrt{nh_n^3} \left(\begin{pmatrix} \hat{f}'_n(t) \\ \hat{f}_n(t) \\ 1 - F_n(t) \end{pmatrix} - \begin{pmatrix} f'(t) \\ f(t) \\ 1 - F(t) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \begin{pmatrix} T_1 \\ 0 \\ 0 \end{pmatrix},$$

où $T_1 \stackrel{\mathcal{L}}{=} N(0, b_1(t))$ avec $b_1(t)$ donné par (3.5). La méthode Delta F.3 avec le nouveau choix $\phi(u, v, w) = \frac{u}{w} + \frac{v^2}{w^2}$ peut ensuite être appliquée pour obtenir la loi asymptotique de l'estimateur $\hat{\alpha}'_n(t)$:

$$\sqrt{nh_n^3} (\hat{\alpha}'_n(t) - \alpha'(t)) \xrightarrow{\mathcal{L}} N\left(0, \frac{b_1(t)}{(1 - F(t))^2}\right). \quad (4.6)$$

Analysons enfin la dérivée seconde du taux de hasard

$$\alpha''(t) = \frac{f''(t)(1 - F(t)) + f'(t)f(t)}{(1 - F(t))^2} + 2 \frac{f'(t)f(t)(1 - F(t)) + f(t)^3}{(1 - F(t))^3}$$

dont l'estimateur empirique $\hat{\alpha}''_n(t)$ s'écrit

$$\hat{\alpha}''_n(t) = \frac{\hat{f}''_n(t)(1 - F_n(t)) + \hat{f}'_n(t)\hat{f}_n(t)}{(1 - F_n(t))^2} + 2 \frac{\hat{f}'_n(t)\hat{f}_n(t)(1 - F_n(t)) + \hat{f}_n(t)^3}{(1 - F_n(t))^3},$$

où \hat{f}''_n correspond à l'estimateur à noyau de la densité f donné par (3.1) pour $m = 2$.

Introduisons la suite (h_n) telle que $nh_n^5 \rightarrow +\infty$ et $nh_n^7 \rightarrow 0$ en vertu du Théorème 3.2.2. D'après le Lemme 4.2.1, nous avons

$$\sqrt{nh_n^5} \left(\begin{pmatrix} \hat{f}''_n(t) \\ \hat{f}'_n(t) \\ \hat{f}_n(t) \\ 1 - F_n(t) \end{pmatrix} - \begin{pmatrix} f''(t) \\ f'(t) \\ f(t) \\ 1 - F(t) \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \begin{pmatrix} T_1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

où $T_1 \stackrel{\mathcal{L}}{=} N(0, b_2(t))$ avec $b_2(t)$ donné par (3.5). La méthode Delta F.3 avec le choix $\phi(u, v, w, x) = \frac{ux+vw}{x^2} + 2 \frac{vwx+w^3}{x^3}$ permet de déduire que :

$$\sqrt{nh_n^5} (\hat{\alpha}''_n(t) - \alpha''(t)) \xrightarrow{\mathcal{L}} N\left(0, \frac{b_2(t)}{(1 - F(t))^2}\right). \quad (4.7)$$

La procédure décrite pour obtenir la normalité asymptotique de $\hat{\alpha}_n(t)$, $\hat{\alpha}'_n(t)$ et $\hat{\alpha}''_n(t)$ semble se généraliser aux estimateurs empiriques des dérivées d'ordre supérieur à 2 du taux de hasard. Finalement, si $m \in \mathbb{N}$ et $\hat{\alpha}_n^{(m)}(t)$ représente la version empirique de la

dérivée m^e du taux de hasard $\alpha^{(m)}(t)$ alors grâce au Théorème 3.2.2, au Lemme 4.2.1 et à la méthode Delta F.3, il vient que :

$$\sqrt{nh_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right) \xrightarrow{\mathcal{L}} \text{N} \left(0, \frac{b_m(t)}{(1-F(t))^2} \right), \quad (4.8)$$

avec $b_m(t)$ donné par (3.5), à condition que $nh_n^{2m+1} \rightarrow +\infty$ et $nh_n^{2m+3} \rightarrow 0$.

4.4 Estimation du taux de hasard et de ses dérivées successives dans le modèle à intensité multiplicative

4.4.1 Cadre d'étude

Le taux de hasard n'est pas exclusivement défini pour une variable aléatoire (dont la fonction de répartition est de support $]0, +\infty[$), mais également pour un processus de comptage dont l'intensité prend la forme multiplicative. Nous nous plaçons désormais dans le modèle multiplicatif de Nelson-Aalen, une situation plus générale où les observations peuvent être dépendantes et censurées.

Nous présentons brièvement (pour plus de détails, voir [11]) les notions nécessaires à la compréhension des résultats obtenus. Soit $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espace probabilisé satisfaisant les conditions habituelles de continuité à droite et de complétion de \mathbf{F} . On note \mathbb{E} l'espérance par rapport à la probabilité \mathbb{P} et t représente un réel positif. Soit N un processus de comptage, de compensateur Λ tel que

$$\Lambda(t) = \int_0^t \lambda(s) ds,$$

où le processus intensité λ est de forme multiplicative, c'est-à-dire :

$$\lambda(t) = Y(t) \alpha(t),$$

avec Y un processus prévisible et α une fonction déterministe appelée *taux de hasard* ou *taux de risque*. Supposons que $\mathbb{E} \Lambda(t) < +\infty$ pour tout $t \geq 0$. Le processus de comptage compensé, noté M , est défini par $M(0) = 0$ et grâce à la décomposition de Doob-Meyer, on a

$$M(t) = N(t) - \Lambda(t),$$

d'où

$$dM(t) = dN(t) - Y(t) \alpha(t) dt. \quad (4.9)$$

Le processus M est une martingale (voir définition dans l'Annexe D) de carré intégrable, c'est-à-dire

$$\mathbb{E} M(t)^2 < +\infty,$$

d'espérance nulle ($\mathbb{E} M(t) = \mathbb{E} M(0) = 0$) et de variance

$$\text{Var} M(t) = \mathbb{E} M(t)^2 = \mathbb{E} \int_0^t d\langle M, M \rangle(s),$$

où le processus de variation prévisible est donné par

$$d\langle M, M \rangle(t) = Y(t) \alpha(t) dt. \quad (4.10)$$

Introduisons l'intégrale stochastique

$$(H \cdot M)(t) := \int_0^t H(s) dM(s), \quad (4.11)$$

où H est un processus prévisible tel que pour tout $t \geq 0$,

$$\mathbb{E}(H \cdot M)(t)^2 < +\infty.$$

Alors $(H \cdot M)$ est une martingale de carré intégrable, d'espérance nulle et de variance

$$\text{Var},(H \cdot M)(t) = \mathbb{E}(H \cdot M)(t)^2 = \mathbb{E} \int_0^t d\langle (H \cdot M), (H \cdot M) \rangle(s), \quad (4.12)$$

où le processus de variation prévisible s'écrit

$$d\langle (H \cdot M), (H \cdot M) \rangle(t) = H^2(t) d\langle M, M \rangle(t). \quad (4.13)$$

Notons J le processus tel que

$$J(t) := \mathbf{1}_{[Y(t) > 0]}.$$

On en déduit que

$$dN(t) = J(t) dN(t). \quad (4.14)$$

presque sûrement. Ceci semble évident par intuition car les sauts du processus N ont seulement lieu aux temps où le processus Y est positif. Voici plus bas un argument rigoureux.

Tout d'abord, l'expression (4.9) implique que :

$$M(t) = N(t) - \int_0^t Y(s) \alpha(s) ds = \int_0^t J(s) \left[dN(s) - Y(s) \alpha(s) ds \right]$$

et ensuite, puisque J est un processus prévisible et borné,

$$\int_0^t J(s) \left[dN(s) - Y(s) \alpha(s) ds \right]$$

est une martingale. Ainsi, leur différence est une martingale et comme $Y(t) = J(t) Y(t)$,

$$\int_0^t (1 - J(s)) dN(s)$$

est une martingale. Étant de variation bornée et strictement croissante, elle doit donc être nulle, d'où (4.14).

En combinant (4.9) et (4.14), nous écrivons

$$\int_0^t dN(s) = \int_0^t J(s) dM(s) + \int_0^t J(s) Y(s) \alpha(s) ds,$$

puis en convenant que $\frac{J(s)}{Y(s)}$ s'annule lorsque $Y(s) = 0$, nous pouvons diviser par $Y(s)$

sous les intégrales et obtenir :

$$\int_0^t \frac{dN(s)}{Y(s)} = \int_0^t \frac{J(s)}{Y(s)} dM(s) + \int_0^t J(s) \alpha(s) ds. \quad (4.15)$$

Remarquons que le premier terme du membre de droite est de la forme (4.11), avec $\frac{J}{Y}$ un processus prévisible borné et M une martingale, donc son espérance est nulle. Il résulte que

$$\int_0^t \frac{dN(s)}{Y(s)}, \quad (4.16)$$

appelé l'*estimateur de Nelson-Aalen* de la fonction intensité cumulée $A(t) := \int_0^t \alpha(s) ds$, est un estimateur sans biais de

$$\mathbb{E} \int_0^t J(s) \alpha(s) ds.$$

Deux estimateurs lisses des dérivées du taux de hasard α sont proposés dans ce chapitre : l'estimateur à noyau classique dans la Section 4.4.3 et l'estimateur polynomial local dans la Section 4.4.4. Leurs propriétés asymptotiques sont examinées dans le cadre d'une suite de modèles multiplicatifs, où les objets appartenant au n^e modèle sont indexés par n , c'est-à-dire $N_n, Y_n, J_n, M_n, \dots$. Afin de les énoncer clairement, il est nécessaire de fixer plusieurs hypothèses.

4.4.2 Hypothèses et notations

Dans la suite du chapitre, t est un réel positif, et m et r sont deux entiers tels que $0 \leq m \leq r$.

Hypothèse (H1)

α est $(m + 1)$ fois continûment différentiable en t .

Hypothèse (H1')

α est $(r + 1)$ fois continûment différentiable en t .

Hypothèse (H2)

(h_n) est une suite décroissante de nombres réels strictement positifs telle que $h_n \rightarrow 0$ lorsque $n \rightarrow +\infty$.

Hypothèse (H3)

$\sup_{s \in [t-c, t+c]} \mathbb{P}[Y_n(s) = 0] = o(h_n^{m+1})$ pour un certain $c > 0$.

Hypothèse (H3')

$\sup_{s \in [t-c, t+c]} \mathbb{P}[Y_n(s) = 0] = o(h_n^{r+1})$ pour un certain $c > 0$.

Hypothèse (H3'')

Il existe une suite croissante (a_n) de nombres strictement positifs telle que $a_n \rightarrow +\infty$ lorsque $n \rightarrow +\infty$ et vérifiant $\sup_{s \in [t-c, t+c]} \mathbb{P}[Y_n(s) = 0] = o(a_n^{-2})$ pour un certain $c > 0$.

Les hypothèses (H3), (H3') et (H3'') sont satisfaites dans la situation de durées de vie i.i.d. non censurées pour $Y_n(s) \sim \mathcal{B}(n, \bar{F}(s))$ avec $\bar{F}(s) = 1 - F(s)$. On a :

$$\sup_{s \in [t-c, t+c]} \mathbb{P}[Y_n(s) = 0] \leq F(t+c)^n,$$

où le membre de droite converge rapidement et de manière exponentielle vers 0.

Hypothèse (H4)

Il existe une suite croissante (a_n) de nombres strictement positifs telle que $a_n \rightarrow +\infty$ lorsque $n \rightarrow +\infty$ et une fonction y positive et continue en t vérifiant

$$\sup_{s \in [t-c, t+c]} \left| \mathbb{E} \left[\frac{a_n^2 J_n(s)}{Y_n(s)} \right] - \frac{1}{y(s)} \right| \xrightarrow{\mathbb{P}} 0$$

lorsque $n \rightarrow +\infty$ pour un certain $c > 0$.

Dans la situation de durées de vie i.i.d. non censurées, définissons la variable aléatoire $Y_n(s)$ telle que $Y_n(s) \sim \mathcal{B}(n, \bar{F}(s))$ avec $\bar{F}(s) = 1 - F(s)$. Alors une condition suffisante pour l'hypothèse (H4) est : $\mathbb{E} \left[\frac{J_n(s)}{Y_n(s)} \right] = O\left(\frac{1}{n}\right)$. En effet, nous avons

$$\begin{aligned} \mathbb{E} \left[\frac{J_n(s)}{Y_n(s)} \right] &= \mathbb{E} \left[\mathbf{1}_{\{0 < Y_n(s) \leq \frac{n\bar{F}(s)}{2}\}} \frac{1}{Y_n(s)} \right] + \mathbb{E} \left[\mathbf{1}_{\{Y_n(s) > \frac{n\bar{F}(s)}{2}\}} \frac{1}{Y_n(s)} \right] \\ &\leq \mathbb{E} \left[\mathbf{1}_{\{0 < Y_n(s) \leq \frac{n\bar{F}(s)}{2}\}} \right] + \frac{2}{n\bar{F}(s)} \\ &\leq \mathbb{P} \left[\left| \frac{Y_n(s)}{n} - \bar{F}(s) \right| \geq \frac{\bar{F}(s)}{2} \right] + \frac{2}{n\bar{F}(s)} \\ &\leq \frac{\bar{F}(s)F(s)}{n\left(\frac{\bar{F}(s)}{2}\right)^2} + \frac{2}{n\bar{F}(s)} \\ &\leq \frac{4\bar{F}(t-c)F(t+c)}{n\bar{F}(t+c)^2} + \frac{2}{n\bar{F}(t+c)} \\ &= O\left(\frac{1}{n}\right). \end{aligned}$$

Hypothèse (H5)

Il existe une suite croissante (a_n) de nombres strictement positifs telle que $a_n \rightarrow +\infty$ et une fonction y positive et continue en t vérifiant

$$\sup_{s \in [t-c, t+c]} \left| a_n^{-2} Y_n(s) - y(s) \right| \xrightarrow{\mathbb{P}} 0$$

lorsque $n \rightarrow +\infty$ pour un certain $c > 0$.

Nous devons insister sur le fait que la suite (a_n) et la fonction y sont respectivement identiques pour les hypothèses (H3''), (H4) et (H5) puis pour les hypothèses (H4) et (H5).

Par ailleurs, dans la situation de durées de vie i.i.d. non censurées, si $Y_n(s) \sim \mathcal{B}(n, \bar{F}(s))$ avec $\bar{F}(s) = 1 - F(s)$, alors $Y_n(s)$ est une suite de Bernoulli indépendantes. D'après la loi faible des grands nombres, on a

$$\frac{Y_n(s)}{n} \xrightarrow{\mathbb{P}} \bar{F}(s),$$

et en appliquant la fonction inverse, continue sur \mathbb{R}^* , nous en déduisons que

$$\frac{n}{Y_n(s)} \xrightarrow{\mathbb{P}} \frac{1}{\bar{F}(s)}.$$

De plus,

$$J_n(s) = \mathbf{1}_{[Y_n(s) > 0]} \xrightarrow{\mathbb{P}} 1$$

car nous avons démontré un peu plus haut que

$$\sup_{s \in [t-c, t+c]} \mathbb{P}[Y_n(s) = 0] \rightarrow 0$$

était vrai dans ce cadre.

On conclut grâce aux règles de Cramér :

$$\frac{nJ_n(s)}{Y_n(s)} \xrightarrow{\mathbb{P}} \frac{1}{\bar{F}(s)},$$

d'où pour s dans un voisinage de t ,

$$\frac{nJ_n(s)}{Y_n(s)} \xrightarrow{\mathbb{P}} \frac{1}{\bar{F}(t)}.$$

Les hypothèses (H3), (H3'), (H4) et (H5) sont très faibles car elles peuvent être également vérifiées dans d'autres situations, avec des données pas nécessairement i.i.d. ni non censurées. Par exemple, dans un modèle à risques concurrents, comme cela est expliqué à travers l'Exemple IV.2.2 dans [11].

4.4.3 Estimation à noyau classique du taux de hasard et de ses dérivées successives

Nous nous intéressons dans cette section à l'estimation du taux de hasard et de ses dérivées successives par un noyau classique. Pour cela, nous avons besoin d'une fonction de répartition K de support $[-1, 1]$ de densité $K' = k$ satisfaisant les conditions suivantes :

Hypothèse (H6)

(i) k est m fois continûment différentiable ,

(ii) Le support de $k^{(m)}$ est $[-1, 1]$,

(iii)

$$\sup_{u \in [-1, 1]} |k^{(m)}(u)| < +\infty .$$

Définissons l'estimateur de Nelson-Aalen lissé à l'aide d'une convolution classique par

$$\hat{A}_n(t) = \int K\left(\frac{t-s}{h_n}\right) \frac{dN_n(s)}{Y_n(s)} ,$$

alors grâce à l'hypothèse (H6(i)), nous pouvons différentier $(m+1)$ fois l'estimateur $\hat{A}_n(t)$ et obtenir l'expression de l'estimateur de $\hat{\alpha}_n^{(m)}(t)$, la dérivée m^e de α , au point t :

$$\hat{\alpha}_n^{(m)}(t) = \frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) \frac{dN_n(s)}{Y_n(s)} . \quad (4.17)$$

En posant

$$\hat{M}_n^{(m)}(t) := \frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) \frac{J_n(s)}{Y_n(s)} dM_n(s) , \quad (4.18)$$

qui est une intégrale d'un processus prévisible par rapport à une martingale M_n donc un "terme d'erreur" et en utilisant (4.15), on écrit

$$\begin{aligned} \hat{\alpha}_n^{(m)}(t) &= \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) J_n(s) \alpha(s) ds \\ &= \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) [J_n(s) - 1] \alpha(s) ds \\ &+ \frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) \alpha(s) ds . \end{aligned}$$

En appliquant le changement de variable $u = \frac{s-t}{h_n}$ sur les deux dernières lignes puis en

utilisant les propriétés de la convolution sur la troisième ligne, cela entraîne que :

$$\begin{aligned}\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \\ &+ \int_{-1}^1 k(-u) \left(\alpha^{(m)}(t + h_n u) - \alpha^{(m)}(t) \right) du.\end{aligned}$$

Nous verrons que la décomposition précédente est précieuse pour l'étude des propriétés asymptotiques de notre estimateur $\hat{\alpha}_n^{(m)}(t)$. Nous commençons immédiatement par un résultat sur le biais de l'estimateur $\hat{\alpha}_n^{(m)}(t)$.

Lemme 4.4.1. *Soient t un réel positif et m un entier. Alors sous les hypothèses (H1)-(H3)-(H6), on a*

$$\mathbb{E} \hat{\alpha}_n^{(m)}(t) = \alpha^{(m)}(t) + O(h_n). \quad (4.19)$$

Nous poursuivons par un lemme décrivant la variance exacte asymptotique de l'estimateur $\hat{\alpha}_n^{(m)}(t)$.

Lemme 4.4.2. *Soient $t \geq 0$ et m un entier. Si les hypothèses (H1)-(H3'')-(H4)-(H6) sont satisfaites alors :*

$$\text{Var} \hat{\alpha}_n^{(m)}(t) = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \quad (4.20)$$

Les conditions requises pour les convergences en moyenne quadratique et en probabilité de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ sont indiquées dans le théorème suivant et découlent des deux lemmes précédents.

Théorème 4.4.1. *Supposons que t est un réel positif, que m un entier et que les hypothèses (H1)-(H3)-(H3'')-(H4)-(H6) sont vérifiées. De plus, si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (4.21)$$

suffisamment lentement tel que

$$a_n^2 h_n^{2m+1} \rightarrow +\infty, \quad (4.22)$$

alors l'estimateur $\hat{\alpha}_n^{(m)}(t)$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right]^2 \rightarrow 0, \quad (4.23)$$

donc également convergent en probabilité,

$$\hat{\alpha}_n^{(m)}(t) \xrightarrow{\mathbb{P}} \alpha^{(m)}(t). \quad (4.24)$$

Remarque :

Nous pouvons rechercher des conditions plus souples pour la convergence en probabilité de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ et obtenir un théorème semblable au Théorème IV.2.1 dans [11].

Nous exposons enfin la normalité asymptotique de l'estimateur $\hat{\alpha}_n^{(m)}(t)$.

Théorème 4.4.2. *Soient $t \geq 0$ et m un entier. Si les hypothèses (H1)-(H3)-(H4)-(H5)-(H6) sont satisfaites et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$a_n^2 h_n \rightarrow +\infty, \quad (4.25)$$

et suffisamment vite tel que

$$a_n^2 h_n^{2m+3} \rightarrow 0, \quad (4.26)$$

alors

$$\sqrt{a_n^2 h_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right)$$

converge vers une loi normale centrée et de variance

$$b_m(t) = \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du. \quad (4.27)$$

Les conditions de consistance (4.21) et (4.22) sont convenables : la première est identique pour tout m , alors que la seconde devient moins souple quand m augmente. Les conditions (4.25) et (4.26) pour la normalité asymptotique sont moins intuitives : la première est la même pour tout m , alors que la deuxième devient plus souple quand m augmente. Ces observations servent à illustrer que les deux modes de convergence – en moyenne quadratique et en loi – sont très différents. Nous retrouvons les conditions du Chapitre 3 pour l'estimation de $f^{(m)}$ dans la situation i.i.d. avec $a_n^2 = n$. Par ailleurs, les conditions et la loi asymptotique coïncident avec celles de l'estimation du taux de hasard et de sa dérivée première dans la situation i.i.d. avec $a_n^2 = n$ et $y = 1 - F$.

Les conditions de consistance et de normalité asymptotique dépendent de m donc si l'on s'intéresse à l'estimation conjointe des dérivées, on ne peut appliquer la même fenêtre pour tous les ordres de dérivation.

Les hypothèses posées dans cette section, pourtant peu nombreuses, vont pouvoir être assouplies grâce à l'estimation polynomiale locale. Nous traitons cela dans la section suivante.

4.4.4 Estimation polynomiale locale du taux de hasard et de ses dérivées successives

Nous utilisons désormais un noyau d'ordre supérieur afin d'estimer le taux de hasard et ses dérivées successives par des polynômes locaux. Si K_0 représente une densité de probabilité sur $[-1, 1]$ et si le noyau reproduisant de l'espace des polynômes de degré inférieur à r est noté $\mathcal{K}_r^{(0)}(\cdot, \cdot)$, alors le noyau d'ordre $(m, r + 1)$ a pour expression

$$K_r^{(m)}(x) = \frac{\partial^m \mathcal{K}_r^{(0)}(x, v)}{\partial v^m} \Big|_{v=0} \cdot K_0(x)$$

et satisfait les conditions suivantes :

Hypothèse (H6')

(i) Le support de $K_r^{(m)}$ est $[-1, 1]$,

(ii)

$$\int_{-1}^1 u^i K_r^{(m)}(u) du = \begin{cases} m!, & i = 0, \\ 0, & i \neq m \text{ et } i \in \{1, \dots, r\}, \\ C_{r+1}, & i > r, \end{cases}$$

avec C_{r+1} une constante finie ,

(iii)

$$\sup_{u \in [-1, 1]} |K_r^{(m)}(u)| < +\infty .$$

L'estimateur de Nelson-Aalen lissé à l'aide de l'approximation polynomiale locale s'écrit :

$$\hat{A}_n(t) = \int K_0\left(\frac{s-t}{h_n}\right) \frac{dN_n(s)}{Y_n(s)}$$

et l'estimateur polynomial local noté $\hat{\alpha}_n^{(m)}(t)$ de la dérivée m^e de α est défini par :

$$\hat{\alpha}_n^{(m)}(x) = \frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{s-t}{h_n}\right) \frac{dN_n(s)}{Y_n(s)}. \quad (4.28)$$

L'estimation polynomiale locale du taux de hasard et de ses dérivées successives pour des processus de comptage a été considérée par Chen, Huggins, Yip, Lam ([42]) mais ils emploient des noyaux dont l'expression mathématique est différente.

Introduisons

$$\hat{M}^{(m)}(t) := \frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{s-t}{h_n}\right) \frac{J_n(s)}{Y_n(s)} dM_n(s), \quad (4.29)$$

qui est une intégrale d'un processus prévisible par rapport à une martingale M_n donc un "terme d'erreur", alors grâce à (4.15), nous avons :

$$\begin{aligned} \hat{\alpha}_n^{(m)}(t) &= \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{s-t}{h_n}\right) J_n(s) \alpha(s) ds \\ &= \hat{M}^{(m)}(t) \\ &+ \frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{s-t}{h_n}\right) [J_n(s) - 1] \alpha(s) ds \\ &+ \frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{s-t}{h_n}\right) \alpha(s) ds, \end{aligned}$$

puis à l'aide du changement de variable $u = \frac{s-t}{h_n}$ sur les deux dernières lignes, nous

déduisons que

$$\begin{aligned} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \\ &+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t). \end{aligned}$$

Une telle décomposition nous aidera à décrire le comportement asymptotique de notre estimateur $\hat{\alpha}_n^{(m)}(t)$. Le lemme suivant donne le biais de l'estimateur $\hat{\alpha}_n^{(m)}(t)$.

Lemme 4.4.3. *Soient t un réel positif et m un entier compris entre 0 et r . Alors sous les hypothèses (H1')-(H3')-(H6'), on a*

$$\mathbb{E} \hat{\alpha}_n^{(m)}(t) = \alpha^{(m)}(t) + O(h_n^{r+1-m}). \quad (4.30)$$

Le Lemme 4.4.3 illustre l'avantage de l'estimation polynomiale locale par rapport à l'estimation à noyau classique. Les noyaux $K_r^{(m)}$ impliqués dans l'expression des estimateurs $\hat{\alpha}_n^{(m)}(t)$ ont pour effet de faire baisser l'ordre du biais. Puisque $m \leq r$, l'ordre du biais dans le Lemme 4.4.3 est inférieur à celui mentionné dans le Lemme 4.4.1.

Nous poursuivons l'étude de notre estimateur $\hat{\alpha}_n^{(m)}(t)$ avec deux résultats similaires à ceux de la section précédente. Nous fournissons tout d'abord la variance exacte asymptotique de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ qui nous rappelle celle du Lemme 4.4.2.

Lemme 4.4.4. *Soient $t \geq 0$ et m un entier compris entre 0 et r . Si les hypothèses (H1')-(H3'')-(H4)-(H6') sont satisfaites alors :*

$$\text{Var} \hat{\alpha}_n^{(m)}(t) = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \quad (4.31)$$

On remarque que (4.31) a la même forme que (4.20). L'unique différence vient du noyau $K_r^{(m)}$ qui se trouve à la place du noyau $k^{(m)}$.

Nous présentons maintenant la convergence dans \mathbb{L}^2 et en probabilité de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ obtenue sous les mêmes conditions que celles du Théorème 4.4.1.

Théorème 4.4.3. *Supposons que t est un réel positif, que m un entier compris entre 0 et r et que les hypothèses (H1')-(H3')-(H3'')-(H4)-(H6') sont vérifiées. De plus, si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (4.32)$$

suffisamment lentement tel que

$$a_n^2 h_n^{2m+1} \rightarrow +\infty, \quad (4.33)$$

alors l'estimateur $\hat{\alpha}_n^{(m)}(t)$ est convergent en moyenne quadratique,

$$\mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right]^2 \rightarrow 0, \quad (4.34)$$

donc également convergent en probabilité,

$$\hat{\alpha}_n^{(m)}(t) \xrightarrow{\mathbb{P}} \alpha^{(m)}(t). \quad (4.35)$$

Remarque :

Nous pouvons rechercher des conditions plus souples pour la convergence en probabilité de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ et obtenir un théorème semblable au Théorème IV.2.1 dans [11].

Nous établissons finalement un résultat sur la normalité asymptotique de l'estimateur $\hat{\alpha}_n^{(m)}(t)$.

Théorème 4.4.4. *Soient $t \geq 0$ et m un entier compris entre 0 et r . Si les hypothèses (H1')-(H3')-(H4)-(H5)-(H6') sont satisfaites et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que*

$$a_n^2 h_n \rightarrow +\infty, \quad (4.36)$$

et suffisamment vite tel que

$$a_n^2 h_n^{2r+3} \rightarrow 0, \quad (4.37)$$

alors

$$\sqrt{a_n^2 h_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right)$$

converge vers une loi normale centrée et de variance

$$b_m(t) = \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du. \quad (4.38)$$

Les conditions de consistance (4.32) et (4.33) sont les mêmes que (4.21) et (4.22) obtenues dans le cadre de l'estimation à noyau classique. La première ne dépend pas de m alors que la deuxième est plus contraignante lorsque m augmente. La première condition de normalité asymptotique (4.36), indépendante de m , est analogue à (4.25). Quant à la seconde condition (4.37), elle s'exprime uniquement en fonction de r et est plus souple que la condition (4.26) car $m \leq r$. Remarquons que la condition (4.37) est différente de (2.24) donc nous pouvons maintenant abandonner l'idée que cette dernière condition serait commune à toutes les fonctionnelles de la fonction de répartition pour obtenir la normalité asymptotique de leurs dérivées successives.

Pour terminer, nous constatons que l'expression des variances (4.27) et (4.38) ne diffèrent que par le noyau employé : $k^{(m)}$ dans la première et $K_r^{(m)}$ dans la seconde.

Nous achevons l'estimation du taux de hasard et de ses dérivées successives en écrivant quelques mots sur la fenêtre (h_n). Pour obtenir la consistance et la normalité asymptotique des estimateurs, des hypothèses sur h_n ont été fixées. Dès lors, nous sommes en mesure de nous demander s'il est possible de trouver une fenêtre (h_n) satisfaisant toutes ces conditions. C'est l'objet de la section suivante.

4.4.5 Choix de fenêtre

Soient r un entier fixé et m un entier tel que $0 \leq m \leq r$.

Convolution :

Obtenir la consistance des estimateurs du taux de hasard et de ses dérivées successives n'est possible qu'en supposant les hypothèses (H2) et (4.22) pour tout m . La normalité asymptotique des estimateurs et la construction d'intervalles de confiance nécessitent quant à eux, de poser les hypothèses supplémentaires (4.25) et (4.26) pour tout m .

Nous constatons que (4.26) est de plus en plus souple lorsque m augmente, contrairement à (4.22). Par ailleurs, la condition (4.25) n'est pas à considérer puisqu'elle découle de (4.22). Les conditions (4.22) et (4.26) s'expriment toutes les deux en fonction de m , la fenêtre (h_n) doit donc absolument dépendre de l'ordre de dérivation m .

Fixons de manière arbitraire un réel h strictement positif et définissons ensuite pour tout m :

$$h_n = a_n^{-\frac{1}{m+1}} h. \quad (4.39)$$

Les hypothèses décrites précédemment sont validées avec le choix de fenêtre (h_n).

Approximation polynomiale locale :

Les conditions (H2) et (4.33) pour tout m entraînent la consistance des estimateurs du taux de hasard et de ses dérivées successives. Poser en plus (4.36) et (4.37) pour tout m , permet de conclure à la normalité asymptotique de ces estimateurs et d'obtenir des intervalles de confiance.

Remarquons que l'hypothèse (4.33) entraîne l'hypothèse (4.36), donc la fenêtre (\tilde{h}_n) à déterminer doit seulement vérifier (H2), (4.33) et (4.37). De plus, lorsque m augmente, (4.33) devient de plus en plus contraignante (pendant que (4.37) le devient de moins en moins) donc il suffit de supposer $a_n^2 h_n^{2r+1} \rightarrow +\infty$ pour satisfaire (4.33) pour tout m , puisque $m \leq r$. Par suite, nous pouvons choisir une fenêtre indépendante de l'ordre de dérivation m .

4.5 Un mot sur les effets de bord

Nielsen et Tanggaard ([125]) ont mentionné que les estimateurs décrits dans [134] étaient sujets aux effets de bord. La méthode développée dans le Section 3.4 pour remédier aux effets de bord des estimateurs à noyau de la densité et de ses dérivées, se généralise à l'analyse de survie lorsque les données sont censurées à droite. En remplaçant f par α puis n par a_n^2 , nous obtenons la convergence et la normalité asymptotique des estimateurs corrigés des effets de bord du taux de hasard et de ses dérivées successives.

4.6 Preuves

Dans les preuves de ce chapitre, t et m désignent respectivement un réel positif et un entier compris entre 0 et r , avec $r \in \mathbb{N}$.

De plus, nous avons besoin de travailler avec une martingale définie sur l'intervalle $[-1, 1]$, construite à l'aide d'un changement de variable. La fonction utilisée dans le changement de variable satisfait plusieurs conditions que le Lemme suivant explicite.

Lemme 4.6.1. *Soient I un intervalle de la forme $[a, b]$ (ou $[a, b[$), a et b étant deux réels positifs avec b pouvant être égal à $+\infty$, et J un intervalle de la forme $[c, d]$ (ou $]c, d]$), avec $c \in \mathbb{R}_-$ pouvant être égal à $-\infty$ et $d \in \mathbb{R}_+$. Nommons $(N(t))_{t \in I}$ un processus de comptage d'intensité $(\lambda(t))_{t \in I}$ par rapport à sa filtration naturelle $\mathbf{F}^N = (\mathcal{F}^N(t))_{t \in I}$ et $\phi : I \rightarrow J$ une fonction strictement croissante et différentiable. Notons sa fonction inverse ϕ^{-1} vérifiant $\phi^{-1}(c) = a$ et $\phi^{-1}(v) \rightarrow +\infty$ lorsque $v \rightarrow +\infty$. Alors $(\tilde{N}(v))_{v \in J}$, défini par $\tilde{N}(v) = N(\phi^{-1}(v))$, est un processus de comptage d'intensité $(\tilde{\lambda}(v))_{v \in J}$ par rapport à sa filtration naturelle $\mathbf{F}^{\tilde{N}} = (\mathcal{F}^{\tilde{N}}(v))_{v \in J}$, où $\mathcal{F}_v^{\tilde{N}} = \mathcal{F}_{\phi^{-1}(v)}^N$ et*

$$\tilde{\lambda}(v) = \lambda(\phi^{-1}(v)) \cdot \frac{\partial \phi^{-1}(v)}{\partial v}. \quad (4.40)$$

L'unique résultat non trivial (4.40) découle du fait que $\tilde{N}(\cdot) - \int_c^{\cdot} \tilde{\lambda}(y) dy$ est une $\mathbf{F}^{\tilde{N}}$ -martingale. Soient u et v deux réels appartenant à l'intervalle J tels que $u \leq v$ alors en appliquant deux fois le changement de variable $y = \phi^{-1}(w)$, nous obtenons :

$$\begin{aligned} \mathbb{E} \left[\tilde{N}(v) - \int_c^v \tilde{\lambda}(w) dw \middle| \mathcal{F}_u^{\tilde{N}} \right] &= \mathbb{E} \left[N(\phi^{-1}(v)) - \int_c^v \lambda(\phi^{-1}(w)) \cdot \frac{\partial \phi^{-1}(w)}{\partial w} dw \middle| \mathcal{F}_u^{\tilde{N}} \right] \\ &= \mathbb{E} \left[N(\phi^{-1}(v)) - \int_a^{\phi^{-1}(v)} \lambda(y) dy \middle| \mathcal{F}_{\phi^{-1}(u)}^N \right] \\ &= N(\phi^{-1}(u)) - \int_a^{\phi^{-1}(u)} \lambda(y) dy \\ &= \tilde{N}(u) - \int_c^u \tilde{\lambda}(w) dw. \end{aligned}$$

Remarque :

On rencontre la situation fréquente $a = c = 0$. Cependant, à la page 48 dans [11], il est indiqué que l'on peut choisir une valeur arbitraire au lieu de 0.

Soient $t \geq 0$ et $h_n > 0$. Posons la fonction ϕ telle que

$$\begin{aligned} \phi : [t - h_n, t + h_n] &\rightarrow [-1, 1] \\ s &\mapsto \frac{s - t}{h_n}. \end{aligned}$$

Cette fonction entre dans le contexte du Lemme 4.6.1 et sa fonction inverse ϕ^{-1} s'écrit :

$$\begin{aligned} \phi^{-1} : [-1, 1] &\rightarrow [t - h_n, t + h_n] \\ u &\mapsto t + h_n u. \end{aligned}$$

Donc pour tout $u \in [-1, 1]$,

$$\tilde{N}_n(u) = N_n(t + h_n u)$$

est un processus de comptage d'intensité

$$\tilde{\lambda}_n(u) = Y_n(t + h_n u) \alpha(t + h_n u) h_n ,$$

d'après le Lemme 4.6.1.

Nous adoptons désormais la notation \tilde{M}_n pour définir la martingale associée au processus de comptage \tilde{N}_n . Ainsi, dans l'ensemble des preuves :

$$\tilde{M}_n(u) := N_n(t + h_n u) - \int_{-1}^u Y_n(t + h_n w) \alpha(t + h_n w) h_n dw \quad (4.41)$$

est une martingale définie sur $[-1, 1]$.

4.6.1 Preuve du Lemme 4.4.1

Reprenons la décomposition de $\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t)$:

$$\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) = \hat{M}_n^{(m)}(t) \quad (4.42)$$

$$+ \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \quad (4.43)$$

$$+ \int_{-1}^1 k(-u) (\alpha^{(m)}(t + h_n u) - \alpha^{(m)}(t)) du . , \quad (4.44)$$

avec

$$\begin{aligned} \hat{M}_n^{(m)}(t) &= \frac{1}{h_n^{m+1}} \int_{-1}^1 k^{(m)}(-u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} [dN_n(t + h_n u) - Y_n(t + h_n u) \alpha(t + h_n u) du] \\ &= \frac{1}{h_n^{m+1}} \int_{-1}^1 k^{(m)}(-u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} d\tilde{M}_n(u) , \end{aligned} \quad (4.45)$$

d'après (4.18) et le changement de variable $u = \frac{s-t}{h_n}$.

Par la linéarité de l'espérance, nous avons :

$$\begin{aligned} \mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \mathbb{E} \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) \mathbb{E} [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \\ &+ \int_{-1}^1 k(-u) (\alpha^{(m)}(t + h_n u) - \alpha^{(m)}(t)) du , \end{aligned}$$

la seconde ligne résultant du théorème de Fubini. Les hypothèses (H1) et (H6(iii)) entraînent que la fonction α est continue en t donc est bornée dans un voisinage de t puis le noyau $k^{(m)}$ est borné sur $[-1, 1]$. Comme J_n est une indicatrice, les fonctions sous l'intégrale dans (4.43) sont bornées. Étant intégrées sur un intervalle fermé, elles sont intégrables.

Soit v un réel compris entre -1 et 1 puis introduisons :

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) := \int_{-1}^v \tilde{H}_n(u) d\tilde{M}_n(u),$$

avec $\tilde{H}_n(u) := \frac{1}{h_n^{m+1}} k^{(m)}(-u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)}$ pour tout $u \in [-1, v]$. Alors $(\tilde{H}_n \cdot \tilde{M}_n)$ est une martingale et on a :

$$\hat{M}_n^{(m)}(t) := (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1},$$

d'où :

$$\mathbb{E} \hat{M}_n^{(m)}(t) = \mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1} = 0.$$

La définition du processus J_n puis la formule de Taylor-Lagrange pour u' compris entre 0 et $h_n u$, avec $u \in [-1, 1]$ conduisent à :

$$\begin{aligned} \mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) \mathbb{P}[Y_n(t+h_n u) = 0] \alpha(t+h_n u) du \\ &\quad + \int_{-1}^1 k(-u) \left(\alpha^{(m)}(t+h_n u) - \alpha^{(m)}(t) \right) du \\ &= \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) \mathbb{P}[Y_n(t+h_n u) = 0] \alpha(t+h_n u) du \\ &\quad + \int_{-1}^1 k(-u) \alpha^{(m+1)}(t+u') h_n u du. \end{aligned}$$

Grâce à (H1), (H3) et (H6(iii)), il vient que

$$\frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) \mathbb{P}[Y_n(t+h_n u) = 0] \alpha(t+h_n u) du = o(h_n). \quad (4.46)$$

Par ailleurs, l'hypothèse (H1) implique que la dérivée $(m+1)^e$ de α est bornée sur un voisinage de t donc en utilisant (H6(iii)), on a :

$$\int_{-1}^1 k(-u) \alpha^{(m+1)}(t+u') h_n u du = O(h_n). \quad (4.47)$$

Nous déduisons de (4.46) et (4.47) le résultat suivant :

$$\mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) = o(h_n) + O(h_n) = O(h_n). \quad \blacksquare \quad (4.48)$$

4.6.2 Preuve du Lemme 4.4.2

Par définition :

$$\begin{aligned} \text{Var} \hat{\alpha}_n^{(m)}(t) &= \mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \mathbb{E} \hat{\alpha}_n^{(m)}(t) \right]^2 \\ &= \mathbb{E} \left[\frac{1}{h_n^{m+1}} \int k^{(m)} \left(\frac{t-s}{h_n} \right) \frac{dN_n(s)}{Y_n(s)} - \frac{1}{h_n^{m+1}} \int k^{(m)} \left(\frac{t-s}{h_n} \right) \mathbb{E} J_n(s) \alpha(s) ds \right]^2 \\ &= \mathbb{E} T_1^2 + 2E [T_1 T_2] + \mathbb{E} T_2^2, \end{aligned} \quad (4.49)$$

avec

$$\begin{aligned} T_1 &:= \frac{1}{h_n^{m+1}} \int_{-1}^1 k^{(m)}(-u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)} d\tilde{M}_n(u), \\ T_2 &:= \frac{1}{h_n^m} \int_{-1}^1 k^{(m)}(-u) (J_n(t+h_n u) - \mathbb{E} J_n(t+h_n u)) \alpha(t+h_n u) du, \end{aligned}$$

obtenus en retranchant et ajoutant la quantité $\frac{1}{h_n^{m+1}} \int k^{(m)}\left(\frac{t-s}{h_n}\right) J_n(s) \alpha(s) ds$ puis en effectuant le changement de variable $u = \frac{s-t}{h_n}$.

Calculons tout d'abord $\mathbb{E} T_1^2$. On constate que :

$$T_1 := (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1}, \quad (4.50)$$

avec \tilde{H}_n le processus défini dans la preuve précédente. Donc d'après (4.12), (4.13) puis le théorème de Fubini (intégration de fonctions bornées sur un intervalle fermé grâce à (H1), (H6(iii)) et le fait que $\frac{J}{Y}$ soit un processus borné), on a :

$$\begin{aligned} \mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2 &= \mathbb{E} \int_{-1}^v \frac{1}{h_n^{2(m+1)}} k^{(m)}(-u)^2 \frac{J_n^2(t+h_n u)}{Y_n^2(t+h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v \frac{1}{a_n^2 h_n^{2m+1}} k^{(m)}(-u)^2 \mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] \alpha(t+h_n u) du \\ &= \frac{1}{a_n^2 h_n^{2m+1}} \int_{-1}^v k^{(m)}(-u)^2 \left(\mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] - \frac{1}{y(t+h_n u)} \right) \alpha(t+h_n u) du \\ &\quad + \frac{1}{a_n^2 h_n^{2m+1}} \int_{-1}^v k^{(m)}(-u)^2 \frac{\alpha(t+h_n u)}{y(t+h_n u)} du. \end{aligned}$$

Un majorant de $\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2$ s'écrit sous la forme d'une somme :

$$\begin{aligned} &\int_{-1}^1 |k^{(m)}(-u)|^2 \left| \mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] - \frac{1}{y(t+h_n u)} \right| |\alpha(t+h_n u)| du \\ &\quad + \int_{-1}^1 |k^{(m)}(-u)|^2 \frac{|\alpha(t+h_n u)|}{y(t+h_n u)} du, \end{aligned}$$

de deux termes bornés grâce aux hypothèses (H1), (H4) et (H6(iii)), ce qui permet d'appliquer le théorème de convergence dominée. En réutilisant les hypothèses (H1) et (H4), on en déduit que :

$$\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2 = o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right) + \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^v k^{(m)}(-u)^2 du,$$

donc en particulier pour $v = 1$:

$$\mathbb{E} T_1^2 = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \quad (4.51)$$

Intéressons-nous ensuite à l'espérance de T_2^2 . L'inégalité de Jensen F.12 permet d'écrire :

$$\mathbb{E} T_2^2 \leq \mathbb{E} \left[\frac{1}{h_n^{2m}} \int_{-1}^1 k^{(m)}(-u)^2 (J_n(t+h_n u) - \mathbb{E} J_n(t+h_n u))^2 \alpha(t+h_n u)^2 du \right].$$

Les fonctions sous l'intégrale sont bornées grâce aux hypothèses (H1), (H6(iii)) et au fait que J_n soit un processus borné puis l'intégration a lieu sur un intervalle fermé donc nous pouvons appliquer le théorème de Fubini. Puisque :

$$\mathbb{E} [J_n(t+h_n u) - \mathbb{E} J_n(t+h_n u)]^2 = \text{Var} J_n(t+h_n u) \leq \mathbb{P}[Y_n(t+h_n u) = 0],$$

nous obtenons :

$$\begin{aligned} \mathbb{E} T_2^2 &\leq \frac{1}{h_n^{2m}} \int_{-1}^1 k^{(m)}(-u)^2 \mathbb{P}[Y_n(t+h_n u) = 0] \alpha(t+h_n(u))^2 du \\ &= o\left(\frac{1}{a_n^2 h_n^{2m}}\right). \end{aligned} \quad (4.52)$$

par les hypothèses (H1), (H3'') et (H6(iii)).

Déterminons enfin le terme $2\mathbb{E}[T_1 T_2]$. D'après l'inégalité de Cauchy-Schwartz F.13, on a :

$$\mathbb{E}[|T_1 T_2|] \leq (\mathbb{E} T_1^2)^{\frac{1}{2}} (\mathbb{E} T_2^2)^{\frac{1}{2}},$$

donc avec (4.51) et (4.52), nous écrivons :

$$\begin{aligned} 2\mathbb{E}[T_1 T_2] &\leq \left[O\left(\frac{1}{a_n^2 h_n^{2m+1}}\right) o\left(\frac{1}{a_n^2 h_n^{2m}}\right) \right]^{\frac{1}{2}} \\ &= \left[o\left(\frac{1}{a_n^4 h_n^{4m+1}}\right) \right]^{\frac{1}{2}} \\ &= o\left(\frac{1}{a_n^2 h_n^{\frac{4m+1}{2}}}\right). \end{aligned} \quad (4.53)$$

Grâce aux expressions (4.51), (4.52) et (4.53), on remarque que les termes $\mathbb{E} T_2^2$ et $2\mathbb{E}[T_1 T_2]$ sont négligeables par rapport à $\mathbb{E} T_1^2$ donc d'après la formule (4.49), il résulte que :

$$\text{Var} \hat{\alpha}_n^{(m)}(t) = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \blacksquare \quad (4.54)$$

4.6.3 Preuve du Théorème 4.4.1

Pour $m \geq 0$, l'erreur en moyenne quadratique est une somme d'un terme de variance et d'un terme de biais :

$$\begin{aligned} \mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right]^2 &= \text{Var} \hat{\alpha}_n^{(m)}(t) + \left(\mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right)^2 \\ &= \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du + o \left(\frac{1}{a_n^2 h_n^{2m+1}} \right) \\ &\quad + O(h_n), \end{aligned}$$

en reprenant les résultats (4.48) et (4.54) démontrés sous les hypothèses (H1), (H3), (H3''), (H4) et (H6). Les termes de variance et de biais convergent vers 0 à condition que $h_n \rightarrow 0$ et $a_n^2 h_n^{2m+1}$ lorsque $n \rightarrow +\infty$. La règle de Cramér F.10-1(i) entraîne la convergence \mathbb{L}^2 de $\hat{\alpha}_n^{(m)}(t)$ vers $\alpha^{(m)}(t)$, qui implique la convergence en probabilité de $\hat{\alpha}_n^{(m)}(t)$ vers $\alpha^{(m)}(t)$. ■

4.6.4 Preuve du Théorème 4.4.2

La décomposition de $\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t)$ en trois termes (4.42), (4.43) et (4.44), entraîne que

$$\begin{aligned} \sqrt{a_n^2 h_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right) \\ = \sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) \end{aligned} \quad (4.55)$$

$$+ \sqrt{a_n^2 h_n} \int_{-1}^1 k^{(m)}(-u) \mathbf{1}_{[Y_n(t+h_n u)=0]} \alpha^{(m)}(t+h_n u) du \quad (4.56)$$

$$+ \sqrt{a_n^2 h_n^{2m+1}} \int_{-1}^1 k(-u) \left(\alpha^{(m)}(t+h_n u) - \alpha^{(m)}(t) \right) du. \quad (4.57)$$

D'après (4.47), le terme dans (4.57) peut être remplacé. D'où :

$$\begin{aligned} \sqrt{a_n^2 h_n^{2m+1}} \int_{-1}^1 k(-u) \left(\alpha^{(m)}(t+h_n u) - \alpha^{(m)}(t) \right) du &= \sqrt{a_n^2 h_n^{2m+1}} O(h_n) \\ &= O \left(\sqrt{a_n^2 h_n^{2m+3}} \right). \end{aligned} \quad (4.58)$$

Ensuite, la fonction α est continue en t donc bornée dans un voisinage de t puis le noyau $k^{(m)}$ est borné sur $[-1, 1]$ selon les hypothèses (H1) et (H6(iii)). On en déduit que le terme dans (4.56) converge en probabilité vers 0, c'est-à-dire :

$$\sqrt{a_n^2 h_n} \int_{-1}^1 k^{(m)}(-u) \mathbf{1}_{[Y_n(t+h_n u)=0]} \alpha^{(m)}(t+h_n u) du \xrightarrow{\mathbb{P}} 0 \quad (4.59)$$

si on a :

$$\sup_{s \in [t-c, t+c]} \sqrt{a_n^2 h_n} \mathbf{1}_{[Y_n(s)=0]} \xrightarrow{\mathbb{P}} 0. \quad (4.60)$$

Lemme 4.6.2. Soient (X_n) et (Y_n) deux suites de variables aléatoires. Si pour tout $n \in \mathbb{N}$, $X_n > 0$ presque sûrement et Y_n est une variable de Bernoulli avec $\mathbb{P}[Y_n = 1] =$

$1 - \mathbb{P}[Y_n = 0] = p_n$ tel que $p_n \rightarrow 0$ lorsque $n \rightarrow +\infty$, alors pour tout $\epsilon > 0$,

$$\mathbb{P}[|X_n Y_n| > \epsilon] \leq \mathbb{P}[Y_n = 1] \rightarrow 0,$$

d'où

$$X_n Y_n \xrightarrow{\mathbb{P}} 0.$$

Appliquons le Lemme 4.6.2 pour $X_n = \sqrt{a_n^2 h_n} > 0$ et $Y_n = \mathbf{1}_{[Y_n(s)=0]}$. La convergence dans (4.60) est donc obtenue en considérant que $\mathbb{P}[\mathbf{1}_{[Y_n(s)=0]} = 1] = \mathbb{P}[Y_n(s) = 0] \rightarrow 0$, lorsque $n \rightarrow +\infty$. Cette dernière condition est remplie grâce aux hypothèses (H2) et (H3).

Enfin, intéressons-nous au terme dans (4.55). De (4.45), il vient que :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) = \sqrt{\frac{a_n^2}{h_n}} \int_{-1}^1 k^{(m)}(-u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} d\tilde{M}_n(u).$$

Soit v un réel compris entre -1 et 1 puis posons :

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) := \int_{-1}^v \tilde{H}_n(u) d\tilde{M}_n(u),$$

avec $\tilde{H}_n(u) := \sqrt{\frac{a_n^2}{h_n}} k^{(m)}(-u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)}$ pour tout $u \in [-1, v]$. Alors on a :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) := (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1}. \quad (4.61)$$

L'idée est de démontrer la normalité asymptotique de $(\tilde{H}_n \cdot \tilde{M}_n)(v)$ pour tout $v \in [-1, 1]$, à l'aide du théorème de la limite centrale pour les martingales dit de Rebolledo F.14, afin d'en déduire celle pour $v = 1$ en particulier, c'est-à-dire de $\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t)$.

Prouvons tout d'abord la condition (i) du Théorème F.14. (\tilde{H}_n) est une suite de processus prévisibles telle que $\mathbb{E}(\tilde{H}_n \cdot \tilde{M}_n)(v)^2 < +\infty$. En effet, (4.12), (4.13) puis le théorème de Fubini (intégration de fonctions bornées sur un intervalle fermé grâce à (H1), (H6(iii)) et le fait que $\frac{J}{Y}$ soit un processus borné) impliquent que :

$$\begin{aligned} \mathbb{E}(\tilde{H}_n \cdot \tilde{M}_n)(v)^2 &= \mathbb{E} \int_{-1}^v \frac{a_n^2}{h_n} k^{(m)}(-u)^2 \frac{J_n^2(t + h_n u)}{Y_n^2(t + h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v k^{(m)}(-u)^2 \mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] \alpha(t + h_n u) du \\ &= \int_{-1}^v k^{(m)}(-u)^2 \left(\mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] - \frac{1}{y(t + h_n u)} \right) \alpha(t + h_n u) du \\ &\quad + \int_{-1}^v k^{(m)}(-u)^2 \frac{\alpha(t + h_n u)}{y(t + h_n u)} du. \end{aligned}$$

Donc $\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2$ est majorée par une somme :

$$\int_{-1}^1 |k^{(m)}(-u)|^2 \left| \mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] - \frac{1}{y(t + h_n u)} \right| |\alpha(t + h_n u)| du \\ + \int_{-1}^1 |k^{(m)}(-u)|^2 \frac{|\alpha(t + h_n u)|}{y(t + h_n u)} du ,$$

bornée grâce aux hypothèses (H1), (H4) et (H6(iii)).

Montrons ensuite la condition (ii) du Théorème F.14. Par (4.13), nous avons :

$$\begin{aligned} & \langle (\tilde{H}_n \cdot \tilde{M}_n), (\tilde{H}_n \cdot \tilde{M}_n) \rangle(v) \\ &= \int_{-1}^v \frac{a_n^2}{h_n} k^{(m)}(-u)^2 \frac{J_n^2(t + h_n u)}{Y_n^2(t + h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v k^{(m)}(-u)^2 J_n(t + h_n u) \frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)} \alpha(t + h_n u) du \quad (4.62) \\ &+ \int_{-1}^v k^{(m)}(-u)^2 \frac{\alpha(t + h_n u)}{y(t + h_n u)} du . \end{aligned} \quad (4.63)$$

D'une part, notons que le terme

$$\frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)}$$

est borné sous l'hypothèse (H4). Sous la même hypothèse et puisque (H2) est admise, son numérateur converge en probabilité vers 0 et son dénominateur converge vers $y(t)^2$. De plus, J_n est un processus borné convergeant en probabilité vers 1 selon (H3) puis les hypothèses (H1) et (H6(iii)) garantissent que le noyau $k^{(m)}$ est borné sur $[-1, 1]$ et α est continue en t donc bornée dans un voisinage de t . Donc d'après le théorème de convergence dominée, le terme (4.62) converge en probabilité vers 0. D'autre part, les mêmes hypothèses nous assurent que le terme (4.63) est borné car la fonction y est notamment continue en t donc bornée dans un voisinage de t . Le théorème de convergence dominée implique que le terme (4.63) converge en probabilité vers $\frac{\alpha(t)}{y(t)} \int_{-1}^v k^{(m)}(-u)^2 du$.

Il résulte alors de la règle de Cramér F.10-1(i) que :

$$\langle (\tilde{H}_n \cdot \tilde{M}_n), (\tilde{H}_n \cdot \tilde{M}_n) \rangle(v) \xrightarrow{\mathbb{P}} \frac{\alpha(t)}{y(t)} \int_{-1}^v k^{(m)}(-u)^2 du . \quad (4.64)$$

Enfin, démontrons la condition (iii) du Théorème F.14. Pour tout $\epsilon > 0$ fixé, on a :

$$\begin{aligned} & \int_{-1}^v \tilde{H}_n^2(u) \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} d\langle \tilde{M}_n, \tilde{M}_n \rangle(u) \\ &= \int_{-1}^v \frac{a_n^2}{h_n} k^{(m)}(-u)^2 \frac{J_n^2(t+h_n u)}{Y_n^2(t+h_n u)} \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v k^{(m)}(-u)^2 J_n(t+h_n u) \frac{y(t+h_n u) - a_n^{-2} Y_n(t+h_n u)}{a_n^{-2} Y_n(t+h_n u) y(t+h_n u)} \\ & \quad \times \alpha(t+h_n u) \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} du \end{aligned} \tag{4.65}$$

$$+ \int_{-1}^v k^{(m)}(-u)^2 \frac{\alpha(t+h_n u)}{y(t+h_n u)} \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} du. \tag{4.66}$$

Par définition, nous avons :

$$\begin{aligned} & \{|\tilde{H}_n(u)| > \epsilon\} \\ &= \left\{ \left| k^{(m)}(-u) \frac{J_n(t+h_n u)}{a_n^{-2} Y_n(t+h_n u)} \right| > \epsilon \sqrt{a_n^2 h_n} \right\} \\ &= \left\{ \left| k^{(m)}(-u) J_n(t+h_n u) \left(\frac{y(t+h_n u) - a_n^{-2} Y_n(t+h_n u)}{a_n^{-2} Y_n(t+h_n u) y(t+h_n u)} + \frac{1}{y(t+h_n u)} \right) \right| > \epsilon \sqrt{a_n^2 h_n} \right\}. \end{aligned}$$

On remarque, dans le dernier ensemble, que le membre de gauche est borné, en vertu des hypothèses (H1), (H4) et (H6(iii)), donc en faisant tendre le membre de droite vers $+\infty$ (possible si l'on pose (4.25)), il vient que :

$$\{|\tilde{H}_n(u)| > \epsilon\} \rightarrow \emptyset,$$

d'où

$$\mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} \xrightarrow{\mathbb{P}} 0.$$

En réutilisant les hypothèses (H1), (H4) et (H6(iii)) puis le théorème de convergence dominée, nous déduisons que :

$$\int_{-1}^v \tilde{H}_n^2(u) \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} d\langle \tilde{M}_n, \tilde{M}_n \rangle(u) \xrightarrow{\mathbb{P}} 0. \tag{4.67}$$

d'après la règle de Cramér F.10-1(i) appliquée aux termes (4.65) et (4.66).

Toutes les conditions du théorème de Rebolledo F.14 sont réunies. Avec (4.64) et (4.67), on conclut que

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) \xrightarrow{\mathcal{L}} \mathbf{N}\left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^v k^{(m)}(-u)^2 du\right)$$

pour tout $v \in [-1, 1]$. Ceci est donc vrai en particulier pour $v = 1$, d'où :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) \xrightarrow{\mathcal{L}} \mathbf{N}\left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du\right), \tag{4.68}$$

par la définition (4.61).

Reprenons (4.58), (4.59) et (4.68) puis appliquons la règle de Cramér F.10-2, en considérant que $a_n^2 h_n^{2m+3} \rightarrow 0$ lorsque $n \rightarrow +\infty$. Ainsi,

$$\sqrt{a_n^2 h_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right) \xrightarrow{\mathcal{L}} \text{N} \left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^1 k^{(m)}(u)^2 du \right). \blacksquare \quad (4.69)$$

Dans les prochaines démonstrations, nous retrouvons en grande majorité les étapes des preuves précédentes. La différence essentielle concerne l'utilisation d'un noyau satisfaisant maintenant l'hypothèse (H6'), au lieu de l'hypothèse (H6).

4.6.5 Preuve du Lemme 4.4.3

Rappelons que :

$$\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) = \hat{M}_n^{(m)}(t) \quad (4.70)$$

$$+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \quad (4.71)$$

$$+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t), \quad (4.72)$$

avec

$$\begin{aligned} \hat{M}_n^{(m)}(t) &= \frac{1}{h_n^{m+1}} \int_{-1}^1 K_r^{(m)}(u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} [dN_n(t + h_n u) - Y_n(t + h_n u) \alpha(t + h_n u) du] \\ &= \frac{1}{h_n^{m+1}} \int_{-1}^1 K_r^{(m)}(u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} d\tilde{M}_n(u), \end{aligned} \quad (4.73)$$

d'après (4.29) et le changement de variable $u = \frac{s-t}{h_n}$.

La linéarité de l'espérance mène à :

$$\begin{aligned} \mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \mathbb{E} \hat{M}_n^{(m)}(t) \\ &+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \mathbb{E} [J_n(t + h_n u) - 1] \alpha(t + h_n u) du \\ &+ \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t), \end{aligned}$$

grâce au théorème de Fubini. En effet, les hypothèses (H1') et (H6'(iii)) impliquent que la fonction α est continue en t donc est bornée dans un voisinage de t puis le noyau $K_r^{(m)}$ est borné sur $[-1, 1]$. Comme J_n est une indicatrice, les fonctions sous l'intégrale dans (4.71) sont bornées. Étant intégrées sur un intervalle fermé, elles sont intégrables.

Soit v un réel compris entre -1 et 1 puis notons :

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) := \int_{-1}^v \tilde{H}_n(u) d\tilde{M}_n(u),$$

avec $\tilde{H}_n(u) := \frac{1}{h_n^{m+1}} K_r^{(m)}(u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)}$ pour tout $u \in [-1, v]$. Alors $\hat{M}_n^{(m)}(t)$ est la valeur en $v = 1$ de la martingale $(\tilde{H}_n \cdot \tilde{M}_n)$ donc son espérance est nulle. En utilisant la

définition du processus J_n , la formule de Taylor-Lagrange pour u' compris entre 0 et $h_n u$, avec $u \in [-1, 1]$ puis (H6'(ii)), nous écrivons :

$$\begin{aligned}
 \mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) &= \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \mathbb{P}[Y_n(t + h_n u) = 0] \alpha(t + h_n u) du \\
 &\quad + \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t) \\
 &= \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \mathbb{P}[Y_n(t + h_n u) = 0] \alpha(t + h_n u) du \\
 &\quad - \frac{1}{h_n^m} \int_{-1}^1 \sum_{i=0}^r \alpha^{(i)}(t + u', F) \frac{(h_n u)^i}{i!} K_r^{(m)}(u) du \\
 &\quad + \frac{1}{h_n^m} \int_{-1}^1 \alpha^{(r+1)}(t + u', F) \frac{(h_n u)^{r+1}}{(r+1)!} K_r^{(m)}(u) du - \alpha^{(m)}(t) \\
 &= \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \mathbb{P}[Y_n(t + h_n u) = 0] \alpha(t + h_n u) du \\
 &\quad + \frac{1}{h_n^m} \int_{-1}^1 \alpha^{(r+1)}(t + u', F) \frac{(h_n u)^{r+1}}{(r+1)!} K_r^{(m)}(u) du.
 \end{aligned}$$

Grâce à (H1'), (H3') et (H6'(iii)), il résulte que

$$\frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \mathbb{P}[Y_n(t + h_n u) = 0] \alpha(t + h_n u) du = o(h_n^{r+1-m}). \quad (4.74)$$

Puisque l'hypothèse (H1') entraîne que la dérivée $(r+1)^e$ de α est bornée sur un voisinage de t , nous avons par l'hypothèse (H6'(ii)) :

$$\frac{1}{h_n^m} \int_{-1}^1 \alpha^{(r+1)}(t + u', F) \frac{(h_n u)^{r+1}}{(r+1)!} K_r^{(m)}(u) du = O(h_n^{r+1-m}). \quad (4.75)$$

Le résultat provient des égalités (4.74) et (4.75) :

$$\mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) = o(h_n^{r+1-m}) + O(h_n^{r+1-m}) = O(h_n^{r+1-m}). \blacksquare \quad (4.76)$$

4.6.6 Preuve du Lemme 4.4.4

La variance de l'estimateur $\hat{\alpha}_n^{(m)}(t)$ s'exprime comme suit :

$$\begin{aligned}
 \text{Var} \hat{\alpha}_n^{(m)}(t) &= \mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \mathbb{E} \hat{\alpha}_n^{(m)}(t) \right]^2 \\
 &= \mathbb{E} \left[\frac{1}{h_n^{m+1}} \int K_r^{(m)} \left(\frac{t-s}{h_n} \right) \frac{dN_n(s)}{Y_n(s)} - \frac{1}{h_n^{m+1}} \int K_r^{(m)} \left(\frac{t-s}{h_n} \right) \mathbb{E} J_n(s) \alpha(s) ds \right]^2 \\
 &= \mathbb{E} T_1^2 + 2E [T_1 T_2] + \mathbb{E} T_2^2, \quad (4.77)
 \end{aligned}$$

avec

$$\begin{aligned} T_1 &:= \frac{1}{h_n^{m+1}} \int_{-1}^1 K_r^{(m)}(u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)} d\tilde{M}_n(u), \\ T_2 &:= \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) (J_n(t+h_n u) - \mathbb{E} J_n(t+h_n u)) \alpha(t+h_n u) du, \end{aligned}$$

obtenus en retranchant et ajoutant le terme $\frac{1}{h_n^{m+1}} \int K_r^{(m)}\left(\frac{t-s}{h_n}\right) J_n(s) \alpha(s) ds$ puis en substituant $u = \frac{s-t}{h_n}$.

Déterminons $\mathbb{E} T_1^2$ dans un premier temps. On observe que :

$$T_1 := (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1}, \quad (4.78)$$

grâce à la définition de \tilde{H}_n donnée dans la preuve précédente.

Donc d'après (4.12), (4.13) puis le théorème de Fubini (intégration de fonctions bornées sur un intervalle fermé grâce à (H1'), (H6'(iii)) et le fait que $\frac{J}{Y}$ soit un processus borné), on a :

$$\begin{aligned} \mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2 &= \mathbb{E} \int_{-1}^v \frac{1}{h_n^{2(m+1)}} K_r^{(m)}(u)^2 \frac{J_n^2(t+h_n u)}{Y_n^2(t+h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v \frac{1}{a_n^2 h_n^{2m+1}} K_r^{(m)}(u)^2 \mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] \alpha(t+h_n u) du \\ &= \frac{1}{a_n^2 h_n^{2m+1}} \int_{-1}^v K_r^{(m)}(u)^2 \left(\mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] - \frac{1}{y(t+h_n u)} \right) \alpha(t+h_n u) du \\ &\quad + \frac{1}{a_n^2 h_n^{2m+1}} \int_{-1}^v K_r^{(m)}(u)^2 \frac{\alpha(t+h_n u)}{y(t+h_n u)} du. \end{aligned}$$

Cette expression est majorée par :

$$\begin{aligned} &\int_{-1}^1 |K_r^{(m)}(u)|^2 \left| \mathbb{E} \left[\frac{a_n^2 J_n(t+h_n u)}{Y_n(t+h_n u)} \right] - \frac{1}{y(t+h_n u)} \right| |\alpha(t+h_n u)| du \\ &\quad + \int_{-1}^1 |K_r^{(m)}(u)|^2 \frac{|\alpha(t+h_n u)|}{y(t+h_n u)} du. \end{aligned}$$

En vertu des hypothèses (H1'), (H4) et (H6'(iii)), le théorème de convergence dominée peut être utilisé et il entraîne que :

$$\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2 = o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right) + \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^v K_r^{(m)}(u)^2 du,$$

d'où en particulier pour $v = 1$:

$$\mathbb{E} T_1^2 = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \quad (4.79)$$

Concentrons-nous à l'espérance de T_2^2 dans un second temps. Par l'inégalité de Jensen

F.12, nous avons :

$$\mathbb{E} T_2^2 \leq \mathbb{E} \left[\frac{1}{h_n^{2m}} \int_{-1}^1 K_r^{(m)}(u)^2 (J_n(t + h_n u) - \mathbb{E} J_n(t + h_n u))^2 \alpha(t + h_n u)^2 du \right].$$

Nous pouvons appliquer le théorème de Fubini puisque l'intégrale est définie sur l'intervalle fermé $[-1, 1]$ et les fonctions se trouvant sous cette intégrale sont bornées d'après les hypothèses (H1), (H6(iii)) et la définition du processus J_n . Comme on a :

$$\mathbb{E} [J_n(t + h_n u) - \mathbb{E} J_n(t + h_n u)]^2 = \text{Var} J_n(t + h_n u) \leq \mathbb{P}[Y_n(t + h_n u) = 0],$$

nous écrivons :

$$\begin{aligned} \mathbb{E} T_2^2 &\leq \frac{1}{h_n^{2m}} \int_{-1}^1 K_r^{(m)}(u)^2 \mathbb{P}[Y_n(t + h_n u) = 0] \alpha(t + h_n(u))^2 du \\ &= o\left(\frac{1}{a_n^2 h_n^{2m}}\right). \end{aligned} \quad (4.80)$$

sous les hypothèses (H1), (H3'') et (H6(iii)).

Analysons dans un dernier temps le terme $2\mathbb{E}[T_1 T_2]$. Grâce à l'inégalité de Cauchy-Schwartz F.13, on a :

$$\mathbb{E}[|T_1 T_2|] \leq (\mathbb{E} T_1^2)^{\frac{1}{2}} (\mathbb{E} T_2^2)^{\frac{1}{2}},$$

donc avec (4.79) et (4.80), nous obtenons :

$$\begin{aligned} 2\mathbb{E}[T_1 T_2] &\leq \left[O\left(\frac{1}{a_n^2 h_n^{2m+1}}\right) o\left(\frac{1}{a_n^2 h_n^{2m}}\right) \right]^{\frac{1}{2}} \\ &= \left[o\left(\frac{1}{a_n^4 h_n^{4m+1}}\right) \right]^{\frac{1}{2}} \\ &= o\left(\frac{1}{a_n^2 h_n^{\frac{4m+1}{2}}}\right). \end{aligned} \quad (4.81)$$

Constatons que les termes $\mathbb{E} T_2^2$ et $2\mathbb{E}[T_1 T_2]$ sont négligeables par rapport à $\mathbb{E} T_1^2$ d'après les expressions (4.79), (4.80) et (4.81), d'où :

$$\text{Var} \hat{\alpha}_n^{(m)}(t) = \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du + o\left(\frac{1}{a_n^2 h_n^{2m+1}}\right). \quad (4.82)$$

par la formule (4.77). ■

4.6.7 Preuve du Théorème 4.4.3

Pour $m \geq 0$, l'erreur en moyenne quadratique se décompose en un terme de variance et un terme de biais :

$$\begin{aligned} \mathbb{E} \left[\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right]^2 &= \text{Var} \hat{\alpha}_n^{(m)}(t) + \left(\mathbb{E} \hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right)^2 \\ &= \frac{1}{a_n^2 h_n^{2m+1}} \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du + o \left(\frac{1}{a_n^2 h_n^{2m+1}} \right) \\ &\quad + O \left(h_n^{r+1-m} \right), \end{aligned}$$

grâce aux résultats (4.76) et (4.82) vérifiés sous les hypothèses (H1'), (H3'), (H3''), (H4) et (H6). En supposant (H2) et (4.33), les termes de variance et de biais convergent respectivement vers 0. La convergence en moyenne quadratique de $\hat{\alpha}_n^{(m)}(t)$ vers $\alpha^{(m)}(t)$ est obtenue à l'aide de la règle de Cramér F.10-1(i).

La convergence en probabilité de $\hat{\alpha}_n^{(m)}(t)$ vers $\alpha^{(m)}(t)$ découle de la convergence en moyenne quadratique de $\hat{\alpha}_n^{(m)}(t)$ vers $\alpha^{(m)}(t)$. ■

4.6.8 Preuve du Théorème 4.4.4

En reprenant la décomposition de $\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t)$ en trois termes (4.70), (4.71) et (4.72), on a :

$$\begin{aligned} &\sqrt{a_n^2 h_n^{2m+1}} \left(\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t) \right) \\ &= \sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) \end{aligned} \quad (4.83)$$

$$+ \sqrt{a_n^2 h_n} \int_{-1}^1 K_r^{(m)}(u) \mathbf{1}_{[Y_n(t+h_n u)=0]} \alpha^{(m)}(t + h_n u) du \quad (4.84)$$

$$+ \sqrt{a_n^2 h_n^{2m+1}} \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t). \quad (4.85)$$

Utilisons tout d'abord (4.75). Alors le terme dans (4.85) devient :

$$\begin{aligned} \sqrt{a_n^2 h_n^{2m+1}} \frac{1}{h_n^m} \int_{-1}^1 K_r^{(m)}(u) \alpha(t + h_n u) du - \alpha^{(m)}(t) &= \sqrt{a_n^2 h_n^{2m+1}} O \left(h_n^{r+1-m} \right) \\ &= O \left(\sqrt{a_n^2 h_n^{2r+3}} \right). \end{aligned} \quad (4.86)$$

Ensuite, le noyau $K_r^{(m)}$ est borné sur $[-1, 1]$ et la fonction α est continue en t donc bornée dans un voisinage de t sous les hypothèses (H1') et (H6'(iii)). Il résulte que le terme dans (4.84) converge en probabilité vers 0, c'est-à-dire :

$$\sqrt{a_n^2 h_n} \int_{-1}^1 K_r^{(m)}(u) \mathbf{1}_{[Y_n(t+h_n u)=0]} \alpha^{(m)}(t + h_n u) du \xrightarrow{\mathbb{P}} 0 \quad (4.87)$$

pourvu que :

$$\sup_{s \in [t-c, t+c]} \sqrt{a_n^2 h_n} \mathbf{1}_{[Y_n(s)=0]} \xrightarrow{\mathbb{P}} 0. \quad (4.88)$$

Nous avons démontré, dans la Preuve du Théorème 4.4.2, la convergence dans (4.88) à condition que $\mathbb{P}[Y_n(s) = 0] \rightarrow 0$, lorsque $n \rightarrow +\infty$. Ceci est vrai sous les hypothèses (H2) et (H3').

Enfin, intéressons-nous au terme dans (4.83). Grâce à (4.73), nous écrivons :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) = \sqrt{\frac{a_n^2}{h_n}} \int_{-1}^1 K_r^{(m)}(u) \frac{J_n(t + h_n u)}{Y_n(t + h_n u)} d\tilde{M}_n(u).$$

Soit v un réel compris entre -1 et 1 puis introduisons :

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) := \int_{-1}^v \tilde{H}_n(u) d\tilde{M}_n(u),$$

avec $\tilde{H}_n(u) := \sqrt{\frac{a_n^2}{h_n}} K_r^{(m)}(u) \frac{J_n(t+h_n u)}{Y_n(t+h_n u)}$ pour tout $u \in [-1, v]$. Alors on a :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) := (\tilde{H}_n \cdot \tilde{M}_n)(v) \Big|_{v=1}. \quad (4.89)$$

Comme dans la preuve du Théorème 4.4.2, nous allons utiliser le théorème de Rebolledo (F.14) pour montrer la normalité asymptotique de $(\tilde{H}_n \cdot \tilde{M}_n)(v)$ pour tout $v \in [-1, 1]$, puis en déduire celle de $\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t)$ (correspondant au cas particulier $v = 1$).

Commençons par démontrer la condition (i) du Théorème F.14. Avec (4.12), (4.13) puis le théorème de Fubini (intégration de fonctions bornées sur un intervalle fermé grâce à (H1'), (H6'(iii))) et le fait que $\frac{J}{Y}$ soit un processus borné, nous avons :

$$\begin{aligned} \mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2 &= \mathbb{E} \int_{-1}^v \frac{a_n^2}{h_n} K_r^{(m)}(u)^2 \frac{J_n^2(t + h_n u)}{Y_n^2(t + h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v K_r^{(m)}(u)^2 \mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] \alpha(t + h_n u) du \\ &= \int_{-1}^v K_r^{(m)}(u)^2 \left(\mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] - \frac{1}{y(t + h_n u)} \right) \alpha(t + h_n u) du \\ &\quad + \int_{-1}^v K_r^{(m)}(u)^2 \frac{\alpha(t + h_n u)}{y(t + h_n u)} du. \end{aligned}$$

Ainsi, la quantité $\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(v)^2$ est majorée par une somme :

$$\begin{aligned} &\int_{-1}^1 |K_r^{(m)}(u)|^2 \left| \mathbb{E} \left[\frac{a_n^2 J_n(t + h_n u)}{Y_n(t + h_n u)} \right] - \frac{1}{y(t + h_n u)} \right| |\alpha(t + h_n u)| du \\ &\quad + \int_{-1}^1 |K_r^{(m)}(u)|^2 \frac{|\alpha(t + h_n u)|}{y(t + h_n u)} du. \end{aligned}$$

Les hypothèses (H1'), (H4) et (H6'(iii)) nous assurent que cette somme est bornée, donc (\tilde{H}_n) est une suite de processus prévisibles telle que $\mathbb{E} (\tilde{H}_n \cdot \tilde{M}_n)(t)^2 < +\infty$.

Prouvons ensuite la condition (ii) du Théorème F.14. Par (4.13), il vient que :

$$\begin{aligned} & \langle (\tilde{H}_n \cdot \tilde{M}_n), (\tilde{H}_n \cdot \tilde{M}_n) \rangle(v) \\ &= \int_{-1}^v \frac{a_n^2}{h_n} K_r^{(m)}(u)^2 \frac{J_n^2(t + h_n u)}{Y_n^2(t + h_n u)} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v K_r^{(m)}(u) J_n(t + h_n u) \frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)} \alpha(t + h_n u) du \quad (4.90) \end{aligned}$$

$$+ \int_{-1}^v K_r^{(m)}(u)^2 \frac{\alpha(t + h_n u)}{y(t + h_n u)} du. \quad (4.91)$$

Nous avons obtenu, dans la preuve du Théorème 4.4.2, que le terme

$$\frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)}$$

est borné et converge en probabilité vers 0 sous les hypothèses (H2) et (H4), que J_n est un processus borné convergeant en probabilité vers 1 d'après (H3'), que la fonction α est bornée dans un voisinage de t grâce à (H1) et que la fonction y est bornée dans un voisinage de t par (H4). Puisque le noyau $K_r^{(m)}$ est borné sur $[-1, 1]$ d'après (H6'(iii)), nous pouvons appliquer le théorème de convergence dominée et obtenir que le terme (4.90) converge en probabilité vers 0 et que le terme (4.91) converge en probabilité vers $\frac{\alpha(t)}{y(t)} \int_{-1}^v K_r^{(m)}(u)^2 du$. La règle de Cramér F.10-1(i) permet alors d'écrire :

$$\langle (\tilde{H}_n \cdot \tilde{M}_n), (\tilde{H}_n \cdot \tilde{M}_n) \rangle(v) \xrightarrow{\mathbb{P}} \frac{\alpha(t)}{y(t)} \int_{-1}^v K_r^{(m)}(u)^2 du. \quad (4.92)$$

Montrons enfin la condition (iii) du Théorème F.14. Pour tout $\epsilon > 0$ fixé, nous avons :

$$\begin{aligned} & \int_{-1}^v \tilde{H}_n^2(u) \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} d\langle \tilde{M}_n, \tilde{M}_n \rangle(u) \\ &= \int_{-1}^v \frac{a_n^2}{h_n} K_r^{(m)}(u)^2 \frac{J_n^2(t + h_n u)}{Y_n^2(t + h_n u)} \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} \tilde{\lambda}_n(u) du \\ &= \int_{-1}^v K_r^{(m)}(u)^2 J_n(t + h_n u) \frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)} \end{aligned} \quad (4.93)$$

$$\alpha(t + h_n u) \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} du \quad (4.94)$$

$$+ \int_{-1}^v K_r^{(m)}(u)^2 \frac{\alpha(t + h_n u)}{y(t + h_n u)} \mathbf{1}_{\{|\tilde{H}_n(u)| > \epsilon\}} du. \quad (4.95)$$

Par définition, on a :

$$\begin{aligned} & \{|\tilde{H}_n(u)| > \epsilon\} \\ &= \left\{ \left| K_r^{(m)}(u) \frac{J_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u)} \right| > \epsilon \sqrt{a_n^2 h_n} \right\} \\ &= \left\{ \left| K_r^{(m)}(u) J_n(t + h_n u) \left(\frac{y(t + h_n u) - a_n^{-2} Y_n(t + h_n u)}{a_n^{-2} Y_n(t + h_n u) y(t + h_n u)} + \frac{1}{y(t + h_n u)} \right) \right| > \epsilon \sqrt{a_n^2 h_n} \right\}. \end{aligned}$$

Le membre de gauche du dernier ensemble est borné sous les hypothèses (H1'), (H4)

et (H6'(iii)). En considérant (4.36), le membre de droite converge vers $+\infty$ et nous déduisons que :

$$\{|\tilde{H}_n(u)| > \epsilon\} \rightarrow \emptyset,$$

d'où

$$\mathbf{1}_{[|\tilde{H}_n(u)| > \epsilon]} \xrightarrow{\mathbb{P}} 0.$$

En réutilisant les hypothèses (H1'), (H4) et (H6'(iii)) puis le théorème de convergence dominée, nous obtenons :

$$\int_{-1}^v \tilde{H}_n^2(u) \mathbf{1}_{[|\tilde{H}_n(u)| > \epsilon]} d\langle \tilde{M}_n, \tilde{M}_n \rangle(u) \xrightarrow{\mathbb{P}} 0. \quad (4.96)$$

d'après la règle de Cramér F.10-1(i) appliquée aux termes (4.94) et (4.95).

Toutes les conditions du théorème de Rebolledo F.14 sont satisfaites. Avec (4.92) et (4.96), il vient que

$$(\tilde{H}_n \cdot \tilde{M}_n)(v) \xrightarrow{\mathcal{L}} \mathbf{N}\left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^v K_r^{(m)}(u)^2 du\right)$$

pour tout $v \in [-1, 1]$. Ceci est donc vrai en particulier pour $v = 1$, d'où par la définition (4.89) :

$$\sqrt{a_n^2 h_n^{2m+1}} \hat{M}_n^{(m)}(t) \xrightarrow{\mathcal{L}} \mathbf{N}\left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du\right). \quad (4.97)$$

En supposant (4.37), avec (4.86), (4.87) et (4.97) puis la règle de Cramér F.10-2, on conclut :

$$\sqrt{a_n^2 h_n^{2m+1}} (\hat{\alpha}_n^{(m)}(t) - \alpha^{(m)}(t)) \xrightarrow{\mathcal{L}} \mathbf{N}\left(0, \frac{\alpha(t)}{y(t)} \int_{-1}^1 K_r^{(m)}(u)^2 du\right). \blacksquare \quad (4.98)$$

4.7 Application au taux d'accumulation du surplus en actuariat

4.7.1 Cadre

"Un *contrat d'assurance vie* est un contrat d'assurance dont le risque, objet du contrat, dépend de la durée de la vie humaine.", d'après l'Institut National de la Consommation (INC). L'*assureur* ou la compagnie d'assurances s'engage à verser à un *bénéficiaire* un capital ou une rente (*prestation*) en contrepartie de *primes* payées par le *souscripteur* pendant une période déterminée lorsque le risque survient pour la personne désignée au contrat (l'*assuré*, souvent le souscripteur). La prestation est versée à une date déterminée dans le contrat si l'assuré est en vie. En cas de décès, elle est versée au décès de l'assuré.

Au cours de sa vie, un assuré est susceptible de passer par plusieurs états (actif, handicapé, décès...). Hoem ([85]) a proposé l'emploi des processus multi-états afin de décrire la trajectoire de vie d'un assuré.

Considérons une police d'assurance vie avec r états possibles $\{1, \dots, r\}$, débutant à

la date 0 dans l'état initial 1, et se terminant à une date finie n . Son développement est modélisé par une chaîne de Markov continue Z avec des probabilités de transition $p_{jk}(t, u)$ et des intensités de transition $\mu_{jk}(t)$, $j \neq k$, $t \neq u$. L'intensité totale de transition de l'état j à la date t est $\mu_{j\cdot}(t) = \sum_{k; k \neq j} \mu_{jk}(t)$.

Nous nous plaçons dans le cadre simplifié, où la durée passée par l'assuré dans un état n'est pas prise en compte. Ainsi, nous supposons que les prestations de l'assureur sont indépendantes de la durée.

Décrivons les prestations de l'assureur grâce au processus $Prest$ et les primes payées par l'assuré par le processus $Prim$. Nous désignons par B le processus de la différence des flux entrants et sortants du contrat de l'assuré dont la dynamique est

$$dB(t) = dPrest(t) - dPrim(t). \quad (4.99)$$

Les prestations honorées par l'assureur sont de deux types : un montant $\beta_j(t)$ versé si l'assuré occupe l'état j à la date t et un montant $b_{jk}(t)$ versé en cas de transition à la date t de l'état j vers l'état k . Les primes payées par l'assuré lorsqu'il se trouve dans l'état j à la date t sont notées $\Pi_j(t)$. En supposant toutes ces fonctions déterministes et en introduisant $dB_j(t) = d(\beta_j(t) - \Pi_j(t))$, la dynamique des flux entrants et sortants de la police d'assurance est donnée par

$$dB(t) = \sum_j I_j(t) dB_j(t) + \sum_{j \neq k} b_{jk}(t) dN_{jk}(t), \quad (4.100)$$

où le processus $I_j(t)$ est égal à 1 ou 0 selon que l'assuré se trouve dans l'état j ou non à la date t , et le processus de comptage $N_{jk}(t)$ donne le nombre de transitions de l'état j vers l'état k ($k \neq j$) pendant la période $[0, t]$.

Nous ne considérons pas les taux d'intérêts identiques à chaque instant mais préférons nous placer dans la situation générale où les taux d'intérêts sont décrits par une courbe $(\delta(s))_{s \geq 0}$.

Le coefficient d'actualisation sur la période $[0, t]$ sera alors défini par

$$v(t) = e^{-\int_0^t \delta(s) ds}. \quad (4.101)$$

Sur la période $[t, \tau]$ lorsque $t < \tau$, le coefficient d'actualisation sera donc donné par

$$v(t, \tau) = e^{-\int_t^\tau \delta(s) ds} = \frac{v(\tau)}{v(t)}. \quad (4.102)$$

Remarquons que $v(t) = v(0, t)$.

Soit \mathcal{G}_t l'historique complet de la police d'assurance vie. Le montant mis en réserve par l'assureur est une estimation basée sur l'information disponible à cet instant. La réserve à la date t est définie précisément comme une espérance conditionnelle sachant l'historique de la police.

La réserve prospective à la date t est égale à l'espérance de la valeur actuelle des prestations futures diminuée de l'espérance de la valeur actuelle des primes restant à payer par l'assuré, d'où :

$$V(t) = \mathbb{E} \left[\int_t^n v(t, \tau) dB(\tau) \mid \mathcal{G}_t \right]. \quad (4.103)$$

La réserve peut également être vue comme la différence entre la valeur acquise des primes et de la valeur acquise des prestations, c'est-à-dire

$$U(t) = \int_0^t v(t, \tau) d(-B)(\tau). \quad (4.104)$$

On la nomme dans ce cas *réserve rétrospective*.

Afin de gérer le risque non-diversifiable, les flux du contrat d'assurance vie sont tout d'abord spécifiés en valeur nominale et déterminés sur une base technique, appelée également *base prudente* ou *base de premier ordre*, représentant le pire scénario possible. On désigne par δ^* le taux technique et les intensités de transition, la mesure de probabilité et l'espérance sur cette base seront respectivement notées μ_{jk}^* , \mathbb{P}^* et \mathbb{E}^* . Le coefficient d'actualisation sur la période $[0, t]$ sera alors défini par

$$v^*(t) = e^{-\int_0^t \delta^*(s) ds} \quad (4.105)$$

et sur la période $[t, \tau]$ lorsque $t < \tau$, il sera donné par

$$v^*(t, \tau) = e^{-\int_t^\tau \delta^*(s) ds}. \quad (4.106)$$

La législation impose que les réserves et les primes soient calculées sur la base prudente afin que des bénéfices émergent systématiquement du scénario réalisé, autrement dit *base observable* ou encore *base de second ordre*. La base observable sera supposée stochastique et ses éléments seront désignés par δ , μ_{jk} , \mathbb{P} et \mathbb{E} . Un contrat géré de cette manière résulte d'un dispositif appelé la *participation aux bénéfices*.

La réserve prospective sur la base prudente à la date t est donc définie par

$$V^*(t) = \mathbb{E}^* \left[\int_t^n v^*(t, \tau) dB(\tau) \mid \mathcal{G}_t \right] \quad (4.107)$$

et lorsque l'assuré se trouve dans l'état j , elle s'écrit :

$$V_j^*(t) = \mathbb{E}^* \left[\int_t^n v^*(t, \tau) dB(\tau) \mid Z(t) = j \right]. \quad (4.108)$$

Les primes sont fixées selon le *principe d'équivalence* (égalité en moyenne entre la valeur actuelle des primes payées par l'assuré sur la période $[0, n]$ et la valeur actuelle des prestations versées par l'assureur sur la période $[0, n]$, avec n la durée du contrat d'assurance vie), d'où

$$\mathbb{E}^* \left[\int_0^n v^*(\tau) dB(\tau) \right] = 0, \quad (4.109)$$

ou de manière équivalente,

$$V_1^*(0) = 0. \quad (4.110)$$

Les bénéfices à chaque date t s'obtiennent par différence de la réserve rétrospective sur la base observable et de la réserve prospective sur la base prudente, c'est-à-dire

$$S(t) = U(t) - V^*(t) \quad (4.111)$$

$$= U(t) - \sum_j I_j(t) V_j^*(t). \quad (4.112)$$

En particulier, on a

$$S(0) = U(0) - \sum_j I_j(0) V_j^*(0) \quad (4.113)$$

$$= -I_1(0) V_1^*(0) \quad (4.114)$$

$$= 0 \quad (4.115)$$

d'après (4.110).

Afin d'étudier la dynamique des bénéficiaires, il nous faut différencier $S(t)$. Grâce à (4.112), on a :

$$\begin{aligned} dS(t) &= dU(t) - \sum_j dI_j(t) V_j^*(t) - \sum_j I_j(t) dV_j^*(t) \\ &= \delta(t) dt U(t) - dB(t) - \sum_j dI_j(t) V_j^*(t) - \sum_j I_j(t) dV_j^*(t) \\ &= \delta(t) dt S(t) + \delta(t) dt \sum_j I_j(t) V_j^*(t) - dB(t) - \sum_j dI_j(t) V_j^*(t) - \sum_j I_j(t) dV_j^*(t). \end{aligned}$$

L'égalité $dI_j(t) = \sum_{k; k \neq j} (dN_{jk}(t) - dN_{kj}(t))$ et la formule (4.100) entraînent que

$$\begin{aligned} dS(t) &= \delta(t) dt S(t) + \delta(t) dt \sum_j I_j(t) V_j^*(t) - \sum_j I_j(t) dB_j(t) - \sum_{j \neq k} b_{jk}(t) N_{jk}(t) \\ &\quad - \sum_{j \neq k} (V_k^*(t) - V_j^*(t)) dN_{jk}(t) - \sum_j I_j(t) dV_j^*(t). \end{aligned}$$

Introduisons $R_{jk}^*(t) = b_{jk}(t) + V_k^*(t) - V_j^*(t)$ le flux (positif ou négatif) représentant le *capital sous risque* en cas de transition de l'état j vers l'état k à la date t . En injectant l'équation de Thiele backward :

$$dV_j^*(t) = \delta^*(t) V_j^*(t) dt - dB_j(t) - \sum_{k; k \neq j} R_{jk}^*(t) \mu_{jk}^*(t) dt, \quad (4.116)$$

puis en ajoutant et retranchant la quantité $\sum_j I_j(t) \sum_{k; k \neq j} R_{jk}^*(t) \mu_{jk}(t) dt$, on obtient

$$\begin{aligned} dS(t) &= \delta(t) dt S(t) + \delta(t) dt \sum_j I_j(t) V_j^*(t) - \sum_j I_j(t) dB_j(t) - \sum_{j \neq k} R_{jk}^*(t) N_{jk}(t) \\ &\quad - \sum_j I_j(t) \delta^*(t) V_j^*(t) dt + \sum_j I_j(t) dB_j(t) \\ &\quad + \sum_j I_j(t) \sum_{k; k \neq j} R_{jk}^*(t) \mu_{jk}^*(t) dt \\ &\quad + \sum_j I_j(t) \sum_{k; k \neq j} R_{jk}^*(t) \mu_{jk}(t) dt - \sum_j I_j(t) \sum_{k; k \neq j} R_{jk}^*(t) \mu_{jk}(t) dt \\ &= \delta(t) dt S(t) + \delta(t) dt + \sum_j I_j(t) \left[(\delta(t) - \delta^*(t)) V_j^*(t) + R_{jk}^*(t) (\mu_{jk}^*(t) - \mu_{jk}(t)) \right] dt \\ &\quad - \sum_{j \neq k} R_{jk}^*(t) (dN_{jk}(t) - I_j(t) \mu_{jk}(t) dt). \end{aligned}$$

Posons

$$c_j(t) := (\delta(t) - \delta^*(t)) V_j^*(t) + \sum_{k; k \neq j} R_{jk}^*(t) (\mu_{jk}^*(t) - \mu_{jk}(t)). \quad (4.117)$$

qui correspond au *bénéfice instantané généré par la police d'assurance lorsque l'assuré se trouve dans l'état j* obtenu à partir du gain sur les taux d'intérêt et les intensités de transition, comme l'a défini Ramlau-Hansen dans [136]. Norberg ([127]) est arrivé à la définition (4.117) à partir du concept de bénéfice d'une police individuelle.

L'objectif de la participation aux bénéfices est d'éliminer le risque non diversifiable des compagnies d'assurances. Pour cela, le taux technique δ^* ainsi que les intensités de transition μ_{jk}^* doivent être choisis judicieusement afin de dégager des bénéfices (positifs). Les conditions requises découlent de la formule (4.117) et sont décrites par

$$\delta(t) \geq \delta^*(t) \quad (4.118)$$

en imposant l'hypothèse réaliste que toutes les réserves $V_j^*(t)$ soient positives et

$$\text{sign} \{ \mu_{jk}^*(t) - \mu_{jk}(t) \} = \text{sign} R_{jk}^*(t). \quad (4.119)$$

Les conditions (4.118) et (4.119) constituent les *marges de sécurité* de la compagnie d'assurances.

4.7.2 Estimation du taux d'accumulation du surplus

Dans [128], Norberg a montré que

$$S(t) = \int_0^t v(t, \tau) \sum_j p_{1j}(0, \tau) c_j(\tau) d\tau, \quad (4.120)$$

ce qui permet de définir le taux d'accumulation des bénéfices ou *taux d'accumulation du surplus* à toute date $t \geq 0$ par

$$c(t) := \sum_j p_{1j}(0, t) c_j(t). \quad (4.121)$$

Définissons l'estimateur du taux d'accumulation du surplus par

$$\hat{c}_n(t) = \sum_j \hat{p}_{n,1j}(0, t) \hat{c}_{n,j}(t), \quad (4.122)$$

où $\hat{p}_{n,1j}(0, t)$ correspond à la première ligne de l'estimateur produit-limite $\hat{\mathbf{P}}(s, t)$ donné par (E.4) pour $s = 0$, et

$$\hat{c}_{n,j}(t) = (\delta(t) - \delta^*(t)) V_j^*(t) + \sum_{k; k \neq j} R_{jk}^*(t) (\mu_{jk}^*(t) - \hat{\mu}_{n,jk}(t)), \quad (4.123)$$

avec $\hat{\mu}_{n,jk}(t)$ représentant un estimateur à noyau de $\mu_{jk}(t)$, l'intensité de transition de l'état j vers l'état k . Le travail effectué dans la Section 4.4 permet d'envisager deux choix possibles pour $\hat{\mu}_{n,jk}(t)$: l'estimateur à noyau classique donné par (4.17) et l'estimateur polynomial local donné par (4.28) pour $m = 0$, dont les distributions limites sont obtenues respectivement grâce aux Théorème 4.4.2 et Théorème 4.4.4 pour $m = 0$, en remplaçant $\hat{\alpha}_n$ par $\hat{\mu}_{n,jk}$, α par μ_{jk} et y par y_j .

Remarque :

Le problème des effets de bord de $\hat{\mu}_{n,jk}(t), t > 0$, pourrait être traité à l'aide de la Section 3.3 .

Nous énonçons à présent deux théorèmes décrivant les propriétés de l'estimateur $\hat{c}_n(t)$ basé sur l'estimation à noyau classique des intensités de transition $\mu_{jk}(t), t > 0$. Nous avons besoin au préalable de modifier légèrement les hypothèses (H3), (H3''), (H4) et (H5) en indexant tous les objets par j (Y_n, J_n et y se transforment en $Y_{n,j}, J_{n,j}$ et y_j). Voici maintenant un résultat sur la consistance de l'estimateur $\hat{c}_n(t)$.

Théorème 4.7.1. *Supposons que t est un réel strictement positif puis que les hypothèses (H1)-(H3) pour $m = 0$ et (H3'')-(H4)-(H6) sont vérifiées. Sous les conditions (E.5) et (E.6) pour $s = 0$ et si lorsque $n \rightarrow +\infty$,*

$$h_n \rightarrow 0 \quad (4.124)$$

suffisamment lentement tel que

$$a_n^2 h_n \rightarrow +\infty, \quad (4.125)$$

alors l'estimateur $\hat{c}_n(t)$ est convergent en probabilité,

$$\hat{c}_n(t) \xrightarrow{\mathbb{P}} c(t). \quad (4.126)$$

Nous exposons, pour finir, la normalité asymptotique de l'estimateur du taux d'accumulation du surplus.

Théorème 4.7.2. *Soit $t > 0$. Supposons que les hypothèses (H1)-(H3) pour $m = 0$ et (H4)-(H5)-(H6) sont satisfaites. Si les conditions (E.9), (E.10) et (E.11) sont vérifiées et si lorsque $n \rightarrow +\infty, h_n \rightarrow 0$ suffisamment lentement tel que*

$$a_n^2 h_n \rightarrow \infty, \quad (4.127)$$

et suffisamment vite tel que

$$a_n^2 h_n^3 \rightarrow 0, \quad (4.128)$$

alors

$$\sqrt{a_n^2 h_n} (\hat{c}_n(t) - c(t)) \xrightarrow{\mathcal{L}} N\left(0, \sum_j \sum_{k; k \neq j} R_{jk}^*(t)^2 \sigma_{jk}^2(t) p_{1j}^2(0, t)\right), \quad (4.129)$$

avec

$$\sigma_{jk}^2(t) = \frac{\mu_{jk}(t)}{y_j(t)} \int_{-1}^1 k^{(0)}(u)^2 du. \quad (4.130)$$

Le dernier théorème peut cependant être amélioré en utilisant l'estimateur polynomial local des intensités de transition $\mu_{jk}(t), t > 0$. En effet, en supposant (H1')-(H3') pour $m = 0$ puis l'hypothèse (H6') au lieu de (H1)-(H3) pour $m = 0$ puis (H6), (4.128) est remplacée par la condition (4.37), plus simple à vérifier. La variance (4.130) devient alors

$$\sigma_{jk}^2(t) = \frac{\mu_{jk}(t)}{y_j(t)} \int_{-1}^1 K_r^{(0)}(u)^2 du. \quad (4.131)$$

4.7.3 Preuve du Théorème 4.7.1

D'une part, le Théorème E.2.1 établit la convergence uniforme en probabilité de l'estimateur produit-limite $\hat{\mathbf{P}}(s, t)$ donné par (E.4) vers la matrice des probabilités de transition $\mathbf{P}(s, t)$ donnée par (E.1) pour tout $s \leq t$, sous les conditions (E.5) et (E.6). En considérant uniquement la première ligne de l'estimateur produit-limite $\hat{\mathbf{P}}(s, t)$ pour $s = 0$ et en la notant $\hat{p}_{n,1j}(0, t)$, il vient que

$$\hat{p}_{n,1j}(0, t) \xrightarrow{\mathbb{P}} p_{1j}(0, t), \quad (4.132)$$

où $p_{1j}(0, t)$ correspond à la première ligne de la matrice des probabilités de transition $\mathbf{P}(s, t)$ pour $s = 0$.

D'autre part, nous prouvons la convergence en probabilité de $\hat{\mu}_{n,jk}(t)$ vers $\mu_{jk}(t)$. Il suffit d'utiliser le Théorème 4.4.1 pour $m = 0$ puis d'insérer respectivement $\hat{\mu}_{n,jk}$ et μ_{jk} à la place de $\hat{\alpha}_n$ et de α . Selon (4.123), $\hat{c}_{j,n}(t)$ s'exprime en fonction de $\hat{\mu}_{n,jk}(t)$ donc les règles de Cramér F.10-1(i)-(ii) impliquent que

$$\hat{c}_{j,n}(t) \xrightarrow{\mathbb{P}} c_j(t). \quad (4.133)$$

Ainsi, nous obtenons pour tout j ,

$$\hat{p}_{n,1j}(0, t) \hat{c}_{n,j}(t) \xrightarrow{\mathbb{P}} p_{1j}(0, t) c_j(t),$$

en appliquant la règle de Cramér F.10-1(ii) à (4.132) et à (4.133). D'où :

$$\hat{c}_n(t) \xrightarrow{\mathbb{P}} c(t),$$

d'après la définition (4.122) et la règle de Cramér F.10-1(i). ■

4.7.4 Preuve du Théorème 4.7.2

Tout d'abord, supposons que les conditions (E.9), (E.10) et (E.11) soient satisfaites. Alors le Théorème E.2.1 entraîne que

$$\sqrt{n} (\hat{p}_{n,1j}(0, t) - p_{1j}(0, t)) \xrightarrow{\mathcal{L}} N(0, \sigma_j^2(t)), \quad (4.134)$$

où $\hat{p}_{n,1j}(0, t)$ représente la première ligne de l'estimateur produit-limite $\hat{\mathbf{P}}(s, t)$ donné par (E.4) pour $s = 0$ et

$$\sigma_j^2(t) = \int_0^t p_{1j}(0, u) dU_{1j}(u) p_{1j}(u, t), \quad (4.135)$$

avec des martingales U_{1j} Gaussiennes indépendantes.

Par ailleurs, sous les hypothèses (H1)-(H3) pour $m = 0$ et (H4)-(H5)-(H6) et si lorsque $n \rightarrow +\infty$, $h_n \rightarrow 0$ suffisamment lentement tel que $a_n^2 h_n \rightarrow +\infty$ et suffisamment vite tel que $a_n^2 h_n^3 \rightarrow 0$, on a pour $j \neq k$:

$$\sqrt{a_n^2 h_n} (\hat{\mu}_{n,jk}(t) - \mu_{jk}(t)) \xrightarrow{\mathcal{L}} N(0, \sigma_{jk}^2(t)), \quad (4.136)$$

avec $\sigma_{jk}^2(t)$ donné par (4.130) d'après le Théorème 4.4.2 pour $m = 0$, puis en substituant

$\hat{\mu}_{n,jk}$ à $\hat{\alpha}_n$, μ_{jk} à α et y_j à y . L'utilisation conjointe de (4.117), (4.123), (4.136) et de l'indépendance des intensités de transition conduit à :

$$\sqrt{a_n^2 h_n} (\hat{c}_{n,j}(t) - c_j(t)) \xrightarrow{\mathcal{L}} \text{N} \left(0, \sum_{k; k \neq j} R_{jk}^*(t)^2 \sigma_{jk}^2(t) \right). \quad (4.137)$$

Ainsi, nous avons grâce à (4.134) et (4.137) :

$$\left(\sqrt{n} (\hat{p}_{n,1j}(0, t) - p_{1j}(0, t)), \sqrt{a_n^2 h_n} (\hat{c}_{n,j}(t) - c_j(t)) \right) \xrightarrow{\mathcal{L}} \mathbf{T},$$

pour un certain vecteur aléatoire $\mathbf{T} = (T_1, T_2)$. Puisque $\sqrt{a_n^2 h_n}$ converge vers $+\infty$ moins vite que \sqrt{n} , nous obtenons par le Lemme 4.2.1, nous obtenons :

$$\sqrt{a_n^2 h_n} (\hat{p}_{n,1j}(0, t) - p_{1j}(0, t), \hat{c}_{n,j}(t) - c_j(t)) \xrightarrow{\mathcal{L}} (0, T_2),$$

avec $T_2 \stackrel{\mathcal{L}}{=} \text{N} \left(0, \sum_{k; k \neq j} R_{jk}^*(t)^2 \sigma_{jk}^2(t) \right)$. En appliquant la méthode Delta F.3 pour la fonction $\phi(u, v) = u.v$, il vient que

$$\sqrt{a_n^2 h_n} (\hat{p}_{n,1j}(0, t) \hat{c}_{n,j}(t) - p_{1j}(0, t) c_j(t)) \xrightarrow{\mathcal{L}} \text{N} \left(0, \sum_{k; k \neq j} R_{jk}^*(t)^2 \sigma_{jk}^2(t) p_{1j}^2(0, t) \right),$$

puis

$$\sqrt{a_n^2 h_n} (\hat{c}_n(t) - c(t)) \xrightarrow{\mathcal{L}} \text{N} \left(0, \sum_j \sum_{k; k \neq j} R_{jk}^*(t)^2 \sigma_{jk}^2(t) p_{1j}^2(0, t) \right), \quad (4.138)$$

grâce aux définitions (4.121), (4.122) et à la propriété d'une somme de lois normales indépendantes. ■

Conclusion

L'exploration du sujet de thèse a pu mettre en lumière plusieurs atouts. Tout d'abord, il se situe au carrefour des mathématiques, des sciences économiques et de gestion (fiabilité, science actuarielle). Ensuite, il englobe des fonctions appartenant à une même classe, les fonctionnelles de la fonction de répartition. Enfin, il fournit une étude complète grâce à l'estimation de certaines fonctions, mais également de leurs dérivées.

Le premier chapitre s'est focalisé sur l'estimation non paramétrique de la courbe de Lorenz et du temps total de test normalisé dans le cadre de variables aléatoires indépendantes et identiquement distribuées. Le challenge, qui consistait à proposer une estimation lisse de ces deux fonctions, a été rempli. Un estimateur lisse du temps total de test normalisé n'avait pas encore été proposé. La théorie des espaces de Hilbert à noyaux reproduisants décrite dans [4] a permis de construire des estimateurs continus avec de bonnes propriétés asymptotiques (convergence en probabilité, convergence en moyenne quadratique, convergence presque sûre, normalité asymptotique et intervalles de confiance). Deux applications, l'estimation de l'indice de Gini et la détermination du moment optimal de remplacement des composants non réparables à partir de l'estimation du graphique du temps total de test normalisé, ont été étudiées puis des simulations ont illustré la performance des estimateurs.

L'estimation des dérivées des fonctions L_F et T_F n'a jamais été examinée. Le second chapitre a permis de compléter l'étude du chapitre précédent, en construisant les estimateurs lisses des dérivées successives de la fonction L_F et des dérivées d'ordre supérieur ou égal à 2 de la fonction T_F à l'aide de l'approximation polynomiale locale et en prouvant leur bon comportement asymptotique (convergence en probabilité, normalité asymptotique, intervalles de confiance), toujours dans la situation de variables aléatoires indépendantes et identiquement distribuées. Dans ce contexte, l'estimation lisse de la dérivée première du temps total de test normalisé s'obtient grâce à l'estimateur à noyau de la fonction de répartition et sa convergence presque sûre ainsi que sa normalité asymptotique sont prouvées. Nous nous sommes aventurés sur l'étude globale de l'estimation des fonctionnelles linéaires de la fonction de répartition, malheureusement sans succès. L'ordre de convergence des estimateurs semble fluctuer avec l'expression de la fonctionnelle de la fonction de répartition, d'où l'impossibilité de fournir des conditions précises de convergence et de normalité asymptotique pour les estimateurs de toutes les fonctionnelles linéaires de la fonction de répartition. Le biais asymptotique a cependant pu être fourni. Nous avons par ailleurs obtenu les estimateurs polynomiaux locaux de la densité et de toutes ses dérivées successives, qui se sont avérés d'une grande aide pour l'estimation des dérivées d'ordre supérieur ou égal à 2 de la fonction T_F . Enfin, les dérivées première et seconde des fonctions L_F et T_F ont été représentées par des simulations.

Le troisième chapitre s'est concentré sur l'estimation à noyau classique de la densité et de ses dérivées successives. Afin de réduire le biais de ces estimateurs, nous avons montré l'avantage d'employer des noyaux d'ordre supérieur. La convergence en probabilité et la normalité asymptotique de ces nouveaux estimateurs ont été prouvées. Ce chapitre a ensuite motivé l'estimation à noyau des dérivées successives de la densité (ou d'une autre fonction) en montrant qu'elle est utile pour remédier aux effets de bord de l'estimateur de toute dérivée de la densité (ou d'une autre fonction). Nous avons développé une méthode de correction des effets de bord apparaissant notamment lorsque le support d'estimation est compact et qui pourrait être appliquée en analyse de survie.

Le dernier chapitre a traité de l'estimation du taux de hasard et de ses dérivées successives dans le cadre des variables aléatoires indépendantes et identiquement distribuées puis dans le modèle à intensité multiplicative, un cadre général permettant d'inclure la censure et la dépendance des données. Nous avons achevé la procédure décrite dans ([134]) à l'aide de la méthode de convolution et nous avons réussi à adopter la méthode d'approximation locale. La convergence en moyenne quadratique (et donc la convergence en probabilité) et la normalité asymptotique ont été démontrées, d'où la construction possible d'intervalles de confiance pour le taux de hasard et de ses dérivées successives. À partir de ces travaux d'estimation, les bénéfices (taux d'accumulation du surplus) émanant du dispositif de la participation aux bénéfices ont pu être estimés non paramétriquement. En effet, il n'était pas légitime de fixer des hypothèses paramétriques et cette application est innovante dans le cadre de l'assurance vie.

Nous pouvons imaginer quelques extensions au contenu de notre thèse.

Des courbes de concentration s'expriment en fonction de la courbe de Lorenz, comme la courbe de Bonferroni ou la courbe de Leimkuhler utilisée en bibliométrie. Notre travail sur la courbe de Lorenz peut alors s'appliquer à d'autres domaines que les sciences économiques.

Nous aurions pu envisager d'étudier la fiabilité des composants réparables en utilisant le taux de défaillance (taux de hasard dans un modèle à intensité multiplicative).

La correction des effets de bord présentée dans le troisième chapitre est extensible à d'autres fonctions que la densité comme le taux de transition. L'introduction de l'estimation corrigée des effets de bord des taux de transition dans les formules du bénéfice instantané lorsque l'assuré se trouve dans l'état j et du taux d'accumulation du surplus pourrait aboutir à une meilleure estimation de ces deux derniers objets. La méthode de correction peut également être construite à partir de l'approximation polynomiale locale.

De nombreuses interrogations subsistent et pourront faire l'objet de recherches futures. Tout d'abord, l'estimation de la courbe de Lorenz et de l'indice de Gini supposent que les revenus soient positifs, mais dans la réalité les revenus peuvent être négatifs ou nuls. Ensuite, la courbe de Lorenz, l'indice de Gini, le temps total de test normalisé ainsi que les dérivées successives des fonctions L_F et T_F ont été estimées dans la situation de données indépendantes et identiquement distribuées. Comment obtenir une estimation lisse en sortant de ce cadre? Enfin, ne devrait-on pas plutôt employer des processus semi-markoviens dans le dispositif de la participation aux bénéfices? La durée passée par un assuré dans un état a probablement une influence sur l'avenir des personnes mais cela n'est pas pris en compte avec l'usage d'une chaîne de Markov.

Annexe A

Espaces de Hilbert à noyaux reproduisants et approximation polynomiale locale

A.1 Projection

Soit $(\mathbb{H}, \langle \cdot, \cdot \rangle)$ un espace muni d'un produit scalaire et notons sa norme $\| \cdot \|$, c'est-à-dire telle que $\|f\|^2 = \langle f, f \rangle$. Supposons que cet espace soit de Hilbert, ce qui signifie complet pour sa norme (toute suite de Cauchy de cet espace converge en un point dans cet espace). Soit \mathbb{V} un sous-espace fermé de \mathbb{H} . La projection de $f \in \mathbb{H}$ sur \mathbb{V} est l'unique point

$$\Pi_{\mathbb{V}} f \in \mathbb{V} \tag{A.1}$$

satisfaisant les équations normales

$$\langle f - \Pi_{\mathbb{V}} f, g \rangle = 0, \quad \forall g \in \mathbb{V}. \tag{A.2}$$

L'orthogonalité entraîne

$$\|f\|^2 = \|\Pi_{\mathbb{V}} f\|^2 + \|f - \Pi_{\mathbb{V}} f\|^2.$$

A.2 Sous-espaces de dimension finie

Supposons que \mathbb{V} est un espace de dimension finie muni d'une base orthonormale Q_j , $j = 0, \dots, r$, c'est-à-dire telle que

$$\langle Q_i, Q_j \rangle = \delta_{ij}, \tag{A.3}$$

où δ_{ij} est le symbole de Kronecker. Alors la projection (A.1) est de la forme

$$\Pi_{\mathbb{V}} f = \sum_{j=0}^r \alpha_j Q_j, \tag{A.4}$$

et les équations normales (A.2) sont réduites à

$$\langle f - \sum_{j=0}^r \alpha_j Q_j, Q_i \rangle = 0, \quad i = 0, \dots, r.$$

De (A.3), nous obtenons

$$\langle f, Q_i \rangle - \alpha_i \langle Q_i, Q_i \rangle = 0, \quad i = 0, \dots, r,$$

et en déduisons les expressions suivantes pour les coefficients α_j :

$$\alpha_j = \langle f, Q_j \rangle. \quad (\text{A.5})$$

Ainsi, la projection (A.4) a la forme explicite

$$\Pi_{\mathbb{V}} f = \sum_{j=0}^r \langle f, Q_j \rangle Q_j. \quad (\text{A.6})$$

Si $f \in \mathbb{V}$ alors $f = \Pi_{\mathbb{V}} f$, donc la représentation de f par rapport à la base Q_0, \dots, Q_r est

$$f = \sum_{j=0}^r \langle f, Q_j \rangle Q_j. \quad (\text{A.7})$$

Cela signifie que

$$f(v) = \langle f(\cdot), K(\cdot, v) \rangle, \quad (\text{A.8})$$

où $K : \mathbb{V} \times \mathbb{V} \mapsto \mathbb{R}$ est défini par

$$K(u, v) = \sum_{j=0}^r Q_j(u) Q_j(v). \quad (\text{A.9})$$

Définition A.2.1. ([34]) Soient E un espace non vide et \mathbb{C} l'espace des nombres complexes. Une fonction

$$\begin{aligned} K : E \times E &\rightarrow \mathbb{C} \\ (s, t) &\mapsto K(s, t) \end{aligned}$$

est un noyau reproduisant de l'espace de Hilbert \mathcal{H} si et seulement si

- a) $\forall t \in E, K(\cdot, t) \in \mathcal{H}$
- b) $\forall t \in E, \forall \phi \in \mathcal{H}, \langle \phi, K(\cdot, t) \rangle = \phi(t)$.

Un espace de Hilbert de l'espace des fonctions à valeurs complexes possédant un noyau reproduisant est appelé un espace de Hilbert à noyau reproduisant.

Ainsi, \mathbb{V} est un espace de Hilbert à noyau reproduisant noté K .

A.3 Espaces de fonctions

Soient $(\mathbb{X}, \mathcal{X}, \mu)$ un espace mesuré, \mathbb{H} l'espace des fonctions mesurables sur \mathbb{X} à valeurs réelles dont le carré est intégrable et soit le produit scalaire donné par

$$\langle f, g \rangle = \int f g d\mu,$$

donc $\|f\|^2 = \int f^2 d\mu$. Soit \mathbb{V} un sous-espace de dimension finie inclus dans \mathbb{H} et muni d'une base orthonormale Q_0, \dots, Q_r . La projection de $f \in \mathbb{H}$, donnée par (A.6), a

maintenant la représentation

$$\Pi_{\mathbb{V}}f(v) = \sum_{j=0}^r \int f(u) Q_j(u) d\lambda(u) Q_j(v). \quad (\text{A.10})$$

A.4 Sous-espaces de polynômes

Supposons dorénavant que $(\mathbb{X}, \mathcal{X}) = (\mathbb{R}, \mathcal{R})$ (l'espace des réels muni de la tribu borélienne) et

$$d\mu(u) = K_0(u) du,$$

où K_0 est une densité de probabilité. Soit r un entier non nul. En supposant que u^{2r} est intégrable par rapport μ , le sous-espace \mathbb{V} des polynômes de degré inférieur ou égal à r est un sous-espace fermé de \mathbb{H} . Soit Q_j , $j = 0, \dots, r$, une base orthonormale (donc constituée de polynômes orthonormaux). La projection (A.10) devient alors

$$\Pi_{\mathbb{V}}f(v) = \sum_{j=0}^r \int f(u) Q_j(u) d\mu(u) Q_j(v) \quad (\text{A.11})$$

et c'est un polynôme de degré r .

Soit f une fonction \mathbb{C}^{r+1} donc grâce à la formule de Taylor avec reste intégral, on écrit :

$$f(u) = \sum_{j=0}^r \frac{f^{(j)}(0)}{j!} u^j + O(u^{r+1}). \quad (\text{A.12})$$

Dans [4], il est dit que le polynôme de degré r dans le développement de Taylor (A.12) est proche de la projection, d'où :

$$\sum_{j=0}^r \frac{f^{(j)}(0)}{j!} v^j \approx \sum_{j=0}^r \int f(u) Q_j(u) d\mu(u) Q_j(v).$$

En différentiant m fois au point $v = 0$, on a

$$\left. \frac{\partial^m}{\partial v^m} f(v) \right|_{v=0} = f^{(m)}(0) \approx \sum_{j=0}^r \int f(u) Q_j(u) d\mu(u) Q_j^{(m)}(0) \quad (\text{A.13})$$

$$= \int f(u) K_r^{(m)}(u) du, \quad (\text{A.14})$$

où

$$K_r^{(m)}(u) = \sum_{j=0}^r Q_j(u) Q_j^{(m)}(0) K_0(u) = \left. \frac{\partial^m}{\partial v^m} K(u, v) \right|_{v=0} K_0(u). \quad (\text{A.15})$$

Pour $m = 0, \dots, r$,

$$\int u^i K_r^{(m)}(u) du = \begin{cases} m!, & i = m, \\ 0, & i \neq m, \quad i \in \{0, \dots, r\}, \\ C_{r+1}, & i > r, \end{cases} \quad (\text{A.16})$$

avec C_{r+1} une constante finie. Le cas $i > r$ vient du fait que $K_r^{(m)}$ est borné.

A.5 Fonctionnelles de la fonction de répartition

Soit F une fonction de répartition sur \mathbb{R} . La classe des fonctionnelles de la fonction de répartition F est constituée de l'ensemble des fonctions :

$$\begin{aligned} \Phi(\cdot, F) : I \subset \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \Phi(x, F). \end{aligned}$$

Supposons que $\Phi(\cdot, F) \in \mathbb{C}^{r+1}$. Dans le cas particulier $f(\cdot) = \Phi(x+h\cdot, F)$ avec $x \in \mathbb{R}$ et $h > 0$, la formule (A.14) devient

$$\left. \frac{\partial^m}{\partial v^m} \Phi(x+hv, F) \right|_{v=0} \approx \int \Phi(x+hu, F) K_r^{(m)}(u) du,$$

d'où

$$\Phi^{(m)}(x, F) \approx \frac{1}{h^m} \int \Phi(z, F) \frac{1}{h} K_r^{(m)}\left(\frac{z-x}{h}\right) dz, \quad (\text{A.17})$$

d'après le changement de variable $z = x + hu$.

En substituant un estimateur F_n à F , nous obtenons l'estimateur de $\Phi^{(m)}(\cdot, F)$, pour tout $m = 0 \dots, r$, et pour tout $x \in \mathbb{R}$, donné dans [4], que nous notons :

$$\hat{\Phi}_n^{(m)}(x) := \frac{1}{h^{m+1}} \int \Phi(z, F_n) K_r^{(m)}\left(\frac{z-x}{h}\right) dz, \quad (\text{A.18})$$

avec $h = h(n) > 0$ et $K_r^{(m)}$ donné par (A.16).

Annexe B

Calcul de noyaux d'ordre supérieur d'une hiérarchie

Soit \mathbb{H} l'espace des fonctions mesurables et de carré intégrable sur $(\mathbb{R}, \mathcal{R})$ (l'espace des réels muni de la tribu borélienne) à valeurs réelles et

$$d\mu(u) = K_0(u) du,$$

où K_0 est une densité de probabilité. Alors l'espace \mathbb{V} des polynômes de degré inférieur ou égal à r , $r \in \mathbb{N}$, est un (sous-)espace de Hilbert à noyau reproduisant, à condition que u^{2r} est intégrable par rapport à μ .

Soient $(Q_j)_{j=0,\dots,r}$ une base de polynômes orthonormaux et rappelons la formule (A.15) :

$$K_r^{(m)}(u) = \sum_{j=0}^r Q_j(u) Q_j^{(m)}(0) K_0(x), \quad (\text{B.1})$$

pour $m = 0, \dots, r$. Les noyaux $(K_r^{(m)})_{m=0,1,\dots, m \leq r}$, sont d'ordre $(m, r+1)$ et décrivent la hiérarchie de noyaux associée à la densité de probabilité K_0 .

Si K_0 est symétrique, il est mentionné dans [34] que

$$\int u^{r+1} K_r^{(m)}(u) du = C_{r+1} \neq 0$$

uniquement pour $(r+m)$ impair ($r \geq m$). Si $(r+m)$ est impair alors $Q_r^{(m)}(0) = 0$ d'où $K_r^{(m)}(u) = K_{r-1}^{(m)}(u)$ d'après (A.15).

Un choix possible de noyau $K_r^{(m)}$ intervenant dans l'expression (A.18) est alors tel que $r = m + 1$.

Nous nous focalisons maintenant sur la densité d'Epanechnikov, c'est-à-dire la fonction positive, d'intégrale 1, donnée par

$$K_0(x) = \frac{3}{4} (1 - x^2) \mathbf{1}_{[-1,1]}(x)$$

et notre objectif est de déterminer l'expression de $K_r^{(m)}$ pour $m = 0, 1, 2, 3, 4$.

Voici la formule des polynômes de Jacobi :

$$P_j(x) = \frac{1}{2^j} \sum_{i=0}^j C_{j+\alpha}^i C_{j+\beta}^{j-i} (x-1)^{j-i} (x+1)^i.$$

Pour $\alpha = 1$ et $\beta = 1$, il s'agit des polynômes orthogonaux associés à la densité d'Epanchnikov, d'où :

$$P_j(x) = \frac{1}{2^j} \sum_{i=0}^j C_{j+1}^i C_{j+1}^{j-i} (x-1)^{j-i} (x+1)^i.$$

Ainsi,

$$\{P_j(x)\}_{j=0,1,2,3,4} = \left\{ 1 ; 2x ; \frac{3}{4}(5x^2 - 1) ; 7x^3 - 3x ; \frac{5}{8}(21x^4 - 14x^2 + 1) \right\}.$$

Or la norme de P_j s'écrit

$$\|P_j\| = \left(\int Q_j^2(x) K_0(x) \right)^{\frac{1}{2}}$$

donc

$$\{\|P_j\|\}_{j=0,1,2,3,4} = \left\{ 1 ; \frac{2}{\sqrt{5}} ; \frac{3}{\sqrt{14}} ; \sqrt{\frac{8}{15}} ; \sqrt{\frac{5}{11}} \right\}.$$

Les polynômes orthonormaux Q_j s'obtiennent grâce à la formule

$$Q_j(x) = \frac{P_j(x)}{\|P_j\|}$$

d'où

$$\{Q_j(x)\}_{j=0,1,2,3,4} = \left\{ 1 ; \sqrt{5}x ; \frac{\sqrt{14}}{4}(5x^2 - 1) ; \sqrt{\frac{15}{8}}(7x^3 - 3x) ; \sqrt{\frac{11}{5}} \frac{5}{8}(21x^4 - 14x^2 + 1) \right\}.$$

À l'aide de dérivations successives, nous écrivons

$$\{Q_j^{(1)}(x)\}_{j=0,1,2,3,4} = \left\{ 0 ; \sqrt{5} ; \frac{5\sqrt{14}}{2}x ; 7\sqrt{\frac{15}{8}}(3x^2 - 1) ; \sqrt{\frac{11}{5}} \frac{5}{8}(84x^3 - 28x) \right\}$$

$$\{Q_j^{(2)}(x)\}_{j=0,1,2,3,4} = \left\{ 0 ; 0 ; \frac{5\sqrt{14}}{2} ; 42\sqrt{\frac{15}{8}}x ; \sqrt{\frac{11}{5}} \frac{5}{8}(252x^2 - 28) \right\}$$

$$\{Q_j^{(3)}(x)\}_{j=0,1,2,3,4} = \left\{ 0 ; 0 ; 0 ; 42\sqrt{\frac{15}{8}} ; 504\sqrt{\frac{11}{5}} \frac{5}{8}x \right\}$$

$$\{Q_j^{(4)}(x)\}_{j=0,1,2,3,4} = \left\{ 0 ; 0 ; 0 ; 0 ; 504\sqrt{\frac{11}{5}} \frac{5}{8} \right\}$$

et en particulier pour $x = 0$, nous avons

$$\begin{aligned} \{Q_j^{(0)}(0)\}_{i=0,1,2,3,4} &= \left\{ 1 ; 0 ; -\frac{\sqrt{14}}{4} ; 0 ; \sqrt{\frac{11}{5} \frac{5}{8}} \right\} \\ \{Q_j^{(1)}(0)\}_{i=0,1,2,3,4} &= \left\{ 0 ; \sqrt{5} ; 0 ; -7\sqrt{\frac{15}{8}} ; 0 \right\} \\ \{Q_j^{(2)}(0)\}_{i=0,1,2,3,4} &= \left\{ 0 ; 0 ; \frac{5\sqrt{14}}{2} ; 0 ; -7\sqrt{\frac{11}{5} \frac{5}{2}} \right\} \\ \{Q_j^{(3)}(0)\}_{i=0,1,2,3,4} &= \left\{ 0 ; 0 ; 0 ; 42\sqrt{\frac{15}{8}} ; 0 \right\} \\ \{Q_j^{(4)}(0)\}_{i=0,1,2,3,4} &= \left\{ 0 ; 0 ; 0 ; 0 ; 504\sqrt{\frac{11}{5} \frac{5}{8}} \right\}. \end{aligned}$$

Avec (B.1), nous calculons et obtenons :

- pour $m = 0$:

$$\begin{aligned} K_0^{(0)}(x) = K_1^{(0)}(x) &= \frac{3}{4} (1 - x^2) \mathbf{1}_{[-1;1]}(x) = K_0(x) \\ K_2^{(0)}(x) = K_3^{(0)}(x) &= \frac{15}{32} (7x^4 - 10x^2 + 3) \mathbf{1}_{[-1;1]}(x) \\ K_4^{(0)}(x) = K_5^{(0)}(x) &= \frac{15}{256} (-231x^6 + 441x^4 - 245x^2 + 35) \mathbf{1}_{[-1;1]}(x) \end{aligned}$$

- pour $m = 1$:

$$\begin{aligned} K_1^{(1)}(x) = K_2^{(1)}(x) &= \frac{15}{4} (-x^3 - x) \mathbf{1}_{[-1;1]}(x) \\ K_3^{(1)}(x) = K_4^{(1)}(x) &= \frac{15}{32} (147x^5 - 218x^3 + 71x) \mathbf{1}_{[-1;1]}(x) \end{aligned}$$

- pour $m = 2$:

$$\begin{aligned} K_2^{(2)}(x) = K_3^{(2)}(x) &= \frac{105}{16} (-5x^4 + 6x^2 - 1) \mathbf{1}_{[-1;1]}(x) \\ K_4^{(2)}(x) = K_5^{(2)}(x) &= \frac{45}{64} (539x^6 - 945x^4 + 441x^2 - 35) \mathbf{1}_{[-1;1]}(x) \end{aligned}$$

- pour $m = 3$:

$$K_3^{(3)}(x) = K_4^{(3)}(x) = \frac{945}{16} (-7x^5 + 10x^3 - 3x) \mathbf{1}_{[-1;1]}(x)$$

- pour $m = 4$:

$$K_4^{(4)}(x) = K_5^{(4)}(x) = \frac{10395}{32} (-21x^6 + 35x^4 - 15x^2 + 1) \mathbf{1}_{[-1;1]}(x).$$

Remarque :

Les polynômes $(Q_j)_{j=0,\dots,r}$ peuvent s'obtenir sans passer par la formule des polynômes de Jacobi. Dans ce cas, il faut appliquer l'orthonormalisation de Gram-Schmidt.

Annexe C

Convolution

C.1 Convolutions

Soient F et G des mesures finies sur $(\mathbb{R}, \mathcal{R})$, l'espace des réels muni de la tribu borélienne. Nous nous permettons le léger abus de notation $F(t) = F((-\infty, t])$. La mesure produit de F et G est la mesure $F \times G$ définie sur l'ensemble des boréliens de $\mathbb{R} \times \mathbb{R}$ telle que par $(F \times G)(A \times B) := F(A)G(B)$ pour tous $A, B \in \mathcal{R}$. La *convolution* de F et G est la mesure $F * G$ générée sur $(\mathbb{R}, \mathcal{R})$ par l'application $T : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ définie par $T(r, s) = r + s$:

$$(F * G)(A) = (F \times G)(T^{-1}(A)).$$

Comme T est symétrique, on a $F * G = G * F$, ce qui peut également se démontrer par calcul direct :

$$\begin{aligned}(F * G)(t) &= (F \times G)(T^{-1}(t)) \\ &= (F \times G)(\{(r, s) ; r + s \leq t\}) \\ &= \int_{s \in \mathbb{R}} \int_{r \leq t-s} dG(r) dF(s),\end{aligned}$$

d'après le théorème de Fubini. D'où

$$(F * G)(t) = \int G(t-s) dF(s) = \int F(t-s) dG(s) = (G * F)(t). \quad (\text{C.1})$$

Si F et G sont des mesures de probabilité, alors $F \times G$ est la distribution jointe des variables aléatoires indépendantes X et Y ayant pour distributions respectives F et G et $F * G$ est la distribution de $X + Y$.

On sait que si au moins une des mesures F et G est absolument continue par rapport à la mesure de Lebesgue, alors la mesure $F * G$ l'est également et, si $dG(t) = g(t) dt$,

$$d(F * G)(t) = \int g(t-s) dF(s) dt.$$

Ainsi, si F et G sont absolument continues avec densités f et g , respectivement, alors la densité de $(F * G)$ est

$$(F * G)^{(1)}(t) = \int g(t-s) f(s) ds = \int f(t-s) g(s) ds. \quad (\text{C.2})$$

Si, de plus, au moins une des densités f et g est continue, alors $F * G$ est différentiable et (C.2) est sa première dérivée.

Les intégrales dans (C.2) correspondent à la *convolution* de f et g notée $f * g = g * f$. Ainsi,

$$(f * g)(t) = \int g(t - s) f(s) ds = \int f(t - s) g(s) ds \quad (\text{C.3})$$

pour toutes fonctions $f, g : \mathbb{R} \mapsto \mathbb{R}$ telles que les intégrales existent. L'équivalence des deux expressions dans (C.3) s'obtient par substitution $u = t - s$. La terminologie probabiliste et la notation dans (C.1) n'est pas parfaitement conforme à la terminologie mathématique et la notation dans (C.3). Les deux étant pratiques, nous les utiliserons de manière interchangeable.

Supposons que g est r fois continûment différentiable, et notons sa m^e dérivée $g^{(m)}$ avec la convention $g^{(0)} = g$. Supposons de plus que g s'annule hors d'un intervalle fini. Alors $f * g$ est r fois continûment différentiable et

$$(f * g)^{(m)}(t) = \int g^{(m)}(t - s) f(s) ds, \quad (\text{C.4})$$

$m = 0, \dots, r$. La preuve s'obtient par récurrence. Pour $m = 0$, la formule est simplement (C.3). Supposons qu'elle est vraie pour un $m \in \{0, \dots, r - 1\}$ fixé, formons le quotient des différences

$$\begin{aligned} \frac{(f * g)^{(m)}(t + h) - (f * g)^{(m)}(t)}{h} &= \int \frac{g^{(m)}(t + h - s) - g^{(m)}(t - s)}{h} f(s) ds \\ &= \int g^{(m+1)}(t + h' - s) f(s) ds, \end{aligned}$$

la dernière ligne résultant de la formule de Taylor-Lagrange avec h' compris entre 0 et h . Comme $g^{(m+1)}$ est continue et s'annule hors de son support fini, elle est bornée. De plus, grâce au théorème de convergence dominée, la limite quand h tend vers 0 existe, et nous arrivons à (C.4) avec $m + 1$ à la place de m . La continuité de $(f * g)^{(m)}$ est triviale.

Si de plus f est r fois continûment différentiable, alors nous avons également

$$(f * g)^{(m)}(t) = \int f^{(m)}(t - s) g(s) ds, \quad (\text{C.5})$$

$m = 0, \dots, r$. Cela découle également de la convergence dominée puisque $f^{(m)}(t - s)$ (fonction de s pour t fixé) est bornée sur le support fini de g . Constatons que nous avons également

$$(f * g)^{(m)}(t) = \int f^{(k)}(t - s) g^{(m-k)}(s) ds$$

pour tout $k = 0, \dots, m$.

C.2 Estimation à noyau

Soit F une fonction de répartition $(r+1)$ fois continûment différentiable. Soient X_1, X_2, \dots des variables aléatoires i.i.d. de même fonction de répartition F qu'une variable aléa-

toire X , et notons F_n la fonction de répartition empirique du n -échantillon X_1, \dots, X_n .

Soit k une fonction bornée de support $[-1, 1]$, de r^e dérivée continue $k^{(r)}$, et telle que $\int k(u) du = 1$. Si elle est positive, alors $K(u) = \int_{-\infty}^u k(s) ds$ est une fonction de répartition. Soit $(h_n)_{n=1,2,\dots}$ une suite décroissante de nombres strictement positifs. Introduisons la fonction G_n définie par

$$G_n(t) = K\left(\frac{t}{h_n}\right).$$

Sa première dérivée est g_n donnée par

$$g_n(t) = \frac{1}{h_n} k\left(\frac{t}{h_n}\right).$$

Cette dérivée a pour support $[-h_n, h_n]$, son intégrale est égale à 1 (donc est une densité de probabilité si elle est positive), et sa m^e dérivée est

$$g_n^{(m)}(t) = \frac{1}{h_n^{m+1}} k^{(m)}\left(\frac{t}{h_n}\right), \quad (\text{C.6})$$

$m = 0, \dots, r$.

L'estimateur à noyau de F basé sur le n -échantillon X_1, \dots, X_n , le noyau k et la fenêtre h_n , est la fonction de répartition \hat{F}_n obtenue par convolution de F_n et G_n :

$$\hat{F}_n(t) = (F_n * G_n)(t) = \int G_n(t-s) dF_n(s) = \frac{1}{n} \sum_{i=1}^n G_n(t - X_i). \quad (\text{C.7})$$

L'estimateur hérite des propriétés de K et est $(r+1)$ fois continûment différentiable. Cela motive l'utilisation de \hat{F}_n comme estimateur de F car F est supposée $(r+1)$ fois continûment différentiable. Alors que la fonction de répartition empirique F_n n'est pas différentiable sur l'ensemble $\{X_1, \dots, X_n\}$ et a des dérivées triviales égales à 0 hors de cet ensemble, l'estimateur à noyau \hat{F}_n est lisse jusqu'à l'ordre $(r+1)$, et sa m^e dérivée $\hat{F}_n^{(m)}$ est en plus un estimateur candidat de $F^{(m)}$:

$$\hat{F}_n^{(m)}(t) = \frac{1}{n} \sum_{i=1}^n G_n^{(m)}(t - X_i) = \frac{1}{n} \sum_{i=1}^n g_n^{(m-1)}(t - X_i). \quad (\text{C.8})$$

Ainsi, le programme est de déterminer les conditions sur la suite de fenêtres permettant d'obtenir un estimateur consistant, non biaisé, asymptotiquement normal, etc.

Le but premier de l'estimation à noyau est d'obtenir des estimateurs de la densité et de ses dérivées.

La fonction de répartition empirique F_n est déjà un estimateur de la fonction de répartition F puisqu'elle appartient à l'espace des fonctions de répartition, et converge vers F en plusieurs sens. Premièrement, le théorème de Glivenko-Cantelli F.5 affirme que $F_n(t)$ converge uniformément vers $F(t)$ presque sûrement. Il s'ensuit que $F_n(t)$ converge presque sûrement vers $F(t)$ en tout point de continuité t de F (faible convergence), ce qui signifie que $\int \phi(s) dF_n(s)$ converge presque sûrement vers $\int \phi(s) dF(s)$ pour toute fonction continue bornée ϕ . Cependant, comme $\int \phi(s) dF_n(s)$ est juste la moyenne

empirique $\frac{1}{n} \sum_{i=1}^n \phi(X_i)$, la loi forte des grands nombres (F.6) donne le résultat beaucoup plus fort : $\int \phi(s) dF_n(s)$ converge presque sûrement vers $\mathbb{E} \phi(X) = \int \phi(s) dF(s)$ si cette intégrale existe. Deuxièmement, $F_n(t)$ converge vers $F(t)$ au sens \mathbb{L}^2 pour tout $t \in \mathbb{R}$: $\mathbb{E} [F_n(t) - F(t)]^2 \rightarrow 0$. Troisièmement et dernièrement, le processus stochastique $\sqrt{n} (F_n(t) - F(t))$, $t \in \mathbb{R}$, converge faiblement vers un processus Gaussien Y d'espérance 0 et de $\text{Cov}[Y(s), Y(t)] = F(\min(s, t)) [1 - F(\max(s, t))]$, voir [35].

Par conséquent, l'estimateur à noyau \hat{F}_n va seulement servir à lisser l'estimateur F_n afin d'améliorer son aspect. Puisque G_n converge vers la distribution de Dirac δ_0 , \hat{F}_n a les mêmes propriétés asymptotiques que F_n . Pour tout n fini, cela peut néanmoins introduire un léger biais que nous devons expliquer. Si K est une distribution de probabilité, alors G_n l'est également, et $\hat{F}_n^{(m)}$ est la distribution de $X_n^* + Y_n$, où $\mathcal{L}(X_n^* | X_1, X_2, \dots) = F_n$, $\mathcal{L}(Y_n) = G_n$, et Y_n est indépendant de X_n^*, X_1, X_2, \dots . Comme F_n converge vers F et comme Y_n converge vers 0 en probabilité, \hat{F}_n converge vers F . Cependant, pour tout n fixé, la distribution de $X_n^* + Y_n$ est plus dispersée que F : si $\mathbb{E} Y_n = 0$, alors

$$\mathbb{E}[X_n^* + Y_n] = \mathbb{E} X_n^* = \mathbb{E} \mathbb{E}[X_n^* | X_1, \dots, X_n] = \mathbb{E} \bar{X}_n = \mathbb{E} X,$$

et

$$\begin{aligned} \text{Var} [X_n^* + Y_n] &= \text{Var} X_n^* + \text{Var} Y_n \\ &= \text{Var} \mathbb{E}[X_n^* | X_1, \dots, X_n] + \mathbb{E} \text{Var}[X_n^* | X_1, \dots, X_n] + \text{Var} Y_n \\ &= \text{Var} \bar{X}_n + \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \right] + \text{Var} Y_n \\ &= \text{Var} X + \text{Var} Y_n \\ &> \text{Var} X. \end{aligned}$$

Si nous sommes intéressés par l'estimation de $\Phi(F)$, où Φ est une fonctionnelle définie pour toute fonction de répartition, alors nous utiliserons l'estimateur $\Phi(F_n)$.

Si Φ est définie uniquement pour une certaine sous-classe de fonctions de répartition, alors il sera naturel d'utiliser un estimateur $\Phi(\hat{F}_n)$, où \hat{F}_n appartiendra à cette sous-classe et sera proche de F_n en un certain sens. En particulier, si F est supposée lisse jusqu'à un certain ordre, alors l'estimateur à noyau offre l'opportunité d'estimer ses dérivées, c'est-à-dire de la densité ainsi que de ses dérivées successives.

Annexe D

Martingales

Soit $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espace probabilisé filtré satisfaisant les conditions habituelles de continuité à droite et de complétion de \mathbf{F} . On note \mathbb{E} l'espérance par rapport à la probabilité \mathbb{P} et t représente un réel positif.

Un processus stochastique M continu à droite et limité à gauche est une *martingale* s'il est *adapté* à la filtration $\{\mathcal{F}_t, t \geq 0\}$, *intégrable* et vérifiant la *propriété des martingales*. Autrement dit,

- $M(t)$ est \mathcal{F}_t -mesurable pour tout $t \geq 0$
- $\mathbb{E} |M(t)| < \infty$ pour tout $t \geq 0$
- $\mathbb{E}[M(t) | \mathcal{F}_s] = M(s)$ pour tout $0 \leq s \leq t$.

Annexe E

Estimation de probabilités de transition d'un processus de Markov non homogène

E.1 Définition de l'estimateur de Aalen-Johansen

Soient t un réel positif et $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espace probabilisé filtré satisfaisant les conditions habituelles de continuité à droite et de complétion de \mathbf{F} . Nous introduisons un processus de Markov à temps continu avec r états possibles (l'espace d'états sera noté $\mathcal{E} = \{1, \dots, r\}$) et non homogène (ne dépendant pas de la durée passée dans un état) puis le processus de comptage associé $\mathbf{N}(t) = (N_{jk}(t))_{j,k \in \mathcal{E}}$, où $N_{jk}(t)$ représente le nombre de transitions (sauts) du processus de Markov de l'état j vers l'état k à la date $t \geq 0$.

La matrice des probabilités de transition $\mathbf{P}(s, t) = (p_{jk}(s, t))_{j,k \in \mathcal{E}}$ pour $s \leq t$ du processus de Markov peut être décrite par le produit-intégral

$$\mathbf{P}(s, t) = \prod_{u \in]s, t]} (\mathbf{I} + d\mathbf{A}(u)), \quad (\text{E.1})$$

où \mathbf{I} est la matrice identité et $\mathbf{A}(t) = (A_{jk}(t))_{j,k \in \mathcal{E}}$ est la matrice des intensités cumulées de transition à la date t , c'est-à-dire

$$A_{jk}(t) = \int_0^t \mu_{jk}(t) dt, \quad (\text{E.2})$$

avec $\mu_{jk}(t)$ l'intensité de transition de l'état j vers l'état k à la date t .

Désignons par $Y_j(t)$ le nombre d'individus présents dans l'état $j \in \mathcal{E}$ à la date t et définissons le processus $J_j(t) = \mathbf{1}_{[Y_j(t) > 0]}$. Alors l'estimateur de Nelson-Aalen $\hat{A}_{jk}(t)$ de l'intensité cumulée de transition $A_{jk}(t)$ pour tous $j, k \in \mathcal{E}$ s'écrit

$$\hat{A}_{jk}(t) = \int_0^t \frac{J_j(s)}{Y_j(s)} dN_{jk}(s). \quad (\text{E.3})$$

En situation de censure à droite indépendante et de troncature à gauche indépendante, Aalen-Johansen ([2]) a naturellement proposé d'estimer la matrice des probabilités de

transition par une matrice $\hat{\mathbf{P}}(s, t) = (\hat{p}_{jk}(s, t))_{j, k \in \mathcal{E}}$ pour $s \leq t$ de la même forme que (E.1) :

$$\hat{\mathbf{P}}(s, t) = \prod_{u \in]s, t]} (\mathbf{I} + d \hat{\mathbf{A}}(u)), \quad (\text{E.4})$$

où $\hat{\mathbf{A}}(t) = (\hat{A}_{jk}(t))_{j, k \in \mathcal{E}}$. La matrice $\hat{\mathbf{P}}(s, t)$ pour $s \leq t$ se nomme l'*estimateur de Aalen-Johansen* de la matrice des probabilités de transition $\mathbf{P}(s, t)$ pour $s \leq t$.

E.2 Propriétés asymptotiques de l'estimateur de Aalen-Johansen

Nous énonçons deux théorèmes sur le comportement asymptotique de l'estimateur de Aalen-Johansen (E.4) extraits de ([11]). Certaines notations ont été modifiées afin de faciliter l'utilisation des théorèmes dans les Sous-sections 4.7.2, 4.7.3 et 4.7.4.

Soit τ le plus grand réel positif w satisfaisant $\int_0^w \mu_{jk}(u) du < +\infty$ pour $j, k \in \mathcal{E}$ tels que $j \neq k$. La norme $\| \cdot \|$ d'une matrice $B = (b_{jk})_{j, k \in \mathcal{E}}$ est définie par

$$\| B \| = \sup_j \sum_k |b_{jk}|.$$

Le premier théorème présente la convergence uniforme de l'estimateur de Aalen-Johansen.

Théorème E.2.1. (Théorème IV.4.1 dans [11])

Soient $s, v \in [0, \tau]$ avec $s < v$ et supposons que lorsque $n \rightarrow +\infty$, pour tout $j \neq k$,

$$\int_s^v (1 - J_{n,j}(u)) \mu_{jk}(u) du \xrightarrow{\mathbb{P}} 0 \quad (\text{E.5})$$

et

$$\int_s^v \frac{J_{n,j}(u)}{Y_{n,j}(u)} \mu_{jk}(u) du \xrightarrow{\mathbb{P}} 0. \quad (\text{E.6})$$

Alors

$$\sup_{t \in [s, v]} \|\hat{\mathbf{P}}(s, t) - \mathbf{P}(s, t)\| \xrightarrow{\mathbb{P}} 0. \quad (\text{E.7})$$

Le second théorème concerne la convergence en loi de l'estimateur de Aalen-Johansen.

Théorème E.2.2. (Théorème IV.4.2 dans [11])

Soient $s, v \in [0, \tau]$ avec $s < v$ et supposons qu'il existe des fonctions positives y_j définies sur $[s, v]$ telles que $\frac{\mu_{jk}}{y_j}$ est intégrable sur $[s, v]$ pour tout $j \neq k$. Soit

$$\rho_{jk}(t) = \int_s^v \frac{\mu_{jk}(u)}{y_j(u)} du \quad (\text{E.8})$$

pour tout $j \neq k$ et supposons que :

(A) Pour tout $t \in [s, v]$ et tout $j \neq k$,

$$n \int_s^v \frac{J_{n,j}(u)}{Y_{n,j}(u)} \mu_{hj}(u) du \xrightarrow{\mathbb{P}} \rho_{jk}(t) \text{ lorsque } n \rightarrow +\infty \quad (\text{E.9})$$

(B) Pour tout $\epsilon > 0$ et tout $h \neq j$,

$$n \int_s^v \frac{J_{n,j}(u)}{Y_{n,j}(u)} \mu_{jk}(u) \mathbf{1}_{\left[\left|\sqrt{n} \frac{J_{n,j}(u)}{Y_{n,j}(u)}\right| > \epsilon\right]} du \xrightarrow{\mathbb{P}} 0 \text{ lorsque } n \rightarrow +\infty \quad (\text{E.10})$$

(C) Pour tout $h \neq j$,

$$\sqrt{n} \int_s^v (1 - J_{n,j}(u)) \mu_{jk}(u) du \xrightarrow{\mathbb{P}} 0 \text{ lorsque } n \rightarrow +\infty. \quad (\text{E.11})$$

Alors

$$\sqrt{n} (\hat{\mathbf{P}}(s, \cdot) - \mathbf{P}(s, \cdot)) \xrightarrow{\mathcal{D}} \int_s^\cdot \mathbf{P}(s, u) d\mathbf{U}(u) \mathbf{P}(u, \cdot), \quad (\text{E.12})$$

où $\mathbf{U} = (U_{jk})$ est une matrice $r \times r$ telle que pour $j \neq k$, U_{jk} sont des martingales gaussiennes indépendantes avec $U_{jk}(0) = 0$ et $\text{cov}(U_{jk}(t_1), U_{jk}(t_2)) = \rho_{jk}(\max(t_1, t_2))$ et $U_{jj} = -\sum_{k \neq j} U_{jk}$.

Remarque :

Les conditions des Théorème E.2.1 et Théorème E.2.2 sont respectivement identiques à celles de la convergence uniforme et de la convergence en loi de l'estimateur de Nelson-Aalen, voir [11].

Annexe F

Théorèmes et résultats utiles

F.1 Loi faible des grands nombres

Soit (X_n) une suite de n variables aléatoires indépendantes et de même espérance finie. Alors

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mathbb{E}[X_1].$$

F.2 Théorème de la limite centrale multidimensionnel

Soient X_1, \dots, X_n des vecteurs aléatoires indépendants et de même loi dans \mathbb{R}^d , d'espérance μ et de matrice de variance-covariance Σ . Alors

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i - \mu \right) \xrightarrow{\mathcal{L}} N(0, \Sigma).$$

F.3 Méthode Delta

Soit $\phi : \mathbb{R}^k \mapsto \mathbb{R}^m$ une fonction définie sur un sous-espace de \mathbb{R}^k et différentiable en θ . Soient \mathbf{T}_n des vecteurs aléatoires prenant leurs valeurs dans le domaine de ϕ .

Lorsque $\sqrt{n}(\mathbf{T}_n - \theta)$ converge vers une loi normale multivariée $N_k(\mu, \Sigma)$, la suite $\sqrt{n}(\phi(\mathbf{T}_n) - \phi(\theta))$ converge vers la loi $N_m(\phi'_\theta \mu, \phi'_\theta \Sigma (\phi'_\theta)^T)$, où ϕ'_θ correspond à la matrice jacobienne de la fonction ϕ en θ .

Remarque :
Voir [157].

F.4 Théorème de convergence dominée de Lebesgue

Soit (f_n) une suite de fonctions mesurables de E dans \mathbb{R} telle que la suite de fonctions (f_n) converge simplement vers une fonction f sur E et telle qu'il existe une fonction

intégrable g définie sur E vérifiant

$$\forall n \in \mathbb{N}, \forall x \in E, |f_n(x)| \leq g(x).$$

Alors pour tout $n \in \mathbb{N}$, f_n est intégrable sur E , f est intégrable sur E et

$$\lim_{n \rightarrow +\infty} \int_E f_n(x) dx = \int_E f(x) dx.$$

F.5 Théorème de Glivenko-Cantelli

Soient X_1, \dots, X_n , n variables aléatoires i.i.d. de fonction de répartition commune F . La fonction de répartition F_n associée à ce n -échantillon est définie par :

$$F_n(x) = \sum_{i=1}^n \mathbf{1}_{[X_i \leq x]}.$$

Alors :

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{p.s.} 0.$$

F.6 Loi forte des grands nombres

Soit (X_n) une suite de n variables aléatoires i.i.d. telles que $\mathbb{E}|X_1| < +\infty$. Alors

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} \mathbb{E}[X_1].$$

F.7 Théorème des fonctions continues

Soit $g : \mathbb{R}^k \mapsto \mathbb{R}^m$ une fonction continue en tout point d'un ensemble C tel que $\mathbb{P}[X \in C] = 1$. Alors :

- (i) Si $X_n \xrightarrow{\mathcal{L}} X$, alors $g(X_n) \xrightarrow{\mathcal{L}} g(X)$,
- (ii) Si $X_n \xrightarrow{\mathbb{P}} X$, alors $g(X_n) \xrightarrow{\mathbb{P}} g(X)$,
- (iii) Si $X_n \xrightarrow{p.s.} X$, alors $g(X_n) \xrightarrow{p.s.} g(X)$.

Remarque :
Voir [157].

F.8 Convergence en probabilité et convergence \mathbb{L}^p , $p \in \mathbb{N}^*$

Soit $p \in \mathbb{N}^*$. La convergence dans \mathbb{L}^p entraîne la convergence en probabilité.

La réciproque n'est pas vraie en général. Soit (X_n) une suite de variables aléatoires convergeant en probabilité vers une variable aléatoire X . Si $\mathbb{P}[|X_n| \leq b] = 1$ pour tout

n et un certain $b \in \mathbb{R}$, alors la suite (X_n) converge également vers X dans \mathbb{L}^p (on dit aussi que (X_n) converge en moyenne d'ordre p vers X) pour tout $p \in \mathbb{N}^*$.

F.9 Théorème de Lindeberg-Feller

Pour tout $n = 1, 2, \dots$ soit $\mathbf{Y}_{n1}, \dots, \mathbf{Y}_{n,k_n}$ des d -vecteurs aléatoires indépendants de carré intégrable satisfaisant les conditions

(i)

$$\forall \epsilon > 0, \sum_{i=1}^{k_n} \mathbb{E} \left[\|\mathbf{Y}_{ni}\|^2 \mathbf{1}_{\{\|\mathbf{Y}_{ni}\| > \epsilon\}} \right] \rightarrow 0 \quad (\text{Lindeberg-Feller})$$

et

(ii)

$$\sum_{i=1}^{k_n} \text{Cov} \mathbf{Y}_{ni} \rightarrow \Sigma \neq \mathbf{0}$$

lorsque $n \rightarrow +\infty$. Alors le théorème de la limite centrale de Lindeberg-Feller pour tableaux triangulaires affirme que

$$\sum_{i=1}^{k_n} (\mathbf{Y}_{ni} - \mathbb{E} \mathbf{Y}_{ni}) \xrightarrow{\mathcal{L}} \mathbf{N}(\mathbf{0}, \Sigma).$$

Si $k_n = n$ et $\mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nn}$ sont des copies i.i.d. d'une variable aléatoire \mathbf{Y}_n , alors (i) et (ii) se réduisent à :

(iii)

$$\forall \epsilon > 0, n \mathbb{E} \left[\|\mathbf{Y}_n\|^2 \mathbf{1}_{\{\|\mathbf{Y}_n\| > \epsilon\}} \right] \rightarrow 0$$

et

(iv)

$$n \text{Cov} \mathbf{Y}_n \rightarrow \Sigma \neq \mathbf{0}.$$

Nous donnons deux conditions suffisantes pour (iii) faciles à vérifier. La première s'obtient à l'aide du théorème de la convergence dominée.

Lemme F.9.1 La condition de Lindeberg-Feller (iii) est satisfaite si $\|\mathbf{Y}_n\| \xrightarrow{\mathbb{P}} 0$ et s'il existe une variable aléatoire intégrable \hat{Y} telle que $\sup_n (n \|\mathbf{Y}_n\|^2) \leq \hat{Y}$.

La seconde est particulièrement utile dans les situations considérées ici, et sa preuve résulte clairement par inspection de (iii).

Lemme F.9.2 La condition de Lindeberg-Feller (iii) est satisfaite si, pour tout ϵ positif, $\mathbb{P}[\|\mathbf{Y}_n\| > \epsilon] = 0$ pour n suffisamment grand.

F.10 Règles de Cramér

1. Soient X une variable aléatoire et c une constante. (X_n) et (Y_n) sont deux suites de variables aléatoires telles que $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} c$. Alors :

- (i) $X_n + Y_n \xrightarrow{\mathcal{L}} X + c$,
- (ii) $X_n \cdot Y_n \xrightarrow{\mathcal{L}} c \cdot X$,
- (iii) $\frac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \frac{X}{c}$ si $c \neq 0$.

Remarque :

Grâce au théorème des fonctions continues F.7, ces règles sont vraies en remplaçant la convergence en loi par la convergence en probabilité ou la convergence presque sûre.

2. Soit X une variable aléatoire. (X_n) et (Y_n) sont deux suites de variables aléatoires telles que $X_n \xrightarrow{\mathcal{L}} X$ et $|X_n - Y_n| \xrightarrow{\mathbb{P}} 0$. Alors $Y_n \xrightarrow{\mathcal{L}} X$.

F.11 Inégalité de Markov

Soient X une variable aléatoire et $p \in \mathbb{N}^*$. Alors pour tout $\epsilon > 0$, on a

$$\mathbb{P}(|X| \geq \epsilon) \leq \frac{\mathbb{E}[|X|^p]}{\epsilon^p}.$$

F.12 Inégalité de Jensen

Soit f une fonction convexe sur un intervalle I réel. Supposons que X une variable aléatoire à valeurs dans I telle que $\mathbb{E}[X] < +\infty$. Alors

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

F.13 Inégalité de Cauchy-Schwartz

Soient X et Y deux variables aléatoires réelles telles que $\mathbb{E}[X^2] < +\infty$ et $\mathbb{E}[Y^2] < +\infty$. Alors

$$\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2] \mathbb{E}[Y^2]}.$$

F.14 Théorème de Rebolledo

Soient $a \in \mathbb{R}_-$ pouvant être égal à $-\infty$, $v \geq 0$ et (M_n) une suite de martingales associée à (N_n) une suite de processus de comptage, de compensateurs (Λ_n) .

Introduisons

$$I_n(v) = \int_a^v H_n(u) dM_n(u),$$

avec (H_n) une suite de processus prévisibles vérifiant pour tout $v \geq 0$,

(i)

$$\mathbb{E} I_n(v)^2 < +\infty.$$

Définissons pour tout $\epsilon > 0$,

$$I_n^\epsilon(v) = \int_a^v H_n(u) \mathbf{1}_{[|H_n(u)| > \epsilon]} dM_n(u).$$

Alors si les deux conditions suivantes sont satisfaites :

(ii)

$$\langle I_n(v), I_n(v) \rangle \xrightarrow{\mathbb{P}} V(v),$$

(iii)

$$\langle I_n^\epsilon(v), I_n^\epsilon(v) \rangle \xrightarrow{\mathbb{P}} 0,$$

on en déduit que $I_n(v)$ converge dans $\mathcal{D}(\mathbb{R})$ vers un processus Gaussien d'espérance nulle et de variance $V(v)$, où $\mathcal{D}(\mathbb{R})$ désigne l'espace des fonctions à valeurs réelles, continues à droite, ayant une limite à gauche, muni de la topologie de Skorohod.

Table des figures

1	Schéma reliant les revenus primaires et le revenu disponible.	18
2	Un exemple de courbe de Lorenz (en rouge) avec la ligne d'égalité parfaite (en bleu) et la ligne de parfaite inégalité (en turquoise).	20
3	3 courbes de Lorenz : loi de Pareto (3,5) (en noir), loi de Weibull (3,5.5) (en magenta) et loi de Maxwell (en rouge).	20
4	Composantes de la sûreté de fonctionnement.	22
5	Illustration du temps total de test de n ampoules jusqu'à la k^e défaillance, où k est un entier compris entre 1 et n . Les durées de vie ordonnées des n ampoules sont notées $X(1), \dots, X(n)$	24
6	Durée moyenne de fonctionnement en fonction de la proportion de composants défaillants de plusieurs systèmes de 4 composants.	25
7	TTT-plots associés aux 3 premiers systèmes de la figure précédente. La courbe bleue représente l'absence de tendance.	25
8	Temps total de test normalisé traduisant un accroissement de fiabilité (en rouge), une tendance au vieillissement (en magenta) et ne traduisant aucune tendance (en bleu).	26
1.1	Méthode graphique pour déterminer le moment optimal de remplacement des composants non réparables, à partir du graphique du temps total de test normalisé (en magenta).	47
1.2	Courbe de Lorenz théorique (en rouge) avec son estimateur empirique (en turquoise), son estimateur lisse (en bleu) et la ligne d'égalité parfaite (en noir).	63
1.3	Erreur quadratique intégrée de l'estimateur empirique (Emp.) et de l'estimateur lisse (Poly.) de la courbe de Lorenz pour la loi Exponentielle de paramètre 3 (en haut) puis pour la loi de Maxwell (en bas), en fonction de 3 tailles d'échantillon.	64
1.4	Temps total de test normalisé théorique (en rouge) avec son estimateur empirique (en turquoise), son estimateur lisse (en bleu) et la droite d'absence de tendance (en noir).	65
1.5	Erreur quadratique intégrée de l'estimateur empirique (Emp.) et de l'estimateur lisse (Poly.) du temps total de test normalisé pour la loi Exponentielle de paramètre 3 (en haut) puis pour la loi de Weibull de paramètres 2 et 1.2 (en bas), en fonction de 3 tailles d'échantillon.	65
1.6	Estimateur empirique (en turquoise) et estimateur lisse (en bleu) du temps total de test normalisé des ampoules, avec $h = 0.8$	66

2.1	Dérivée première de la courbe de la fonction L_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Maxwell (à droite).	101
2.2	Dérivée seconde de la courbe de la fonction L_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Maxwell (à droite).	102
2.3	Erreur quadratique intégrée de l'estimateur à noyau de la dérivée première (en haut) et de la dérivée seconde (en bas) de la fonction L_F pour la loi Exponentielle de paramètre 3 (à gauche) puis pour la loi de Maxwell (à droite), en fonction de 3 tailles d'échantillon.	102
2.4	Dérivée première de la fonction T_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Weibull de paramètres 2 et 1.2 (à droite).	103
2.5	Dérivée seconde de la fonction T_F (en rouge) avec son estimateur à noyau (en vert) pour la loi Exponentielle de paramètre 3 (à gauche) et la loi de Weibull de paramètres 2 et 1.2 (à droite).	103
2.6	Erreur quadratique intégrée de l'estimateur à noyau de la dérivée première (en haut) et la dérivée seconde (en bas) de la fonction T_F pour la loi Exponentielle de paramètre 3 (à gauche) puis pour la loi de Weibull de paramètres 2 et 1.2 (à droite), en fonction de 3 tailles d'échantillon.	104
2.7	Estimateur à noyau de la dérivée première de la fonction T_F , avec $h = 0.8$	104
2.8	Estimateur à noyau de la dérivée seconde de la fonction T_F , avec $h = 0.8$	105

Bibliographie

- [1] O. O. Aalen, *Nonparametric inference for a family of counting processes*, Ann. Statist., 6, 701-726, 1978.
- [2] O. O. Aalen, S. Johansen, *An empirical transition matrix for nonhomogeneous Markov chains based on censored observations*, Scand. J. Statist., 5, 141-150, 1978.
- [3] B. Abdous, *On a parametric estimation of a class of survival functionals*, Preprint, 2007.
- [4] B. Abdous, B. Berlinet, N. Hentgartner, *A general theory for kernel estimation of smooth functionals of the distribution function and their derivatives*, Revue Roumaine de Mathématiques Appliquées, 48, 217-232, 2003.
- [5] B. Abdous, A. Berred, *Mean residual life estimation*, JSPI, 132, 3-19, 2005.
- [6] E. I. Abdul-Sathar, E. S. Jeeveland, K. R. Muraleedharan Nair, *Bayesian estimation of Lorenz curve, Gini index and variance of logarithms in a Pareto distribution*, Statistica, 65, 193-205, 2005.
- [7] A. M. Abouammoh, A. Khalique, *Some tests for mean residual criteria based on the total time on test transform*, Reliability Engineering, 19, 85-101, 1987.
- [8] A. Al Hussainan, *Lorenz curve estimation by an adjusted form of the empirical Lorenz curve*, J. Econom. Adm. Sciences, 16, 189-199, 2000.
- [9] S. D'Agostino, G. Trombert, *Les inégalités de revenus, les enjeux d'un partage*, Viubert, 1992.
- [10] H. R. Jr Alker, *Mathematics and Politics*, MacMillan Co., New York, 1965.
- [11] P. K. Andersen, Ø. Borgan, R. Gill, N. Keiding, *Statistical models based on counting processes*, Springer-Verlag, New York, 1993.
- [12] D. R. Anderson, D. J. Sweeney, T. A. Williams, *Statistiques pour l'Economie et la Gestion*, De Boeck, 2010.
- [13] A. B. Atkinson, *On the measurement of inequality*, J. Econom. Theory, 2, 244-263 1970.
- [14] R. E. Barlow, *Geometry of the total time on test transform*, Nav. Res. Logist. Quart., 26, 393-402, 1979.
- [15] R. E. Barlow, D. Bartholomew, J. M. Bremner, H. D. Brunk, *Statistical inference under order restrictions*, Wiley, New York, 1972.
- [16] R. E. Barlow, R. Campo, *Total time on test processes and applications to failure data analysis*, In : Barlow et al. (eds.), Reliability and Fault Tree Analysis, SIAM, Philadelphia, 451-481, 1975.

- [17] R. E. Barlow, B. Davies, *Total time on test plots*, In : Proceedings of the 24th Conference on the Design of Experiments in Army Research Development and Testing, 361-380, 1979.
- [18] R. E. Barlow, K. Doksum, *Isotonic tests for convex orderings*, In : Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, 1, 293-323, 1972.
- [19] R. E. Barlow, F. Proschan, *Asymptotic theory of the total time on test processes with applications to life testing*, In : Krishnaiah (Ed.), Multivariate Analysis-IV, North-Holland, Amsterdam., 1975.
- [20] R. E. Barlow, F. Proschan, *Statistical theory of reliability and life testing*, Holt, Rinehart and Winston, New York, 1981.
- [21] R. E. Barlow, R. H. Toland, T. Freeman, *Stress-rupture life of Kevlar/Epoxy spherical pressure vessels UCID-17755, Part 3*, Lawrence Livermore Laboratories, Livermore, California, 94550, 1979.
- [22] R. E. Barlow, W. R. van Zwet, *Asymptotic properties of isotonic estimators for the generalized failure rate function, Part I : Strong consistency*, In : Nonparametric Techniques in Statistical Inference (M. L. Puri, Ed.), 159-173, Cambridge Univ. Press, Cambridge, 1970.
- [23] J. Bartoszewicz, *Dispersive ordering and the total time on test transformation*, Statist. Probab. Lett., 4, 285-288, 1986.
- [24] J. Bartoszewicz, *Stochastic order relations and the total time on test transform*, Statist. Prob. Lett., 22, 103-110, 1995.
- [25] J. Bartoszewicz, *Tail orderings and the total time on test transform*, Appl. Math., 24, 77-86, 1996.
- [26] R. L. Basmann, K. J. Hayes, D. J. Slottje, *The Lorenz curve and the mobility function*, Econom. Lett., 35, 105-111, 1991.
- [27] F. Belzunce, X. Li, J. F. Pinar, J. M. Ruiz, *Test for the total time on test transform order*, J. Statist. Plan. Inf., 133, 11-121, 2002.
- [28] B. Bergman, *Crossings in the total time on test concept*, Scand. J. Statist., 4, 171-177, 1977.
- [29] B. Bergman, *On age replacement and the total time on test concept*, Scand. J. Statist., 6, 161-168. 1979.
- [30] B. Bergman, B. Klefsjö, *The total time on test concept and its use in reliability theory*, Oper. Res., 32, 596-606, 1984.
- [31] A. Berlinet, *Reproducing kernels and finite order kernels*, In : G.G. Roussas, Ed. Nonparametric Functional Estimation and related topics, NATO ASI Series, Kluwer Academic publishers, Dordrecht, 3-18, 1991.
- [32] A. Berlinet, *Hierarchies of higher order kernels*, Probab. Theory Related Fields, 94, 489- 504, 1993.
- [33] A. Berlinet, L. Devroye, *A comparison of kernel density estimates*, Publications de l'Institut de Statistique de l'Université de Paris, 38, 3-59, 1994.
- [34] A. Berlinet, C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in Probability and Statistics*, Kluwer, Boston, 2004.
- [35] P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.

- [36] F. Bourguignon, *La mondialisation de l'inégalité*, Seuil, La République des idées, 2012.
- [37] Q. L. Burell, *The Bradford distribution and the Gini index*, *Scientometrics*, 21, 181-194, 1991.
- [38] Q. L. Burell, *The Gini index and the Leimkuhler curve for bibliometric processes*, *Information Processing and Management*, 28, 19-33, 1992.
- [39] Q. L. Burell, *Symmetry and other transformation features of Lorenz/Leimkuhler representations of informetric data*, *Information Processing and Management*, 41, 1317-1329, 2005.
- [40] M. Chandra, N. D. Singpurwalla, *The Gini index, the Lorenz curve, and the total time on test transforms*, Unpublished technical report, George Washington University, Washington, 1978.
- [41] M. Chandra, N. D. Singpurwalla, *Relationships between some notions which are common to reliability theory and economics*, *Math. Oper. Res.*, 6, 113-121, 1981.
- [42] F. Chen, R. M. Huggins, P. S. P. Yip, K. P. Lam, *Nonparametric estimation of multiplicative counting process intensity functions with an application to the Beijing SARS epidemic*, *Commun. Stat. – Theory and Methods*, 37, 294–306, 2009.
- [43] M. Y. Cheng, J. Fan, J. S. Marron, *On automatic boundary corrections*, *Ann. Statist.*, 25, 1691-1708, 1997.
- [44] D. Chotikapanich, *Modelling income distributions and Lorenz curves*, Springer, New York, 2008.
- [45] D. Chotikapanich, W. E. Griffiths, *Estimating Lorenz curves using a Dirichet distribution*, *J. Business Econom. Statist.*, 20, 290-295, 2002.
- [46] W. S. Cleveland, *Robust Locally Weighted Regression and Smoothing Scatterplots*, *J. Amer. Stat. Assoc.*, 74, 829-836, 1979.
- [47] D. B. H. Cline, J. D. Hart, *Kernel estimation of densities with discontinuities or discontinuous derivatives*, *Statistics*, 22, 69-84, 1991.
- [48] A. Cowling, P. Hall, *On pseudodata methods for removing boundary effects in kernel density estimation*, *J. Roy. Statist. Soc. Ser. B*, 58, 551-563, 1996.
- [49] M. Csörgő, *Quantile processes with statistical applications*, SIAM, Philadelphia, 1983.
- [50] M. Csörgő, S. Csörgő et L. Horváth, *An asymptotic theory for empirical reliability and concentration processes*, *Lecture Notes in Statistics*, 33, Springer, New York, 1986.
- [51] M. Csörgő, S. Csörgő, L. Horváth, *Estimation of total time on test transforms and Lorenz curves under random censorship*, *Statistics*, 18, 77-97, 1987.
- [52] M. Csörgő, L. Horváth, *On the rate of strong consistency of the total time on test statistic*, *J. Multiv. Anal.*, 66, 99-117, 1998.
- [53] M. Csörgő, P. Révész, *Strong approximations of the quantile process*, *Ann. Statist.*, 6, 882-894, 1978.
- [54] M. Csörgő, H. Yu, *Estimation of total time on test transforms for stationary observations*, *Stoch. Proc. and their Appl.*, 68 229-253, 1997.

- [55] M. Csörgő, H. Yu, *Weak approximations for empirical Lorenz curves and their Goldie inverses of stationary observations*, Adv. Appl. Prob., 31, 698-719, 1999.
- [56] M. Csörgő, R. Zitikis, *Strassen's LIL for the Lorenz curve*, J. Multiv. Anal., 59, 1-12, 1996.
- [57] M. Csörgő, R. Zitikis, *Confidence bands for the Lorenz and Goldie curves*, In : A Volume in Honor of Samuel Kotz, Wiley, New York, 1996.
- [58] M. Csörgő, R. Zitikis, *On the rate of strong consistency of the Lorenz curves*, Statist. Probab. Lett., 34, 113-121, 1997.
- [59] M. Csörgő, R. Zitikis, *On the rate of strong consistency of the total time on test statistic*, J. Multiv. Anal., 66, 99-117, 1998.
- [60] H. Dalton, *The measurement of the inequality of incomes*, Econom. J., 30, 348-361, 1920.
- [61] L. Devroye, *A course in density estimation*, Birkhäuser, Boston, 1987.
- [62] G. Demengel, P. Bénichou, R. Bénichou, N. Boy, J. P. Pouget, *Probabilités, Statistique inférentielle, Fiabilité, Outils pour l'ingénieur*, Ellipses Marketing, 1997.
- [63] M. Denuit, C. Robert, *Actuariat des assurances de personnes : modélisation, tarification et provisionnement*, Economica, 2007.
- [64] D. J. Doiron, G. F. Barrett, *Inequality in male and female earnings : the role of hours and earnings*, Rev. Econom. Statist., 78, 410-420, 1996.
- [65] J. D'Eichtal, *La participation aux bénéfices facultative et obligatoire*, Guillaumin, Paris 1892.
- [66] V. A. Epanechnikov, *Nonparametric estimation of a multivariate probability density*, Theory Prob. Appl., 14, 153-158, 1969.
- [67] B. Epstein, M. Sobel, *Life testing*, J. Amer. Stat. Assoc., 48, 486-502, 1953.
- [68] J. Fan, I. Gijbels, *Local polynomial modelling and its applications*, Chapman et Hall, London, 1996.
- [69] T. R. Fleming, D. P. Harrington, *Counting processes and survival analysis*, Wiley, New-York, 1991.
- [70] T. Gasser, H. G. Müller, *Kernel estimation of regression functions*, In : Smoothing techniques for curve estimation, Lecture Notes in Mathematics, 757, 23-68, Springer-Verlag, Berlin, 1979.
- [71] T. Gasser, H. G. Müller, V. Mammitzsch, *Kernels for nonparametric curve estimation*, J. Roy. Statist. Soc. Ser. B, 47, 238-252, 1985.
- [72] M. H. Gail, J. L. Gastwirth, *A scale-free goodness-of-fit for the exponential distribution based on the Gini statistic*, J. Roy. Statist. Soc. Ser. B, 40, 350-357, 1978.
- [73] M. H. Gail, J. L. Gastwirth, *A scale-free goodness-of-fit for the exponential distribution based on the Lorenz curve*, J. Amer. Stat. Assoc., 73, 787-793, 1978.
- [74] J. L. Gastwirth, *A general definition of the Lorenz curve*, Econometrica, 39, 1037-1039, 1971.
- [75] J. L. Gastwirth, *The estimation of the Lorenz curve and Gini index*, Rev. Econom. and Statist., 54, 306-316, 1972.

- [76] J. L. Gastwirth, R. Modarres, E. Bura, *The use of the Lorenz curve, Gini index and related measures of relative inequality and uniformity in securities law*, Int. J. Statist., 63, 451-469, 2005.
- [77] R. D. Gill, *The total time on test plot and the cumulative total on test statistic for a counting process*, Ann. Statist., 14, 1234-1239, 1986.
- [78] C. M. Goldie, *Convergence theorems for empirical Lorenz curves and their inverses*, Adv. Appl. Prob., 9, 765-791, 1977.
- [79] P. Hall, B. U. Park, *Bandwidth choice for local polynomial estimation of smooth boundaries*, J. Multivar. Anal., 91, 240-261, 2004.
- [80] P. Hall, T. E. Wehrly, *A geometrical method for removing edge effects from kernel-type nonparametric regression estimators*, J. Amer. Stat. Assoc., 86, 665-672, 1991.
- [81] P. E. Hart, *Entropy and other measures of concentration*, J. Roy. Statist. Soc. Ser. A, 134, 73-89, 1971.
- [82] P. E. Hart, *Moment distributions in economics : an exposition*, J. Roy. Statist. Soc. Ser. A, 138, 423-434, 1975.
- [83] H. Hasegawa, H. Kozumi, H., *Estimation of Lorenz Curves : a Bayesian nonparametric approach*, J. Econom., 115, 277-291, 2003.
- [84] N. L. Hjort, M. C. Jones, *Locally parametric nonparametric density estimation*, Ann. Statist., 24, 1619-1647, 1997.
- [85] J. M. Hoem, *Markov chain models in life insurance*, Blätter der Deutschen Gesellschaft für Versicherungsmathematik, 9, 91-107, 1969.
- [86] M. Jacobsen, *Statistical analysis of counting processes*, Lecture Notes in Statistics, 12, Springer-Verlag, New York, 1982.
- [87] J. Jiang, K. A. Doksum, *On local polynomial estimation of hazard functions and their derivatives under random censoring* In : Mathematical statistics and applications : Festschrift for Constance van Eeden, IMS Lecture Notes and Monograph Series, 42 (eds. M. Moore, S. Froda & C. Léger), 463-482, 2003.
- [88] M. C. Jones, *On some kernel density estimation bandwidth selectors related to the double kernel method*, Sankhya, 60, 249-264, 1998.
- [89] M. C. Jones, *Simple boundary correction for density estimation kernel*, Statistics and Computing, 3, 135-146, 1993.
- [90] W. D. Kaigh, *Total time on test function principal components*, Statist. Probab. Lett., 44, 337-341, 1999.
- [91] N. C. Kakwani, *Lorenz curves in economic analysis*, Econometrica, 45, 719-728, 1977.
- [92] N. C. Kakwani, N. Podder, *On the estimation of Lorenz curves from grouped observations*, Int. Econom. Rev., 14, 278-292, 1973.
- [93] R. J. Karunamuni, T. Alberts, *On boundary correction in kernel density estimation*, Statistical Methodology, 2, 191-212, 2005.
- [94] M. G. Kendall, A. Stuart, *The Advanced Theory of Statistics*, Vol. 1, Hafner, New York, 1969.
- [95] B. Klefsjö, *On aging properties and total time on test transforms*, Scand. J. Statist., 9, 37-41, 1982.

- [96] B. Klefsjö, *Some tests against aging based on the total time on test transform*, Commun. Statist. Theory Methods, 12, 907-927, 1983.
- [97] B. Klefsjö, *Reliability interpretations of some concepts from economics*, Naval Res. Logist. Quart., 31, 1984.
- [98] B. Klefsjö, *TTT-plotting - A tool for both theoretical and practical problems*, J. Statist. Plan. Inf., 29, 99-110, 1991.
- [99] C. Kleiber, *The Lorenz curve in economics and econometrics*, Invited paper, Gini-Lorenz Centennial Conference, Siena, May 23-26, 2005.
- [100] C. Kochar, X. Li, M. Shaked, *The total time on test transform and the excess wealth stochastic orders of distributions*, Adv. Appl. Prob., 34, 826-845, 2002.
- [101] C. D. Lai, M. Xie, *Stochastic ageing and dependence for reliability*, Springer, New York, 2006.
- [102] N. A. Langberg, R. V. Leon, F. Proschan, *Characterization of nonparametric classes of life distributions*, Ann. Probab., 8, 1163-1170, 1980.
- [103] M. Lejeune, *Estimation non-paramétrique par noyaux : régression polynomiale mobile*, Rev. Statist. Appl., 33, 43-67, 1985.
- [104] M. Lejeune, M., P. Sarda, *Smooth estimators of distribution and density functions*, Computational Statistics Data Analysis, 4, 457-471, 1992.
- [105] E. Lenglart, *Relation de domination entre deux processus*, Ann. Inst. Henri Poincaré, 13, 171-179, 1977.
- [106] X. Li, M. Shaked, *The observed total time on test and the observed excess wealth*, Statist. Probab. Lett., 68, 247-258, 2004.
- [107] R. S. Liptser, A. N. Shiriyayev, *A functional central limit theorem for semimartingales*, Theory Probab. Appl., 25, 667-688, 1980.
- [108] C. R. Loader, *Local likelihood density estimation*, Ann. Statist., 24, 1602-1618, 1996.
- [109] M. C. Lorenz, *Methods of measuring the concentration of wealth*, J. Amer. Statist. Assoc., 9, 209-219, 1905.
- [110] M. Lubrano, *Lorenz curves and parametric distributions*, 2010.
- [111] S. Luo, *New Non-Parametric Methods for Income Distributions*, Dissertation, Georgia State University, 2013.
- [112] A. W. Marshall, F. Proschan, *Classes of distributions applicable in replacement with renewal theory implications*, In : Proceedings of the 6th Berkeley Symposium of Mathematical Statistics and Probability, University of California Press, Berkeley, 395-415, 1965.
- [113] K. Messer, L. Goldstein, *A new class of kernels for nonparametric curve estimation*, Ann. Statist., 1, 179-195, 1993.
- [114] C. Morrisson, *Les inégalités de revenus*, Que sais-je?, Puf, 1986.
- [115] H. G. Müller, *Boundary effects in nonparametric curve estimation models*, Compstat, 84-89, 1984.
- [116] H. G. Müller, *Nonparametric regression analysis of longitudinal data*, Lecture Notes in Statistics, 46, Springer-Verlag, New York, 1988.
- [117] H. G. Müller, *Smooth optimum kernel estimators near endpoints*, Biometrika, 78, 521-530, 1991.

- [118] H. G. Müller, *On the boundary kernel method for nonparametric curve estimation near endpoints*, Scand. J. Statist., 1993.
- [119] H. G. Müller, J. L. Wang, *Locally adaptive hazard smoothing*, Probab. Theory Related Fields, 85, 523-538, 1990.
- [120] H. G. Müller, J. L. Wang, *Hazard rate estimation under random censoring with varying kernels and bandwidths*, Biometrics 50, 61-76, 1994.
- [121] E. A. Nadaraya, *On Estimating Regression*, Theory Probab. Appl., 9, 141-142, 1964.
- [122] N. U. Nair, P. G. Sankaran, *Some new applications of the total time on test transforms*, Statistical Methodology, 10, 93-102, 2013.
- [123] M. Navarro, *Les inégalités de revenus*, Armand Colin, coll "Cursus", 2016.
- [124] W. Nelson, *Hazard plotting for incomplete failure data*, J. Qual. Technol., 1, 27-52, 1969.
- [125] J. P. Nielsen, C. Tanggaard, *Boundary and bias correction in kernel hazard estimation*, Scand. J. Statist., 28, 675-698, 2001.
- [126] R. Norberg, *Reserves in life and pension insurance*, Scand. Actuar. J., 1-22, 1991.
- [127] R. Norberg, *A theory of bonus in life insurance*, Finance and Stochastics, 3, 373-390, 1999.
- [128] R. Norberg, *On bonus and prognoses in life insurance*, Scand. Actuar. J., 126-147, 2001.
- [129] R. Norberg, *Life insurance mathematics*, Wiley StatsRef : Statistics Reference Online, 1-19, 2015.
- [130] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist., 33, 1065-1076, 1962.
- [131] E. Parzen, *Nonparametric statistical data modeling*, J. Amer. Stat. Assoc., 74, 105-121, 1979.
- [132] T. G. Pham, N. Turkkan, *The Lorenz and the scaled total-time-on-test transform curves : a unified approach*, IEEE Trans. Reliab., 43, 1994.
- [133] H. Procaccia, E. Ferton, M. Procaccia, *Fiabilité et maintenance des matériels industriels réparables et non réparables*, Lavoisier, 2011.
- [134] H. Ramlau-Hansen, *Smoothing counting process intensities by means of kernel functions*, Ann. Statist., 11, 453-466, 1983.
- [135] H. Ramlau-Hansen, *The choice of a kernel function in the graduation of counting process intensities*, Scand. Actuar. J., 165-182, 1983.
- [136] H. Ramlau-Hansen, *Distribution of surplus in life insurance*, ASTIN Bulletin, 21, 57-71, 1991.
- [137] C. R. Rao, L. C. Zhao, *Strassen's law of the iterated logarithm for the Lorenz curves*, J. Multiv. Anal., 54, 239-252, 1995.
- [138] R. Rebolledo, *Central limit theorems for local martingales*, Z. Wahrsch. Verw. Geb., 51, 269-286, 1980.
- [139] J. Rice, M. Rosenblatt, *Estimation of the log survivor function and hazard function*, Sankhya Ser. A, 38, 60-78, 1976.

- [140] M. Rosenblatt, *Remarks on some nonparametric estimates of a density function*, Ann. Math. Statist., 27, 832-837, 1956.
- [141] J. M. Sarabia, *A general definition of the Leimkuhler curve*, J. Informetrics, 2, 156-163, 2008.
- [142] E. F. Schuster, *Estimation of a probability density function and its derivatives*, Ann. Math. Statist., 40, 1187-1195, 1969.
- [143] E. F. Schuster, *Incorporating support constraints into nonparametric estimators of densities*, Commun. Statist., 14, 1123-1126, 1985.
- [144] A. Sen, *On Economic Inequality*, Norton, New York, 1973.
- [145] P. K. Sen, *Sequential nonparametrics*, Wiley, New York, 1981.
- [146] G. R. Shorack, J. A. Wellner, *Empirical Processes with Applications to Statistics*, Wiley, New York, 1986.
- [147] B. W. Silverman, *Density estimation for statistics and data analysis*, London, Chapman and Hall, 1986.
- [148] D. J. Slottje, *The structure of earnings and the measurement of income inequality in the US*, North-Holland, Amsterdam, 1989.
- [149] D. Smith, *Fiabilité, maintenance, et risque*, Dunod, 2006.
- [150] . Stone, *Consistent nonparametric regression*, Ann. Math. Statist., 5, 595-645, 1977.
- [151] . Stone, *Optimal global rates of convergence for nonparametric regression*, Ann. Math. Statist., 10, 1040-1053, 1982.
- [152] I. Suoniemi, *Estimating Lorenz curves using locally weighted regression*, Government Institute for Economic Research Finland, 1994.
- [153] G. Szegö, *Orthogonal polynomials*, AMS publications 23, 1939.
- [154] M. A. Tanner, . W. H. Wong, *The estimation of the hazard function from randomly censored data by the kernel method*, Ann. Statist., 11, 989-993, 1983.
- [155] W. A. Jr Thompson, *Fisherman's luck*, Biometrics, 32, 265-271, 1976.
- [156] S. M. Tse, *Lorenz curve for truncated and censored data*, Ann. Inst. Statist. Math., 58, 675-686, 2006.
- [157] A.W. Van der Vaart, *Asymptotic Statistics*, Cambridge series in Statistical and Probabilistic Mathematics, 1998.
- [158] G. S. Watson, M. R. Leadbetter, *On the estimation of the probability density II*, Ann. Math. Statist., 34, 480-491, 1963.
- [159] G. S. Watson, M. R. Leadbetter, *Hazard analysis I*, Biometrika, 51, 175-184, 1964.
- [160] G. S. Watson, M. R. Leadbetter, *Hazard analysis II*, Sankhya Ser. A, 26, 101-116, 1964.
- [161] J. A. Wellner, , *Limit theorems for the ratio of the empirical distribution function to the true distribution*, Z. Wahrsch. Verw. Geb., 45, 73-88, 1978.
- [162] H. Wolthuis, *Prospective and retrospective premium reserves*, Blätter der DGVM, 20, 317-327, 1992.
- [163] H. Wolthuis, J. M. Hoem, *The retrospective premium reserves*, Insurance : Mathematics & Economics, 9, 229-234, 1990.

-
- [164] M. Xie, *Some total time on test quantities useful for testing constant against bathtub-shaped failure rate distributions*, Scand. J. Statist., 6, 137-144, 1989.
- [165] H. Yamato, *Uniform convergence of an estimator of a distribution function*, Bull. Math. Statist., 15, 69-78, 1973.
- [166] B. S. Yandell, *Nonparametric inference for rates with censored survival data*, Ann. Statist., 11, 1119-1135, 1983.
- [167] S. Zhang, R. J. Karunamuni, *On nonparametric density estimation at the boundary*, Nonparametric Statistics, 12, 197-221, 2000.
- [168] S. Zhang, R. J. Karunamuni, M. C. Jones, *An improved estimator of the density function at the boundary*, J. Amer. Stat. Assoc., 448, 1231-1241, 1999.
- [169] Y. Y. Zhang, X. Wu, Q. Li, *A Simple Consistent Nonparametric Estimator of the Lorenz Curve*, Advances in Econometrics, 2015.
- [170] N. Zhao, H. Y. Lu, Y. Song, *Risk assessment of strategies using total time on test transforms*, IEEE, 2006.