Etude du processus empirique composé

Myriam Maumy

THESE DE DOCTORAT DE L'UNIVERSITE PARIS 6

Spécialité
Mathématiques

présentée par

Myriam MAUMY

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITE PARIS 6

Sujet de la thèse :

Étude du processus empirique composé

soutenue le 2 décembre 2002

devant le jury composé de :

Directeur de thèse :	M. Paul DEHEUVELS	Université Paris 6
Rapporteurs :	Jan BEIRLANT	Université catholique de Louvain
	Jean DIEBOLT	Université Marne La Vallée
Examinateurs :	Alain BERLINET	Université Montpellier 2
	Denis BOSQ	Université Paris 6
	Armelle GUILLOU	Université Paris 6
	Zan SHI	Université Paris 6
	Philippe VIEU	Université Toulouse 3

Mis en page avec la classe thloria.

Remerciements

Je souhaite d'abord exprimer ma profonde gratitude à mon directeur de thèse, le Professeur Paul Deheuvels pour son soutien constant durant ces trois années d'étude. Je veux lui témoigner toute mon estime parce qu'il sait créer une ambiance propice à la recherche scientifique : ses encouragements aux initiatives personnelles, son respect profond d'une grande liberté de pensée et d'action. Enfin je tiens à le remercier pour m'avoir inculqué pendant ces trois ans certaines valeurs morales comme la rigueur intellectuelle ou le respect du travail d'autrui.

Je voudrais adresser un salut amical à tous mes collègues du laboratoire et d'ailleurs : Alain Lucas, Jules Maes, Roxanne Jallet, Serge Guillas, Emmanuel Guerre, Jérôme Dedecker, Anne Massiani, Sophie Dabo-Niang, Alexandre Depire, Fateh Chebanah, Amor Keziou, François Pain et tant d'autres avec qui j'ai partagé de bons moments.

Je souhaite remercier Pascal Epron pour m'avoir ouvert pendant ces deux dernières années les portes de sa bibliothèque ainsi que Peggy Deplanque pour avoir eu la gentillesse de me prêter dans les meilleurs conditions les livres de la bibliothèque de mathématiques-recherche de l'université de Paris 6.

J'ai une pensée émue pour Louise Lamart qui fait preuve chaque jour de gentillesse et de patience.

Enfin, je remercie tout particulièrement Dan Fliderbaum pour m'avoir soutenue, aidée et assistée dans le domaine technique de la rédaction. Merci à toi.

Je dédie cette thèse
à ma mère, Jacqueline.

Table des matières

Introduction

L'objectif de cette thèse est l'étude des lois limites du processus empirique composé et l'application de cette étude à l'estimation de la régression non paramétrique par la méthode du noyau. Le processus empirique composé se définit par

$$\alpha_{n,c}(t) = \sqrt{n}\left(\mathbb{U}_{n,c}(t) - \mathbb{E}[\mathbb{U}_{n,c}(t)]\right), \qquad \text{pour } t \in [0,1], \tag{1}$$

où $\mathbb{U}_{n,c}$ est la fonction empirique composée continue à droite définie par

$$\mathbb{U}_{n,c}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq t\}}, \qquad \text{pour } t \in [0,1], \tag{2}$$

où $\{Y_i : i \geq 1\}$ désigne une suite de variables aléatoires réelles indépendantes et identiquement distribuées et $\{U_i : i \geq 1\}$ désigne une suite de variables aléatoires réelles indépendantes et uniformément distribuées sur l'intervalle $[0,1]$.
Les suites de variables aléatoires $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. Pour plus de simplicité dans les calculs, nous supposerons, dans cette thèse que pour tout entier $i \geq 1$,

$$\mathbb{E}[Y_i] = 1 \quad \text{ et } \mathrm{Var}[Y_i] = \sigma^2 < \infty.$$

Notons que le processus empirique uniforme $\{\alpha_n(t) = \sqrt{n}\,(\mathbb{U}_n(t) - t) : 0 \leq t \leq 1\}$ est un cas particulier du processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$. En effet, il suffit de prendre les variables aléatoires Y_i égales à 1 pour obtenir le processus empirique uniforme $\{\alpha_n(t) : 0 \leq t \leq 1\}$.

L'étude du processus empirique et de la fonction de répartition empirique est l'un des thèmes majeurs les plus fréquents dans le développement des statistiques (voir, par exemple le livre de Shorack et Wellner [125] ou encore le livre de Csörgő et Révész [27]). En effet, les applications en sont nombreuses, et ce du fait que de nombreuses statistiques peuvent être exprimées directement à partir du processus empirique. Par exemple, le processus empirique peut intervenir dans des domaines comme les tests d'ajustement (Nikitin [106]), le bootstrap (Shao et Tu [124], Hall [63]), les combinaisons linéaires de statistiques d'ordre (Serfling [123]), les tests de rang (Hájek, Sidák et Sen [61]), les données censurées (Elandt-Johnson et Johnson [52]), l'estimation non paramétrique de la densité par la méthode du noyau ou celle des k points les plus proches (Bosq et Lecoutre [11], Scott [122]). Le processus empirique a donné naissance à plusieurs variantes telles que le processus empirique pondéré par des variables qui peuvent être déterministes ou aléatoires. Lorsque le processus empirique uniforme a été utilisé pour démontrer des lois du logarithme itéré en un point ou des vitesses de convergence pour l'estimateur à noyau de la densité, il est naturel de rechercher des résultats analogues dans le cadre de l'estimation non paramétrique de la régression. On introduit alors un processus empirique composé. Pour établir, par exemple, selon cette approche, des lois du logarithme itéré en un point pour un estimateur de régression du type Nadaraya-Watson (Nadaraya [102], Watson [135]), il faut préalablement établir des lois limites fonctionnelles pour le processus empirique composé.

Le but de cette thèse est précisément d'introduire les outils techniques qui devraient être utiles pour obtenir de telles lois limites.

Le premier chapitre de cette thèse est consacré à l'obtention de la meilleure vitesse possible d'approximation uniforme du processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ par une combinaison linéaire d'un pont brownien et d'un processus de Wiener. Nous montrons dans ce chapitre que la meilleure vitesse possible est de l'ordre de $n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}$ et que cette vitesse semble optimale. Pour établir ce résultat, nous nous sommes inspirés des arguments de Komlós, Major et Tusnády [83]-[84]. Nous obtenons en particulier (voir § 1.2, théorème 1.2.1) :

Théorème 0.0.1. *Supposons que*
$(H.1)$ *les suites de variables aléatoires* $\{U_i : i \geq 1\}$ *et* $\{Y_i : i \geq 1\}$ *soient indépendantes.*
$(H.2)$ *La suite* $\{U_i : i \geq 1\}$ *est une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle* $[0, 1]$.
$(H.3)$ *La suite* $\{Y_i : i \geq 1\}$ *est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire* $Y = Y_1$ *qui est telle que* $\mathbb{E}[Y] = 1$ *et* $\mathrm{Var}[Y] = \sigma^2 < \infty$.
$(H.4)$ $\mathbb{E}[\exp(xY)] < \infty$ *si* $|x| \leq x_0$ *pour une valeur convenable de* $x_0 > 0$.
Alors, il existe une suite de ponts browniens $\{B_n^*(t) : 0 \leq t \leq 1\}$ *et une suite de processus de Wiener* $\{W_n^*(t) : t \geq 0\}$ *telles que*

$$\mathbb{P}\left[\sup_{t\in[0,1]} |\alpha_{n,c}(t) - \sigma W_n^*(t) - B_n^*(t)| > n^{-1/2}(\mathcal{C}_1 \log n + y)\right] \leq \mathcal{C}_2 \exp(-\mathcal{C}_3 y)$$

pour tout réel $y > 0$, *où* $\mathcal{C}_1, \mathcal{C}_2$ *et* $\mathcal{C}_3$ *sont des constantes strictement positives.*

Le théorème suivant (voir § 1.2, théorème 1.2.2) montre que cette vitesse est optimale.

Théorème 0.0.2. *Sous les conditions du théorème 0.0.1, alors pour toute suite de ponts browniens* $\{B_n^*(t) : 0 \leq t \leq 1\}$ *et toute suite de processus de Wiener* $\{W_n^*(t) : t \geq 0\}$ *construites sur le même espace de probabilités que le processus empirique composé* $\alpha_{n,c}$, *il existe une constante* $\mathcal{C}_4 > 0$ *telle que*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{t\in[0,1]} |\alpha_{n,c}(t) - \sigma W_n^*(t) - B_n^*(t)| \geq \mathcal{C}_4 n^{-1/2} \log n\right] = 1.$$

Nous montrons également dans le chapitre 1 (voir § 1.2, théorème 1.2.3) que le processus empirique composé peut être convenablement approximé par une combinaison linéaire d'un processus de Kiefer et d'un processus de Wiener à deux paramètres (voir le livre de Csörgő-Révész [27], p.80 pour de plus amples renseignements sur le processus de Kiefer).

Théorème 0.0.3. *Sous les conditions (H.1–2–3–4), nous pouvons définir un processus de Kiefer* $\{K(t,y) : 0 \leq t \leq 1, 0 \leq y < \infty\}$ *et un processus de Wiener à deux paramètres* $\{\Gamma(t,y) : 0 \leq t \leq 1, 0 \leq y < \infty\}$ *tels que*

$$\sup_{t\in[0,1]} \left| \sum_{i=1}^{n} [Y_i \mathbf{1}_{\{U_i \leq t\}} - t] - \sigma\Gamma(t,n) - K(t,n) \right| = O\left(n^{1/4}(\log n)^{1/2}\right) \qquad p.s.$$

Dans le chapitre 2, nous établirons un résultat similaire à celui de Finkelstein [55]. Soit $\mathcal{E} = \mathcal{E}[0,1]$ l'espace des fonctions définies sur l'intervalle $[0,1]$ muni de la norme uniforme $||f||_\infty = \sup_{0\leq t\leq 1} |f(t)|$. Posons, pour tout $t \in [0,1]$ et pour tout $n \geq 3$,

$$G_{n,c}(t) = \frac{n\left(\mathbb{U}_{n,c}(t) - t\right)}{\sqrt{2\mathbb{E}[Y^2]\, n \log\log n}}. \tag{3}$$

Soit $K_{\mathbb{U}_c}$ l'ensemble (dit de Finkelstein [55]) des fonctions $f_c \in \mathcal{E}[0,1]$ telles que

-i- $f_c(0) = f_c(1) = 0$,

-ii- f_c est absolument continue par rapport à la mesure de Lebesgue, et

-iii- $\int_0^1 [f_c'(s)]^2\, ds \leq 1$,

où la fonction f_c' est la dérivée de Lebesgue de la fonction f_c (déterminée presque partout par rapport à la mesure de Lebesgue). Notons que $K_{\mathbb{U}_c}$ est la boule unité de l'espace de Hilbert autoreproduisant associé au pont brownien (voir Deheuvels et Lifshits [30]-[31], Menneteau [100]).

Théorème 0.0.4. *La suite* $\{G_{n,c}(\cdot) : n \geq 3\}$ *est presque sûrement relativement compacte dans* $\mathcal{E}[0,1]$ *et l'ensemble de ses points limites est l'ensemble compact* $K_{\mathbb{U}_c}$.

Le chapitre 3 donne quelques résultats utiles sur le processus de Poisson composé. Entre autres, nous montrons que ce dernier a des accroissements indépendants et stationnaires, et vérifions que le processus de Poisson composé centré définit une martingale relativement à une certaine filtration. Cette propriété de martingale est commode pour établir des inégalités de type exponentiel qui peuvent être ensuite appliquées au processus empirique composé. En effet, ce dernier n'a pas en général la propriété d'être une martingale par rapport à une filtration adaptée. Mais pour appliquer ces inégalités de type exponentiel obtenues pour le processus de Poisson composé, faut il encore relier le processus empirique composé au processus de Poisson composé. Le résultat correspondant est exprimé dans le théorème suivant (voir § 3.5.1, théorème 3.5.2) :

Théorème 0.0.5. *Soit* $\{Y_i : i \geq 1\}$ *une suite de variables aléatoires réelles indépendantes et identiquement distribuées. Soit* $\{U_i : i \geq 1\}$ *une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle* $[0,1]$. *Les suites de variables* $\{Y_i : i \geq 1\}$ *et* $\{U_i : i \geq 1\}$ *sont supposées indépendantes. Soit* $\{\Pi_\lambda(t) = \Pi(\lambda t) : t \geq 0\}$ *un processus de Poisson homogène, continu à droite et de paramètre* $\lambda > 0$. *Soit*

$$\left\{ \Pi_{\lambda,c}(t) = \sum_{i=1}^{\Pi(\lambda t)} Y_i : t \geq 0 \right\}$$

un processus de Poisson composé. Alors, pour tout réel $\lambda > 0$, *nous avons l'égalité en distribution suivante*

$$\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\} \stackrel{d}{=} \{\Pi_{\lambda,c}(t) : 0 \leq t \leq 1 | \Pi(\lambda) = n\},$$

où la fonction empirique composée $\mathbb{U}_{n,c}(t)$ *est définie en* (2).

Enfin, nous établissons dans le chapitre 3 une inégalité fort commode dans l'esprit du lemme 2.1 de Deheuvels et Mason [35]. En effet, le théorème correspondant permet d'obtenir une “approximation de Poisson”, qui est appropriée pour obtenir des lois limites fortes pour le processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ (voir § 3.5.2, théorème 3.5.3).

Théorème 0.0.6. *Pour tout réel* $a \in (0,1)$, *il existe une constante* $\mathcal{C}_a$, *vérifiant* $0 < \mathcal{C}_a < \infty$, *et telle que, pour tout sous-ensemble* $B \in \mathcal{B}_a$, *nous ayons*

$$\mathbb{P}\big[\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq a\} \in B\big] \leq \mathcal{C}_a \, \mathbb{P}\big[\{\Pi_{n,c}(t) : 0 \leq t \leq a\} \in B\big]. \tag{4}$$

Notons que $I_{RC}[0,a]$ représente l'ensemble des fonctions de répartition continues à droite, positives et de mesures de Radon bornées de support inclus dans l'intervalle $[0,a]$ et que $\mathcal{B}_a$ désigne la σ-algèbre de sous-ensembles boréliens induits sur l'espace $I_{RC}[0,a]$ (muni de la topologie de la convergence en loi).

Les outils de ce chapitre, très techniques, nous seront indispensables pour établir la suite de nos résultats, en particulier pour ceux ayant trait au comportement des oscillations du processus empirique composé. Ce chapitre est résumé dans une note au Comptes-Rendus de l'Académie des Sciences sous la référence :
M.Maumy. Sur les oscillations du processus de Poisson composé.
C.R.Acad.Sci. Paris, Série I, Tome 334, pp.705-708, 2002.

Dans le chapitre 4, nous décrivons le comportement des oscillations du processus empirique composé. Les résultats correspondants sont inspirés par les travaux de Stute [131]. Le module d'oscillation $\omega_{n,c}$ du processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ sur l'intervalle $[0,a]$ est défini par

$$\begin{aligned} \omega_{n,c}(a) &= \omega_{\alpha_{n,c}}(a) = \sup_{|C| \leq a} |\omega_{n,c}(C)| = \sup_{0 \leq t-s \leq a} |\alpha_{n,c}(s,t]| \\ &= \sup_{0 \leq t-s \leq a} |\alpha_{n,c}(t) - \alpha_{n,c}(s)| = \sup_{0 \leq h \leq a} \sup_{0 \leq t \leq 1-h} |\alpha_{n,c}(t+h) - \alpha_{n,c}(t)|. \end{aligned}$$

Nous avons alors (voir § 4.3, théorème 4.3.1) le théorème :

Théorème 0.0.7. *Soit* $(a_n)_{n\geq 1}$ *une suite de constantes à valeurs dans* $(0,1)$ *qui vérifie les conditions de Csörgő-Révész-Stute* $[CRS]$ *(voir par exemple Csörgő-Révész [27], Stute [131]), c'est-à-dire*
(S.1) (i) $a_n \downarrow 0$ *et (ii)* $na_n \uparrow \infty$ *lorsque* $n \uparrow \infty$ *;*
(S.2) $\log(1/a_n)/\log\log n \to \infty$ *lorsque* $n \to \infty$ *;*
(S.3) $\log(1/a_n)/na_n \to 0$ *lorsque* $n \to \infty$*.*
Alors, nous avons

$$\lim_{n\to\infty} \frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} = 1 \quad p.s. \tag{5}$$

Pour démontrer le théorème 0.0.7 nous allons utiliser une technique dite de "poissonisation", consistant à approcher le processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ par un processus de Poisson composé. Cette technique est celle établie au chapitre 3. Ce chapitre fait l'objet d'une note au Comptes-Rendus de l'Académie des Sciences :
M.Maumy. Le comportement des oscillations du processus empirique composé.
C.R.Acad.Sci. Paris, Série I, Tome 333, pp.1101-1104, 2001.

Le chapitre 5 est principalement consacré à l'établissement de lois fonctionnelles du logarithme itéré pour les accroissements du processus empirique composé. Ces résultats s'appuient sur ceux de l'article de Deheuvels et Mason [35]. Deheuvels et Mason établissent une loi fonctionnelle du logarithme itéré pour les accroissements du processus empirique uniforme qui leur permet d'énoncer une loi du logarithme itéré pour les estimateurs non paramétriques de la densité, tels que l'estimateur des k-plus proches voisins ou l'estimateur à noyau. Une version partielle de ce résultat avait déjà été établie par Hall [62].
Énonçons la version que nous obtenons de la loi fonctionnelle du logarithme itéré pour les accroissements $\xi_{n,c}$ du processus empirique composé $\alpha_{n,c}$. Ces derniers sont

$$\xi_{n,c}(a,t;s) = \alpha_{n,c}(t+as) - \alpha_{n,c}(t), \tag{6}$$

où $0 < a < 1$, $0 \leq t \leq 1$ et $0 \leq s \leq 1$. Nous noterons $\mathbb{S}_0$ l'ensemble de Strassen (voir par exemple l'article de Strassen [130]) qui est composé des fonctions f absolument continues sur l'intervalle $[0,1]$ et telles que

$$f(0) = 0 \quad \text{et} \quad \int_0^1 [f'(s)]^2 \, ds \leq 1,$$

où la fonction f' désigne la dérivée au sens de Lebesgue de la fonction f. Nous obtenons alors (voir § 5.2, théorème 5.2.1) :

Théorème 0.0.8. *Soit* $\{U_i : i \geq 1\}$ *une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle* $[0,1]$*. Soit* $\{Y_i : i \geq 1\}$ *une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire* $Y = Y_1$*. Nous supposons que* $\mathbb{E}[Y] = 1$, $\mathrm{Var}[Y] < \infty$ *et* $|Y - 1| \leq \mathcal{M}$ *p.s. pour une constante* $\mathcal{M}$ *convenable. De plus, les suites de variables* $\{U_i : i \geq 1\}$ *et* $\{Y_i : i \geq 1\}$ *sont* *supposées indépendantes. Alors, sous les conditions [CRS] (S.1–2–3), pour tout $\varepsilon > 0$ il existe presque sûrement un entier n'_ε tel que, pour tout entier $n \geq n'_\varepsilon$, nous ayons*

$$\left\{ \frac{\xi_{n,c}(a_n, t; \cdot)}{\sqrt{2\mathbb{E}[Y^2] a_n \log(1/a_n)}} : 0 \leq t \leq 1 - a_n \right\} \subset \mathbb{S}_0^\varepsilon. \tag{7}$$

De plus, pour toute fonction $f \in \mathbb{S}_0$ et pour tout $\varepsilon > 0$, il existe presque sûrement un entier $n''_{\varepsilon,f}$ tel que, pour tout entier $n \geq n''_{\varepsilon,f}$, il existe un réel t, avec $0 \leq t = t_{n,\varepsilon,f} \leq 1 - a_n$ et tel que nous ayons

$$\left\| \frac{\xi_{n,c}(a_n, t; \cdot)}{\sqrt{2\mathbb{E}[Y^2] a_n \log(1/a_n)}} - f \right\| < \varepsilon. \tag{8}$$

Ce résultat est soumis à la revue Statistics and Decisions.

Le chapitre 6 est inspiré par les résultats de l'article de Deheuvels et Mason [36]. Deheuvels et Mason ont établi une généralisation de la loi fonctionnelle du logarithme itéré établie par Mason [94]. Cette généralisation utilise les processus locaux indéxés par des ensembles. Ces derniers serviront à établir des lois du logarithme itéré pour les estimateurs de la densité multivariés.

Pour énoncer le théorème principal de ce chapitre, introduisons quelques notations. Soit $\mathbb{B}$ la classe des parties boréliennes de $[0,1]^d$. Notons $\lambda(\cdot)$ la mesure de Lebesgue sur $\mathbb{R}^d$, et $\lambda_0(\cdot)$ la restriction de la mesure $\lambda(\cdot)$ sur $[0,1]^d$. Introduisons la *mesure empirique composée indéxée par la classe* $\mathbb{B}$ par

$$\lambda_{n,c}(B) = \frac{1}{n} \sum_{i=1}^n Y_i \mathbf{1}_{\{U_i \in B\}}, \quad \text{pour } B \in \mathbb{B}. \tag{9}$$

Pour toute sous-classe $\mathbb{D}$ de la classe $\mathbb{B}$, introduisons le *processus empirique composé indexé par la sous-classe* $\mathbb{D}$, défini par

$$\alpha_{n,c}(D) = \sqrt{n} \left(\lambda_{n,c}(D) - \mathbb{E}[\lambda_{n,c}(D)] \right), \quad \text{pour } D \in \mathbb{D}. \tag{10}$$

Pour établir le résultat principal de ce chapitre, il faut considérer une classe $\mathbb{D}$ particulière. Soient $\mathbf{t} \in [0,1]^d$ et $\mathbb{C}$ la classe de parties boreliennes de $\mathbb{R}^d$ de la forme $[a,b]^d$ où $a < b$ et $b - a = 1$. Soit $\mathbf{a} = (a, \ldots, a) \in \mathbb{R}^d$ et posons

$$\mathbb{D} = \{\mathbf{t} + C : C \in \mathbb{C}\}. \tag{11}$$

De plus, nous supposons que
(C.1) $\mathbf{t} + C \subseteq [0,1]^d$ pour tout $C \in \mathbb{C}$ et, pour tout $h > 0$ suffisamment petit, $\mathbf{t} + h^{1/d} [a,b]^d \subseteq [0,1]^d$.
(C.2) $\mathbb{C}$ est *engendrée par une famille dénombrable.*
(C.3) $\mathbb{C} - \{\mathbf{a}\}$ est une *classe de Donsker pour* λ_0.
(C.4) Pour tout $1/2 \leq h_1 \leq h_2 \leq 1$, $h_1\mathbb{C} \subseteq h_2\mathbb{C} \subseteq \mathbb{C}$.

(C.5) Pour tout sous ensemble fini $\{C_1, \cdots, C_M\} \subseteq \mathbb{C}$, avec $\lambda(C_i) > 0$, la classe $\mathbb{C}$ peut être agrandie, si cela est nécessaire, pour inclure un ensemble fini d'éléments disjoints $\{D_1, \cdots, D_M\} \subseteq \mathbb{C}$, avec $\lambda(D_i) > 0$, tel que pour chaque C_i il existe un sous-ensemble $J \subseteq \{1, \cdots, N\}$ pour lequel $\cup_{j \in J} D_j = C_i$.

Posons, par commodité,

$$\log^+ u = \log(\max(u, e)) \quad \text{ et } \log_2 u = \log^+ \left(\log^+ u\right).$$

Soit une suite $\{k(n) : n \geq 1\}$ de nombres réels vérifiant les conditions suivantes.

(K.1) (i) $0 < k(n) \leq n$, (ii) $k(n) \uparrow \infty$ et (iii) $n^{-1}k(n) \downarrow 0$ lorsque $n \to \infty$;

(K.2) $k(n)/\log_2 n \to \infty$ lorsque $n \to \infty$;

(K.3) $k(n)/\log_2 n \to \mathcal{C} \in (0, \infty)$ lorsque $n \to \infty$.

Soit $\mathbb{C}$ la classe de parties boréliennes de $\mathbb{R}^d$ de la forme $[a, b]^d$. Introduisons le *processus empirique composé local au point* $\mathbf{t}$ *indéxé par la classe* $\mathbb{C}$ défini par

$$\begin{aligned} \Theta_{n,c}(C) &= \Theta_{n,c}\left(C, \frac{k(n)}{n}\right) = \Theta_{n,c}\left(C, \mathbf{t}, \frac{k(n)}{n}\right) \\ &= \left(\frac{k(n)}{n}\mathbb{E}[Y^2]\right)^{-1/2} \alpha_{n,c}\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right), \quad \text{pour } C \in \mathbb{C} \end{aligned} \tag{12}$$

et le processus $\alpha_{n,c}(\cdot)$ est défini en (10) dans sa version indexée par des ensembles.

Remarque 0.0.1. La définition (12) du processus $\Theta_{n,c}$ s'appuie sur celle du processus empirique local Θ_n au point $\mathbf{t}$ indéxé par la classe $\mathbb{C}$. Ce processus Θ_n a été introduit par Deheuvels et Mason [36] p.1620.

Introduisons la classe $B(\mathbb{C})$ qui est composée des fonctions bornées et définies sur la classe $\mathbb{C}$, où $\mathbb{C}$ est munie de la topologie de la convergence uniforme.

Théorème 0.0.9. *Sous les hypothèses (K.1–2), (C.1–2–3–4–5) et (H.1–2), la suite de fonctions*

$$\left\{\frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} : C \in \mathbb{C}\right\}$$

est presque sûrement relativement compacte dans $B(\mathbb{C})$ muni de la topologie de la convergence uniforme. De plus, l'ensemble limite est égal à

$$\mathbb{S}(\mathbb{C}) = \left\{f \in B(\mathbb{C}) : f(C) = \int_C \phi(\mathbf{s})\, d\lambda(\mathbf{s}) \text{ avec } \int_{\mathbb{R}^d} [\phi(\mathbf{s})]^2\, d\lambda(\mathbf{s}) \leq 1\right\}. \tag{13}$$

Dans le chapitre 7, nous introduisons un estimateur non paramètrique de la courbe de régression de Nadaraya-Watson (voir Nadaraya [102], Watson [135], Collomb [21]), dont nous rappelons les propriétés asymptotiques et les propriétés de convergence ponctuelle et asymptotique. L'estimateur de la courbe de régression que nous considérons dans le chapitre 7 est défini par

$$r_n(x) = \frac{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h_n}\right)}{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)} \quad \text{pour } x \in \mathbb{R},$$

où la suite $\{h_n : n \geq 1\}$ vérifie les conditions suivantes

(S.1) (i) $h_n \to 0$, (ii) $h_n \downarrow 0$, (iii) $nh_n \uparrow \infty$;

(S.2) $nh_n/\log n \to \infty$ lorsque $n \to \infty$;

(S.3) $\log(1/h_n)/\log\log n \to \infty$ lorsque $n \to \infty$.

Il faut noter que les propriétés de convergence de l'estimateur $r_n(x)$ vers la courbe de régression $r(x)$ ont fait l'objet de nombreux travaux (voir Collomb [21], le chapitre 5 du livre de Bosq et Lecoutre [11] et le livre de Nadaraya [105]). Il convient de citer tout particulièrement à ce sujet Devroye [39]-[40], Collomb [19], Greblicki et Krzyzak [60] qui ont montré que les conditions $(S.1)(i)$ et $(S.2)$ (respectivement $(S.1)(i-iii)$) étaient essentiellement nécessaires et suffisantes pour obtenir la convergence presque sûre (respectivement la convergence en probabilité)

$$\lim_{n\to\infty}\left(\sup_{x\in[a,b]} |r_n(x) - r(x)|\right) = 0. \tag{14}$$

Le problème de la vitesse de convergence de (14) est resté ouvert pendant de nombreuses décennies avant d'être résolu sur le fond par Einmahl et Mason [51] (rectifiant et précisant des travaux antérieurs de Konakov et Piterbarg [86] et Haerdle, Janssen et Serfling [65]). Tous ces résultats sont rappelés dans le chapitre 7. Enfin, nous concluons le chapitre 7 par une loi du logarithme itéré ponctuelle pour l'estimateur à noyau r_n de la courbe de régression défini ci-dessus.

Chapitre 1

Une approximation forte pour le processus empirique composé

L'essentiel de ce chapitre consacré à l'obtention de la meilleure vitesse possible d'approximation uniforme du processus empirique composé $\alpha_{n,c}$ par une combinaison linéaire d'un pont brownien et d'un processus de Wiener. Nous montrons que la meilleure vitesse possible est de l'ordre de $n^{-1/2}\log n$ et que cette vitesse est optimale. Rappelons les résultats qui existent déjà pour le processus empirique et pour d'autres processus similaires.

1.1. Introduction et rappels

Soit $\{X_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées, de fonction de répartition F continue. La fonction de répartition empirique associée à la suite $(X_1, \ldots, X_n)$ est définie, pour chaque entier $n \geq 1$, par

$$\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}}, \qquad \text{pour } x \in \mathbb{R}. \tag{1.1}$$

Nous savons que le processus empirique $e_n(x) = \sqrt{n}[\mathbb{F}_n(x) - F(x)]$ converge en loi vers $B(F(x))$ dans l'espace $\mathcal{D}[-\infty, \infty]$ (voir Billingsley [5]), où $\{B(x) : 0 \leq x \leq 1\}$ désigne un pont brownien. Nous nous intéressons à une construction de $e_n(\cdot)$ et d'une suite de répliques $B_n(\cdot)$ de $B(\cdot)$ sur le même espace de probabilités de sorte que

$$||e_n - B_n(F)|| \to 0 \quad \text{ p.s.}$$

Nous savons, par Skorohod [126] qu'une telle construction existe, et la vitesse de convergence de $||e_n - B_n(F)||$ qui peut être établie par une construction optimale de B_n est une question importante dans le domaine des statistiques ainsi que dans celui des probabilités. D'ailleurs nombreux sont les articles sur ce sujet (voir par exemple les livres de Csörgő et Révész [27] et de Shorack et Wellner [125]). La meilleure estimation obtenue est dûe à Komlós, Major et Tusnády [83]. De leurs méthodes sont issus beaucoup de travaux (voir par exemple les articles de Mason et van Zwet [96], Massart [97], Bonvalot et Castelle [8]). Nous avons à ce sujet le théorème :

Théorème 1.1.1. *Il existe une suite de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ telle que*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} |e_n(x) - B_n(F(x))| > n^{-1/2}(C\log n + y)\right] \leq K\exp(-\lambda y),$$

pour tout réel $y > 0$, où C, K et λ sont des constantes strictement positives.

Ici et dans la suite de cette thèse, $\mathbb{P}$ et $\mathbb{E}$ désignent respectivement probabilité et espérance.

Démonstration. Voir l'article de Komlós, Major et Tusnády [83]. □

Remarque 1.1.1. Il est commode de se ramener au cas du processus empirique uniforme $\alpha_n(x) = \sqrt{n}(\mathbb{U}_n(x) - x)$, où

$$\mathbb{U}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{U_i \leq x\}} \qquad \text{pour } x \in [0,1],$$

est la fonction de répartition empirique uniforme associée à la suite $\{U_i : i \geq 1\}$ de variables aléatoires uniformes sur l'intervalle $[0,1]$ définie par $U_i = F(X_i)$ pour $i \geq 1$. En observant que $\sup_{x\in[0,1]} |\alpha_n(x) - B_n(x)| = ||e_n - B_n(F)||$ le théorème 1.1.1 peut être raffiné dans le résultat suivant :

Théorème 1.1.2. *Il existe une suite de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ telle que*

$$\mathbb{P}\left[\sup_{x\in[0,1]} |\alpha_n(x) - B_n(x)| > n^{-1/2}(12\log n + y)\right] \leq 2\exp(-y/6), \tag{1.2}$$

pour tout réel $y > 0$ et $n \geq 1$.

Démonstration. Voir l'article de Bretagnolle et Massart [12]. □

Remarque 1.1.2. Notons que le théorème 1.1.1 implique directement (par le lemme de Borel-Cantelli) que

$$\sup_{x\in\mathbb{R}} |e_n(x) - B_n(F(x))| = O(n^{-1/2}\log n) \qquad \text{p.s.},$$

où p.s. signifie presque sûrement.

Le théorème 1.1.3 montre qu'il n'existe pas de ponts browniens $\{B_n^*(x) : 0 \leq x \leq 1\}$ tels que

$$\sup_{x\in\mathbb{R}} |e_n(x) - B_n^*(F(x))| = o(n^{-1/2}\log n) \qquad \text{p.s.}$$

Théorème 1.1.3. *Si la fonction de répartition F est continue, alors pour toute suite de ponts browniens $\{B_n^*(x) : 0 \leq x \leq 1\}$ sur le même espace de probabilités que e_n, nous avons*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{x\in\mathbb{R}} |e_n(x) - B_n^*(F(x))| > \frac{1}{6} n^{-1/2}\log n\right] = 1.$$

Démonstration. Voir l'article de Komlós, Major et Tusnády [83]. □

Remarque 1.1.3. Dans le cas du processus empirique uniforme $\alpha_n(x) = \sqrt{n}(\mathbb{U}_n(x) - x)$, le théorème 1.1.3 implique l'existence d'une constante universelle K_1 telle que, pour toute suite de ponts browniens $\{\widetilde{B}_n(x) : 0 \leq x \leq 1\}$

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{x\in[0,1]} \left|\alpha_n(x) - \widetilde{B}_n(x)\right| > \mathcal{K}_1 n^{-1/2}\log n\right] = 1,$$

où la constante $\mathcal{K}_1$ est telle que $\mathcal{K}_1 \geq \frac{1}{6}$. La démonstration de ce résultat se trouve par exemple dans le livre de Csörgő et Révész [27], p.114.

Plusieurs techniques statistiques, telles que l'estimation de la densité par la méthode du noyau (voir, e.g. Csörgő et Révész [27], Deheuvels et Mason [35]), ou encore, les méthodes statistiques basées sur la fonction caractéristique empirique (voir, e.g. Csörgő [23]) peuvent être commodément décrites à partir de processus empirique $e_n(x)$. Les distributions limites des estimateurs correspondants s'expriment comme fonctionnelles de $B(F(x))$. Ainsi ces limites peuvent dépendre de la fonction de répartition inconnue F. Pour pallier ce défaut, il peut être utile de faire usage de méthodes telles que le bootstrap (voir par exemple Hall [63]), de plus en plus utilisé pour établir des approximations de ces distributions limites. Pour la version bootstrappée des théorèmes 1.1.1 et 1.1.3, nous renvoyons par exemple aux articles de Csörgő, Horváth et Kokoszka [25] et Horváth et Steinebach [72].

Diebolt [43]-[44] et Diebolt et Laïb [45] ont introduit et étudié des méthodes de tests non-paramétriques pour la fonction autorégressive (voir par exemple, le livre de Bosq [10]) dans une classe de processus autorégressifs non linéaires d'ordre un. Les propriétés asymptotiques qu'ils ont obtenues pour leurs méthodes peuvent être déduites du comportement limite des *processus empiriques hybrides* ainsi que celui de *processus de sommes partielles.* Par exemple, le processus empirique bootstrappé construit sur des poids indépendants est un cas particulier de *processus hybride.* De manière générale, le *processus hybride* se définit de la façon suivante. On pose

$$A(x,n) = \sum_{i=1}^{n} \varepsilon_i U(X_i)\mathbf{1}_{\{X_i \leq x\}}, \qquad \text{pour } x \in \mathbb{R}.$$

Ici $\{X_i : i \geq 1\}$ et $\{\varepsilon_i : i \geq 1\}$ composent des suites de variables aléatoires satisfaisant des conditions générales. Sous des hypothèses convenables, Diebolt [43]-[44] et Diebolt et Laïb [45] ont montré que le processus $n^{-1/2}A(x,n)$ converge faiblement vers un processus de Wiener à temps transformé.

Pour établir les propositions qui suivent et rappeler certains résultats utiles, nous allons faire les hypothèses suivantes :

$(H.1)$ les suites $\{X_i : i \geq 1\}$ et $\{\varepsilon_i : i \geq 1\}$ sont mutuellement indépendantes.

$(H.2)$ Les variables aléatoires $\{X_i : i \geq 1\}$ sont indépendantes et identiquement distribuées.

$(H.3)$ Les variables aléatoires $\{\varepsilon_i : i \geq 1\}$ sont indépendantes et identiquement distribuées, telles que $\mathbb{E}[\varepsilon_1] = 0$ et $\mathrm{Var}[\varepsilon_1] = 1$.

$(H.4)$ La fonction $U : \mathbb{R} \to \mathbb{R}$ est à variation bornée sur $\mathbb{R}$.

$(H.5)$ Les variables aléatoires $\{\varepsilon_i : i \geq 1\}$ sont telles que, indépendamment de $i \geq 1$, $\mathbb{E}[\exp(t\varepsilon_i)] < \infty$, pour $|t| \leq t_0$, où t_0 est une constante convenable.

La transformation temporelle du processus de Wiener se définit à partir de

$$G_n(x) = \int_{-\infty}^{x} U^2(s) d\mathbb{F}_n(s) \tag{1.3}$$

et de

$$G(x) = \int_{-\infty}^{x} U^2(s) dF(s), \tag{1.4}$$

où $F(s)$ désigne la fonction de répartition commune des variables $\{X_i : i \geq 1\}$ et $\mathbb{F}_n(s)$ la fonction de répartition empirique des variables $X_1, \cdots, X_n$, définie comme dans (1.1).

Nous supposerons, sans perte de généralité, que les variables aléatoires et les processus introduits ici et par la suite, peuvent être définis sur le même espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$ (cf. l'appendice 2 du livre de Csörgő et Horváth [24], p.418).

Diebolt [44] obtient des bornes supérieures pour la vitesse de convergence du processus hybride $n^{-1/2}A(n,x)$, vers un processus de Wiener, qui sont rappelées dans le théorème 1.1.4 ci-dessous.

Théorème 1.1.4. *Supposons que les hypothèses* $(H.1–2–3–4–5)$ *soient vérifiées. Alors nous pouvons définir une suite de processus de Wiener* $\{W_n(x) : x \geq 0\}$ *telle que :*

$$\text{Les suites } \{W_n(x) : x \geq 0, n \geq 1\} \quad \text{et} \quad \{X_i : i \geq 1\} \quad \text{sont indépendantes,} \tag{1.5}$$

et

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|n^{-1/2}A(x,n) - W_n\left(G_n(x)\right)\right| > n^{-1/2}(C_1 \log n + y)\right] \leq C_2 \exp\left(-C_3 y\right), \tag{1.6}$$

pour tout réel $y > 0$*, où* C_1, C_2 *et* C_3 *sont des constantes strictement positives qui dépendent uniquement de la fonction* U *et de la distribution de la variable aléatoire* ε_1*.*

Si g *désigne une fonctionnelle lipschitzienne d'ordre 1 et si* $g(W(\cdot))$ *a une densité bornée, où* $\{W(x) : x \geq 0\}$ *est un processus de Wiener, alors*

$$\sup_{z\in\mathbb{R}} \left|\mathbb{P}\left[g(n^{-1/2}A(\cdot,n)) \leq z\right] - \mathbb{P}\left[g(W(\cdot)) \leq z\right]\right| = O\left(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}\right). \tag{1.7}$$

Démonstration. Voir l'article de Diebolt [44]. □

En remplaçant l'hypothèse (H.5) par l'hypothèse
(H.6) $\mathbb{E}[|\varepsilon_1|^r] < \infty$ où $r > 2$,
on obtient alors une variante du théorème 1.1.4, ci-dessous.

Théorème 1.1.5. *Supposons que les hypothèses (*H.1–2–3–4*) et (*H.6*) soient vérifiées. Alors nous pouvons définir une suite de processus de Wiener* $\{W_n(x) : x \geq 0\}$ *telle que*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|n^{-1/2}A(x,n) - W_n\left(G_n(x)\right)\right| > n^{-1/2}y\right] \leq a(y)ny^{-r}, \tag{1.8}$$

pour tout réel y *tel que* $n^{1/r} \leq y \leq (n\log n)^{1/2}$*, où* $a(y) \to 0$ *lorsque* $y \to \infty$ *et* $2 < r \leq 3$.

Démonstration. Voir l'article de Diebolt [44]. □

Théorème 1.1.6. *Supposons que les hypothèses (*H.1–2–3–4*) et (*H.6*) soient vérifiées. Alors nous pouvons définir une suite de processus de Wiener* $\{W_n(x) : x \geq 0\}$ *telle que l'hypothèse* (1.5) *soit réalisée et*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|n^{-1/2}A(x,n) - W_n\left(G_n(x)\right)\right| > n^{-1/2}y\right] \leq a(y)ny^{-r}, \tag{1.9}$$

pour tout réel $y > 0$*, où* $a(y) \to 0$ *lorsque* $y \to \infty$.

Remarque 1.1.4. L'approximation (1.9) est plus forte que l'approximation (1.8).

Démonstration. En utilisant l'extension dûe à Einmahl [49]-[50] des inégalités établies par Komlós, Major et Tusnády à la place des inégalités données dans le livre de Shorack et Wellner [125] p.67 (cf le théorème 1.2 et le lemme 1.1 dans le livre de Csörgő et Horváth [24] pp.2-4), nous adaptons la démonstration du théorème 1.1.5 pour obtenir l'approximation (1.9). □

Dans l'article [70], Horváth a établi que la transformation de temps aléatoire définie de (1.3) dans les approximations (1.6) et (1.9) peut être remplacée par la transformation temps non aléatoire définie en (1.4) et ce sans modifier les vitesses de convergence. Les démonstrations établies par Horváth utilisent l'approximation "Poissonnienne" pour le processus empirique, associée à l'approximation gaussienne pour des sommes partielles à temps d'arrêt aléatoire (voir également, e.g. Csörgő, Deheuvels et Horváth [22]).

Théorème 1.1.7. *Si les conditions du théorème 1.1.4 sont satisfaites, alors nous pouvons définir une suite de processus de Wiener* $\{W_n^*(x) : x \geq 0\}$ *telle que*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|n^{-1/2}A(x,n) - W_n^*\left(G(x)\right)\right| > n^{-1/2}\left(y + C_1^*\log n\right)\right] \leq C_2^*\exp(-C_3^*y) \tag{1.10}$$

pour tout réel $y > 0$, où C_1^*, C_2^* et C_3^* sont des constantes strictement positives qui dépendent uniquement des fonctions U, F et de la distribution de la variable aléatoire ε_1.

Si les conditions du théorème 1.1.6 sont satisfaites, alors nous pouvons définir une suite de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ telle que

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|n^{-1/2}A(x,n) - W_n^*\left(G(x)\right)\right| > n^{-1/2}y\right] \leq a^*(y)ny^{-r} \tag{1.11}$$

pour tout réel $y > 0$ et $a^*(y) \to 0$ lorsque $y \to \infty$.

Démonstration. Voir l'article de Horváth [70]. □

Les approximations établies dans (1.10) et dans (1.11) ont la même forme que celles de Komlós, Major et Tusnády [83]-[84] pour le processus empirique et pour les sommes partielles. Nous noterons que l'optimalité de l'approximation dans (1.10) a été prouvée par Horváth, Kokoszka et Steinebach [71] dans le cas particulier où la fonction $U(t) = 1$. Le théorème 1.1.7 peut être utilisé pour établir la vitesse de convergence de (1.7). Soient g une fonctionnelle lipschitzienne d'ordre un et W un processus de Wiener. Nous supposerons que $g(W(\cdot))$ a une densité bornée. Sous les conditions du théorème 1.1.4, nous avons

$$\sup_{z\in\mathbb{R}} \left|\mathbb{P}\left[g(n^{-1/2}A(\cdot,n)) \leq z\right] - \mathbb{P}\left[g(W(G(\cdot))) \leq z\right]\right| = O\left(n^{-1/2}\log n\right).$$

Si les conditions du théorème 1.1.6 sont satisfaites, alors nous avons

$$\sup_{z\in\mathbb{R}} \left|\mathbb{P}\left[g(n^{-1/2}A(\cdot,n)) \leq z\right] - \mathbb{P}\left[g(W(G(\cdot))) \leq z\right]\right| = o\left(n^{-(r-2)/(2(r+1))}\right).$$

Le théorème 1.1.8 étudie l'approximation du processus hybride $A(x,n)$ par un processus de Wiener à deux paramètres et établit les vitesses de convergence presque sûrement. Auparavant, définissons le processus de Wiener à deux paramètres.

Définition 1.1.1. $\{\Gamma(x,y) : 0 \leq x, y < \infty\}$ *est un processus de Wiener à deux paramètres, si* $\Gamma(x,y)$ *est un processus gaussien tel que*

$$\mathbb{E}[\Gamma(x,y)] = 0$$

et

$$\mathbb{E}[\Gamma(x,y)\Gamma(x',y')] = \min(x,x')\min(y,y')$$

pour $x, x', y, y' \geq 0$.

Pour de plus amples détails sur le processus $\Gamma(x,y)$, on pourra consulter le livre de Csörgő et Révész [27].

Théorème 1.1.8. *Si les conditions du théorème 1.1.4 sont satisfaites, alors nous pouvons définir un processus de Wiener à deux paramètres* $\{\Gamma(x,y) : 0 \leq x, y < \infty\}$ *tel que*

$$\sup_{x\in\mathbb{R}} |A(x,n) - \Gamma(G(x),n)| = O\left(n^{1/4}(\log n)^{1/2}\right) \quad p.s. \tag{1.12}$$

Si les hypothèses $(H.1\text{–}2\text{–}3\text{–}4)$ *sont satisfaites et si l'hypothèse* $(H.6)$ *est réalisée avec* $r \geq 5$*, alors nous pouvons définir un processus de Wiener à deux paramètres* $\{\Gamma(x,y) : 0 \leq x, y < \infty\}$ *tel que*

$$\sup_{x\in\mathbb{R}} |A(x,n) - \Gamma(G(x),n)| = O\left(n^{1/2-\delta(r)}(\log n)^{1/2}\right) \quad p.s. \tag{1.13}$$

où $\delta(r) = (r-2)/4(r+1)$*.*

Démonstration. Voir l'article de Horváth [70]. □

Remarque 1.1.5. Dans le cas du processus empirique uniforme $\{\alpha_n(x) : 0 \leq x \leq 1\}$, l'égalité (1.12) devient

$$\sup_{x\in[0,1]} \left|\alpha_n(x) - n^{-1/2}K(x,n)\right| = O(n^{-1/2}(\log n)^2) \quad \text{p.s.}, \tag{1.14}$$

où $K(x,n)$ est un processus de Kiefer (voir le livre de Csörgő et Révész [27], p.80 pour de plus amples renseignements sur le processus $K(x,n)$).

Une conséquence immédiate du théorème 1.1.8 est la loi du logarithme itéré pour le processus hybride $A(x,n)$.

Corollaire 1.1.1. *Si les hypothèses* $(H.1\text{–}2\text{–}3\text{–}4)$ *sont satisfaites et si l'hypothèse* $(H.6)$ *est réalisée avec* $r \geq 5$*, alors nous avons*

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n\log\log n}} \sup_{x\in\mathbb{R}} |A(x,n)| = \sqrt{\int_{\mathbb{R}} U^2(s)dF(s)} \quad p.s. \tag{1.15}$$

et

$$\liminf_{n\to\infty} \sqrt{\frac{\log\log n}{n}} \sup_{x\in\mathbb{R}} |A(x,n)| = \sqrt{\frac{\pi^2}{8}\int_{\mathbb{R}} U^2(s)dF(s)} \quad p.s. \tag{1.16}$$

Démonstration. Si les hypothèses $(H.1\text{–}2\text{–}3\text{–}4)$ sont satisfaites et si l'hypothèse $(H.6)$ est réalisée avec $r \geq 5$, alors nous avons l'égalité (1.13) presque sûrement. De plus la loi fonctionnelle du logarithme itéré pour le processus de Wiener à deux paramètres $\{\Gamma(x,y) : 0 \leq x, y < \infty\}$ (cf. le théorème 1.14.1 p.75 dans le livre de Csörgő et Révész [27]) conduit à ce résultat. □

1.2. Résultats principaux

Nous supposons que
(C.1) les suites de variables aléatoires $\{U_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont indépendantes.
(C.2) La suite $\{U_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0, 1]$.
(C.3) La suite $\{Y_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. De plus, nous supposons que $\mathbb{E}[Y] = 1$ et $\mathrm{Var}[Y] = \sigma^2 < \infty$.
(C.4) $\mathbb{E}[\exp(xY_1)] < \infty$ pour $|x| \leq x_0$ où x_0 est une constante strictement positive convenable.
Parfois, nous remplaçerons la condition (C.4) par la condition suivante
(C.5) $\mathbb{E}[|Y_1|^r] < \infty$ pour $r > 2$.
Pour chaque entier $n \geq 1$, soit

$$\mathbb{U}_{n,c}(x) = \frac{1}{n} \sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}}, \qquad \text{pour } x \in [0, 1], \tag{1.17}$$

la fonction empirique composée continue à droite, définie à partir des n premières variables aléatoires issues des suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$. Ici, $\mathbf{1}_A$ représente la fonction indicatrice de l'événement A et, la lettre "c" est utilisée pour rappeler le terme "composé".

Le processus empirique composé est défini, pour chaque entier $n \geq 1$, par

$$\alpha_{n,c}(x) = \sqrt{n}\big[\mathbb{U}_{n,c}(x) - \mathbb{E}[\mathbb{U}_{n,c}(x)]\big] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[Y_i \mathbf{1}_{\{U_i \leq x\}} - x\right], \qquad \text{pour } x \in [0, 1]. \tag{1.18}$$

Pour chaque entier $n \geq 1$ et pour chaque réel $x \notin [0, 1]$, $\alpha_{n,c}(x) = 0$.

Remarque 1.2.1. Calculons la variance du processus empirique composé $\{\alpha_{n,c}(x) : 0 \leq x \leq 1\}$. Pour tout entier $n \geq 1$, nous avons

$$\mathbb{E}[\alpha_{n,c}(x)] = 0 \qquad \text{pour } x \in [0, 1],$$

et pour $x \in [0, 1]$, nous avons également

$$\begin{aligned}
\mathbb{E}\left[\alpha_{n,c}^2(x)\right] &= n^{-1}\mathbb{E}\left[\left(\sum_{i=1}^{n}(Y_i \mathbf{1}_{\{U_i \leq x\}} - x)\right)^2\right] \\
&= n^{-1}\mathbb{E}\left[\left(\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}}\right)^2 + n^2x^2 - 2nx\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}}\right] \\
&= n^{-1}\left(n\mathbb{E}[Y^2]x + n(n-1)\mathbb{E}^2[Y]x^2 + n^2x^2 - 2n^2x^2\mathbb{E}[Y]\right) \\
&= (\sigma^2 + 1)x - x^2,
\end{aligned}$$

puisque $\mathbb{E}[Y] = 1$. Nous en déduisons que, pour tout entier $n \geq 1$,

$$\mathrm{Var}[\alpha_{n,c}(x)] = (\sigma^2 + 1)x - x^2 = \sigma^2 x + x(1-x), \qquad \text{pour } x \in [0,1].$$

Théorème 1.2.1. *(i) Supposons que les conditions (C.1–2–3–4) soient satisfaites. Alors, il existe une suite de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ et une suite de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ telles que*

$$\mathbb{P}\left[\sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| > n^{-1/2}(\mathcal{C}_1 \log n + y)\right] \leq \mathcal{C}_2 \exp(-\mathcal{C}_3 y) \tag{1.19}$$

pour tout réel $y > 0$, où $\mathcal{C}_1, \mathcal{C}_2$ et $\mathcal{C}_3$ sont des constantes strictement positives.

(ii) Supposons que les conditions (C.1–2–3) et (C.5) soient satisfaites. Alors il existe une suite de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ et une suite de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ telles que

$$\mathbb{P}\left[\sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| > n^{-1/2} y\right] \leq a^*(y) n y^{-r} \tag{1.20}$$

pour tout réel $y > 0$ et $a^*(y) \to 0$ lorsque $y \to \infty$.

Remarque 1.2.2. Notons que le résultat (1.19) implique directement que

$$\sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| = O\left(n^{-1/2}\log n\right) \qquad \text{p.s.}$$

Remarque 1.2.3. Si les variables aléatoires Y_i sont égales à 1, alors $\mathrm{Var}[Y_i] = 0$. Par conséquent l'inégalité (1.19) devient alors l'inégalité (1.2) où $\mathcal{C}_1 = 12$, $\mathcal{C}_2 = 2$ et $\mathcal{C}_3 = 1/6$.

Le théorème suivant montre qu'il n'existe pas de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ et de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ tels que

$$\sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| = o(n^{-1/2}\log n) \qquad \text{p.s.}$$

Théorème 1.2.2. *Si les conditions (C.1–2–3–4) sont satisfaites, alors pour toute suite de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ et toute suite de ponts browniens $\{B_n(x) : 0 \leq x \leq 1\}$ nous avons*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| \geq \mathcal{C}_4 n^{-1/2}\log n\right] = 1, \tag{1.21}$$

où $\mathcal{C}_4$ est une constante strictement positive.

Le théorème suivant établit une approximation du processus empirique composé $\alpha_{n,c}(\cdot)$ par un processus de Wiener à deux paramètres et un processus de Kiefer, et établit les vitesses de convergence presque sûrement.

Théorème 1.2.3. *Si les conditions $(C.1\text{–}2\text{–}3\text{–}4)$ sont satisfaites, alors nous pouvons définir un processus de Wiener à deux paramètres $\{\Gamma(x,y) : 0 \le x, 0 \le y < \infty\}$ et un processus de Kiefer $\{K(x,y) : 0 \le x \le 1, 0 \le y < \infty\}$ tels que*

$$\sup_{x\in[0,1]} \left| \sum_{i=1}^{n} [Y_i \mathbf{1}_{\{U_i \le x\}} - x] - \sigma\Gamma(x,n) - K(x,n) \right| = O\left(n^{1/4}(\log n)^{1/2}\right) \qquad p.s. \tag{1.22}$$

Si les conditions $(C.1\text{–}2\text{–}3)$ sont satisfaites et si la condition $(C.5)$ est réalisée avec $r \ge 5$, alors nous pouvons définir un processus de Wiener à deux paramètres $\{\Gamma(x,y) : 0 \le x, 0 \le y < \infty\}$ et un processus de Kiefer $\{K(x,y) : 0 \le x \le 1, 0 \le y < \infty\}$ tels que

$$\sup_{x\in[0,1]} \left| \sum_{i=1}^{n} [Y_i \mathbf{1}_{\{U_i \le x\}} - x] - \sigma\Gamma(x,n) - K(x,n) \right| = O\left(n^{1/2-\delta(r)}(\log n)^{1/2}\right) \qquad p.s. \tag{1.23}$$

où $\delta(r) = (r-2)/4(r+1)$.

1.3. Démonstrations des théorèmes

1.3.1. Démonstration du théorème 1.2.1

Dans un premier temps, nous allons établir la première partie du théorème 1.2.1 en 3 étapes.
Remarquons que, pour tout entier $n \ge 1$, nous avons

$$\alpha_{n,c}(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - 1)\, \mathbf{1}_{\{U_i \le x\}} + \alpha_n(x), \quad \text{pour } x \in [0,1].$$

Pour des commodités d'écriture, posons :

$$H_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - 1)\mathbf{1}_{\{U_i \le x\}}, \quad \text{pour } x \in [0,1].$$

Première étape. Approximation du processus empirique uniforme par une suite de ponts browniens.
D'après le théorème 1.1.2, il existe une suite de ponts browniens $\{B_n(x) : 0 \le x \le 1\}$ telle que

$$\mathbb{P}\left[\sup_{x\in[0,1]} |\alpha_n(x) - B_n(x)| > n^{-1/2}(12\log n + y)\right] \le 2\exp(-y/6),$$

pour tout réel $y > 0$ et $n \ge 1$.

Deuxième étape. Approximation du processus $\{H_n(x) : 0 \le x \le 1\}$ par une suite de processus de Wiener $\{\sigma W_n^*(x) : x \ge 0\}$.
Dans un premier temps, nous allons établir la proposition suivante.

Proposition 1.3.1. *Supposons que les conditions (C.1–2–3–4) soient satisfaites. Alors il existe sur un espace de probabilités élargi une suite de processus de Wiener* $\{W_n(x) : x \geq 0\}$ *telle que*

$$\text{les suites } \{W_n(x) : x \geq 0, n \geq 1\} \quad \text{et} \quad \{U_i : i \geq 1\} \quad \text{sont indépendantes}$$

et

$$\mathbb{P}\left[\sup_{0\leq x\leq 1} |H_n(x) - \sigma W_n(\mathbb{U}_n(x))| > n^{-1/2}\left(\mathcal{K}_1 \log n + y\right)\right] < \mathcal{K}_2 \exp\left(-\mathcal{K}_3 y\right), \tag{1.24}$$

pour tout réel $y > 0$*, où* $\mathcal{K}_1$*,* $\mathcal{K}_2$ *et* $\mathcal{K}_3$ *sont des constantes strictement positives qui dépendent uniquement de la distribution de la variable aléatoire* Y*. De plus, nous avons*

$$\sup_{0\leq x\leq 1} |H_n(x) - \sigma W_n(\mathbb{U}_n(x))| = O\left(n^{-1/2}\left(\log n\right)\right) \quad \text{p.s. lorsque } n \to \infty. \tag{1.25}$$

Si l'hypothèse $(C.4)$ *est remplacée par l'hypothèse* $(C.5)$*, alors nous obtenons que*

$$\mathbb{P}\left[\sup_{0\leq x\leq 1} |H_n(x) - \sigma W_n(\mathbb{U}_n(x))| > n^{-1/2} y\right] < o(n) y^{-r}, \tag{1.26}$$

pour tout réel y *tel que* $n^{1/r} \leq y \leq \sqrt{n \log n}$*, où* $2 < r \leq 3$*. De plus, nous avons*

$$\sup_{0\leq x\leq 1} |H_n(x) - \sigma W_n(\mathbb{U}_n(x))| = o\left(n^{-1/2+1/r}\right) \quad \text{p.s. lorsque } n \to \infty. \tag{1.27}$$

Remarque 1.3.1. Cette proposition est en fait un corollaire des théorèmes 1.1.4 et 1.1.5 de Diebolt [44].

Démonstration. La démonstration de cette proposition reprend certaines composantes de l'article de Diebolt [44] en les adaptant au contexte de nos hypothèses. Nous montrons uniquement (1.24) et (1.25) sous les hypothèses (C.1–2–3–4). La démonstration de (1.26) et de (1.27), lorsque l'hypothèse $(C.4)$ est remplacée par $(C.5)$, est similaire.
L'idée est de réarranger les variables aléatoires de la somme définissant le processus $\{H_n(x) : 0 \leq x \leq 1\}$ pour appliquer le principe d'approximation forte de Komlós, Major et Tusnády [83]-[84]. Pour cela, introduisons le processus de sommes partielles S_n construit sur les variables aléatoires $(Y_i - 1)$, défini par

$$S_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nx \rfloor} (Y_i - 1), \quad \text{pour } 0 \leq x \leq 1, \tag{1.28}$$

où $\lfloor u \rfloor$ représente la partie entière du réel u et la somme vide est égale à 0. Le processus de sommes partielles nous permet d'utiliser un principe d'approximation forte de type Komlós, Major et Tusnády [83]-[84]. D'après ce principe, il existe sur un espace de probabilités élargi un processus de Wiener W tel que, si nous notons $W_n(x) = n^{-1/2} W(nx)$, pour $x \geq 0$, nous ayons

$$\mathbb{P}\left[\sup_{x\in[0,1]} |S_n(x) - \sigma W_n(x)| > n^{-1/2}\left(\mathcal{K}_1 \log n + y\right)\right] < \mathcal{K}_2 \exp\left(-\mathcal{K}_3 y\right), \tag{1.29}$$

pour tout réel $y > 0$ et $\mathcal{K}_1$, $\mathcal{K}_2$ et $\mathcal{K}_3$ qui sont des constantes strictement positives et qui dépendent uniquement de la distribution de la variable Y. De plus, nous avons également

$$\sup_{x\in[0,1]} |S_n(x) - \sigma W_n(x)| = O\left(n^{-1/2}\log n\right) \quad \text{p.s. lorsque } n \to \infty, \tag{1.30}$$

où S_n est le processus de sommes partielles défini en (1.28). Ceci conduit à estimer la distance pour la norme sup entre le processus H_n et le processus $\mathbb{X}_{n,c}$ défini par

$$\mathbb{X}_{n,c}(x) = \sum_{i=1}^{k(n,x)} \left(W_n\left(\frac{i}{n}\right) - W_n\left(\frac{i-1}{n}\right)\right) \quad \text{pour } x \in [0,1],$$

où $k(n,x) = n\mathbb{U}_n(x)$. La conclusion s'ensuivra alors du fait que le processus gaussien $\{\mathbb{X}_{n,c}(x) : 0 \le x \le 1\}$ est égal (via une transformation d'échelle du processus de Wiener) en distribution au processus $\{W(\mathbb{U}_n(x)) : 0 \le x \le 1\}$.
Soient $U_{1,n} < \cdots < U_{i,n} < \cdots < U_{n,n}$ les n statistiques d'ordre associées à la suite $\{U_1, \ldots, U_n\}$. Nous avons alors

$$H_n(x) \overset{\mathrm{d}}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_{n,i} - 1)\mathbf{1}_{\{U_{n,i} \le x\}} \quad \text{pour } x \in [0,1],$$

où $Y_{n,i}$ désigne le réarrangement des variables aléatoires Y_i correspondant. Ainsi, comme les variables aléatoires Y_i sont indépendantes et identiquement distribuées et de plus indépendantes des variables uniformes U_i qui sont elles-mêmes identiquement distribuées, le processus $\{H_n(x) : 0 \le x \le 1\}$ est égal en distribution au processus $\{S_n \circ \mathbb{U}_n(x) : 0 \le x \le 1\}$ défini par

$$\left(S_n \circ \mathbb{U}_n\right)(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - 1)\mathbf{1}_{\{U_{n,i} \le x\}} \quad \text{pour } x \in [0,1].$$

La conclusion découle alors de (1.29)-(1.30). □

Maintenant, nous allons établir une approximation forte de type Komlós, Major et Tusnády pour le processus $\{W(\mathbb{U}_n(x)) : 0 \le x \le 1\}$. Pour cela, nous avons besoin d'établir trois lemmes. Le premier lemme montre que nous pouvons approcher la fonction de répartition empirique uniforme $\mathbb{U}_n(x)$ par une suite de processus de Poisson.

Lemme 1.3.1. *Nous pouvons définir une suite de processus de Poisson $\{N_n(x) : x \ge 0\}$ telle que $\mathbb{E}[N_n(x)] = x$ et telle que*

$$\mathbb{P}\left[\sup_{x\in[0,1]} \left|\mathbb{U}_n(x) - \frac{N_n(nx)}{N_n(n)}\right| > n^{-1}(\mathcal{C}_5 \log n + y)\right] \le \mathcal{C}_6 \exp(-\mathcal{C}_7 y)$$

pour tout réel $y > 0$ où $\mathcal{C}_5$, $\mathcal{C}_6$ et $\mathcal{C}_7$ sont des constantes strictement positives.

Démonstration. Voir le lemme 3.1 dans l'article de Horváth [70]. □

Nous supposons que
(C.6) $\{W(x) : 0 \leq x \leq 1\}$ est un processus de Wiener, indépendant de la suite $\{U_i : i \geq 1\}$ et du processus de Poisson $\{N_n(x) : x \geq 0, n \geq 1\}$ défini dans le lemme 1.3.1.

Remarque 1.3.2. Nous avons à notre disposition ce processus de Wiener. En effet, nous l'avons construit à partir du processus de sommes partielles $S_n(x)$ défini par la suite de variables $\{Y_i : i \geq 1\}$. De plus, ce processus de Wiener est indépendant de la suite de variables $\{U_i : i \geq 1\}$ par construction et par conséquent indépendant du processus de Poisson $\{N_n(x) : x \geq 0, n \geq 1\}$ défini dans le lemme 1.3.1.

Lemme 1.3.2. *Comme l'hypothèse (C.6) est satisfaite, alors nous pouvons définir des processus de Wiener $\{W_{1,n}(x) : x \geq 0\}$ et des processus de Poisson $\{N_{1,n}(x) : x \geq 0\}$ où $\mathbb{E}[N_{1,n}(x)] = x$ tels que*

les processus $\{W_{1,n}(x) : x \geq 0\}$ et $\{N_{1,n}(x) : x \geq 0\}$ sont indépendants pour chaque n

et

$$\mathbb{P}\left[\sup_{x\in[0,1]} \left|W(\mathbb{U}_n(x)) - n^{-1/2}W_{1,n}(N_{1,n}(nx))\right| > n^{-1/2}(\mathcal{C}_8 \log n + y)\right] \leq \mathcal{C}_9 \exp(-\mathcal{C}_{10} y) \tag{1.31}$$

pour tout réel $y > 0$, où $\mathcal{C}_8$, $\mathcal{C}_9$ et $\mathcal{C}_{10}$ sont des constantes strictement positives.

Démonstration. Voir le lemme 3.2 dans l'article de Horváth [70], dans le cas particulier où la fonction $U(t) = 1$. □

Soit $\{W_{1,n}(x) : x \geq 0\}$ un processus de Wiener et $\{N_{1,n}(x) : x \geq 0\}$ un processus de Poisson avec $\mathbb{E}[N(x)] = x$. Nous supposons que

les processus $\{W_{1,n}(x) : x \geq 0\}$ et $\{N_{1,n}(x) : x \geq 0\}$ sont indépendants.

Remarque 1.3.3. Nous avons également à notre disposition ces 2 processus grâce au lemme 1.3.2.

Lemme 1.3.3. *Nous pouvons définir un processus de Wiener $\{\widehat{W}(x) : x \geq 0\}$ tel que*

$$\mathbb{P}\left[\sup_{x\in[0,1]} \left|W_{1,n}(N_{1,n}(nx)) - \widehat{W}(nx)\right| > \mathcal{C}_{11} \log n + y\right] \leq \mathcal{C}_{12} \exp(-\mathcal{C}_{13} y) \tag{1.32}$$

pour tout réel $y > 0$, où $\mathcal{C}_{11}$, $\mathcal{C}_{12}$ et $\mathcal{C}_{13}$ sont des constantes strictement positives.

Démonstration. Introduisons la suite de points x_i définie par $x_i = i/n$, $1 \leq i \leq n$. Introduisons un lemme nécessaire pour la suite de la démonstration de (1.32).

Lemme 1.3.4. *Pour tout $\varepsilon > 0$ il existe une constante $C = C(\varepsilon)$ telle que l'inégalité*

$$\mathbb{P}\left[\sup_{0\leq s\leq 1-h} \sup_{0<t\leq h} |W(s+t) - W(s)| > v\sqrt{h}\right] \leq \frac{C}{h} \exp\left(-\frac{v^2}{2+\varepsilon}\right)$$

est réalisée pour tout réel positif v et $h < 1$.

Démonstration. Voir le lemme 1.1.1 p.24 du livre de Csörgő et Révész [27]. □

En utilisant le lemme 1.3.4, nous obtenons que

$$\mathbb{P}\left[\max_{1\leq i\leq n}\sup_{x_{i-1}\leq x\leq x_i}|W(N(nx))-W(N(nx_i))|>\mathcal{C}_{14}\log n+y\right]\leq\mathcal{C}_{15}\exp(-\mathcal{C}_{16}y)$$

pour tout réel $y>0$, où $\mathcal{C}_{14}$, $\mathcal{C}_{15}$ et $\mathcal{C}_{16}$ sont des constantes strictement positives. D'autre part, nous remarquons que le processus $W(N(nx_i))$ s'écrit comme une somme de différence de processus de Wiener

$$W(N(nx_i))=\sum_{j=1}^{i}w_j,\ \text{pour } 1\leq i\leq n,$$

où $w_j=W(N(nx_j))-W(N(nx_{j-1}))$, pour $1\leq j\leq n$. Il s'ensuit alors de la propriété de Markov pour le processus de Wiener W que $w_1,\cdots,w_n$ sont des variables aléatoires indépendantes. De plus, les variables aléatoires $w_1,\ldots,w_n$ vérifient l'égalité en distribution suivante

$$\{w_i,1\leq i\leq n\}\stackrel{\mathrm{d}}{=}\left\{\sum_{j=N(i-1)+1}^{N(i)}Z_j,1\leq i\leq n\right\},$$

où $Z_1,Z_2,\ldots$ sont des variables aléatoires indépendantes et identiquement distribuées de loi normale $\mathcal{N}(0,1)$. Il est alors évident de montrer que $w_1,\ldots,w_n$ sont également des variables aléatoires indépendantes et identiquement distribuées qui ont une fonction génératrice des moments finie dans un voisinage de zéro. En utilisant le résultat de Komlós, Major et Tusnády [83] on peut définir un processus de Wiener indépendant $\widehat{W}$ tel que

$$\mathbb{P}\left[\max_{1\leq k\leq n}\left|\sum_{j=1}^{k}w_j-\widehat{W}(k)\right|>\mathcal{C}_{17}\log n+y\right]\leq\mathcal{C}_{18}\exp(-\mathcal{C}_{19}y)$$

pour tout réel $y>0$, où $\mathcal{C}_{17}$, $\mathcal{C}_{18}$ et $\mathcal{C}_{19}$ sont des constantes strictement positives.

En utilisant à nouveau le lemme 1.3.4, nous obtenons que

$$\mathbb{P}\left[\max_{1\leq i\leq n}\sup_{x_{i-1}\leq x\leq x_i}|\widehat{W}([nx_i])-\widehat{W}(nx)|>\mathcal{C}_{20}\log n+y\right]\leq\mathcal{C}_{21}\exp(-\mathcal{C}_{22}y),$$

pour tout réel $y>0$, ce qui complète la démonstration de (1.32). □

En combinant (1.31) et (1.32), nous en déduisons le lemme suivant.

Lemme 1.3.5. *Nous pouvons définir un processus de Wiener* $\{W_n^*(x)=n^{-1/2}\widehat{W}(nx):x\geq 0\}$ *tel que*

$$\mathbb{P}\left[\sup_{x\in[0,1]}|W(\mathbb{U}_n(x))-W_n^*(x)|>n^{-1/2}(\mathcal{C}_{23}\log n+y)\right]\leq\mathcal{C}_{24}\exp(-\mathcal{C}_{25}y)\tag{1.33}$$

pour tout réel $y>0$, *où* $\mathcal{C}_{23}$, $\mathcal{C}_{24}$ *et* $\mathcal{C}_{25}$ *sont des constantes strictement positives.*

Enfin, en rassemblant la proposition 1.3.1 et le lemme 1.3.5, nous en déduisons le théorème suivant.

Théorème 1.3.1. *Il existe une suite de processus de Wiener* $\{W_n^*(x) : x \geq 0\}$ *telle que*

$$\mathbb{P}\left[\sup_{x\in[0,1]} |H_n(x) - \sigma W_n^*(x)| > n^{-1/2}\left(\mathcal{C}_{26}\log n + y\right)\right] \leq \mathcal{C}_{27}\exp(-\mathcal{C}_{28}y) \tag{1.34}$$

pour tout réel $y > 0$, *où* $\mathcal{C}_{26}$, $\mathcal{C}_{27}$ *et* $\mathcal{C}_{28}$ *sont des constantes strictement positives.*

Troisième étape. Pour montrer la première partie du théorème 1.2.1, nous allons procéder de la manière suivante. Soient $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$, 3 expressions et $A = \sup|\widetilde{A}| > \alpha$, $B = \sup|\widetilde{B}| > \beta$ et $C = \sup|\widetilde{C}| > \gamma$, 3 événements. De plus, supposons que

$$\widetilde{C} = \widetilde{A} + \widetilde{B} \quad \text{et} \quad \gamma = \alpha + \beta.$$

Par conséquent, nous avons

$$|\widetilde{C}| \leq |\widetilde{A}| + |\widetilde{B}|.$$

En prenant le supremum, nous obtenons que

$$\sup|\widetilde{C}| \leq \sup|\widetilde{A}| + \sup|\widetilde{B}|.$$

On a

$$\overline{A} \cap \overline{B} \Rightarrow \overline{C}.$$

Nous en déduisons alors

$$1 - \mathbb{P}[C] = \mathbb{P}\left[\,\overline{C}\,\right] \geq \mathbb{P}\left[\,\overline{A}\cap\overline{B}\,\right] = \mathbb{P}\left[\,\overline{A\cup B}\,\right].$$

Ainsi, on a

$$1 - \mathbb{P}\left[\,\overline{A\cup B}\,\right] \geq \mathbb{P}[C].$$

Enfin, nous obtenons que

$$\mathbb{P}[C] \leq \mathbb{P}[A] + \mathbb{P}[B]. \tag{1.35}$$

Ici les 3 événements C, A et B sont définis respectivement par

$$\begin{aligned} C &= \sup_{x\in[0,1]} |\alpha_{n,c}(x) - \sigma W_n^*(x) - B_n(x)| > n^{-1/2}\left(\mathcal{C}_1\log n + y\right), \\ A &= \sup_{x\in[0,1]} |H_n(x) - \sigma W_n^*(x)| > n^{-1/2}\left(\mathcal{C}_{26}\log n + y/2\right) \\ B &= \sup_{x\in[0,1]} |\alpha_n(x) - B_n(x)| > n^{-1/2}\left(12\log n + y/2\right), \end{aligned}$$

où $\mathcal{C}_1 = \mathcal{C}_{26} + 12$. On applique l'inégalité (1.35) aux 3 événements A, B et C et on obtient la première partie du théorème 1.2.1 en utilisant les deux théorèmes 1.1.2 et 1.3.1.
La démonstration de la deuxième partie du théorème 1.2.1 est similaire à celle de la première partie du théorème 1.2.1. Il suffit de remplacer (1.24) par (1.26) et d'utiliser les mêmes arguments.

1.3.2. Démonstration du théorème 1.2.2

Remarquons que, pour tout entier $n \geq 1$, nous avons

$$\alpha_{n,c}(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - 1)\, \mathbf{1}_{\{U_i \leq x\}} + \alpha_n(x), \quad \text{pour } x \in [0,1].$$

D'après la remarque 1.1.3, nous savons que pour toute suite de ponts browniens $\{\widetilde{B}_n(x) : 0 \leq x \leq 1\}$

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{x\in[0,1]} \left|\alpha_n(x) - \widetilde{B}_n(x)\right| > \mathcal{K}_1 n^{-1/2} \log n\right] = 1, \tag{1.36}$$

où la constante $\mathcal{K}_1$ est telle que $\mathcal{K}_1 \geq \frac{1}{6}$.
Maintenant étudions le processus $\{H_n(x) - \sigma W_n^*(x) : 0 \leq x \leq 1\}$, où le processus $H_n(x)$ est défini par

$$H_n(x) = n^{-1/2} \sum_{i=1}^{n} (Y_i - 1) \mathbf{1}_{\{U_i \leq x\}} \quad \text{pour } x \in [0,1].$$

Nous allons établir un théorème pour ce processus.

Théorème 1.3.2. *Si les conditions (C.1–2–3–4) sont satisfaites, alors pour toute suite de processus de Wiener $\{W_n^*(x) : x \geq 0\}$ nous avons*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{x\in[0,1]} |H_n(x) - \sigma W_n^*(x)| \geq \mathcal{C}_{29} n^{-1/2} \log n\right] = 1, \tag{1.37}$$

où $\mathcal{C}_{29}$ est une constante strictement positive.

Démonstration. Pour démontrer ce théorème, nous avons besoin dans un premier temps de 3 lemmes techniques. Soit $\{N(x) : x \geq 0\}$ un processus de Poisson homogène continu à droite et de paramètre égal à 1, indépendant de la suite $\{Y_i : i \geq 1\}$. Nous allons d'abord rappeler deux lemmes et en établir un.

Lemme 1.3.6. *Soit $\{\zeta_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuéees de loi exponentielle telle que $\mathbb{E}[\zeta_i] = 1$ et définissons*

$$S(i) = \sum_{j=1}^{i} \zeta_j.$$

Alors, pour chaque entier $n \geq 1$, nous avons l'égalité en distribution suivante

$$\{U_{1,n}, U_{2,n}, \cdots, U_{n,n}\} \stackrel{d}{=} \{S(1)/S(n+1), S(2)/S(n+1), \cdots, S(n)/S(n+1)\}. \tag{1.38}$$

Démonstration. Voir le livre de Csörgő et Horváth [24], p.112. □

Lemme 1.3.7. *Soit* $\{N(x) : x \geq 0\}$ *un processus de Poisson homogène, continu à droite, tel que* $\mathbb{E}[N(x)] = 1$ *et ayant des sauts aux points* $S(1), S(2), \ldots$ *Alors, pour chaque entier* $n \geq 1$*, nous avons l'égalité en distribution suivante*

$$\{n\mathbb{U}_n(x) : 0 \leq x \leq 1\} \stackrel{d}{=} \{N(xS(n+1)) : 0 \leq x \leq 1\}.$$

Démonstration. La démonstration de ce lemme utilise la représentation exponentielle du lemme 1.3.6. Le livre de Csörgő et Horváth [24] donne la démonstration de ce lemme, p.144. □

On déduit, du lemme 1.3.7, le lemme suivant.

Lemme 1.3.8. *Soit* $\{N(x) : x \geq 0\}$ *un processus de Poisson homogène, continu à droite, tel que* $\mathbb{E}[N(x)] = 1$ *et ayant des sauts aux points* $S(1), S(2), \ldots$ *Alors, pour chaque* $n \geq 1$*, nous avons l'égalité en distribution suivante*

$$\{\sqrt{n}H_n(x) : 0 \leq x \leq 1\} \stackrel{d}{=} \{\Pi_c(N(xS(n+1))) : 0 \leq x \leq 1\},$$

où la somme $\Pi_c(k)$ *est définie, pour tout entier* $k \geq 1$*, par*

$$\Pi_c(k) = \sum_{i=1}^{k}(Y_i - 1).$$

Démonstration. D'abord, pour chaque entier $n \geq 1$, nous avons l'égalité en distribution suivante

$$\left\{\sum_{i=1}^{n}(Y_i - 1)\mathbf{1}_{\{U_i \leq x\}} : 0 \leq x \leq 1\right\} \stackrel{\mathrm{d}}{=} \left\{\sum_{i=1}^{n\mathbb{U}_n(x)}(Y_i - 1) : 0 \leq x \leq 1\right\}. \tag{1.39}$$

D'autre part, nous avons, en utilisant le lemme 1.3.7, pour chaque entier $n \geq 1$, l'égalité en distribution suivante

$$\left\{\sum_{i=1}^{n\mathbb{U}_n(x)}(Y_i - 1) : 0 \leq x \leq 1\right\} \stackrel{\mathrm{d}}{=} \left\{\sum_{i=1}^{N(xS(n+1))}(Y_i - 1) : 0 \leq x \leq 1\right\}. \tag{1.40}$$

Ce qui achève la démonstration du lemme 1.3.8. □

Remarque 1.3.4. Le processus $\{\Pi_c(N(xS(n+1))) : 0 \leq x \leq 1\}$ est un processus de Poisson composé. Nous l'étudierons plus particulièrement dans le chapitre 3.

Maintenant, nous avons à notre disposition tous les outils pour démontrer le théorème 1.3.2. Posons, pour $0 \leq i \leq n$,

$$z(i, n) = \frac{i}{S(n+1)},$$

et

$$a(n) = \frac{\lfloor C \log n \rfloor}{S(n+1)},$$

où $\lfloor u \rfloor$ désigne la partie entière du réel u. D'après (1.39) il existe des variables aléatoires $\tau(i,n)$, $0 \le i \le n$ et $h(n)$ telles que

$$\begin{aligned} &\left\{ \sum_{i=1}^{n} (Y_i - 1)\mathbf{1}_{\{U_i \le x\}} : 0 \le x \le 1,\, \tau(i,n),\, 0 \le i \le n,\, h(n) \right\} \\ &\qquad \stackrel{\mathrm{d}}{=} \left\{ \Pi_c\left(N(xS(n+1))\right) : 0 \le x \le 1,\, z(i,n),\, 0 \le i \le n,\, a(n) \right\}. \end{aligned} \tag{1.41}$$

Introduisons un théorème essentiel pour la suite de la démonstration, dû à Erdős et Rényi [53].

Théorème 1.3.3. *Soit $\{X_i : i \ge 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées. Nous supposons que la fonction génératrice des moments de la suite des variables aléatoires $\{X_i : i \ge 1\}$*

$$M(t) = \mathbb{E}[\exp(X_1)]$$

existe pour $t \in I$ où I est un voisinage ouvert contenant l'origine. De plus, supposons que

$$\mathbb{E}[X_i] = 0 \quad \textit{pour tout entier } i \ge 1.$$

Soit

$$\rho(x) = \inf_t \exp(-tx)\, M(t),$$

où ρ désigne la fonction de Chernoff de la variable aléatoire X_1. Alors pour tout réel $C > 0$ nous avons, lorsque $n \to \infty$,

$$\max_{1 \le k \le n - \lfloor C \log n \rfloor} \frac{V(k + \lfloor C \log n \rfloor) - V(k)}{\lfloor C \log n \rfloor} \stackrel{p.s.}{\to} \beta(C), \tag{1.42}$$

où $V(n) = X_1 + \cdots + X_n$, pour $n \ge 1$ et

$$\beta(C) = \sup\{x : \rho(x) \ge \exp(-1/C)\}. \tag{1.43}$$

De plus, la fonction $\beta(C)$ détermine uniquement la fonction de répartition de la variable aléatoire X.

Démonstration. La démonstration de ce théorème est établie dans l'article de Erdős et Rényi [53]. On peut également consulter le livre de Csörgő et Révész [27] et plus particulièrement les théorèmes 2.4.3 et 2.4.5. p.98-101. □

Remarque 1.3.5. Nous avons, pour tout entier $n \geq 1$ et pour tout réel $x \in [0,1]$,

$$\begin{aligned} \mathrm{Var}\left[\Pi_c\left(N(xS(n+1))\right)\right] &= \mathrm{Var}\left[N(xS(n+1))\right](\mathbb{E}[Y-1])^2 \\ &\quad + \mathbb{E}\left[N(xS(n+1))\right]\mathrm{Var}[Y-1] \\ &= \sigma^2 nx. \end{aligned}$$

Pour calculer la variance d'un processus de Poisson composé, on utilise une identité de Wald (voir par exemple le livre de Rolski, Schmidli, Schmidt et Teugels [114]).

Il s'ensuit d'après la loi des grands nombres de Erdős et Rényi (théorème 1.3.3) que pour toute constante $C > 0$

$$\begin{aligned} &\frac{1}{\sigma\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\big[\Pi_c\left(N((z(i,n)+a(n))S(n+1))\right) \\ &\qquad - \Pi_c\left(N(z(i,n)S(n+1))\right)\big] \\ &= \frac{1}{\sigma\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\big[\Pi_c\left(N\left(i+\lfloor C\log n\rfloor\right)\right) - \Pi_c\left(N(i)\right)\big] \\ &\overset{\text{p.s.}}{\to} \beta^*(C) \text{ lorsque } n\to\infty, \end{aligned} \tag{1.44}$$

où $\beta^*(C)$ est définie en (1.43).

Erdős et Rényi [53] ont également montré qu'il existe une relation entre la fonction $\beta^*(C)$ et la fonction génératrice des moments de $\Pi_c(N(1))$. En rassemblant (1.41) et (1.44), lorsque $n \to \infty$ nous obtenons que

$$\frac{\sqrt{n}}{\sigma\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\left[H_n(\tau(i,n)+h(n)) - H_n(\tau(i,n))\right] \overset{\mathbb{P}}{\to} \beta^*(C) \tag{1.45}$$

pour toute constante $C > 0$. Soit $\{W(x) : 0 \leq x \leq 1\}$ un processus de Wiener. Nous pouvons définir un processus de Wiener $\{\widehat{W}(x) : 0 \leq x \leq 1\}$ tel que

$$\begin{aligned} &\{W(x) : 0\leq x\leq 1,\ \tau(i,n),\ 0\leq i\leq n,\ h(n)\} \\ &\qquad \overset{\text{d}}{=} \{\widehat{W}(x) : 0\leq x\leq 1,\ z(i,n),\ 0\leq i\leq n,\ a(n)\}. \end{aligned} \tag{1.46}$$

D'après la loi du logarithme itéré pour une somme partielle nous avons que $|S(n+1)-n| = O(\sqrt{\log\log n})$. De plus, d'après le théorème 1.2.1 p.30 du livre de Csörgő et Révész [27], nous obtenons que

$$\begin{aligned} &\frac{\sqrt{n}}{\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\left[\widehat{W}(z(i,n)+a(n)) - \widehat{W}(z(i,n))\right] \\ &= \frac{\sqrt{n}}{\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\left[\widehat{W}\left(\frac{i}{n}+\frac{\lfloor C\log n\rfloor}{n}\right) - \widehat{W}\left(\frac{i}{n}\right)\right] + o_{\mathbb{P}}(1). \end{aligned} \tag{1.47}$$

La transformation d'échelle du processus de Wiener $\widehat{W}$ conduit à

$$\begin{aligned} &\frac{\sqrt{n}}{\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\left[\widehat{W}\left(\frac{i}{n}+\frac{\lfloor C\log n\rfloor}{n}\right) - \widehat{W}\left(\frac{i}{n}\right)\right] \\ &\overset{\text{d}}{=} \frac{1}{\lfloor C\log n\rfloor}\max_{0\leq i\leq n-\lfloor C\log n\rfloor}\left[\widehat{W}\left(i+\lfloor C\log n\rfloor\right) - \widehat{W}(i)\right]. \end{aligned} \tag{1.48}$$

En utilisant le résultat 4.1 de l'article de Erdős et Rényi [53] nous obtenons que

$$\frac{1}{\lfloor C\log n\rfloor}\max_{0\le i\le n-\lfloor C\log n\rfloor}\left[\widehat{W}\left(i+\lfloor C\log n\rfloor\right)-\widehat{W}\left(i\right)\right]\overset{\text{p.s.}}{\to}\sqrt{\frac{2}{C}}\text{ lorsque } n\to\infty, \qquad (1.49)$$

pour toute constante $C>0$. En combinant (1.46) et (1.49), nous en déduisons que pour tout processus de Wiener $\{W(x):0\le x\le 1\}$ nous avons

$$\frac{\sqrt{n}}{\lfloor C\log n\rfloor}\max_{0\le i\le n-\lfloor C\log n\rfloor}\left[W(\tau(i,n)+h(n))-W(\tau(i,n))\right]\overset{\mathbb{P}}{\to}\sqrt{\frac{2}{C}}\text{ lorsque } n\to\infty, \quad (1.50)$$

pour toute constante $C>0$.
Si $M(t)$ représente la fonction génératrice des moments de la variable aléatoire $Y-1$, alors la fonction génératrice des moments de la variable aléatoire $\Pi_c(N(1))$ est égale à $\exp(M(t)-1)$. Il est facile de vérifier que $\exp(M(t)-1)$ ne peut pas être la fonction génératrice des moments d'une variable aléatoire de loi normale $\mathcal{N}(0,1)$. Comme il y a une relation entre les limites établies par Erdős et Rényi [53] et les fonctions génératrices des moments, il existe une constante $C^*>0$ telle que $\beta^*(C^*)\neq\sqrt{(2/C^*)}$. Alors le théorème 1.3.2 découle de (1.45) et de (1.50) avec $C_{10}\ge C^*|\beta^*(C^*)-\sqrt{(2/C^*)}|$. □

En combinant (1.36) et (1.37) on en déduit la conclusion du théorème 1.2.2.

1.3.3. Démonstration du théorème 1.2.3

Démontrons d'abord (1.22). Remarquons que, pour tout entier $n\ge 1$, nous avons

$$\begin{aligned}\sqrt{n}\alpha_{n,c}(x) &= n\left(\mathbb{U}_{n,c}-x\right)\\ &= \sum_{i=1}^{n} Y_i\mathbf{1}_{\{U_i\le x\}}-nx\\ &= \sum_{i=1}^{n}\left(Y_i-1\right)\mathbf{1}_{\{U_i\le x\}}+\sqrt{n}\alpha_n(x), \quad \text{pour } x\in[0,1].\end{aligned}$$

D'après la remarque 1.1.5, nous avons le théorème suivant.

Théorème 1.3.4. *Si les conditions* (C.1–2–3) *et* (C.5) *sont satisfaites, alors nous pouvons définir un processus de Kiefer* $\{K(x,y):0\le x\le 1, 0\le y<\infty\}$ *tel que*

$$\sup_{x\in[0,1]}\left|\sqrt{n}\alpha_n(x)-K(x,n)\right|=O((\log n)^2) \quad p.s. \qquad (1.51)$$

Démonstration. La démonstration de ce théorème se fait en deux étapes. On établit une approximation de type Komlós, Major et Tusnády entre le processus empirique uniforme et le processus de Kiefer et on conclut en utilisant le lemme de Borel-Cantelli. Pour de plus amples détails sur cette démonstration, nous renvoyons au livre de Csörgő et Révész [27], p.150. □

Maintenant étudions le comportement du processus $\{\sqrt{n}H_n(x) - \sigma\Gamma(x,n) : 0 \leq x \leq 1\}$, où $\{\Gamma(x,y) : 0 \leq x \leq 1, 0 \leq y < \infty\}$ est un processus de Wiener à deux paramètres, lorsque $n \to \infty$. Nous pouvons appliquer le théorème 1.1.8 de Horvàth. Il suffit de remplacer le processus $A(x,n)$ par le processus $\sqrt{n}H_n(x)$ (qui est centré comme le processus $A(x,n)$), de prendre la fonction U égale à 1 et la fonction G égale à la fonction identité. Ce qui donne le résultat suivant.

Théorème 1.3.5. *Si les conditions (*C.1–2–3*) et* (C.5) *sont satisfaites, alors nous pouvons définir un processus de Wiener à deux paramètres* $\{\Gamma(x,y) : 0 \leq x \leq 1, 0 \leq y < \infty\}$ *tel que*

$$\sup_{x\in[0,1]} \left|\sqrt{n}H_n(x) - \sigma\Gamma(x,n)\right| = O\left(n^{1/4}(\log n)^{1/2}\right) \qquad p.s. \tag{1.52}$$

Par souci du détail, nous allons rappeler la démonstration du théorème 1.3.5.

Démonstration. Pour des commodités d'écriture, posons

$$\mathcal{U}_c(x,i,j) = \sum_{k=i+1}^{j} (Y_k - 1)\mathbf{1}_{\{U_k \leq x\}}, \qquad \text{pour } x \in [0,1] \text{ et } 1 < i < j < \infty.$$

L'idée de la démonstration de (1.52) est la suivante. Introduisons une suite d'entiers telle que $0 = n_0 < n_1 < n_2 < \cdots$. Les processus $\{\mathcal{U}_c(x, n_k, n_{k-1}) : 0 \leq x \leq 1\}$ sont alors indépendants. D'après l'approximation (1.34), il existe une suite de processus de Wiener qui approche $\mathcal{U}_c(x, n_k, n_{k-1})$. Nous sommons ces k approximations et nous obtenons l'approximation de $\mathcal{U}_c(x, 0, n_k)$ par une suite de processus de Wiener. Ensuite, nous montrons que les accroissements de $\mathcal{U}_c(x,0,n_k) - \mathcal{U}_c(x,0,n)$ sont uniformément petits, si $0 \leq x \leq 1$ et $n_{k-1} \leq n < n_k$. De façon analogue, nous estimons les accroissements du processus de Wiener à deux paramètres.

Nous commençons la démonstration de (1.22) en établissant un lemme qui sera utile par la suite.

Lemme 1.3.9. *Si* $\mathbb{E}[|Y_1|^r] < \infty$ *avec* $r > 2$ *et* $0 \leq a < b \leq 1$, *alors nous avons, pour tout réel* r *tel que* $r > 2$,

$$\begin{aligned} &\mathbb{P}\left[\left|\sum_{k=i+1}^{j} \left(|Y_k|\mathbf{1}_{\{a\leq Y_k \leq b\}} - (b-a)\mathbb{E}[|Y_k|]\right)\right| > y\right] \\ &\quad \leq 2(1+(2/r))^r(b-a)\gamma(r)y^{-r} + 2\exp\left(-4(r+2)^{-2}\exp(-r)y^2/(n(b-a)\gamma(2))\right) \end{aligned} \tag{1.53}$$

pour tout réel $y > 0$, *où* $\gamma(2) = \mathrm{Var}|Y_1|$ *et* $\gamma(r) = 2^{r+1}\mathbb{E}[|Y_1|^r]$, *si* $r > 2$.

Démonstration. D'abord remarquons que

$$\mathrm{Var}[|Y_k|\mathbf{1}_{a\leq U_k\leq b}] \geq \frac{1}{2}(b-a)\gamma(2)$$

et

$$\mathbb{E}[|Y_k|\mathbf{1}_{a\leq U_k\leq b} - (b-a)\mathbb{E}[|Y_k|]]^r \leq \gamma(r).$$

Ensuite, nous utilisons un résultat de Fuk et Nagaev [81] (voir également le livre de Petrov [111] p.78) et nous obtenons (1.53). □

Maintenant introduisons la suite d'entiers $0 = n_0 < n_1 < n_2 < \cdots$, où $n_k = k^2$ et posons $m_k = n_k - n_{k-1}$. D'après (1.34) il existe une suite de processus de Wiener indépendants $\{\widetilde{W}_k(x) : 0 \le x \le 1\}$ telle que

$$\mathbb{P}\left[\sup_{x\in[0,1]}\left|\mathcal{U}_c(x,n_k,n_{k-1}) - \sigma\sqrt{m_k}\,\widetilde{W}_k(x)\right| > (1+2/\mathcal{C}_3)\log m_k\right] \le \mathcal{C}_2/m_k^2. \tag{1.54}$$

En utilisant les processus $\widetilde{W}_1, \widetilde{W}_2, \ldots$ nous pouvons définir un processus gaussien

$$\widetilde{\Gamma}(x,n_i) = \sum_{k=1}^{i}\sqrt{m_k}\,\widetilde{W}_k(x), \text{ pour } x\in[0,1]. \tag{1.55}$$

Il existe un processus de Wiener à deux paramètres temps $\{\Gamma(x,y) : 0 \le x \le 1, y \ge 0\}$ (cf. Section 1.11 du livre de Csörgő et Révész [27]) tel que $\Gamma(x,n_i) = \widetilde{\Gamma}(x,n_i)$, $0 \le x \le 1$, $i \ge 1$. L'inégalité (1.54) conduit à

$$\begin{aligned}&\mathbb{P}\left[\sup_{x\in[0,1]}\left|\mathcal{U}_c(x,n_k,n_{k-1}) - \sigma(\Gamma(x,n_k)-\Gamma(x,n_{k-1}))\right| > (1+2/c_3^*)\log m_k\right]\\ &\qquad\le c_2^*/k^2.\end{aligned} \tag{1.56}$$

Pour tout entier K posons $x_i = x_{i,K} = i/n_K$, $0 \le i \le n_K = K^2$. Il s'ensuit alors que

$$\sup_{x_{i-1}\le x\le x_i}|x - x_i| = x_i - x_{i-1} = \frac{1}{n_K}.$$

Maintenant établissons une approximation pour $\mathcal{U}_c(x,0,n_K)$. Nous avons

$$\begin{aligned}&\sup_{0\le x\le 1}|\mathcal{U}_c(x,0,n_K) - \sigma\Gamma(x,n_K)|\\ &\quad\le \max_{1\le i\le n_K}|\mathcal{U}_c(x_i,0,n_K) - \sigma\Gamma(x_i,n_K)|\\ &\qquad+ \max_{1\le i\le n_K}\sup_{x_{i-1}\le x\le x_i}\left|\mathcal{U}_c(x,0,n_K) - \mathcal{U}_c(x_i,0,n_K)\right|\\ &\qquad+ \max_{1\le i\le n_K}\sup_{x_{i-1}\le x\le x_i}|\sigma\Gamma(x,n_K) - \sigma\Gamma(x_i,n_K)|\\ &\quad= \mathcal{U}_{c,1}(n_K) + \mathcal{U}_{c,2}(n_K) + \mathcal{U}_{c,3}(n_K).\end{aligned} \tag{1.57}$$

Posons

$$\zeta_k^{(i)} = \mathcal{U}_c(x_i,n_k,n_{k-1}) - \sigma\{\Gamma(x_i,n_k) - \Gamma(x_i,n_{k-1})\},$$

si

$$|\mathcal{U}_c(x_i,n_k,n_{k-1}) - \sigma\{\Gamma(x_i,n_k) - \Gamma(x_i,n_{k-1})\}| \le C\log m_k$$

où $C = 1 + 2/c_3^*$ et 0 sinon.

Maintenant montrons que

$$\max_{1\leq i\leq n_K}\left|\sum_{k=1}^{K}\zeta_k^{(i)}\right| = O\left(\sqrt{K\log n_K}\log m_K\right) \quad \text{p.s.} \tag{1.58}$$

En appliquant l'inégalité de Hoeffding [67], nous obtenons que

$$\begin{aligned}
&\mathbb{P}\left[\max_{1\leq i\leq n_K}\left|\sum_{k=1}^{K}\zeta_k^{(i)}\right| > \sqrt{\frac{3}{2}\mathcal{C}^2\sum_{k=1}^{K}\log^2 m_k\log n_k}\right]\\
&\qquad\leq \sum_{i=1}^{n_K}\mathbb{P}\left[\left|\sum_{k=1}^{K}\zeta_k^{(i)}\right| > \sqrt{\frac{3}{2}\mathcal{C}^2\sum_{k=1}^{K}\log^2 m_k\log n_k}\right]\\
&\leq 2/n_K^2 = 2/K^4
\end{aligned}$$

où $\mathcal{C}$ est une constante strictement positive. De plus, le lemme de Borel-Cantelli implique (1.58). D'après (1.56) nous avons

$$\mathcal{U}_{c,1}(n_K) = O\left(n_K^{1/4}(\log n_K)^{3/2}\right) \text{ p.s.} \tag{1.59}$$

Maintenant nous allons établir l'ordre de $\mathcal{U}_{c,2}(n_K)$. Nous avons

$$\mathcal{U}_{c,2}(n_K) \leq \max_{1\leq i\leq n_k}\sum_{k=1}^{n_K}\left(|Y_k-1|\mathbf{1}_{\{x_{i-1,K}\leq U_k\leq x_{i,K}\}}\right). \tag{1.60}$$

Le lemme 1.3.9 implique, en remarquant que $\mathbb{E}[|Y_k-1|]=0$, que

$$\begin{aligned}
&\mathbb{P}\left[\max_{1\leq i\leq n_K}\left|\sum_{k=1}^{n_K}\left(|Y_k-1|\mathbf{1}_{\{x_{i-1,K}\leq U_k\leq x_{i,K}\}}\right)\right| > n_K^{1/r}K^{2/r}\right]\\
&\leq n_K\mathbb{P}\left[\left|\sum_{k=1}^{n_K}\left(|Y_k-1|\mathbf{1}_{\{0\leq U_k\leq 1/n_K\}}\right)\right| > n_K^{1/r}K^{2/r}\right]\\
&\leq C(r)/K^2
\end{aligned} \tag{1.61}$$

où la constante $C(r)$ dépend uniquement du réel r et de $\mathbb{E}|Y|^r$. D'après le lemme de Borel-Cantelli, et d'après (1.61) nous concluons à

$$\mathcal{U}_{c,2}(n_K) = O\left(n_K^{2/r}\right) \text{ p.s.} \tag{1.62}$$

Afin d'obtenir l'ordre de $\mathcal{U}_{c,3}(n_K)$, nous remarquons que

$$\mathcal{U}_{c,3}(n_K) \leq \sup_{x\in[0,1]}\sup_{y\in[0,1/n_K]}\sigma|\Gamma(x+y,n_K)-\Gamma(x,n_K)|. \tag{1.63}$$

De plus, nous avons

$$\begin{aligned}&\sup_{0\le x\le 1}\sup_{0\le y\le 1/n_K}\sigma|\Gamma(x+y,n_K)-\Gamma(x,n_K)|\\&\stackrel{\mathrm{d}}{=}\sup_{0\le x\le 1}\sup_{0\le y\le 1/n_K}\sigma\sqrt{n_K}|W(x+y,n_K)-W(x,n_K)|,\end{aligned}$$

où W est un processus de Wiener. En utilisant le lemme 1.3.4 et le lemme de Borel-Cantelli, nous avons

$$\mathfrak{U}_{c,3}(n_K)=O\left(\sqrt{\log n_K}\right)\ \text{p.s.}\tag{1.64}$$

En rassemblant (1.57), (1.59), (1.62) et (1.64) nous obtenons que

$$\sup_{x\in[0,1]}|\mathfrak{U}_c(x,0,n_K)-\sigma\Gamma(x,n_K)|=O\left(n_K^{1/4}(\log n_K)^{3/2}+n_K^{2/r}\right)\ \text{p.s.}\tag{1.65}$$

Maintenant nous allons montrer que l'ordre de $\mathfrak{U}_c(x,0,n)$ est légérement différent de l'ordre de $\mathfrak{U}_c(x,0,n_K)$. Soit $n_{K-1}<n\le n_K$. Alors, nous avons

$$\begin{aligned}&\sup_{x\in[0,1]}|\mathfrak{U}_c(x,0,n)-\sigma\Gamma(x,n)|\\&\le\sup_{x\in[0,1]}\max_{n_{k-1}<n\le n_K}|\mathfrak{U}_c(x,0,n_K)-\mathfrak{U}_{n,c}(x,0,n)|\\&\quad+\sup_{x\in[0,1]}\max_{n_{k-1}<n\le n_K}|\sigma[\Gamma(x,n_K)-\Gamma(x,n)]|\\&\quad+\sup_{x\in[0,1]}|\mathfrak{U}_c(x,0,n_K)-\sigma\Gamma(x,n_K)|\\&=\mathfrak{U}_{c,4}(n_K)+\mathfrak{U}_{c,5}(n_K)+\mathfrak{U}_{c,6}(n_K).\end{aligned}\tag{1.66}$$

D'abord notons l'égalité en distribution suivante

$$\mathfrak{U}_{c,4}(n_K)\stackrel{\mathrm{d}}{=}\mathfrak{U}_{c,4}(m_K),$$

où

$$\mathfrak{U}_{c,4}(m_K)=\sup_{x\in[0,1]}\max_{1\le i\le m_K}|\mathfrak{U}_c(x,0,i)|.$$

Pour estimer $\mathfrak{U}_{c,4}(m_K)$ introduisons $y_j=j/m_K$, $1\le j\le m_K$. Il est facile de vérifier que

$$\begin{aligned}\mathfrak{U}_{c,4}(m_K)\ &\le\ \max_{1\le j\le m_K}\max_{1\le i\le m_K}|\mathfrak{U}_c(y_j,0,i)|\\&\quad+\max_{1\le j\le m_K}\max_{1\le i\le m_K}\left|\sum_{l=1}^{i}|Y_l-1|\mathbf{1}_{\{y_{j-1}\le U_l\le y_j\}}\right|.\end{aligned}$$

D'après un résultat de Fuk et Nagaev [81], il existe une constante $C(r)$ strictement positive telle que pour tout $1\le j\le m_K$ nous avons

$$\mathbb{P}\left[|\mathfrak{U}_c(y_j,0,m_K)|>\sqrt{m_K\log m_K}\right]\le C(r)m_K^{1-r/2}\tag{1.67}$$

si $r \geq 4$. D'après le théorème 2.4 du livre de Petrov [111], nous avons

$$\mathbb{P}\left[\max_{1\leq i\leq m_K} |\mathcal{U}_c(y_j,0,i)| > 4\sqrt{m_K \log m_K}\right] \leq 4C(r)m_K^{1-r/2} \tag{1.68}$$

pour tout $1 \leq j \leq m_K$. De plus, nous avons

$$\mathbb{P}\left[\max_{1\leq j\leq m_K}\max_{1\leq i\leq m_K} |\mathcal{U}_c(y_j,0,i)| > 4\sqrt{m_K \log m_K}\right] \leq 4C(r)m_K^{2-r/2} \leq C^*(r)/K^2, \tag{1.69}$$

si $r > 8$, où $C^*(r)$ est une constante strictement positive. En associant le théorème 2.4 du livre de Petrov [111] et le lemme 1.3.9, nous obtenons que

$$\mathbb{P}\left[\max_{1\leq j\leq m_K}\max_{1\leq i\leq m_K}\left|\sum_{l=1}^{i}|Y_l-1|\mathbf{1}_{\{y_{j-1}\leq U_l\leq y_j\}}-\frac{1}{m_K}\right| > 4m_K^{1/r}K^{2/r}\right] \leq C(r)/K^2.$$

D'après le lemme de Borel-Cantelli, nous pouvons conclure que

$$\mathcal{U}_{c,4}(n_K) = O\left((n_K)^{1/4}(\log n_K)^{1/2}\right) \qquad \text{p.s.} \tag{1.70}$$

En associant le lemme 2 de l'article de Lai [87] et le lemme de Borel-Cantelli, nous obtenons que

$$\mathcal{U}_{c,5}(n_K) = O\left((n_K)^{1/4}(\log n_K)^{1/2}\right) \qquad \text{p.s.} \tag{1.71}$$

En ce qui concerne l'ordre de $\mathcal{U}_{c,6}(n_K)$, nous l'avons déjà établi dans (1.65).
Enfin, nous avons à notre disposition tous les outils pour établir l'approximation de $\mathcal{U}_c(x,0,n)$. On associe alors (1.65), (1.66), (1.70), (1.71), et nous obtenons que

$$\sup_{x\in[0,1]} |\mathcal{U}_c(x,0,n) - \sigma\Gamma(x,n)| = O\left(n^{1/4}(\log n)^{1/2} + n^{2/r}\right) \qquad \text{p.s.}$$

Nous pouvons choisir le réel r aussi grand que nous voulons et la démonstration de (1.52) est ainsi établie. □

Pour établir la démonstration de (1.22), il suffit de rassembler les théorèmes 1.3.4 et 1.3.5.

Démontrons maintenant (1.23). Pour cela, nous allons suivre la démonstration de (1.22) avec les modifications nécessaires. Il suffit en fait de modifier le théorème 1.3.5 et sa démonstration comme suit. Introduisons une suite n_k telle que $n_k = k^R$ où $R = 2(r+1)/(r-2)$ et posons $m_k = n_k - n_{k-1}$. D'après (1.34), il existe une suite de processus de Wiener indépendants $\{\widetilde{W}(x) : 0 \leq x \leq 1\}$ telle que

$$\mathbb{P}\left[\sup_{x\in[0,1]}\left|\mathcal{U}_c\left(x,n_k,n_{k-1}\right) - \sigma\sqrt{m_k}\widetilde{W}_k(x)\right| > m_k^{1/r}k^{2/r}\right] \leq C/k^2 \tag{1.72}$$

où C est une constante strictement positive. Définissons maintenant $\Gamma(x,z)$ le processus de Wiener à deux paramètres comme celui défini dans (1.55). D'après (1.72), nous avons

$$\begin{aligned}\mathbb{P}\Bigg[&\sup_{x\in[0,1]}\left|\mathfrak{U}_c\left(x,n_k,n_{k-1}\right)-\sigma\left(\Gamma_k(x,n_k)-\Gamma_k(x,n_{k-1})\right)\right| \\ &\quad > m_k^{1/r}k^{2/r}\Bigg] \leq C/k^2.\end{aligned} \tag{1.73}$$

Maintenant, nous allons estimer les expressions $\mathfrak{U}_{c,1}$, $\mathfrak{U}_{c,2}$ et $\mathfrak{U}_{c,3}$ comme dans (1.57). Posons

$$\zeta_k^{(i)}=\mathfrak{U}_c(x_i,n_k,n_{k-1})-\sigma\{\Gamma(x_i,n_k)-\Gamma(x_i,n_{k-1})\},$$

si

$$|\mathfrak{U}_c(x_i,n_k,n_{k-1})-\sigma\{\Gamma(x_i,n_k)-\Gamma(x_i,n_{k-1})\}|\leq m_k^{1/r}k^{2/r}$$

et 0 sinon. Montrons que

$$\max_{1\leq i\leq n_K}\left|\sum_{k=1}^{K}\zeta_k^{(i)}\right|=O\left(\sqrt{\log Kn_K\sum_{k=1}^{K}m_k^{2/r}k^{4/r}}\right)\qquad\text{p.s.} \tag{1.74}$$

L'inégalité de Hoeffding [67] conduit à

$$\mathbb{P}\left[\max_{1\leq i\leq n_K}\left|\sum_{k=1}^{K}\zeta_k^{(i)}\right|>\sqrt{\log Kn_K\sum_{k=1}^{K}m_k^{2/r}k^{4/r}}\right]\leq 2/K^2,$$

ce qui prouve (1.74). Ainsi nous avons,

$$\mathfrak{U}_{c,1}=O\left(n_K^{(2+2R+r)/(2rR)}\sqrt{\log n_K}\right)\qquad\text{p.s.} \tag{1.75}$$

De façon analogue à (1.60) et (1.61), nous avons

$$\mathfrak{U}_{c,2}=O\left(n_K^{(2+R)/(rR)}\right)\qquad\text{p.s.} \tag{1.76}$$

D'après (1.63) et (1.64), nous avons

$$\mathfrak{U}_{c,3}=O\left(\sqrt{\log n_K}\right)\qquad\text{p.s.} \tag{1.77}$$

En rassemblant (1.73)-(1.77) nous obtenons que

$$\begin{aligned}&\sup_{x\in[0,1]}|\mathfrak{U}_c(x,0,n_K)-\sigma\Gamma(x,n_K)| \\ &\quad =O\left(n_k^{(2R+2+r)/(2rR)}\left(\log n_K\right)^{1/2}+n_K^{(2+R)/(rR)}\right)\ \text{p.s.}\end{aligned} \tag{1.78}$$

Comme dans la démonstration du théorème 1.3.5, nous allons montrer que l'ordre de $\mathcal{U}_c(x,0,n)$ est légèrement différent de l'ordre de $\mathcal{U}_c(x,0,n_K)$. Pour cela, nous avons besoin de bornes supérieures pour $\mathcal{U}_{c,4}(x,0,n_K)$, $\mathcal{U}_{c,5}(x,0,n_K)$ et $\mathcal{U}_{c,6}(x,0,n_K)$. En utilisant le résultat de Fuk et Nagaev [81], comme dans (1.67)- (1.69), nous obtenons que

$$\mathbb{P}\left[\mathcal{U}_{c,4}(x,0,n_K) > 4\sqrt{m_K \log m_K}\right] \geq 4C(r) m_K^{2-(r/2)}. \tag{1.79}$$

Comme $m_K^{2-(r/2)} = O\left(K^{(R-1)(4-r)/2}\right)$ et sous l'hypothèse que $r \geq 5$, nous avons $(R-1)(4-r)/2 < -1$. De plus, (1.79) et le lemme de Borel-Cantelli impliquent que

$$\mathcal{U}_{c,4}(x,0,n_K) = O\left(n_K^{(R-1)/(2R)}\sqrt{\log n_K}\right) \quad \text{p.s.} \tag{1.80}$$

Le lemme 2 de l'article de Lai [87] entraine que

$$\mathcal{U}_{c,5}(x,0,n_K) = O\left(n_K^{(R-1)/(2R)}\sqrt{\log n_K}\right) \quad \text{p.s.}$$

De plus, d'après (1.77) et (1.80) nous avons

$$\begin{aligned} &\sup_{x\in[0,1]} |\mathcal{U}_c(x,0,n) - \sigma\Gamma(x,n)| \\ &\quad = O\left(n^{(2R+2+r)/(2rR)}\sqrt{\log n} + n^{(R+2)/(rR)} + n^{(R-1)/(2R)}\sqrt{\log n}\right) \quad \text{p.s.} \end{aligned}$$

Le choix de $R = 2(r+2)/(r-2)$ établit l'approximation (1.23).□

1.4. Remarques

Lorsqu'on établit une approximation de type KMT, pour le processus empirique uniforme, des applications en découlent. Nous citons ici la plus importante : la loi fonctionnelle du logarithme itéré pour le processus empirique uniforme.

Théorème 1.4.1. *La suite* $\left\{\dfrac{\alpha_n(x)}{\sqrt{2\log\log n}} : 0 \leq x \leq 1\right\}$ *est presque sûrement relativement compacte dans* $\mathcal{D}[0,1]$ *et l'ensemble de ses points limites est* $\mathcal{F}$*, où* $\mathcal{F}$ *est l'ensemble des fonctions absolument continues* f *telles que*

$$f(0) = f(1) = 0 \quad \text{et} \quad \int_0^1 [f'(t)]^2 dt \leq 1. \tag{1.81}$$

De plus, nous avons

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2\log\log n}} \sup_{x\in[0,1]} |\alpha_n(x)| = \frac{1}{2} \quad p.s. \tag{1.82}$$

et

$$\liminf_{n\to\infty} \sqrt{\log\log n} \sup_{x\in[0,1]} |\alpha_n(x)| = \sqrt{\frac{\pi^2}{8}} \quad p.s. \tag{1.83}$$

Remarque 1.4.1. $(\mathcal{D}[0,1], \mathbf{S})$ est l'espace des fonctions $f(x)$ de $[0,1]$ vers $[0,1]$, continues à droite et ayant une limite à gauche, muni de la topologie de Skorohod. Le chapitre 3 du livre de Billingsley [5] donne une étude détaillée de l'espace $(\mathcal{D}[0,1], \mathbf{S})$. De plus, sous certaines conditions (voir Jacod et Shiryaev [73] p.287) la convergence dans $(\mathcal{D}[0,1], \mathbf{S})$ est équivalente à la la convergence dans $(\mathcal{D}[0,1], \mathbf{U})$, $\mathbf{U}$ étant la topologie uniforme.

Démonstration. La première partie du théorème 1.4.1 se démontre de deux façons. La première méthode est celle de Finkelstein [55]. La seconde méthode se déroule en quatre étapes.
Première étape.

Théorème 1.4.2. *Nous pouvons définir un processus de Kiefer* $\{K(x,z) : 0 \leq x \leq 1, 0 \leq z < \infty\}$ *tel que*

$$\mathbb{P}\left[\max_{1\leq k\leq n}\sup_{0\leq x\leq 1}\left|k^{1/2}\alpha_k(x) - K(x,k)\right| > (\mathcal{C}_{32}\log n + y)\log n\right] \leq \mathcal{C}_{33}\exp\left(-\mathcal{C}_{34}y\right)$$

pour tout réel $y > 0$*, où* $\mathcal{C}_{32}$*,* $\mathcal{C}_{33}$ *et* $\mathcal{C}_{34}$ *sont des constantes strictement positives.*

Démonstration. On pourra consulter l'article de Komlós, Major et Tusnády [83] ainsi que celui de Bonvalot et Castelle [7] où les auteurs ont établi également ce résultat en précisant $\mathcal{C}_{32} = 76$, $\mathcal{C}_{33} = 2.028$ et $\mathcal{C}_{34} = 1/41$. □

Deuxième étape. Le théorème 1.4.2 et le lemme Borel-Cantelli donnent

$$\sup_{x\in[0,1]}\left|\alpha_n(x) - n^{-1/2}K(x,n)\right| = O(n^{-1/2}(\log n)^2) \quad \text{p.s.} \tag{1.84}$$

Troisième étape.

Théorème 1.4.3. *La suite* $\left\{\dfrac{K(x,z)}{\sqrt{2z\log\log z}} : 0 \leq x \leq 1, 0 \leq z < \infty\right\}$ *est presque sûrement relativement compacte dans* $\mathcal{C}[0,1]$ *et l'ensemble de ses points limites est* $\mathcal{F}$*, où* $\mathcal{F}$ *est l'ensemble des fonctions absolument continues* f *telles que*

$$f(0) = f(1) = 0 \quad \text{et} \quad \int_0^1 [f'(t)]^2 dt \leq 1. \tag{1.85}$$

Démonstration. Voir le théorme 1.15.1 du livre de Csörgő et Révész [27]. □

Quatrième étape. Cette loi fonctionnelle du logarithme itéré pour le processus de Kiefer associée à l'approximation (1.84) conduisent à la première partie du théorème 1.4.1 ainsi qu'aux lois du logarithme itéré (1.82) et (1.83). □

Remarque 1.4.2. On peut montrer directement (1.82). Pour cela, on pourra consulter les articles de Chung [14] et de Smirnov [127]. En ce qui concerne une démonstration directe de (1.83), on renvoie à l'article de Mogul'skii [101].

Dans notre cas, on ne peut pas se servir de l'approximation obtenue sur le processus empirique composé pour établir des lois du logarithme itéré. En effet, la combinaison linéaire d'un processus de Wiener et d'un pont brownien ne permet pas d'utiliser les techniques classiques, c'est-à-dire appliquer une approximation de type Komlós, Major et Tusnády pour obtenir une loi du logarithme itéré. Par conséquent, nous sommes obligés d'utiliser une autre méthode, celle de Finkelstein.

Chapitre 2

Loi fonctionnelle du logarithme itéré

2.1. Introduction

2.1.1. Définitions et notations

Ci-dessous, $(\mathcal{C}[0,1], ||.||_\infty)$ désigne l'espace des fonctions continues sur l'intervalle $[0,1]$, muni de la topologie uniforme. En particulier

$$\forall f \in \mathcal{C}[0,1], \quad ||f||_\infty = \sup_{x\in[0,1]} |f(x)|. \tag{2.1}$$

Pour tous f, g éléments de l'espace $(\mathcal{C}[0,1], ||.||_\infty)$ et pour tout sous-ensemble $\mathcal{K}$ de $(\mathcal{C}[0,1], ||.||_\infty)$, nous notons

$$d(f,g) = ||f-g||_\infty \quad \text{et} \quad d(f,K) = \inf_{g\in\mathcal{K}} ||f-g||_\infty. \tag{2.2}$$

Les lois classiques fonctionnelles du logarithme itéré $(\mathcal{C}[0,1], ||.||_\infty)$ peuvent s'exprimer le plus souvent de la façon suivante.

Définition 2.1.1. *Une suite de fonctions aléatoires $\{f_n : n \geq 1\}$ de l'espace $\mathcal{C}[0,1]$ satisfait la loi fonctionnelle du logarithme itéré sur l'espace $(\mathcal{C}[0,1], ||.||_\infty)$ si*

- $(C.1)$ *la suite $\{f_n : n \geq 1\}$ est presque sûrement relativement compacte dans l'espace $\mathcal{C}[0,1]$;*
- $(C.2)$ *l'ensemble de tous les points limites possibles de la suite $\{f_n : n \geq 1\}$ est presque sûrement égal à un ensemble compact donné $\mathcal{K}$ dans l'espace $\mathcal{C}[0,1]$.*

Remarque 2.1.1. Puisque l'espace $\mathcal{C}[0,1]$ est un espace métrique complet, la condition $(C.1)$ est équivalente à la proposition suivante. On peut extraire de toute suite de fonctions aléatoires une sous-suite uniformément convergente.

La définition 2.1.1 de la loi fonctionnelle du logarithme itéré est celle formulée par Taqqu et Czàdo [133]. Ces derniers ont établi une synthèse des principaux résultats sur les lois fonctionnelles du logarithme itéré des processus self-similaires. Taqqu et Czàdo énoncent d'autres formulations équivalentes de la loi fonctionnelle du logarithme itéré, la plus maniable étant, à notre avis, la suivante

$$\begin{aligned}(C.2) \quad &\Leftrightarrow \quad \mathbb{P}\left[\liminf_{n\to\infty} d(f_n, \mathcal{K}) = 0\right] = 1;\\ (C.3) \quad &\Leftrightarrow \quad \mathbb{P}\left[\liminf_{n\to\infty} ||f_n - g||_\infty, \text{ pour tout } g \in \mathcal{K}\right] = 1.\end{aligned}$$

2.1.2. Résultats de type Strassen

Soient $X_1, \cdots, X_n$ des variables aléatoires indépendantes et identiquement distribuées définies sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, et de fonction de répartition $F(x)$ définie sur un intervalle compact $[a, b]$. Soit $\mathcal{E} = \mathcal{E}[a, b]$ l'espace des fonctions définies sur l'intervalle $[a, b]$ muni de la norme uniforme $||f||_\infty$ définie en (2.1) et de la distance $d(f, g)$, définie en (2.2) selon les notations de Finkelstein [55]. Supposons que X_1 a une espérance $\mathbb{E}[X_1] < \infty$ et une variance $\mathrm{Var}[X_1] = \sigma^2 < \infty$. On définit alors S_n la fonction de l'espace $\mathcal{E}$, pour tout entier $n \geq 1$, par

$$S_n\left(\frac{i}{n}\right) = \sum_{k=1}^{i} \left[\frac{X_k - \mathbb{E}[X_1]}{\sigma}\right] \tag{2.3}$$

pour $i = 1, \cdots, n$. On remarque que la fonction S_n est linéaire sur les intervalles de la forme $[(i-1)/n, i/n]$ pour $i = 1, 2, \cdots, n$. Le théorème suivant sur la convergence en distribution de la fonction S_n est dû à Donsker [46].

Théorème 2.1.1. *La fonction aléatoire S_n définie en* (2.3) *vérifie la convergence en distribution suivante*

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} W, \qquad \textit{lorsque } n \to \infty, \tag{2.4}$$

où W est un processus de Wiener standard de l'espace $\mathcal{E}$.

Démonstration. On peut consulter, par exemple le livre de Billingsley [5], en particulier pour donner un sens précis à la convergence (2.4). □

Strassen prouve dans [130] un théorème fondamental qui permet d'établir une loi fonctionnelle du logarithme itéré pour le processus de sommes partielles. Pour énoncer ce théorème, nous avons besoin d'introduire des notations. Soient $Z_1, Z_2, \cdots$ des variables aléatoires indépendantes et identiquement distribuées. Supposons que pour tout entier $i \geq 1$, on a $\mathbb{E}[Z_i] = 0$. Pour $n \geq 1$, on définit $S(nt)$ par

$$S(nt) = \sum_{i=1}^{[nt]} Z_i, \quad \text{et} \quad \mathbb{S}_n(t) = \frac{S(nt)}{\sqrt{2n \log\log n}} \quad \text{pour } t \in [0, 1].$$

On désigne par $\mathcal{K}$ l'ensemble des fonctions f absolument continues sur $[0, 1]$ telles que $f(0) = 0$ et $\int_0^1 [f'(s)]^2\, ds \leq 1$, où la fonction f' est la dérivée de Lebesgue de la fonction

f définie presque partout par rapport à la mesure de Lebesgue. L'ensemble $\mathcal{K}$ est alors équivalent à

$$\mathcal{K} = \left\{ f \in \mathcal{C}[0,1] : f(0) = 0, f(t) = \int_0^t f'(s)\,ds, \int_0^1 [f'(s)]^2\,ds \leq 1 \right\}.$$

Théorème 2.1.2. *La suite* $\{\mathbb{S}_n : n \geq 3\}$ *est presque sûrement relativement compacte dans l'espace* $\mathcal{C}[0,1]$ *et l'ensemble de tous les points limites est égal à l'ensemble compact* $\mathcal{K}$.

Démonstration. On peut consulter, pour la démonstration du théorème 2.1.2, l'article de Strassen [130] (voir Deheuvels et Lifshits [30]-[31] pour les raffinements de ce résultat). □

Remarque 2.1.2. Le théorème 2.1.2 équivaut à

$$\mathbb{P}\left[\limsup_{n\to\infty} ||Z_n - \mathcal{K}|| = 0\right] = 1 \quad \text{et} \quad \mathbb{P}\left[\bigcap_{f\in\mathcal{K}} \left(\liminf_{n\to\infty} ||Z_n - f|| = 0\right)\right] = 1.$$

Pour $x \in [0,1]$, on note $\mathbb{F}_n(x)$ la fonction de répartition empirique. Ainsi, $n\mathbb{F}_n(x)$ est égal au nombre de variables aléatoires $X_1, \cdots, X_n$ qui sont inférieures ou égales à x.

Soit $\mathcal{D}$ l'espace des fonctions définies sur $[0,1]$ qui sont continues à droite et qui ont une limite à gauche [càdlàg]. On munit l'espace $\mathcal{D}$ de la topologie de Skorohod. On définit la distance entre deux éléments, f et g, de l'espace $\mathcal{D}$ par

$$\inf_{\lambda\in\Lambda} (||\lambda(f) - \lambda(g)|| + ||\lambda - \mathcal{I}||)$$

où Λ est l'ensemble des applications continues strictement croissantes de $[0,1]$ sur lui-même et l'application $\mathcal{I} \in \Lambda$ est l'application identité.

Remarque 2.1.3. Pour de plus amples renseignements et références sur la topologie de Skorohod, on pourra consulter le chapitre 3 du livre de Billingsley [5].

Remarque 2.1.4. Si les variables aléatoires $U_1, \cdots, U_n$ indépendantes suivent une loi uniforme sur l'intervalle $[0,1]$, alors l'équivalent du théorème 2.1.1 de Donsker dans le cas de la fonction de répartition empirique uniforme $\mathbb{U}_n(x)$ est, pour $x \in [0,1]$,

$$\sqrt{n}[\mathbb{U}_n(x) - x] \xrightarrow{\mathrm{d}} B(x) \text{ lorsque } n \to \infty, \tag{2.5}$$

où $B(x)$ est le pont brownien et pour $n \geq 1$, $\mathbb{U}_n(x)$ est définie par

$$\mathbb{U}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{U_i \leq x\}}, \text{ pour } x \in [0,1]. \tag{2.6}$$

($B(x) \overset{\mathrm{d}}{=} W(x) - xW(1)$ où $W(.)$ est un processus de Wiener standard et $\overset{\mathrm{d}}{=}$ représente égalité en distribution).

2.1.3. La loi fonctionnelle du logarithme itéré pour des distributions empiriques uniformes

Finkelstein s'est intéressée dans [55], à une version du théorème de Strassen dans le cas de distributions empiriques. Le théorème 2.1.3 ci-dessous établit le résultat pour des variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0,1]$. Le paragraphe suivant sera consacré à la généralisation du théorème 2.1.3 pour des variables de fonction de répartition quelconque.

Soient $U_1, \cdots, U_n$ des variables aléatoires indépendantes et uniformément distribuées sur $[0,1]$. Posons, pour tout entier $n \geq 3$, pour tout $x \in [0,1]$,

$$G_n(x) = \frac{n\mathbb{U}_n(x) - nx}{\sqrt{2n \log\log n}}. \tag{2.7}$$

Selon les notations de Finkelstein dans [55], K désigne l'ensemble des fonctions f de $\mathcal{E}[0,1]$, telles que

-i- $f(0) = f(1) = 0$,

-ii- f est absolument continue par rapport à la mesure de Lebesgue, et

-iii- $\int_0^1 [f'(s)]^2\, ds \leq 1$,

où la fonction f' est la dérivée de Lebesgue de la fonction f (déterminée presque partout par rapport à la mesure de Lebesgue).

Théorème 2.1.3. *La suite $\{G_n(.) : n \geq 3\}$ est presque sûrement relativement compacte dans l'espace $\mathcal{E}[0,1]$ et l'ensemble de ses points limites est l'ensemble compact K.*

Démonstration. Pour la démonstration du théorème 2.1.3, on pourra consulter l'article de Finkelstein [55] p.609. □

Remarque 2.1.5. La compacité de l'ensemble K est une conséquence du lemme suivant.

Lemme 2.1.1. *Soit f une fonction à valeurs réelles sur l'intervalle $[0,1]$. Les deux conditions suivantes sont équivalentes :*

1. *La fonction f est absolument continue par rapport à la mesure de Lebesgue et*

$$\int_0^1 [f'(s)]^2\, ds \leq C < \infty.$$

2. $\displaystyle\sum_{i=1}^{s} \frac{(f(x_i) - f(x_{i-1}))^2}{x_i - x_{i-1}} \leq C$ *pour tout partition finie $\{x_0, x_1, \cdots, x_s\}$ de l'intervalle* $[0,1]$.

Démonstration. Pour la démonstration du lemme 2.1.1, on peut consulter le livre de Riesz et Nagy [113], p.75. □

Remarque 2.1.6. Chung [14] et Smirnov [127] ont montré indépendamment que si la fonction de répartition de la variable aléatoire X_i est continue alors nous avons (voir (2.7) pour la définition de G_n)

$$\limsup_n \left(\sup_x G_n(x) \right) \leq \frac{1}{4} \quad \text{presque partout.} \tag{2.8}$$

Le facteur $\frac{1}{4}$ s'explique par la propriété suivante de l'ensemble K, qui se déduit du lemme 2.1.1 :

$$(f(s+t) - f(s))^2 \leq t(1-t) \leq \frac{1}{4} \tag{2.9}$$

pour $0 \leq s \leq s+t \leq 1$ et pour tout $f \in K$. De plus, la borne supérieure est atteinte.

Démonstration. Transformons la fonction f comme suit. Posons

$$\begin{array}{rcll} g(x) & = & f(x+s) - f(s) & \text{pour } 0 \leq x \leq t, \\ & = & f(x-t) + f(t+s) - f(s) & \text{pour } t \leq x \leq s+t, \\ & = & f(x) & \text{pour } s+t \leq x \leq 1. \end{array}$$

Ainsi la fonction g est un élément de l'ensemble K ; on peut donc appliquer le lemme 2.1.1 à la fonction g ; ce qui donne : $[g(t)]^2 \leq t(1-t)$. □

2.1.4. Une généralisation du théorème précédent

Soient $X_1, \cdots, X_n$ des variables aléatoires indépendantes de fonction de répartition $F(x)$ définie sur un intervalle compact $[a, b]$.
Soit K_F l'ensemble des fonctions f de l'espace $\mathcal{E}[a, b]$ telles que

-i- $f(a) = f(b) = 0$,

-ii- f est absolument continue par rapport à la fonction F, et

-iii- $\int_a^b [df/dF]^2 \, dF \leq 1$,

où la fonction (df/dF) est la dérivée de la fonction f par rapport à la fonction F déterminée presque partout par rapport à la fonction F.

Théorème 2.1.4. *La suite $\{G_n(.) : n \geq 3\}$ est presque sûrement relativement compacte dans $\mathcal{E}[a, b]$ et l'ensemble de ses points limites est l'ensemble compact K_F.*

Démonstration. Pour la démonstration du théorème 2.1.4, on peut consulter l'article de Finkelstein ([55], page 615). □

Remarque 2.1.7. La compacité de l'ensemble K_F est également une conséquence du lemme 2.1.1.

Dans la suite de ce chapitre, nous allons nous intéresser aux distributions empiriques composées. Plus particulièrement notre but est d'établir une loi du logarithme itéré pour ces distributions. Pour cela, nous allons suivre la même méthode de démonstration que celle de Finkelstein [55]. Notons au passage que cette méthode est antique mais efficace. Il est possible de démontrer plus simplement des lois fonctionnelles du logarithme itéré grâce à des techniques telles que celle de Lai [87] (voir également Deheuvels et Lifshits [30]-[31]). Toutefois, une méthodologie du type de celle de Finkelstein peut s'avérer tout aussi efficace dans le cadre par exemple de lois indexées (voir par exemple Deheuvels et Mason [37]). Introduisons auparavant quelques notations.

2.2. Notations

Soit une suite $\{Y_i : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées (v.a.i.i.d.) de même loi que la variable aléatoire $Y = Y_1$ et de fonction de répartition F continue sur $\mathbb{R}$. De plus, supposons que $\mathbb{E}[Y] = 1$ et $\mathrm{Var}[Y] = \sigma^2 < \infty$. Soit une suite $\{U_i : i \geq 1\}$ de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0, 1]$. Posons, pour tout $n \geq 1$,

$$\mathbb{U}_{n,c}(x) = \frac{1}{n} \sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}} \qquad \text{pour } x \in [0, 1], \tag{2.10}$$

la fonction de répartition empirique composée. La lettre "c" est utilisée ici et par la suite pour rappeler "composé".

2.3. Résultats préliminaires

Théorème 2.3.1. *Soit $\{Y_i : i \geq 1\}$ une suite de v.a.i.i.d. Supposons que pour tout $i \geq 1$, $\mathbb{E}[Y_i] = 1$. Soit $\{U_i : i \geq 1\}$ une suite de v.a. indépendantes et uniformément distribuées sur l'intervalle $[0, 1]$. Les suites de v.a. $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. Alors, nous avons l'égalité en distribution suivante*

$$\Big\{\mathbb{U}_{n,c}(x) : 0 \leq x \leq 1, n \geq 1\Big\} \stackrel{d}{=} \left\{\frac{1}{n} \sum_{i=1}^{n\mathbb{U}_n(x)} Y_i : 0 \leq x \leq 1, n \geq 1\right\}, \tag{2.11}$$

où $\mathbb{U}_n(x)$ est la fonction de répartition empirique uniforme définie en (2.6).

Démonstration. Pour tout entier $n \geq 1$, pour tous réels $x \in [0,1]$ et $\mu \geq 0$, nous avons

$$\begin{aligned}
\mathbb{P}\left[\mathbb{U}_{n,c}(x) \leq \mu\right] &= \mathbb{P}\left[\{\mathbb{U}_{n,c}(x) \leq \mu\} \cap \left\{\bigcup_{i=0}^{n}(n\mathbb{U}_n(x) = i)\right\}\right] \\
&= \mathbb{P}\left[\bigcup_{i=0}^{n}\{\mathbb{U}_{n,c}(x) \leq \mu\} \cap \{n\mathbb{U}_n(x) = i\}\right] \\
&= \sum_{i=1}^{n}\mathbb{P}\left[\{\mathbb{U}_{n,c}(x) \leq \mu\} \cap \{n\mathbb{U}_n(x) = i\}\right] \\
&= \sum_{i=1}^{n}\mathbb{P}\left[\{\mathbb{U}_{n,c}(x) \leq \mu\}\big| \{n\mathbb{U}_n(x) = i\}\right]\mathbb{P}\left[n\mathbb{U}_n(x) = i\right] \\
&= \sum_{i=1}^{n}\mathbb{P}\left[n\mathbb{U}_n(x) = i\right]\mathbb{P}\left[n^{-1}\sum_{j=1}^{n}Y_j\mathbf{1}_{\{U_j \leq x\}} \leq \mu \middle| \left\{\sum_{k=1}^{n}\mathbf{1}_{\{U_k \leq x\}} = i\right\}\right]
\end{aligned}$$

Comme les variables aléatoires Y_i sont indépendantes et identiquement distribuées et indépendantes des variables U_j, alors les variables aléatoires Y_i sont échangeables. Pour des détails sur cette notion, on peut consulter le livre de Chow et Teicher [13]. Ainsi il existe une permutation π de $\{1, \cdots, n\}$ dans $\{1, \cdots, n\}$ telle que nous ayons l'égalité en distribution suivante

$$(Y_1, Y_2, \cdots, Y_n) \stackrel{\mathrm{d}}{=} (Y_{\pi(1)}, Y_{\pi(2)}, \cdots, Y_{\pi(n)}).$$

D'autre part, on a également pour $x \in [0,1]$, l'inégalité en distribution

$$\sum_{k=1}^{n}\mathbf{1}_{\{U_k \leq x\}} \stackrel{\mathrm{d}}{=} \sum_{k=1}^{n}\mathbf{1}_{\{U_{k:n} \leq x\}},$$

où $U_{k:n}$ désigne la k$^{\text{e}}$ statistique d'ordre et $U_{1:n} < U_{2:n} < \cdots < U_{n:n}$, les statistiques d'ordre de $U_1, \ldots, U_n$ p.s. distinctes. Alors, pour tout entier $n \geq 1$ et pour tous réels $x \in [0,1]$, nous obtenons que

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{U}_{n,c}(x) \leq \mu\right] \\
&= \sum_{i=1}^{n}\mathbb{P}\left[n\mathbb{U}_n(x) = i\right]\mathbb{P}\left[n^{-1}\sum_{j=1}^{n}Y_{\pi(j)}\mathbf{1}_{\{U_{\pi(j)} \leq x\}} \leq \mu \middle| \left\{\sum_{k=1}^{n}\mathbf{1}_{\{U_{k:n} \leq x\}} = i\right\}\right] \\
&= \sum_{i=1}^{n}\mathbb{P}\left[n\mathbb{U}_n(x) = i\right]\mathbb{P}\left[n^{-1}\sum_{j=1}^{n}Y_{\pi(j)}\mathbf{1}_{\{U_{j:n} \leq x\}} \leq \mu \middle| \left\{\sum_{k=1}^{n}\mathbf{1}_{\{U_{k:n} \leq x\}} = i\right\}\right].
\end{aligned}$$

Ainsi nous avons, pour $n \geq 1$ et pour $x \in [0,1]$,

$$\begin{aligned}\mathbb{P}\left[\mathbb{U}_{n,c}(x) \leq \mu\right] &= \sum_{i=1}^{n} \mathbb{P}\left[n\mathbb{U}_n(x) = i\right] \mathbb{P}\left[n^{-1} \sum_{j=1}^{i} Y_\pi(j) \leq \mu\right] \\ &= \sum_{i=1}^{n} \mathbb{P}\left[n\mathbb{U}_n(x) = i\right] \mathbb{P}\left[n^{-1} \sum_{j=1}^{n\mathbb{U}_n(x)} Y_j \leq \mu \middle| n\mathbb{U}_n(x) = i\right] \\ &= \mathbb{P}\left[\frac{1}{n} \sum_{j=1}^{n\mathbb{U}_n(x)} Y_j \leq \mu\right].\end{aligned}$$

□

Théorème 2.3.2. *Soit $\{Y_i : i \geq 1\}$ une suite de v.a.i.i.d. telle que pour tout $i \geq 1$, $\mathbb{E}[Y_i] = 1$ et $\mathrm{Var}[Y_i] = \sigma^2 < \infty$. Soit $\{U_i : i \geq 1\}$ une suite de v.a. indépendantes et uniformément distribuées sur $[0,1]$. Les suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. Nous avons alors, pour tout $n \geq 1$ et pour tout $x \in [0,1]$:*

$$\mathbb{E}\big[\mathbb{U}_{n,c}(x)\big] = x \tag{2.12}$$

et

$$\mathrm{Var}\big[\mathbb{U}_{n,c}(x)\big] = n^{-1}\Big[\mathbb{E}[Y^2]x - x^2\Big]. \tag{2.13}$$

Démonstration. Démontrons d'abord (2.12). Pour tout $n \geq 1$ et pour tout $x \in [0,1]$, nous avons

$$\begin{aligned}\mathbb{E}\big[\mathbb{U}_{n,c}(x)\big] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}}\right] = \mathbb{E}\left[Y_1 \mathbf{1}_{\{U_1 \leq x\}}\right] = \mathbb{E}[Y_1]\mathbb{E}\left[\mathbf{1}_{\{U_1 \leq x\}}\right] \\ &= 1 \times \mathbb{P}[U_1 \leq x] = x,\end{aligned}$$

car les suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes et $\mathbb{E}[Y_1] = 1$.

Démontrons maintenant (2.13). *Première méthode.* Pour tout $n \geq 1$ et pour tout $x \in [0,1]$, nous avons

$$\begin{aligned}\mathrm{Var}\big[\mathbb{U}_{n,c}(x)\big] &= \mathbb{E}\big[\mathbb{U}^2_{n,c}(x)\big] - \mathbb{E}^2\big[\mathbb{U}_{n,c}(x)\big] \\ &= \mathbb{E}\left[n^{-2}\sum_{i=1}^{n} Y_i^2 \mathbf{1}_{\{U_i \leq x\}} + n^{-2}\sum_{i \neq j} Y_i Y_j \mathbf{1}_{\{U_i \leq x\}}\mathbf{1}_{\{U_j \leq x\}}\right] - x^2 \\ &= \frac{1}{n} \times \mathbb{E}[Y_1^2] \times x + \frac{n-1}{n} \times x^2 - x^2 \\ &= \frac{1}{n}\Big[\mathbb{E}[Y^2]x - x^2\Big],\end{aligned}$$

car les suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes, $\mathbb{E}[Y_1] = 1$ et $\mathbb{E}[Y_1^2] = \mathbb{E}[Y^2]$.

Deuxième méthode. Nous avons, d'après le théorème 2.3.1,

$$\{\mathbb{U}_{n,c}(x) : 0 \leq x \leq 1, n \geq 1\} \stackrel{\mathrm{d}}{=} \left\{ \frac{1}{n} \sum_{i=1}^{n\mathbb{U}_n(x)} Y_i : 0 \leq x \leq 1, n \geq 1 \right\}.$$

Calculons la fonction génératrice des moments de $X_{n,c}(x) = \sum_{i=1}^{n\mathbb{U}_n(x)} Y_i$. Nous avons pour tout $t \geq 0$

$$\begin{aligned}
g_{X_{n,c}(x)}(t) &= \mathbb{E}\left[\exp\left(tX_{n,c}(x)\right)\right] = \mathbb{E}\left[\exp\left(t \sum_{i=1}^{n\mathbb{U}_n(x)} Y_i\right)\right] \\
&= \sum_{j=0}^{n} \mathbb{E}\left[\exp\left(t \sum_{i=1}^{n\mathbb{U}_n(x)} Y_i\right) \middle| n\mathbb{U}_n(x) = j\right] \mathbb{P}[n\mathbb{U}_n(x) = j] \\
&= \sum_{j=0}^{n} \mathbb{E}\left[\exp\left(t \sum_{i=1}^{j} Y_i\right)\right] \mathbb{P}[n\mathbb{U}_n(x) = j] \\
&= \sum_{j=0}^{n} \mathbb{E}\left[\prod_{i=1}^{j} \exp(tY_i)\right] \mathbb{P}[n\mathbb{U}_n(x) = j] \\
&= \sum_{j=0}^{n} C_n^j x^j (1-x)^{n-j} \left(\mathbb{E}\left[\exp(tY)\right]\right)^j \\
&= [xg_Y(t) + (1-x)]^n,
\end{aligned}$$

où $g_Y(t)$ désigne la fonction génératrice des moments de Y. Calculons maintenant $\mathrm{Var}[X_{n,c}(x)]$. Nous avons, par définition,

$$\mathrm{Var}[X_{n,c}(x)] = \mathbb{E}[X_{n,c}^2(x)] - \mathbb{E}^2[X_{n,c}(x)].$$

Calculons d'abord $\mathbb{E}[X_{n,c}(x)]$. Nous avons

$$\begin{aligned}
\mathbb{E}[X_{n,c}(x)] &= \left\{ \frac{dg_{X_{n,c}(x)}(t)}{dt} \right\}_{t=0} \\
&= \left\{ nxg_Y'(t)[xg_Y(t) + (1-x)]^{n-1} \right\}_{t=0} \\
&= nxg_Y'(0)[xg_Y(0) + (1-x)]^{n-1} \\
&= nx,
\end{aligned}$$

où $g_Y(0) = 1$ et $g'_Y(0) = 1$. Calculons d'abord $\mathbb{E}[X^2_{n,c}(x)]$. Nous avons

$$\begin{aligned}
\mathbb{E}[X^2_{n,c}(x)] &= \left\{ \frac{d^2 g_{X_{n,c}(x)}(t)}{dt^2} \right\}_{t=0} \\
&= \left\{ \frac{d(nxg'_Y(t)[xg_Y(t) + (1-x)]^{n-1})}{dt} \right\}_{t=0} \\
&= \Big\{ nxg''_Y(t)[xg_Y(t) + (1-x)]^{n-1} \\
&\qquad + n(n-1)(xg'_Y(t))^2[xg_Y(t) + (1-x)]^{n-2} \Big\}_{t=0} \\
&= nxg''_Y(0)[xg_Y(0) + (1-x)]^{n-1} + n(n-1)(xg'_Y(0))^2[xg_Y(0) + (1-x)]^{n-2} \\
&= nx\mathbb{E}[Y^2] + n(n-1)x^2.
\end{aligned}$$

Ainsi, nous avons

$$\begin{aligned}
\mathrm{Var}[X_{n,c}(x)] &= nx\mathbb{E}[Y^2] + n(n-1)x^2 - (nx)^2 \\
&= nx\mathbb{E}[Y^2] - nx^2 \\
&= n[x\mathbb{E}[Y^2] - x^2].
\end{aligned}$$

Comme nous avons, pour tout $n \geq 1$ l'égalité en distribution

$$\mathbb{U}_{n,c}(x) \stackrel{\mathrm{d}}{=} \frac{1}{n} X_{n,c}(x), \qquad \text{pour tout réel } x \in [0,1],$$

nous en déduisons, pour tout $n \geq 1$ et pour tout $x \in [0,1]$, que

$$\begin{aligned}
\mathrm{Var}[\mathbb{U}_{n,c}(x)] &= n^{-2} \times n[x\mathbb{E}[Y^2] - x^2] \\
&= n^{-1}[x\mathbb{E}[Y^2] - x^2].
\end{aligned}$$

□

Remarque 2.3.1. Ces calculs sont licites d'après un corollaire du livre de Lukacs [91] qui s'énonce ainsi :
Si la fonction caractéristique $g_X(t)$ de la variable aléatoire X de fonction de répartition $F(x)$ admet des dérivées jusqu'à l'ordre k à l'origine, alors tous les moments jusqu'à l'ordre k existent si k est pair, respectivement jusqu'à l'ordre $(k-1)$ si k est impair.

2.4. Résultat principal : la loi fonctionnelle du logarithme itéré

Soit une suite $\{Y_i : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$ et de fonction de répartition F

continue et définie sur $\mathbb{R}$. Nous supposons que $\mathbb{E}[Y] = 1$ et $\mathrm{Var}[Y] = \sigma^2 < \infty$. Soit une suite $\{U_i : i \geq 1\}$ de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0,1]$ de fonction de répartition U. Les suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes.

Posons, pour tout $x \in [0,1]$ et pour tout $n \geq 3$,

$$G_{n,c}(x) = \frac{n\{\mathbb{U}_{n,c}(x) - x\}}{\sqrt{2\mathbb{E}[Y^2]\, n \log\log n}}. \tag{2.14}$$

Soit $K_{\mathbb{U}_c}$ l'ensemble des fonctions $f_c \in \mathcal{E}[0,1]$ telles que

-i- $f_c(0) = f_c(1) = 0$,

-ii- f_c est absolument continue par rapport à la mesure de Lebesgue, et

-iii- $\int_0^1 [f_c'(s)]^2\, ds \leq 1$,

où la fonction f_c' est la dérivée de Lebesgue de la fonction f_c (déterminée presque partout par rapport à la mesure de Lebesgue).

Théorème 2.4.1. *La suite $(G_{n,c}(\cdot))_{n=3,4,\cdots}$ est presque sûrement relativement compacte dans $\mathcal{E}[0,1]$ et l'ensemble de ses points limites est l'ensemble compact $K_{\mathbb{U}_c}$.*

Remarque 2.4.1. La compacité de $K_{\mathbb{U}_c}$ est une conséquence du lemme 2.1.1.

2.5. Démonstration du théorème 2.4.1

La preuve du théorème 2.4.1 dépend de la borne (2.8) et de la généralisation de la loi du logarithme itére pour les sommes de variables aléatoires indépendantes à valeurs réelles (voir le lemme 2.5.1).

Lemme 2.5.1. *Soit $\mathbf{Z_1}, \mathbf{Z_2}, \cdots$ des vecteurs aléatoires idépendants et identiquement distribués à valeurs dans l'espace euclidien $\mathbb{R}^m$ de dimension m, tels que*

$$\mathbb{E}[\mathbf{Z_1}] = \mathbf{0}$$

et

$$\mathbb{E}[\mathbf{Z_1}\mathbf{Z_1}^T] = \mathbf{I}_m, \text{ où } \mathbf{I}_m \text{ désigne la matrice identité de dimension } m.$$

Soit

$$\mathbf{\Sigma_n} = \frac{\sum_{i=1}^{n} \mathbf{Z_i}}{\sqrt{2n \log\log n}}.$$

Alors, la suite $\{\mathbf{\Sigma_n} : n \geq 3\}$ est presque sûrement relativement compacte et l'ensemble de ses points limites est

$$B_m = \{x \in \mathbb{R}^m : ||x|| \leq 1\}$$

où $||.||$ est la norme euclidienne de $\mathbb{R}^m$.

Démonstration. On pourra consulter la démonstration du lemme 2.5.1 dans l'article de Finkelstein ([55], page 609-610). □

Soit m un entier suffisamment grand. Divisons alors l'intervalle $[0,1]$ en m sous-intervalles égaux $I_i = [(i-1)/m, i/m]$ pour $i = 1, 2, \cdots, m$. Pour tout entier $n = 1, 2, \cdots$ et pour tout entier $i = 1, 2, \cdots, m$ définissons

$$\begin{aligned} W_{ni} &= Y_n \quad \text{si } U_n \in I_i; \\ &= 0 \quad \text{si } U_n \notin I_i. \end{aligned}$$

Les vecteurs $\mathbf{W_n} = (W_{n1}, W_{n2}, \cdots, W_{nm})$ de $\mathbb{R}^m$ sont des vecteurs aléatoires indépendants et identiquement distribués, tels que

$$\begin{aligned} \mathbb{E}[\mathbf{W_1}] &= \left(\frac{1}{m}, \cdots, \frac{1}{m}\right), \\ \mathbb{E}\big[(\mathbf{W_1} - \mathbb{E}[\mathbf{W_1}])(\mathbf{W_1} - \mathbb{E}[\mathbf{W_1}])^T\big] &= \mathbf{\Gamma} \end{aligned}$$

où $\mathbf{\Gamma}$ est la matrice de coefficients $(\gamma_{ij})_{i=1,\cdots,m,\ j=1,\cdots,m}$, où γ_{ij} est défini par

$$\begin{aligned} \gamma_{ij} &= \mathbb{V}_c\left(\frac{1}{m}\right) - \frac{1}{m^2} \quad \text{si } i = j; \\ &= \quad - \frac{1}{m^2} \quad \text{si } i \neq j, \end{aligned}$$

où

$$\mathbb{V}_c\left(\frac{1}{m}\right) = \mathbb{E}\left[Y_1^2 \mathbf{1}_{\{U_1 \in I_i\}}\right] = \mathbb{E}\left[Y_1^2\right] \mathbb{E}\left[\mathbf{1}_{\{U_1 \in I_i\}}\right] = \frac{\mathbb{E}[Y^2]}{m}. \tag{2.15}$$

Lemme 2.5.2. *La suite*

$$\left\{ \frac{\sum_{i=1}^n \left(\mathbf{W_i} - \mathbb{E}[\mathbf{W_i}]\right)}{\sqrt{2n \log\log n}} : n \geq 3 \right\}$$

est presque sûrement relativement compacte et l'ensemble de ses points limites est

$$C_m = \left\{ \mathbf{x} \in \mathbb{R}^m, \mathbf{x} = (x_1, \cdots, x_m) : \sum_{i=1}^m x_i = 0 \ \ et \ \sum_{i=1}^m \frac{x_i^2}{\mathbb{V}_c\,(1/m)} \leq 1 \right\},$$

où $\mathbb{V}_c\,(1/m)$ *aété définie en* (2.15).

Démonstration. L'étendue de $\mathbf{W_1} - \mathbb{E}[\mathbf{W_1}]$ est l'hyperplan $\mathcal{H}$ défini par

$$\sum_{i=1}^m x_i = 0$$

pour $\mathbf{x} = (x_1, \cdots, x_m) \in \mathbb{R}^m$. Il existe une transformation linéaire $T : \mathbb{R}^{m-1} \to \mathcal{H}$ et des vecteurs aléatoires indépendants et identiquement distribués $\mathbf{Z_1}, \mathbf{Z_2}, \cdots$ de $\mathbb{R}^{m-1}$ tels que $\mathbb{E}[\mathbf{Z_1}] = 0$, $\mathbb{E}[\mathbf{Z_1}{\mathbf{Z_1}}^T] = \mathbf{I}_{m-1}$ et

$$\mathbf{W_i} - \mathbb{E}[\mathbf{W_i}] = \mathrm{T}\mathbf{Z_i}$$

pour $i = 1, 2, \cdots$. Pour tout vecteur $\mathbf{a} \in \mathcal{H}$, on a

$$\mathbf{a\Gamma} = \mathbb{V}_c\left(\frac{1}{m}\right)\mathbf{a}.$$

On peut alors choisir la transformation T comme la composition d'une isométrie et de l'homothétie de rapport $\mathbb{V}_c(1/m)$. On obtient ainsi $\mathrm{T}(B_{m-1}) = C_m$. On peut donc appliquer le lemme 2.5.1 aux variables aléatoires $\mathbf{Z_i}$, et l'application de la transformation T à ce résultat donne le lemme 2.5.2. □

Soit $H_{n,m,c}(x)$ l'interpolation de la fonction $G_{n,c}(x)$ sur $[0,1]$, définie de la façon suivante : pour $x = i/m$: $H_{n,m,c}(x) = G_{n,c}(x)$, et pour $x \in I_i = [(i-1)/m, i/m]$:

$$H_{n,m,c}(x) = G_{n,c}\left(\frac{i-1}{m}\right) + m\left(x - \left(\frac{i-1}{m}\right)\right)\left(G_{n,c}\left(\frac{i}{m}\right) - G_{n,c}\left(\frac{i-1}{m}\right)\right).$$

Lemme 2.5.3. *Pour tout $m \geq 1$ fixé, la suite $(H_{n,m,c}(x))_{n=3,4,\cdots}$ est relativement compacte dans $\mathcal{E}$ et l'ensemble de ses points limites est égal à*

$$J_{m,c} = \{f_c \in K_c : f_c \text{ est linéaire sur } I_i \text{ pour } i = 1, 2, \cdots, m\}.$$

Démonstration. D'abord remarquons que pour $i = 1, 2, \cdots, m$

$$H_{n,m,c}\left(\frac{i}{m}\right) - H_{n,m,c}\left(\frac{i-1}{m}\right) = \frac{\sum_{k=1}^{n}\left(W_{ki} - \mathbb{E}[W_{ki}]\right)}{\sqrt{2\mathbb{E}[Y^2] n \log\log n}}. \tag{2.16}$$

Soit $\mathcal{L}_m$ l'espace des fonctions continues sur $[0,1]$ qui valent 0 au point 0 et linéaires sur les intervalles I_i pour $i = 1, 2, \cdots, m$. Munissons $\mathcal{L}_m$ de la topologie uniforme. Alors $H_{n,m,c}(\cdot)$ est un élément de $\mathcal{L}_m$.
Soit V l'application de $\mathcal{L}_m$ vers $\mathbb{R}^m$ qui à $f \in \mathcal{L}_m$ associe le vecteur de coordonnées $(f(i/m) - f((i-1)/m))$ pour $i = 1, \cdots, m$. Alors l'égalité (2.16) et le lemme 2.5.2 impliquent que la suite $\{VH_{n,m,c}(\cdot) : n \geq 3\}$ est relativement compacte dans $\mathbb{R}^m$ et l'ensemble de ses points limites est C_m. Comme l'application V est $1-1$ et bicontinue, la suite $\{H_{n,m,c}(\cdot) : n \geq 3\}$ est relativement compacte dans $\mathcal{E}$ et l'ensemble de ses points limites est $V^{-1}(C_m) = J_m$. □

Soit $f_c \in K_c$ et soit $g \in J_{m,c}$ son approximation linéaire : $g_c\left(\frac{i}{m}\right) = f_c\left(\frac{i}{m}\right)$ pour $i = 0, 1, \cdots, m$. De (2.9), il s'ensuit que pour $x \in I_i$

$$\begin{aligned} |f_c(x) - g_c(x)| &\leq \left|f_c(x) - f_c\left(\frac{i-1}{m}\right)\right| + \left|f_c\left(\frac{i}{m}\right) - f_c\left(\frac{i-1}{m}\right)\right| \\ &\leq \frac{2}{\sqrt{m}}. \end{aligned}$$

Alors $J_{m,c}$ est une bonne approximation de K_c dans le sens où pour tout fonction $f_c \in K_c$ il existe une fonction $g_c \in J_{m,c}$ telle que $||f_c - g_c|| \leq 2/m^{-1/2}$. Alors le lemme 2.5.3 implique le corollaire suivant.

Corollaire 2.5.1. *-i- La suite $\{H_{n,m,c}(\cdot) : n \geq 3\}$ est relativement compacte,*

-ii- L'ensemble de ses points limites est inclus dans K_c,

-iii- Pour toute fonction $f_c \in K_c$ et pour tout $\varepsilon > 0$, si $m > \frac{1}{2}\varepsilon^2$ alors, infiniment souvent

$$\sup_{x\in[0,1]} \left|H_{n,m,c}(x) - f_c(x)\right| < \varepsilon \qquad \text{quand } n \to \infty.$$

Lemme 2.5.4. *Pour tout entier $m \geq 1$ et pour tout $i = 1, 2, \cdots, m$ il existe un entier positif N tel que si $n > N$*

$$\sup_{x\in I_i} \left|H_{n,m,c}(x) - G_{n,c}(x)\right| < \frac{1}{\sqrt{m}}.$$

Démonstration. Fixons les entiers m et i. Posons

$$d_c(n) = \sqrt{2\mathbb{E}[Y^2]n \log\log n} \sup_{x\in I_i} \left|H_{n,m,c}(\omega, x) - G_{n,c}(\omega, x)\right|.$$

Définissons ν_c une fonction de $\mathbb{N}$ à valeurs réelles définies par

$$\nu_c(0) = 0 \quad \text{ et } \quad \nu_c(n) = n\mathbb{U}_{n,c}\left(\frac{i}{m}\right) - n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right) \qquad \text{pour } n \geq 1.$$

Pour tout $n \geq 1$, on a

$$\begin{aligned} \nu_c(n) &= \nu_c(n-1) + Y_n \quad \text{si } U_n \in I_i; \\ &= \nu_c(n-1) \quad \text{si } U_n \notin I_i. \end{aligned} \tag{2.17}$$

$\nu_c(n)$ est une somme de variables aléatoires Y_i où le nombre de variables aléatoires Y_i de cette somme dépend du nombre de variables alátoires de $U_1, \cdots, U_n$ qui tombent dans l'intervalle I_i.

Par calcul, nous obtenons que

$$\begin{aligned} d_c(n) &= \sup_{x\in I_i} \Bigg| n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right) - n\left(\frac{i-1}{m}\right) - n\mathbb{U}_{n,c}(x) + nx + m\left(x - \frac{i-1}{m}\right) \\ &\qquad \times \left[n\mathbb{U}_{n,c}\left(\frac{i}{m}\right) - n\left(\frac{i}{m}\right) - n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right) + n\left(\frac{i-1}{m}\right)\right]\Bigg| \\ &= \sup_{x\in I_i} \Bigg| n\mathbb{U}_{n,c}(x) - n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right) - m\left(x - \frac{i-1}{m}\right) \\ &\qquad \times \left[n\mathbb{U}_{n,c}\left(\frac{i}{m}\right) - n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right)\right]\Bigg| \\ &= \sup_{x\in I_i} \left| n\mathbb{U}_{n,c}(x) - n\mathbb{U}_{n,c}\left(\frac{i-1}{m}\right) - \nu_c(n)m\left(x - \frac{i-1}{m}\right)\right|. \end{aligned} \tag{2.18}$$

On remarque que, d'après (2.18), $d_c(n)$ dépend uniquement des variables aléatoires U_i parmi les variables aléatoires $U_1, \cdots, U_n$ qui tombent dans l'intervalle I_i.

Soient $k \leq j_1 < j_2 < \cdots < j_p \leq K$ des entiers positifs. Soit A l'ensemble de $\mathcal{F}$ sur lequel, excepté les variables $U_k, U_{k+1}, \cdots, U_K$, toutes les variables de l'ensemble $U_{j_1}, U_{j_2}, \cdots, U_{j_p}$, mais aucune autre, prennent leurs valeurs dans I_i. La distribution jointe conditionnelle des variables alátoires $U_{j_1}, U_{j_2}, \cdots, U_{j_p}$ sachant A est égale à la distribution de p variables aléatoires indépendantes et uniformément distribuées sur I_i .

Afin de déterminer la distribution de d_c, définissons $U_1^*, U_2^*, \cdots$ des variables aléatoires indépendantes et uniformément distribuées sur I_i. Soit $\mathbb{U}_{n,c}^*(x)$ la fonction de répartition empirique composée des variables aléatoires Y_i et U_i^*. Posons

$$d_c^*(n) = \sup_{x \in I_i} \left| n\mathbb{U}_{n,c}^*(x) - nm\left(x - \frac{i-1}{m}\right)\right|.$$

Posons $\lambda(z) = \sqrt{2\mathbb{E}[Y^2] z \log\log z}$.

Si $(s_k, s_{k+1}, \cdots, s_K)$ une suite de nombres réels des valeurs possibles que peut prendre la suite $(\nu_c(k), \nu_c(k+1), \cdots, \nu_c(K))$, alors selon 2.17, l'ensemble

$$B = \{\nu_c(k) = s_k, \nu_c(k+1) = s_{k+1}, \cdots, \nu_c(K) = s_K\}$$

semblable à l'ensemble A. Alors la probabilité conditionnelle

$$\begin{aligned} \mathbb{P} \quad \Big[\quad & d_c(n) > \lambda(\nu_c(n)) \quad \text{ pour } n \text{ tel que } : k \leq n \leq K \Big| B\Big] \\ = \quad & \mathbb{P}\Big[d_c^*(s) > \lambda(s) \quad \text{ pour } s \text{ tel que } : s_k \leq s \leq s_K\Big]. \end{aligned} \tag{2.19}$$

Soit S l'ensemble de toutes les suites possibles des valeurs que peut prendre $\nu_c(n)$. Pour $\mathbf{s} \in S$ posons $\mathbf{s}= (s_i)_{i=1,2,\cdots}$. Alors (2.19) implique

$$\begin{aligned} \mathbb{P} \quad \Big[\quad & d_c(n) > \lambda(\nu_c(n)) \quad \text{ pour } n \text{ tel que } : k \leq n \leq K\Big] \\ = \quad & \int_S \mathbb{P}\Big[d_c^*(s) > \lambda(s) \quad \text{ pour } s \text{ tel que } : s_k \leq s \leq s_K\Big]\mathbb{P}_{\nu_c}(d\mathbf{s}) \end{aligned} \tag{2.20}$$

où $\mathbb{P}_{\nu_c}$ est la distribution de $(\nu_c(n))_{n=1,2,\cdots}$.

Posons $C = \{\mathbf{s} \in S : \lim_n s_n = \infty\}$. D'après la loi forte des grands nombres, on a $\mathbb{P}_{\nu_c}(C) = 1$. Alors le théorème de la convergence monotone appliqué à (2.20) en faisant tendre K vers ∞ et par conséquent $k \uparrow \infty$, conduit à

$$\begin{aligned} \mathbb{P} \quad \big[\quad & \limsup_n \{d_c(n) > \lambda(\nu_c(n))\}\big] \\ = \quad & \int_C \mathbb{P}\big[\limsup_{s_n} \{d_c^*(s_n) > \lambda(s_n))\}\big]\mathbb{P}_{\nu_c}(d\mathbf{s}) \\ = \quad & \int_C \mathbb{P}\big[\limsup_{s} \{d_c^*(s) > \lambda(s))\}\big]\mathbb{P}_{\nu_c}(d\mathbf{s}). \end{aligned}$$

Pour conclure la démonstration de ce lemme, nous avons besoin de rappeler un résultat établi au chapitre précédent.

Théorème 2.5.1. *Soit une suite* $\{Y_i : i \geq 1\}$ *de v.a.i.i.d. de même loi que la variable aléatoire* $Y = Y_1$. *Supposons que* $\mathbb{E}[Y] = 1$, $\mathrm{Var}[Y] < \sigma^2 < \infty$ *et que pour* $r \geq 5$, $\mathbb{E}[|Y|^r] < \infty$. *Soit une suite* $\{U_i : i \geq 1\}$ *de v.a. uniformément distribuées sur l'intervalle* $[0,1]$. *Les suites* $\{Y_i : i \geq 1\}$ *et* $\{U_i : i \geq 1\}$ *sont supposées indépendantes. Alors, nous avons*

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2\mathbb{E}[Y^2] n \log\log n}} \sup_{x\in[0,1]} \left| \sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq x\}} - nx \right| = 1 p.s. \tag{2.21}$$

(2.21) implique que

$$\mathbb{P}\left[\limsup_{s} (d_c^*(s) > \lambda(s))\right] = 0.$$

De plus, nous avons

$$\mathbb{P}\left[\limsup_{n} (d_c(n) > \lambda(\nu_c(n)))\right] = 0. \tag{2.22}$$

La loi forte des grands nombres implique

$$\frac{\lambda(\nu_c(n))}{\lambda(n)} \to \frac{1}{\sqrt{m}} \text{ p.s.} \qquad \text{lorsque } n \to \infty. \tag{2.23}$$

Comme,

$$\sup_{x\in I_i} |H_{n,m,c}(x) - G_{n,c}(x)| = \ ,$$

(2.22) et (2.23) impliquent le lemme 2.5.4. □

Il s'ensuit du lemme 2.5.4 et du corollaire 2.5.1 que

-i- la suite $(G_{n,c}(\cdot))_{n\geq 3}$ est relativement compacte,

-ii- l'ensemble de ses points limites est contenue dans l'ensemble compact $K_{\mathbb{U}_c}$, et

-iii- pour toute fonction $f \in K$, et pour tout $\varepsilon > 0$, $\sup_{x\in[0,1]} |G_{n,c}(x) - f(x)| < \varepsilon$ infiniment souvent lorsque $n \to \infty$. Ce qui achève la démonstration du théorème 2.4.1.

Chapitre 3

Quelques résultats sur le processus de Poisson composé

3.1. Définition et exemples

Dans ce chapitre, nous définissons un processus de Poisson composé comme un processus stochastique défini, pour tout réel $t \geq 0$, par

$$\Pi_c(t) = \sum_{i=1}^{\Pi(t)} Y_i, \tag{3.1}$$

où $\{\Pi(t) : t \geq 0\}$ désigne un processus de Poisson homogène continu à droite et de paramètre $\lambda > 0$, et la famille $\{Y_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. Le processus de Poisson $\{\Pi(t) : t \geq 0\}$ et la suite de variables aléatoires $\{Y_i : i \geq 1\}$ sont supposés indépendants. Posons

$$\sum_{\emptyset}(.) = 0$$

pour donner un sens à la définition (3.1) lorsque $\Pi(t) = 0$. On rappelle que la lettre "c" est utilisée ici pour évoquer le caractère composé.

Remarque 3.1.1. Le membre droit de (3.1) est une somme d'un nombre aléatoire de variables aléatoires indépendantes et identiquement distribuées.

Exemple 3.1.1. Des clients entrent dans un grand magasin selon un processus de Poisson $\{\Pi(t) : t \geq 0\}$ de paramètre $\lambda > 0$. On suppose que les montants $\{Y_i : i \geq 1\}$ des dépenses faites par chaque client forment une suite de variables aléatoires indépendantes et identiquement distribuées. De plus, on suppose que la suite de variables aléatoires $\{Y_i : i \geq 1\}$ est indépendante du processus de Poisson $\{\Pi(t) : t \geq 0\}$. Avec la notation (3.1), le processus $\{\Pi_c(t) : t \geq 0\}$ est le montant des dépenses faites par les clients dans le magasin pendant l'intervalle de temps $[0, t]$.

Exemple 3.1.2. Soit $\Pi(t)$ le nombre de sinistres enregistrés par une compagnie d'assurance durant l'intervalle de temps $[0, t]$. Le processus $\{\Pi(t) : t \geq 0\}$ peut être considéré comme un processus de Poisson de paramètre $\lambda > 0$. Supposons que chaque sinistre est classé. Précisons alors le classement de ces sinistres : un sinistre est de type k, si son coût est égal à k unités de monnaie. Alors

$$\Pi(t) = \sum_{k\geq 1} \Pi_k(t) \text{ est le nombre total de sinistres pendant l'intervalle } [0,t]$$

et

$$S(t) = \sum_{k\geq 1} k\Pi_k(t) \text{ est le coût total des sinistres pendant l'intervalle } [0,t].$$

On peut imaginer une autre expression pour le coût total des sinistres. Dans l'intervalle de temps $[0, t]$ se produisent $\Pi(t)$ sinistres, qui coûtent $Y_1, Y_2, \ldots, Y_{\Pi(t)}$ unités de monnaie, respectivement. Le coût total de ces sinistres est donc égal à $S'(t) = Y_1 + \ldots + Y_{\Pi(t)}$. Supposons que les coûts Y_i sont des variables aléatoires, équidistribuées, de loi $\sum_k p_k \varepsilon_k$ et que $(\Pi(t), Y_1, Y_2, \ldots)$ est une suite de variables aléatoires indépendantes, la variable $\Pi(t)$ suivant la loi de Poisson de paramètre $\lambda t > 0$. Chaque variable aléatoire Y_i est ainsi interprétée comme le coût d'un sinistre non personnalisé. Sous ces hypothèses, on a le théorème suivant.

Théorème 3.1.1. *Les deux sommes $S(t)$ et $S'(t)$ ont même loi de probabilité. L'expression de $S'(t)$ montre que cette loi est une loi de Poisson composée.*

Démonstration. Notons

$$G_Y(u) = \sum_{k\geq 1} p_k u^k$$

la fonction génératrice de chaque variable aléatoire Y_i. Par ailleurs, les variables de Poisson $\Pi(t)$ et $\Pi_k(t)$ ont pour fonctions génératrices respectives

$$G_{\Pi(t)}(u) = \exp\left(\lambda t(u-1)\right)$$

et

$$G_{\Pi_k(t)}(u) = \exp\left(\lambda p_k t(u-1)\right).$$

Comme $G_{k\Pi_k(t)}(u) = \mathbb{E}[u^{k\Pi_k(t)}] = G_{\Pi_k(t)}(u^k)$, la fonction génératrice de $S(t)$ est égale à

$$\begin{aligned}
G_{S(t)}(u) &= \prod_{k\geq 1} G_{k\Pi_k(t)}(u) = \prod_{k\geq 1} G_{\Pi_k(t)}(u^k) = \exp\left(\lambda t \sum_{k\geq 1} p_k\left(u^k - 1\right)\right) \\
&= \exp\left(\lambda t\left(\sum_{k\geq 1} p_k u^k - 1\right)\right) = \exp\left(\lambda t\left(G_Y(u) - 1\right)\right) \\
&= G_{\Pi(t)} \circ G_Y(u),
\end{aligned}$$

qui est précisément la fonction génératrice de $S'(t)$. □

3.2. Des résultats classiques sur le processus de Poisson composé

Pour obtenir une information sur la loi du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$, lorsque la suite de variables $\{Y_i : i \geq 1\}$ est quelconque (dans l'exemple 3.1.2 la suite de variables $\{Y_i : i \geq 1\}$ est supposée à valeurs dans $\mathbb{N}$), on détermine la fonction caractéristique du processus $\{\Pi_c(t) : t \geq 0\}$, au lieu de sa fonction génératrice.

3.2.1. Stationnarité des accroissements et fonction caractéristique du processus de Poisson composé

Théorème 3.2.1. *Un processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ a des accroissements indépendants et stationnaires. De plus, sa fonction caractéristique est définie, pour tous réels u et $t \geq 0$, par*

$$\varphi_{\Pi_c(t)}(u) = \exp\big(-\lambda t(1 - \varphi_Y(u))\big), \tag{3.2}$$

où $\varphi_Y(u) = \mathbb{E}[\exp(iuY)]$ désigne la fonction caractéristique de la variable aléatoire Y et $\lambda > 0$ représente le paramètre du processus de Poisson $\{\Pi(t) : t \geq 0\}$.

Si $\mathbb{E}[Y^2] < \infty$, alors le processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ a des moments d'ordre 2, qui sont donnés, pour tout couple de réels t et $s \geq 0$, par

$$\begin{aligned}
\mathbb{E}[\Pi_c(t)] &= \lambda t \mathbb{E}[Y] && (3.3)\\
\mathrm{Var}[\Pi_c(t)] &= \lambda t \mathbb{E}[Y^2] && (3.4)\\
\mathrm{Cov}[\Pi_c(t), \Pi_c(s)] &= \lambda \mathbb{E}[Y^2] \min(t, s). && (3.5)
\end{aligned}$$

Remarque 3.2.1. Les égalités (3.3) et (3.4) sont des cas particuliers de l'identité de Wald.

Démonstration. Ce résultat est classique (voir par exemple le livre de Karlin et Taylor [80] pp.427-429). Nous en donnons néanmoins une démonstration par souci du détail. Pour montrer que le processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ a des accroissements indépendants et stationnaires, nous allons procéder en deux étapes.

Première étape. Montrons d'abord que les accroissements du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$, qui ne se chevauchent pas, sont indépendants.

Soit une suite de réels $\{t_0, \ldots, t_n\}$, telle que $0 \leq t_0 < t_1 < \ldots < t_n$. Comme les variables aléatoires $Y_1, Y_2, \ldots$ sont mutuellement indépendantes et aussi indépendantes du processus de Poisson $\{\Pi(t) : t \geq 0\}$, qui lui-même a des accroissements indépendants, les variables aléatoires

$$\begin{array}{lrcccr}
\{ & Y_1, & \ldots, & Y_{\Pi(t_0)}, & \Pi(t_0) & \}, \\
\{ & Y_{\Pi(t_0)+1}, & \ldots, & Y_{\Pi(t_1)}, & \Pi(t_1) - \Pi(t_0) & \}, \\
\{ & \vdots & \vdots & \vdots & \vdots & \}, \\
\{ & Y_{\Pi(t_{n-1})+1}, & \ldots, & Y_{\Pi(t_n)}, & \Pi(t_n) - \Pi(t_{n-1}) & \}
\end{array}$$

sont indépendantes.

Ainsi les accroissements du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$

$$\begin{array}{rcl} \Pi_c(t_0) - \Pi_c(0) & = & Y_1 + \ldots + Y_{\Pi(t_0)}, \\ \Pi_c(t_1) - \Pi_c(t_0) & = & Y_{\Pi(t_0)+1} + \ldots + Y_{\Pi(t_1)}, \\ & \vdots & \\ \Pi_c(t_n) - \Pi_c(t_{n-1}) & = & Y_{\Pi(t_{n-1})+1} + \ldots + Y_{\Pi(t_n)} \end{array}$$

sont indépendants.

Deuxième étape. Montrons maintenant que la distribution des accroissements du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ est stationnaire. Soient deux réels $t > 0$ et $h > 0$. Alors $\Pi(t+h) - \Pi(t)$ et $\Pi(h)$ ont même loi de probabilité. De même $Y_1, \ldots, Y_{\Pi(h)}$ et $Y_{\Pi(t)+1}, \ldots, Y_{\Pi(t+h)}$ sont de même loi. Ainsi

$$\Pi_c(t+h) - \Pi_c(t) = Y_{\Pi(t)+1} + \ldots + Y_{\Pi(t+h)}$$

et

$$\Pi_c(h) = Y_1 + \ldots + Y_{\Pi(h)}$$

ont la même distribution. Ceci complète la preuve de l'indépendance et de la stationnarité des accroissements du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$.

Pour établir l'inégalité (3.2), il suffit de prouver, pour tout couple de réels t et s tels que $t > s \geq 0$, et pour tout réel u, que

$$\varphi_{\Pi_c(t)-\Pi_c(s)}(u) = \exp\left[\lambda(t-s)\big\{\varphi_Y(u) - 1\big\}\right]. \tag{3.6}$$

Observons que, pour tout entier $n \geq 0$, nous avons

$$\mathbb{E}\left[\exp\Big(iu\big\{\Pi_c(t) - \Pi_c(s)\big\}\Big)\Big|\Pi(t) - \Pi(s) = n\right] = \{\varphi_Y(u)\}^n,$$

sachant qu'il y a n événements réalisés dans l'intervalle $(s,t]$, que $\Pi_c(t) - \Pi_c(s)$ est la somme de n variables aléatoires indépendantes et identiquement distribuées, de même loi que la variable aléatoire Y. Donc, pour tout réel u et pour tout couple de réels t et s tels que $t > s \geq 0$, nous avons

$$\begin{aligned} \varphi_{\Pi_c(t)-\Pi_c(s)}(u) &= \sum_{n=0}^{\infty} \mathbb{E}\left[\exp\Big(iu\{\Pi_c(t) - \Pi_c(s)\}\Big)\Big|\Pi(t) - \Pi(s) = n\right] \\ &\qquad \times \mathbb{P}\left[\Pi(t) - \Pi(s) = n\right] \\ &= \sum_{n=0}^{\infty} \{\varphi_Y(u)\}^n \exp\left(-\lambda(t-s)\right) \frac{\{\lambda(t-s)\}^n}{n!} \\ &= \exp\left(-\lambda(t-s)\right)\exp\left(\lambda(t-s)\varphi_Y(u)\right) \\ &= \exp\Big(\lambda(t-s)\big\{\varphi_Y(u) - 1\big\}\Big). \end{aligned}$$

La preuve de l'égalité (3.6) est ainsi achevée.

Pour prouver les inégalités (3.3), (3.4) et (3.5), on peut soit dériver l'équation (3.2), soit utiliser les formules pour la moyenne et la variance d'une somme d'un nombre aléatoire de variables aléatoires indépendantes et identiquement distribuées, ce qui donne

$$
\begin{aligned}
\mathbb{E}[\Pi_c(t)] &= \mathbb{E}[\Pi(t)]\mathbb{E}[Y] \\
&= \lambda t \mathbb{E}[Y], \\
\mathrm{Var}[\Pi_c(t)] &= \mathbb{E}[\Pi(t)]\mathrm{Var}[Y] + \mathrm{Var}[\Pi(t)]\mathbb{E}^2[Y] \\
&= \mathbb{E}\left[\Pi(t)\right]\left(\mathrm{Var}[Y] + \mathbb{E}^2[Y]\right) \\
&= \lambda t \mathbb{E}[Y^2], \\
\mathrm{Cov}[\Pi_c(t), \Pi_c(s)] &= \mathbb{E}[\Pi_c(t)\Pi_c(s)] - \mathbb{E}[\Pi_c(t)]\mathbb{E}[\Pi_c(s)] \\
&= \mathbb{E}\left[\sum_{i=1}^{\Pi(t)} Y_i \sum_{j=1}^{\Pi(s)} Y_j\right] - \lambda^2 st \mathbb{E}^2[Y] \\
&= \mathbb{E}\left[\sum_{i=1}^{\min(\Pi(t),\Pi(s))} Y_i^2\right] + \mathbb{E}\left[\sum_{i \neq j} Y_i Y_j\right] - \lambda^2 st \mathbb{E}^2[Y] \\
&= \min(t,s)\lambda \mathbb{E}[Y^2],
\end{aligned}
$$

puisque la variable aléatoire $\Pi(t)$ suit une loi de Poisson de paramètre $\lambda > 0$. Ceci achève la démonstration du théorème 3.2.1. □

3.2.2. Fonction de répartition du processus de Poisson composé

La fonction de répartition du processus de Poisson composé $\{\Pi_c(t) : t \geq\}$ peut être représentée explicitement en conditionnant par les valeurs entières prises par le processus de Poisson $\{\Pi(t) : t \geq 0\}$ de paramètre $\lambda > 0$. Ainsi, pour tous réels x et $t \geq 0$, nous obtenons que

$$
\begin{aligned}
\mathbb{P}\left[\Pi_c(t) \leq x\right] &= \mathbb{P}\left[\sum_{i=1}^{\Pi(t)} Y_i \leq x\right] \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n \exp\left(-\lambda t\right)}{n!} \mathbb{P}\left[\sum_{i=1}^{\Pi(t)} Y_i \leq x \middle| \Pi(t) = n\right] \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n \exp\left(-\lambda t\right)}{n!} F^{(n)}(x),
\end{aligned}
\tag{3.7}
$$

puisque le processus de Poisson $\{\Pi(t) : t \geq 0\}$ et la suite de variables $\{Y_i : i \geq 1\}$ sont supposés indépendants et où

$$F^{(n)}(x) = \mathbb{P}\left[Y_1 + \ldots + Y_n \leq x\right]$$

avec

$$F^{(0)}(x) = \begin{cases} 1 & \text{pour } x \geq 0, \\ 0 & \text{pour } x < 0. \end{cases}$$

Exemple 3.2.1. Soient $\Pi(t)$ le nombre de chocs reçus par un système jusqu'au temps t et Y_i le dégât ou la détérioration produit par le $i$ème choc. Nous supposons que la variable aléatoire "dégât" est positive, que $\mathbb{P}[Y_i \geq 0] = 1$ et que les dégâts sont additifs, c'est-à-dire que le dégât total à l'instant t est donné par la somme des dégâts dûs aux chocs jusqu'à l'instant t. Avec nos notations, le dégât total à l'instant t est égal au processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$. Supposons que le sytème continue à fonctionner aussi longtemps que le dégât total est strictement plus petit qu'une certaine valeur critique a, sinon le système s'arrête. Soit T le temps où le système tombe en panne. Alors nous avons

$$\{T > t\} \quad \text{si et seulement si} \quad \{\Pi_c(t) < a\}. \tag{3.8}$$

D'après (3.7) et (3.8), pour tout réel $t \geq 0$, nous obtenons que

$$\mathbb{P}\left[T > t\right] = \sum_{n=0}^{\infty} \frac{(\lambda t)^n \exp\left(-\lambda t\right)}{n!} F^{(n)}(a). \tag{3.9}$$

Calculons la moyenne de la variable aléatoire T. Pour tout réel $\lambda > 0$, nous avons

$$\begin{aligned}
\mathbb{E}[T] &= \int_0^{\infty} \mathbb{P}\left[T > t\right] dt \quad \text{par définition,} \\
&= \sum_{n=0}^{\infty} \left(\int_0^{\infty} \frac{(\lambda t)^n \exp\left(-\lambda t\right)}{n!} dt \right) F^{(n)}(a) \quad \text{d'après (3.9),} \\
&= \frac{1}{\lambda} \sum_{n=0}^{\infty} F^{(n)}(a),
\end{aligned}$$

tous les termes étant positifs nous pouvons permuter le signe $\int$ et le signe $\sum$.

Considérons le cas particulier où chaque variable aléatoire $Y_1, Y_2, \ldots$ suit une loi exponentielle de paramètre $\mu > 0$, c'est-à-dire

$$\mathbb{P}\left[Y_i \leq a\right] = 1 - \exp\left(-\mu a\right), \quad \text{où } a > 0.$$

La somme $Y_1 + \ldots + Y_n$ suit alors une loi gamma, et par conséquent nous avons

$$F^{(n)}(a) = 1 - \sum_{k=0}^{n-1} \frac{(\mu a)^k \exp\left(-\mu a\right)}{k!} = \sum_{k=n}^{\infty} \frac{(\mu a)^k \exp\left(-\mu a\right)}{k!}, \quad n \geq 0.$$

Par conséquent, nous en déduisons que

$$\begin{aligned}
\sum_{n=0}^{\infty} F^{(n)}(a) &= \sum_{n=0}^{\infty}\sum_{k=n}^{\infty} \frac{(\mu a)^k \exp(-\mu a)}{k!} \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{k} \frac{(\mu a)^k \exp(-\mu a)}{k!} \\
&= \sum_{k=0}^{\infty}(k+1) \frac{(\mu a)^k \exp(-\mu a)}{k!} \\
&= 1+\mu a.
\end{aligned}$$

Dans ce cas particulier, pour tout réel $\lambda > 0$, la moyenne de la variable aléatoire T est égale à

$$\mathbb{E}[T] = \frac{1}{\lambda}\sum_{n=0}^{\infty} F^{(n)}(a) = \frac{1+\mu a}{\lambda}.$$

3.2.3. Le processus de Poisson composé centré est une martingale

Théorème 3.2.2. *Le processus de Poisson composé centré* $\{\Pi_c(t) - \mathbb{E}[\Pi_c(t)] : t \geq 0\}$ *est une martingale relativement à la filtration* $\mathcal{A}_s = \sigma\{\Pi_c(u) : u \leq s\}, s \geq 0$.

Démonstration. La démonstration s'appuie sur le fait que le processus de Poisson composé $\{\Pi_c(t) - \mathbb{E}[\Pi_c(t)] : t \geq 0\}$ est centré, avec des accroissements indépendants (voir le théorème 3.2.1 de ce chapitre). Introduisons la σ-algèbre $\mathcal{A}_s$ définie par : $\mathcal{A}_s = \sigma\{\Pi_c(u) : u \leq s\}, s \geq 0$. Pour tout couple de réels $t \geq 0$ et $s \geq 0$, remarquons que

$$\Pi_c(t) - \mathbb{E}[\Pi_c(t)] = \Pi_c(t) - \Pi_c(s) + \Pi_c(s) - \mathbb{E}[\Pi_c(s)] + \mathbb{E}[\Pi_c(s)] - \mathbb{E}[\Pi_c(t)]. \tag{3.10}$$

En faisant usage du fait que l'accroissement $\Pi_c(t) - \Pi_c(s)$ est indépendant de l'accroissement $\Pi_c(s)$, nous obtenons alors

$$\begin{aligned}
\mathbb{E}\big[\Pi_c(t) - \mathbb{E}[\Pi_c(t)]\big|\mathcal{A}_s\big] &= \mathbb{E}\big[\Pi_c(t) - \Pi_c(s) + \Pi_c(s) - \mathbb{E}[\Pi_c(s)] \\
&\qquad + \mathbb{E}[\Pi_c(s)] - \mathbb{E}[\Pi_c(t)]\big|\mathcal{A}_s\big] \\
&= \mathbb{E}\big[\Pi_c(t) - \Pi_c(s)\big|\mathcal{A}_s\big] + \mathbb{E}\big[\Pi_c(s) - \mathbb{E}[\Pi_c(s)]\big|\mathcal{A}_s\big] \\
&\qquad + \mathbb{E}\big[\mathbb{E}[\Pi_c(s)] - \mathbb{E}[\Pi_c(t)]\big|\mathcal{A}_s\big] \\
&= \mathbb{E}[\Pi_c(t) - \Pi_c(s)] + \Pi_c(s) - \mathbb{E}[\Pi_c(s)] + \mathbb{E}[\Pi_c(s)] - \mathbb{E}[\Pi_c(t)] \\
&= \mathbb{E}[\Pi_c(t)] - \mathbb{E}[\Pi_c(s)] + \Pi_c(s) - \mathbb{E}[\Pi_c(s)] \\
&\qquad + \mathbb{E}[\Pi_c(s)] - \mathbb{E}[\Pi_c(t)] \\
&= \Pi_c(s) - \mathbb{E}[\Pi_c(s)].
\end{aligned}$$

Le processus de Poisson composé centré $\{\Pi_c(t) - \mathbb{E}[\Pi_c(t)] : t \geq 0\}$ est donc une martingale relativement à la filtration $\mathcal{A}_s = \sigma\{\Pi_c(u) : u \leq s\}, s \geq 0$. □

3.3. Inégalités de type exponentiel pour le processus de Poisson composé

Le résultat principal de ce paragraphe est basé sur une inégalité de type Bennett pour le processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$. Dans un premier temps, nous rappelons l'inégalité classique de Bennett [2] dans le théorème 3.3.1. Pour cela, introduisons la fonction

$$\psi(u) = \frac{2h(1+u)}{u^2} = \frac{2}{u^2}\{(1+u)\log(1+u) - u\} \quad \text{pour} \quad u > 0, \tag{3.11}$$

où

$$h(u) = \begin{cases} u\log u - u + 1 & \text{pour } u > 0, \\ 1 & \text{pour } u = 0, \\ \infty & \text{pour } u < 0. \end{cases} \tag{3.12}$$

Nous allons énoncer quelques propriétés de la fonction ψ qui seront utiles par la suite.

Proposition 3.3.1. *Les propriétés les plus importantes de la fonction ψ sont*

$$\psi(\zeta) \text{ est décroissante pour } \zeta \geq -1, \text{ et } \psi(0) = 1. \tag{3.13}$$

$$\psi(\zeta) \sim \frac{2\log\zeta}{\zeta} \text{ lorsque } \zeta \to \infty. \tag{3.14}$$

$$\psi(\zeta) \geq 1 - \delta \quad \text{si} \quad 0 \leq \zeta \leq 3\delta \quad \text{et} \quad 0 \leq \delta \leq 1. \tag{3.15}$$

Démonstration. Pour la démonstration de cette proposition, on pourra consulter par exemple la proposition 1 p.441 du livre de Shorack et Wellner [125]. □

Théorème 3.3.1. *Soient $Y_1, \ldots, Y_n$ des variables aléatoires centrées et indépendantes telles que $|Y_i| \leq \mathcal{M}$ pour $i = 1, \ldots, n$. Alors, pour tout $x \geq 0$ et tout $V \geq V_n$, nous avons*

$$\mathbb{P}\left[|Y_1 + \ldots + Y_n| \geq x\right] \leq 2\exp\left(-\frac{1}{2}\left\{\frac{x^2}{V}\right\}\psi\left(\frac{\mathcal{M}x}{V}\right)\right), \tag{3.16}$$

où $V_n = \mathrm{Var}[Y_1 + \ldots + Y_n]$, et où la fonction ψ est définie en (3.11).

Démonstration. Il suffit de montrer que

$$\mathbb{P}\left[Y_1 + \ldots + Y_n \geq x\right] \leq \exp\left(-\frac{1}{2}\left\{\frac{x^2}{V}\right\}\psi\left(\frac{\mathcal{M}x}{V}\right)\right). \tag{3.17}$$

La version négative en découlera en faisant le changement formel de la variable Y_i en la variable $-Y_i$ pour $i = 1, \ldots, n$. Soit une variable aléatoire centrée Y telle que $|Y| \leq \mathcal{M}$. Posons $\mathrm{Var}[Y] = \sigma^2 < \infty$. Alors, pour tout réel $t \geq 0$, nous avons

$$\begin{aligned} \mathbb{E}[\exp(tY)] &= \mathbb{E}\left[1 + tY + t^2Y^2\sum_{k=2}^{\infty}\frac{t^{k-2}}{k!}Y^{k-2}\right] \\ &\leq 1 + t^2\sigma^2\sum_{k=2}^{\infty}\frac{(\mathcal{M}t)^{k-2}}{k!} = 1 + \sigma^2\left\{\frac{\exp(\mathcal{M}t) - 1 - \mathcal{M}t}{\mathcal{M}^2}\right\} \\ &\leq \exp\left(\sigma^2\phi(\mathcal{M}t)\right), \end{aligned}$$

où la fonction $\phi(\cdot)$ est définie par

$$\phi(u) = \frac{\exp(u) - 1 - u}{\mathcal{M}^2}.$$

Ainsi, en posant $\sigma_n^2 = \mathrm{Var}[Y_n]$, $V_n = \sigma_1^2 + \ldots + \sigma_n^2$ et $S_n = Y_1 + \ldots + Y_n$, pour tout réel $t \geq 0$ et pour $V \geq V_n$, nous obtenons que

$$\mathbb{E}[\exp(tS_n)] = \prod_{k=1}^{n} \mathbb{E}[\exp(tY_k)] \leq \prod_{k=1}^{n} \exp\left(\sigma_k^2 \phi(\mathcal{M}t)\right) = \exp\left(V_n \phi(\mathcal{M}t)\right) \leq \exp\left(V\phi\left(\mathcal{M}t\right)\right).$$

Nous concluons en remarquant que, pour tout réel $x \geq 0$, nous avons

$$\mathbb{P}\left[S_n \geq x\right] \leq \exp\left(-\sup_{t\geq 0}\left\{tx - \log \mathbb{E}[\exp(tS_n)]\right\}\right) \leq \exp\left(-\sup_{t\geq 0}\left\{tx - V\phi(\mathcal{M}t)\right\}\right).$$

La solution de l'équation

$$\frac{d}{dt}\left(tx - V\phi(\mathcal{M}t)\right) = x - \frac{V\left(\exp(\mathcal{M}t) - 1\right)}{\mathcal{M}} = 0,$$

est obtenue pour

$$t = t_0 = \frac{1}{\mathcal{M}} \log\left(1 + \frac{\mathcal{M}x}{V}\right) > 0.$$

Par conséquent, nous avons

$$\begin{aligned}
t_0 x - V\phi\left(\mathcal{M}t_0\right) &= \frac{x}{\mathcal{M}} \log\left(1 + \frac{\mathcal{M}x}{V}\right) - \frac{V}{\mathcal{M}^2}\left(1 + \frac{\mathcal{M}x}{V} - 1 - \log\left(1 + \frac{\mathcal{M}x}{V}\right)\right) \\
&= \left\{\frac{V}{\mathcal{M}^2}\right\} \times \left\{\left(1 + \frac{\mathcal{M}x}{V}\right)\log\left(1 + \frac{\mathcal{M}x}{V}\right) - \frac{\mathcal{M}x}{V}\right\} \\
&= \frac{1}{2}\left\{\frac{x^2}{V}\right\} \times 2 \times \left\{\frac{V}{\mathcal{M}x}\right\}^2 \left\{\left(1 + \frac{\mathcal{M}x}{V}\right)\log\left(1 + \frac{\mathcal{M}x}{V}\right) - \frac{\mathcal{M}x}{V}\right\} \\
&= \frac{1}{2}\left\{\frac{x^2}{V}\right\} \psi\left(\frac{\mathcal{M}x}{V}\right).
\end{aligned}$$

Auparavant nous avons obtenu l'inégalité suivante

$$\mathbb{P}\left[S_n \geq x\right] \leq \exp\left(-\left\{t_0 x - V\phi\left(\mathcal{M}t_0\right)\right\}\right).$$

Comme $t_0 x - V\phi\left(\mathcal{M}t_0\right) = \frac{1}{2}\left\{\frac{x^2}{V}\right\}\psi\left(\frac{\mathcal{M}x}{V}\right)$, nous en concluons que

$$\mathbb{P}\left[Y_1 + \ldots + Y_n \geq x\right] \leq \exp\left(-\frac{1}{2}\left\{\frac{x^2}{V}\right\}\psi\left(\frac{\mathcal{M}x}{V}\right)\right).$$

Ce qui achève la démonstration du théorème 3.3.1. □

Maintenant, introduisons les notations nécessaires pour énoncer le résultat principal de ce paragraphe, le théorème 3.3.2. Soit un réel $b > 0$. Soit $\Pi_c(b)$ une variable de Poisson composée définie par

$$\Pi_c(b) = \sum_{i=1}^{\Pi(b)} Y_i, \tag{3.18}$$

où $\Pi(b)$ désigne une variable de Poisson de paramètre $b > 0$ et la famille $\{Y_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$.

Théorème 3.3.2. *Soit un réel $b > 0$. Soit $\Pi_c(b)$ une variable de Poisson composée définie en* (3.18)*. Supposons que la variable aléatoire Y vérifie $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable et que $0 < \mathbb{E}[Y^2] < \infty$. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[|\Pi_c(b) - \mathbb{E}[\Pi_c(b)]| \geq \mu\right] \leq 2\exp\left(-\frac{\mu^2}{2b\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}[Y^2]}\right)\right), \tag{3.19}$$

où la fonction ψ est définie en (3.11)*.*

Remarque 3.3.1. Si, pour tout entier $i \geq 1$, les variables aléatoires $Y_i = 1$, alors nous avons l'égalité suivante

$$\Pi_c(b) = \sum_{i=1}^{\Pi(b)} 1 = \Pi(b).$$

Par conséquent, la variable $\Pi_c(b)$ suit une loi de Poisson de paramètre $b > 0$. Alors, pour tout réel $\mu \geq 0$, l'inégalité (3.19) devient

$$\mathbb{P}\left[\pm(\Pi(b) - b) \geq \mu\right] \leq \exp\left(-\frac{\mu^2}{2b}\psi\left(\frac{\pm\mu}{b}\right)\right), \tag{3.20}$$

où la fonction ψ est définie en (3.11). Cette inégalité est classique (voir par exemple le livre de Shorack et Wellner [125] pp.485-486). Nous en donnons néanmoins une démonstration de l'inégalité (3.20) par souci du détail.

Démonstration. Pour tout réel $\mu \geq 0$, nous avons successivement

$$\begin{aligned}
\mathbb{P}\left[\pm(\Pi(b) - b) \geq \mu\right] &= \mathbb{P}\left[\exp(\pm t[\Pi(b) - b]) \geq \exp(t\mu)\right] \quad \text{pour tout réel } t > 0 \\
&\leq \inf_{t>0} \exp\left(-t\mu + b\left(\exp(\pm t) - 1\right) \pm tb\right) \\
&\leq \begin{cases} \exp\left(\mu - (b+\mu)\log\left(\dfrac{b+\mu}{b}\right)\right) & \text{dans le cas “+”} \\ \exp\left(-\mu + (b-\mu)\log\left(\dfrac{b}{b-\mu}\right)\right) & \text{dans le cas “−”} \end{cases} \\
&\leq \exp\left(-\frac{\mu^2}{2b}\psi\left(\frac{\pm\mu}{b}\right)\right),
\end{aligned}$$

le minimum est obtenu en dérivant l'exponentielle par rapport à la variable t. $\square$

Démonstration. La démonstration du théorème 3.3.2 se fait en cinq étapes.

Première étape. Soit un réel $b > 0$. Soit $\mathbb{X}_{m,c}(b)$ une variable aléatoire définie, pour tout entier $m \geq 1$, par

$$\mathbb{X}_{m,c}(b) = \sum_{i=1}^{m} Y_i \mathbf{1}_{\{U_i \leq b/m\}}, \tag{3.21}$$

où $\{U_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et uniformément distribuées et $\{Y_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. De plus, les suites de variables $\{U_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont supposées indépendantes. Nous allons établir une inégalité de type Bennett pour la variable aléatoire $\mathbb{X}_{m,c}(b)$.

Inégalité 3.3.1. *Soit un réel $b > 0$. Soit $\mathbb{X}_{m,c}(b)$ la variable aléatoire définie en* (3.21). *Supposons que la variable aléatoire Y vérifie $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable et que $0 < \mathbb{E}[Y^2] < \infty$. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[|\mathbb{X}_{m,c}(b) - \mathbb{E}[\mathbb{X}_{m,c}(b)]| \geq \mu\right] \leq \exp\left(-\frac{\mu^2}{2b\mathbb{E}\left[Y^2\right]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}[Y^2]}\right)\right), \tag{3.22}$$

où la fonction ψ est définie en (3.11).

Démonstration. Le cas positif et le cas négatif sont similaires. Nous traiterons ici uniquement le cas positif. En appliquant l'inégalité de Bennett (3.17) avec $V = V_n = m\mathrm{Var}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right]$, pour tout réel $\mu \geq 0$, nous obtenons que

$$\begin{aligned}
\mathbb{P}\left[\mathbb{X}_{m,c}(b) - \mathbb{E}[\mathbb{X}_{m,c}(b)] \geq \mu\right] &= \mathbb{P}\left[\sum_{i=1}^{m}\left(Y_i\mathbf{1}_{\{U_i \leq b/m\}} - \mathbb{E}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right]\right) \geq \mu\right] \\
&\leq \exp\left\{-\frac{\mu^2}{2m\mathrm{Var}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right]}\psi\left(\frac{\mathcal{M}\mu}{m\mathrm{Var}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right]}\right)\right\}.
\end{aligned}$$

D'autre part, comme les suites $\{U_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont indépendantes, pour tout réel $b > 0$ et pour tout entier $m \geq 1$, nous avons

$$\begin{aligned}
\mathrm{Var}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right] &= \mathbb{E}\left[Y^2\mathbf{1}_{\{U_1 \leq b/m\}}\right] - \left(\mathbb{E}\left[Y\mathbf{1}_{\{U_1 \leq b/m\}}\right]\right)^2 \\
&= \mathbb{E}\left[\mathbf{1}_{\{U_1 \leq b/m\}}\right]\mathbb{E}\left[Y^2\right] - \left(\mathbb{E}\left[\mathbf{1}_{\{U_1 \leq b/m\}}\right]\right)^2\mathbb{E}^2\left[Y\right] \\
&= \frac{b}{m}\mathbb{E}\left[Y^2\right] - \left(\frac{b}{m}\right)^2\mathbb{E}^2\left[Y\right] \\
&\leq \frac{b}{m}\mathbb{E}\left[Y^2\right].
\end{aligned}$$

Finalement, pour tout réel $\mu \geq 0$, nous en déduisons que

$$\begin{aligned}\mathbb{P}\left[\mathbb{X}_{m,c}(b) - \mathbb{E}[\mathbb{X}_{m,c}(b)] \geq \mu\right] &\leq \exp\left(-\frac{\mu^2}{2m\frac{b}{m}\mathbb{E}\left[Y^2\right]}\psi\left(\frac{\mathcal{M}\mu}{m\frac{b}{m}\mathbb{E}\left[Y^2\right]}\right)\right) \\ &\leq \exp\left(-\frac{\mu^2}{2b\mathbb{E}\left[Y^2\right]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}\left[Y^2\right]}\right)\right).\end{aligned}$$

□

Deuxième étape. Maintenant établissons un lemme qui sera très utile dans létape suivante.

Lemme 3.3.1. *Soit un réel $b > 0$. Alors, pour tout entier $m \geq 1$, nous avons*

$$\sum_{i=1}^{m} Y_i \mathbf{1}_{\{U_i \leq b/m\}} \stackrel{d}{=} \sum_{i=1}^{B(m,b/m)} Y_i,$$

où $\stackrel{d}{=}$ signifie l'égalité en loi ou en distribution, et où $B(m,\, b/m)$ désigne une variable qui suit une loi binomiale de paramètres m et b/m.

Démonstration. La démonstration de ce lemme est évidente puisque les variables aléatoires $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont indépendantes et qu'une somme de m indicatrices de variables uniformes sur l'intervalle $[0, 1]$ suit une loi binomiale de paramètres m et b/m. □

Troisième étape. Soit un réel $b > 0$. Introduisons la variable aléatoire composée suivante

$$\mathbb{Y}_{m,c}(b) = \sum_{i=1}^{B(m,b/m)} Y_i. \tag{3.23}$$

Nous avons l'inégalité suivante.

Inégalité 3.3.2. *Soit un réel $b > 0$. Soit $\mathbb{Y}_{m,c}(b)$ la variable aléatoire composée définie en* (3.23)*. Supposons que la variable aléatoire Y vérifie $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable et que $0 < \mathbb{E}[Y^2] < \infty$. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[|\mathbb{Y}_{m,c}(b) - \mathbb{E}[\mathbb{Y}_{m,c}(b)]| \geq \mu\right] \leq \exp\left(-\frac{\mu^2}{2b\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}[Y^2]}\right)\right), \tag{3.24}$$

où la fonction ψ est définie en (3.11)*.*

Démonstration. La démonstration de cette inégalité s'appuie sur le lemme 3.3.1 et sur l'inégalité 3.3.1. □

Quatrième étape. Nous avons le lemme suivant.

Lemme 3.3.2. *Soit un réel $b > 0$. Soit $\mathbb{Y}_{m,c}(b)$ la variable aléatoire composée définie en* (3.23)*. Alors la variable $\mathbb{Y}_{m,c}(b)$ converge en loi vers la variable de Poisson composée $\Pi_c(b)$ où la variable $\Pi_c(b)$ est définie en* (3.18)*.*

Démonstration. Pour démontrer cette inégalité, il suffit de calculer la fonction caractéristique de la variable aléatoire composée $\mathbb{Y}_{m,c}(b)$ et d'utiliser le fait qu'une variable aléatoire qui suit une loi binomiale de paramètres m et b/m converge en loi vers une variable aléatoire qui suit une loi de Poisson de paramètre $b > 0$. □

Cinquième étape. D'après le lemme 3.3.2, la variable aléatoire composée $\mathbb{Y}_{m,c}(b)$ converge en loi vers la variable composée de Poisson $\Pi_c(b)$. Alors, pour tout réel $\mu \geq 0$, nous obtenons que

$$\mathbb{P}\left[|\Pi_c(b) - \mathbb{E}[\Pi_c(b)]| \geq \mu\right] \leq \exp\left(-\frac{\mu^2}{2b\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}[Y^2]}\right)\right). \tag{3.25}$$

Ceci achève la démonstration du théorème 3.3.2. □

Introduisons maintenant $\{\Pi_{n,c}(t) : t \geq 0\}$ un processus de Poisson composé défini, pour tout entier $n \geq 1$, par

$$\Pi_{n,c}(t) = \Pi_c(nt), \tag{3.26}$$

où $\{\Pi_c(t) : t \geq 0\}$ désigne le processus de Poisson composé défini en (3.1). Dans la suite, nous posons sans perte de généralité (au prix du changement d'échelle $t \to t/\lambda$) $\lambda = 1$. Ce processus $\{\Pi_{n,c}(t) : t \geq 0\}$ sera très utile dans le chapitre suivant.

Nous allons énoncer une version uniforme du théorème 3.3.2 pour le processus de Poisson composé $\{\Pi_{n,c}(t) : t \geq 0\}$.

Théorème 3.3.3. *Soit un réel $a > 0$. Soit $\{\Pi_{n,c}(t) : t \geq 0\}$ un processus de Poisson composé défini en* (3.26)*. Supposons que la variable aléatoire Y vérifie $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable et que $0 < \mathbb{E}[Y^2] < \infty$. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[\sup_{0\leq t\leq a}\frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] \leq 32\exp\left(-\frac{\mu^2}{32a\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4a\sqrt{n}\mathbb{E}[Y^2]}\right)\right), \tag{3.27}$$

où la fonction ψ est définie en (3.11)*.*

Remarque 3.3.2. Pour un processus de Poisson, nous avons une inégalité semblable qui s'énonce ainsi.
Soit un réel $b > 0$. Soit $\{\Pi(t) : t \geq 0\}$ un processus de Poisson homogène, continu à droite et de paramètre $\lambda = 1$. Alors, pour tout réel $\mu \geq 0$, nous avons

$$\mathbb{P}\left[\sup_{0\leq t\leq b}\frac{\Pi(t) - t}{\sqrt{b}} \geq \mu\right] \leq \exp\left(-\frac{\mu^2}{2}\psi\left(\frac{\mu}{\sqrt{b}}\right)\right).$$

La démonstration de cette inégalité se trouve par exemple, dans le livre de Shorack et Wellner [125], p.570.

Remarque 3.3.3. Pour un processus de Poisson multivarié, nous avons également une inégalité similaire.
Soit $\{\Pi(nt) : t \in [0,1]^d\}$ un processus de Poisson de paramètre $\lambda = n > 0$. Alors, pour tout réel $\mu \geq 0$, nous avons

$$\mathbb{P}\left[\sup_{S \subset R} \frac{\Pi(nS) - nF(S)}{\sqrt{n}} \geq \mu\right] \leq 2^{2d+3} \exp\left(-\frac{\mu^2}{32F(R)}\psi\left(\frac{\mu}{4F(R)\sqrt{n}}\right)\right),$$

où R est un rectangle d'une classe $\mathcal{R}$ de rectangles semi-ouverts définie par $\mathcal{R} = \{R(x,y) : R(x,y) \subset [0,1]^d\}$ et F désigne la fonction de répartition de la suite de variables aléatoires $\{X_i : i \geq 1\}$.
La démonstration de cette inégalité se trouve dans l'article de Ruymgaart et Wellner [118].

Démonstration. Pour démontrer l'inégalité (3.27), nous allons établir quelques résultats préliminaires.

Définition 3.3.1. *Une variable aléatoire X est symétrique si X a la même loi que $-X$. Soient X et X' deux variables aléatoires réelles indépendantes et de même loi. On appelle symétrisation de la variable aléatoire X la variable X^S définie par*

$$X^S = X - X'. \tag{3.28}$$

Inégalité 3.3.3. *Si la variable aléatoire X est intégrable, alors nous avons*

$$|\mu X - \mathbb{E}[X]| \leq \sqrt{2\mathrm{Var}[X]}, \tag{3.29}$$

où μX et $\mathrm{Var}[X]$ représentent respectivement la médiane et la variance de la variable X.

Démonstration. D'après l'inégalité de Tchebychev, nous avons

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \geq \sqrt{2\mathrm{Var}[X]}\right] \leq \frac{\mathbb{E}[|X - \mathbb{E}[X]|^2]}{2\mathrm{Var}[X]} = \frac{1}{2}.$$

Nous distinguons deux cas.
Premier cas. Si

$$\mathbb{P}\left[X \geq \mathbb{E}[X] + \sqrt{2\mathrm{Var}[X]}\right] \leq \frac{1}{2}$$

alors cette inégalité implique que

$$\mu X \leq \mathbb{E}[X] + \sqrt{2\mathrm{Var}[X]},$$

car par définition de la médiane, nous avons

$$\mathbb{P}[X \geq \mu X] \geq \frac{1}{2}.$$

Deuxième cas. Si

$$\mathbb{P}\left[X \leq \mathbb{E}[X] - \sqrt{2\mathrm{Var}[X]}\right] \leq \frac{1}{2}$$

alors cette inégalité implique que

$$\mu X \geq \mathbb{E}[X] - \sqrt{2\mathrm{Var}[X]},$$

car par définition de la médiane, nous avons

$$\mathbb{P}\left[X \leq \mu X\right] \geq \frac{1}{2}.$$

Ce qui achève la démonstration de l'inégalité 3.3.3. □

Inégalité 3.3.4 (Inégalité de symétrisation.). *Soit la variable aléatoire X^S définie en* (3.28)*. Alors, pour tout couple de réels a et $\lambda > 0$, nous avons*

$$\frac{1}{2}\mathbb{P}\left[|X - \mu X| \geq \lambda\right] \leq \mathbb{P}\left[|X^S| \geq \lambda\right] \leq 2\mathbb{P}\left[|X - a| \geq \frac{\lambda}{2}\right]. \tag{3.30}$$

Démonstration. Pour la première inégalité, il suffit de montrer que $\mathbb{P}\left[X^S \geq \lambda\right] \geq 2^{-1}\mathbb{P}\left[X - \mu X \geq \lambda\right]$. Le cas négatif se déduira du cas positif en faisant le changement formel de la variable X en la variable $-X$. Posons $m = \mu X$. La première inégalité provient du fait que

$$\begin{aligned}\mathbb{P}\left[X^S \geq \lambda\right] &= \mathbb{P}\left[(X - m) - (X' - m) \geq \lambda\right] \\ &\geq \mathbb{P}\left[X - m \geq \lambda, X' - m \leq 0\right] = \mathbb{P}\left[X - m \geq \lambda\right]\mathbb{P}\left[X' - m \leq 0\right] \\ &\geq \frac{1}{2}\mathbb{P}\left[X - m \geq \lambda\right].\end{aligned}$$

La seconde inégalité se démontre ainsi. Pour tous réels a et $\lambda > 0$, nous avons

$$\begin{aligned}\mathbb{P}\left[\left|X^S\right| \geq \lambda\right] &= \mathbb{P}\left[|(X - a) - (X' - a)| \geq \lambda\right] \\ &\leq \mathbb{P}\left[|X - a| \geq \frac{\lambda}{2}\right] + \mathbb{P}\left[|X' - a| \leq \frac{\lambda}{2}\right] \\ &= 2\mathbb{P}\left[|X - a| \geq \frac{\lambda}{2}\right].\end{aligned}$$

Ceci achève la démonstration de l'inégalité 3.3.4. □

Lemme 3.3.3. *Soit X un processus de Wiener à N paramètres à valeurs dans l'espace euclidien $\mathbb{R}^d$. Soit $S = \Delta(u, v)$ un intervalle de temps. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[\sup_{R \subset S} |X(R)| > \mu\right] \leq 4^N \mathbb{P}\left[|X(S)| > \mu\right],$$

où R décrit tous les intervalles inclus dans S.

Démonstration. La démonstration de ce lemme se trouve dans l'article de Orey et Pruitt [108]. □

Remarque 3.3.4. On peut établir la même inégalité pour un processus symétrique à accroissements indépendants.

Soit $\{\Pi^*_{n,c}(t) : t \geq 0\}$ un processus de Poisson composé de même loi que le processus de Poisson composé $\{\Pi_{n,c}(t) : t \geq 0\}$. De plus, le processus $\{\Pi^*_{n,c}(t) : t \geq 0\}$ est indépendant du processus $\{\Pi_{n,c}(t) : t \geq 0\}$. Pour tout réel $t \geq 0$, notons $\mu\Pi_{n,c}(t)$ la médiane de la variable de Poisson composée $\Pi_{n,c}(t)$. Posons

$$\Pi^\mu_{n,c} = \Pi_{n,c} - \mu\Pi_{n,c}$$

et

$$\Pi^S_{n,c} = \Pi_{n,c} - \Pi^*_{n,c}.$$

Remarquons que les variables aléatoires $\Pi^S_{n,c}(t)$ sont symétriquement distribuées pour tout réel $t \geq 0$. En appliquant l'inégalité (3.29) à la variable de Poisson composée $\Pi_{n,c}(t)$, nous obtenons que

$$|\mu\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]| \leq \sqrt{2\mathrm{Var}[\Pi_n(t)]\mathbb{E}[Y^2]} = \sqrt{2nt\mathbb{E}[Y^2]}, \quad \text{pour } t \geq 0. \tag{3.31}$$

Par conséquent, pour tout réel $t \in (0, a]$, où a est un réel strictement positif, nous avons

$$\begin{aligned}
\frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} &= \frac{|\Pi_{n,c}(t) - \mu\Pi_{n,c}(t) + \mu\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \\
&\leq \frac{\left|\Pi^\mu_{n,c}(t)\right|}{\sqrt{n}} + \frac{\sqrt{2nt\mathbb{E}[Y^2]}}{\sqrt{n}} \\
&\leq \frac{\left|\Pi^\mu_{n,c}(t)\right|}{\sqrt{n}} + \sqrt{2a\mathbb{E}[Y^2]}, \qquad (3.32)
\end{aligned}$$

d'après l'inégalité (3.31).

Ces remarques établies, nous pouvons maintenant établir l'inégalité (3.27). D'après (3.32), pour tout réel $\mu \geq 0$, nous avons

$$\begin{aligned}
\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] &\leq \mathbb{P}\left[\sup_{0\leq t\leq a} \frac{\left|\Pi^\mu_{n,c}(t)\right|}{\sqrt{n}} \geq \mu - \sqrt{2a\mathbb{E}[Y^2]}\right] \\
&\leq \mathbb{P}\left[\sup_{0\leq t\leq a} \frac{\left|\Pi^\mu_{n,c}(t)\right|}{\sqrt{n}} \geq \frac{\mu}{2}\right],
\end{aligned}$$

si $\mu \geq \sqrt{8a\mathbb{E}[Y^2]}$. D'après la première partie de l'inégalité (3.30), nous avons

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] \leq 2\,\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{\left|\Pi^S_{n,c}(t)\right|}{\sqrt{n}} \geq \frac{\mu}{2}\right].$$

D'après le lemme 3.3.3 appliqué au processus $\{\Pi^S_{n,c}(t) : t \geq 0\}$ qui est un processus symétrique à accroissements indépendants, pour tout réel $\mu \geq \sqrt{8a\mathbb{E}[Y^2]}$, nous avons

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] \leq 2 \times 4\,\mathbb{P}\left[\frac{\left|\Pi^S_{n,c}(a)\right|}{\sqrt{n}} \geq \frac{\mu}{2}\right].$$

Comme $\left|\Pi^S_{n,c}(a)\right| \leq |\Pi_{n,c}(a) - \mathbb{E}[\Pi_{n,c}(a)]| + \left|\Pi^*_{n,c}(a) - \mathbb{E}[\Pi_{n,c}(a)]\right|$, où $\Pi^*_{n,c} \overset{\mathrm{d}}{=} \Pi_{n,c}$, alors nous obtenons que

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] \;\; \leq \;\; 8 \times 2\,\mathbb{P}\left[\frac{|\Pi_{n,c}(a) - \mathbb{E}[\Pi_{n,c}(a)]|}{\sqrt{n}} \geq \frac{\mu}{4}\right].$$

En appliquant l'inégalité (3.19) à la variable de Poisson composée $\Pi_{n,c}(a)$, pour tout réel $\mu \geq \sqrt{8a\mathbb{E}[Y^2]}$, nous concluons à

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \geq \mu\right] \;\; \leq \;\; 16\; \times 2\exp\left(-\frac{\mu^2}{32a\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4a\sqrt{n}\mathbb{E}[Y^2]}\right)\right).$$

Cependant, pour $0 \leq \mu \leq \sqrt{8a\mathbb{E}[Y^2]}$, la borne est plus grande que 1, ce qui implique que l'inégalité (3.19) est une borne supérieure pour tout réel $\mu \geq 0$, ce qui reste vrai pour tout réel $a > 0$. Ceci achève la démonstration du théorème 3.3.3. □

3.4. Approximation forte pour le processus de Poisson composé

Dans ce paragraphe, nous allons établir la meilleure vitesse d'approximation du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ par un processus de Wiener $\{W(t) : t \geq 0\}$.

Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. De plus, nous supposons que $\mathrm{Var}[Y] = \sigma^2 < \infty$. Soit $\{\Pi(t) : t \geq 0\}$ un processus de Poisson homogène, continu à droite et de paramètre $\lambda > 0$. Rappelons des notations déjà introduites dans le chapitre 1 dont nous allons avoir besoin dans ce paragraphe. Soit $\{S(t) : t \geq 0\}$ le processus de sommes partielles défini par

$$S(t) = \sum_{i=1}^{\lfloor t\rfloor} Y_i, \text{ pour } t > 0,$$

où $\lfloor t\rfloor$ désigne la partie entière de t ($\lfloor t\rfloor \leq t < \lfloor t\rfloor + 1$) et

$$S(0) = 0.$$

Le processus de Poisson $\{\Pi(t) : t \geq 0\}$ et le processus de sommes partielles $\{S(t) : t \geq 0\}$ ont été étudiés dans de nombreux livres et articles aussi bien d'un point de vue théorique que d'un point de vue appliqué. Il est bien connu que, via une normalisation, ces processus convergent faiblement vers un processus de Wiener-Lévy ou encore mouvement Brownien. Ici nous allons utiliser des résultats récents d'approximation forte pour les processus $\{\Pi(t) : t \geq 0\}$ et $\{S(t) : t \geq 0\}$ afin d'établir la meilleure approximation pour le processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$.

Afin de donner les vitesses de convergence pour les principes d'invariance, nous avons besoin de conditions supplémentaires sur la suite de variables $\{Y_i : i \geq 1\}$. Nous supposerons tantôt que

$(C.1)$ $\mathbb{E}[\exp(tY)] < \infty$ dans un voisinage de 0,

tantôt l'existence de moments d'ordre r

$(C.2)$ $\mathbb{E}[|Y|^r] < \infty$ pour $r > 2$.

De plus, nous supposons, comme dans le chapitre 1, que les variables aléatoires et les processus introduits dans ce paragraphe, peuvent être définis sur le même espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$ (cf. l'appendice 2 du livre de Csörgő et Horváth [24], p.418).

Rappelons les résultats d'approximation de Komlós, Major and Tusnády [83]-[84] et Major [93] pour le processus de sommes partielles $\{S(t) : t \geq 0\}$.

Théorème 3.4.1. *Nous pouvons définir un processus de Wiener $\{W_1(t) : t \geq 0\}$ tel que (i) si la condition $(C.1)$ est satisfaite, alors nous avons*

$$\mathbb{P}\left[\sup_{0\leq t\leq T} |(S(t) - \mathbb{E}[Y]t)/\sigma - W_1(t)| > A_1 \log T + y\right] \leq B_1 \exp(-C_1 y) \tag{3.33}$$

pour tout réel $y > 0$, où A_1, B_1 et C_1 sont des constantes strictement positives. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0\leq t\leq T} |(S(t) - \mathbb{E}[Y]t)/\sigma - W_1(t)| = O(\log T) \ \ p.s. \tag{3.34}$$

(ii) Si la condition $(C.2)$ est satisfaite, alors nous avons

$$\mathbb{P}\left[\sup_{0\leq t\leq T} |(S(t) - \mathbb{E}[Y]t)/\sigma - W_1(t)| > y\right] \leq B_2(T) T y^{-r} \tag{3.35}$$

pour tout $T^{1/r} \leq y \leq A_2\sqrt{T \log T}$, où A_2 est une constante strictement positive et $B_2(T) \to 0$ lorsque $T \to \infty$. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0\leq t\leq T} |(S(t) - \mathbb{E}[Y]t)/\sigma - W_1(t)| = O\left(T^{1/r}\right) \ \ p.s. \tag{3.36}$$

Maintenant énonçons des résultats similaires pour le processus de Poisson $\{\Pi(t) : t \geq 0\}$ de paramètre $\lambda > 0$. Nous supposerons tantôt que

$(C.3)$ $\mathbb{E}[\exp(t\Pi(1))] < \infty$ dans un voisinage de 0,

tantôt l'existence de moments d'ordre r

$(C.4)$ $\mathbb{E}[|\Pi(1)|^r] < \infty$ pour $r > 2$.

Théorème 3.4.2. *Nous pouvons définir un processus de Wiener $\{W_2(t) : t \geq 0\}$ tel que (i) si la condition $(C.3)$ est satisfaite, alors nous avons*

$$\mathbb{P}\left[\sup_{0\leq t\leq T} \left|(\Pi(t) - \lambda t)/\sqrt{\lambda} - W_2(t)\right| > A_3 \log T + y\right] \leq B_3 \exp(-C_3 y) \tag{3.37}$$

pour tout réel $y > 0$, où A_3, B_3 et C_3 sont des constantes strictement positives. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0 \le t \le T} \left| (\Pi(t) - \lambda t) / \sqrt{\lambda} - W_2(t) \right| = O(\log T) \ \ p.s. \tag{3.38}$$

(ii) Si la condition $(C.4)$ est satisfaite, alors nous avons

$$\mathbb{P}\left[\sup_{0 \le t \le T} \left| (\Pi(t) - \lambda t) / \sqrt{\lambda} - W_2(t) \right| > y \right] \le B_4(T) T y^{-r} \tag{3.39}$$

pour tout $A_4 T^{1/r} \le y \le C_4 \sqrt{T \log T}$, où A_4, C_4 sont des constantes strictement positives et $B_4(T) \to 0$ lorsque $T \to \infty$. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0 \le t \le T} \left| (\Pi(t) - \lambda t) / \sqrt{\lambda} - W_2(t) \right| = O\left(T^{1/r}\right) \ \ p.s. \tag{3.40}$$

Démonstration. Comme le processus de Poisson $\{\Pi(t) : t \ge 0\}$ de paramètre $\lambda > 0$ est un cas particulier de processus de renouvellement, nous pouvons appliquer les résultats obtenus pour le processus de renouvellement par Csörgő, Horváth et Steinebach [26]. □

Remarque 3.4.1. Comme le processus de Poisson $\{\Pi(t) : t \ge 0\}$ et la suite de variables $\{Y_i : i \ge 1\}$ sont supposés indépendants, par conséquent les processus de Wiener $\{W_1(t) : t \ge 0\}$ et $\{W_2(t) : t \ge 0\}$ sont aussi indépendants.

Nous allons maintenant établir le résultat d'approximation forte pour le processus de Poisson composé $\{\Pi_c(t) : t \ge 0\}$ défini en (3.1).

Théorème 3.4.3. *Nous pouvons définir un processus de Wiener $\{W(t) : t \ge 0\}$ tel que (i) si les conditions $(C.1)$ et $(C.3)$ sont satisfaites, alors nous avons*

$$\mathbb{P}\left[\sup_{0 \le t \le T} \left| (\Pi_c(t) - \lambda \mathbb{E}[Y] t) / \sqrt{\lambda \mathbb{E}[Y^2]} - W(t) \right| > A_5 \log T + y \right] \le B_5 \exp(-C_5 y) \tag{3.41}$$

pour tout réel $y > 0$, où A_5, B_5 et C_5 sont des constantes strictement positives. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0 \le t \le T} \left| (\Pi_c(t) - \lambda \mathbb{E}[Y] t) / \sqrt{\lambda \mathbb{E}[Y^2]} - W(t) \right| = O(\log T) \ \ p.s. \tag{3.42}$$

(ii) Si les conditions $(C.2)$ et $(C.4)$ sont satisfaites, alors nous avons

$$\mathbb{P}\left[\sup_{0 \le t \le T} \left| (\Pi_c(t) - \lambda \mathbb{E}[Y] t) / \sqrt{\lambda \mathbb{E}[Y^2]} - W(t) \right| > y \right] \le B_6(T) T y^{-r} \tag{3.43}$$

pour tout $A_6 T^{1/r} \le y \le C_6 \sqrt{T \log T}$, où A_6, C_6 sont des constantes strictement positives et $B_6(T) \to 0$ lorsque $T \to \infty$. Lorsque $T \to \infty$, alors nous avons

$$\sup_{0 \le t \le T} \left| (\Pi_c(t) - \lambda \mathbb{E}[Y] t) / \sqrt{\lambda} - W(t) \right| = O\left(T^{1/r}\right) \ \ p.s. \tag{3.44}$$

Démonstration. Pour démontrer le théorème 3.4.3, nous avons besoin de trois lemmes. Les deux premiers lemmes établissent une approximation du processus de Poisson composé $\{\Pi_c(t) : t \geq 0\}$ par une combinaison linéaire des processus de Wiener $\{W_1(t) : t \geq 0\}$ et $\{W_2(t) : t \geq 0\}$ définis auparavant. Le troisième lemme donne également une approximation. L'approximation a lieu entre la combinaison linéaire des processus de Wiener $\{W_1(t) : t \geq 0\}$ et $\{W_2(t) : t \geq 0\}$ et le processus de Wiener $\{W(t) : t \geq 0\}$ défini dans le théorème 3.4.3 multiplié par le coefficient $\sqrt{\lambda \mathbb{E}[Y^2]}$.

Lemme 3.4.1. *Supposons que les conditions* $(C.1)$ *et* $(C.3)$ *sont satisfaites. Alors, avec les processus de Wiener* $\{W_1(t) : t \geq 0\}$ *et* $\{W_2(t) : t \geq 0\}$ *des théorèmes 3.4.1 et 3.4.2, nous avons*

$$\begin{aligned}\mathbb{P}\Bigg[\sup_{0\leq t\leq T}\Big|(\Pi_c(t)-\lambda\mathbb{E}[Y]t)-\Big(\sigma W_1\left(\lambda t+\sqrt{\lambda}W_2(t)\right)+\mathbb{E}[Y]\sqrt{\lambda}W_2(t)\Big)\Big| \\ \geq A_7\log T+y\Bigg]\leq B_7\exp\left(-C_7 y\right),\end{aligned} \tag{3.45}$$

pour tout réel $y > 0$*, où* A_7*,* B_7 *et* C_7 *sont des constantes strictement positives.*

Démonstration. Pour tous réels $t \geq 0$ et $\lambda > 0$, nous avons

$$\Pi_c(t) - \lambda\mathbb{E}[Y]t = \Pi_c(t) - \mathbb{E}[Y]\Pi(t) + \mathbb{E}[Y]\left(\Pi(t) - \lambda t\right).$$

D'après le théorème 3.4.2, il suffit de montrer que

$$\begin{aligned}\mathbb{P}\Bigg[\sup_{0\leq t\leq T}\Big|(\Pi_c(t)-\mathbb{E}[Y]\Pi(t))-\sigma W_1\left(\lambda t+\sqrt{\lambda}W_2(t)\right)\Big|\geq A_8\log T+y\Bigg] \\ \leq B_8\exp\left(-C_8 y\right),\end{aligned}$$

pour tout réel $y > 0$, où A_8, B_8 et C_8 sont des constantes strictement positives. En utilisant le théorème 3.4.2, pour tous réels $\lambda > 0$ et $y > 0$, nous avons

$$\begin{aligned}\mathbb{P}\left[\Pi(T) > 2\lambda T + 2y\right] &\leq \mathbb{P}\left[\sqrt{\lambda}W_2(T) > \lambda T + 2y - (A_3\log T + y)\right] + B_3\exp\left(-C_3 y\right) \\ &\leq \mathbb{P}\left[\sqrt{\lambda}W_2(1) > \lambda\sqrt{T} + T^{-1/2}y - T^{-1/2}A_3\log T\right] \\ &\qquad + B_3\exp\left(-C_3 y\right) \\ &\leq B_9\exp\left(-C_9 y\right).\end{aligned} \tag{3.46}$$

D'après le théorème 3.4.1 et l'inégalité (3.46), pour tout réel $y > 0$, nous obtenons que

$$\begin{aligned}
\mathbb{P} & \left[\sup_{0\le t\le T} |(\Pi_c(t) - \mathbb{E}[Y]\Pi(t)) - \sigma W_1(\Pi(t))| \ge A_{10}\log T + y \right] \\
& \le \mathbb{P}\left[\sup_{0\le t\le \Pi(T)} |(S(t) - \mathbb{E}[Y]t) - \sigma W_1(t)| \ge A_{10}\log T + y \right] \\
& \le \mathbb{P}\left[\sup_{0\le t\le 2\lambda T+2y} |(S(t) - \mathbb{E}[Y]t) - \sigma W_1(t)| \ge A_{10}\log T + y \right] + B_9\exp(-C_9 y) \\
& \le B_{10}\exp(-C_{10}y), \qquad (3.47)
\end{aligned}$$

pour tout réel $y > 0$, où A_{10}, B_{10} et C_{10} sont des constantes strictement positives. En combinant le lemme 1.2.1 du livre de Csörgő and Révész [27], l'inégalité (3.46) et le théorème 3.4.2, nous en déduisons que

$$\begin{aligned}
\mathbb{P} & \left[\sup_{0\le t\le T} \left| W_1(\Pi(t)) - W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right)\right| > A_{11}\log T + y \right] \\
& \le \mathbb{P}\left[\sup_{0\le t\le 2T+2\lambda^{-1}y}\ \sup_{0\le s\le \lambda^{-1}A_3\log T+\lambda^{-1}y} |W_1(t+s) - W_1(t)| > A_{11}\log T + y \right] \\
& \quad + B_9\exp(-C_9y) + B_3\exp(-C_3y) \\
& \le B_{11}\exp(-C_{11}y),
\end{aligned}$$

pour tout réel $y > 0$, où A_{11}, B_{11} et C_{11} sont des constantes strictement positives. Ce qui achève la démonstration du lemme 3.4.1. □

Maintenant établissons le deuxième lemme.

Lemme 3.4.2. *Supposons que les conditions* $(C.2)$ *et* $(C.4)$ *sont satisfaites. Alors, avec les processus de Wiener* $\{W_1(t) : t \ge 0\}$ *et* $\{W_2(t) : t \ge 0\}$ *des théorèmes 3.4.1 et 3.4.2, nous avons*

$$\begin{aligned}
\mathbb{P} & \left[\sup_{0\le t\le T} \left| (\Pi_c(t) - \lambda\mathbb{E}[Y]t) - \left(\sigma W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t)\right)\right| > y \right] \\
& \le B_{12}(T)Ty^{-r},
\end{aligned} \qquad (3.48)$$

et

$$\sup_{0\le t\le T} \left| (\Pi_c(t) - \lambda\mathbb{E}[Y]t) - \left(\sigma W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t)\right)\right| = o\left(T^{1/r}\right) \quad p.s. \qquad (3.49)$$

pour tout $A_{12}T^{1/r} \le y \le C_{12}\sqrt{T\log T}$*, où* A_{12} *et* C_{12} *sont des constantes strictement positives et* $B_{12}(T) \to 0$ *lorsque* $T \to \infty$*.*

Démonstration. La démonstration de (3.48) est similaire à celle de l'inégalité (3.45). La démonstration de l'égalité (3.49) est également semblable à celle de l'inégalité (3.45), à l'exception que les inégalités de probabilité sont remplacées par les événements correspondants presque sûrement. Nous allons donner les grandes lignes de la démonstration de l'égalité (3.49). D'abord nous allons rappeler un résultat classique.

Proposition 3.4.1. *Soit* $\{\Pi(t) : t \geq 0\}$ *un processus de Poisson homogène, continu à droite et de paramètre* $\lambda > 0$. *Alors, nous avons*

$$\lim_{t\to\infty} \frac{\Pi(t)}{t} = \lambda \quad p.s. \tag{3.50}$$

Démonstration. Pour la démonstration de cette proposition, on pourra consulter, par exemple, le livre de Cocozza-Thivent [15] p.29. □

L'égalité (3.50) et le théorème 3.4.1 impliquent que

$$\sup_{0\leq t\leq T} |(\Pi_c(t) - \mathbb{E}[Y]\Pi(t)) - \sigma W_1(\Pi(t))| = o\left(T^{1/r}\right) \quad \text{p.s.}$$

D'après le théorème 1.2.1 du livre de Csörgő and Révész [27] et le théorème 3.4.2, nous avons

$$\sup_{0\leq t\leq T} \left| W_1\left(\Pi(t)\right) - W_1\left(\lambda t + \sqrt{\lambda} W_2(t)\right)\right| = o\left(T^{1/(2r)}\sqrt{\log T}\right) \quad \text{p.s.}$$

En utilisant à nouveau le théorème 3.4.2, l'égalité (3.49) s'ensuit immédiatement. □

Lemme 3.4.3. *Il existe un processus de Wiener* $\{W(t) : t \geq 0\}$ *tel que*

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\sigma W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t) - \sqrt{\lambda\mathbb{E}[Y^2]}W(t)\right| > A_{13}\log T + y\right]$$
$$\leq B_{13}\exp\left(-C_{13}y\right),$$

pour tout réel $y >$, *où* A_{13}, B_{13} *et* C_{13} *sont des constantes strictement positives.*

Démonstration. Pour tous réels $t \geq 0$ et $\lambda > 0$, posons

$$Z(t) = \sigma W_1\left(t + \sqrt{\lambda}W_2(t/\lambda)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t/\lambda).$$

Par conséquent nous avons l'égalité en distribution suivante

$$\{Z(t) : t \geq 0\} \stackrel{\text{d}}{=} \{\sigma W_1(t + W_2(t)) + \mathbb{E}[Y]W_2(t) : t \geq 0\}. \tag{3.51}$$

Considérons une suite $\{Z_i : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées telle que

$$\mathbb{E}[\exp(tZ_1)] < \infty \text{ dans un voisinage de } 0$$

et

$$\mathbb{E}[Z_1] = \mathbb{E}[Y] \quad \text{et} \quad \mathrm{Var}[Z_1] = \sigma^2 < \infty.$$

Soit $\{\nu(t) : t \geq 0\}$ un processus de Poisson construit sur des variables aléatoires qui suivent une loi exponentielle de paramètre 1. De plus, nous supposons que le processus $\{\nu(t) : t \geq 0\}$ est indépendant de la suite de variables $\{Z_i : i \geq 1\}$. Le processus $\sum_{1\leq i\leq \nu(t)} Z_i$ est, par définition alors un processus de Poisson composé à accroissements indépendants d'après le théorème 3.2.1 et sa fonction génératrice des moments existe. D'après l'approximation de Komlós, Major et Tusnády [83]-[84] :
il existe un processus de Wiener $\{\widetilde{W}(t) : t \geq 0\}$ tel que

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\left(\sum_{i=1}^{\nu(t)} Z_i - \mathbb{E}[Z_1]t\right) \Big/ \sqrt{\mathrm{Var}[Z_1] + \mathbb{E}^2[Z_1]} - \widetilde{W}(t)\right| > A_{14}\log T + y\right] \leq B_{14}\exp\left(-C_{14}y\right), \tag{3.52}$$

pour tout réel $y > 0$, où A_{14}, B_{14} et C_{14} sont des constantes strictement positives. Comme le processus de Poisson $\{\nu(t) : t \geq 0\}$ est un processus de renouvellement, d'après les théorèmes 3.4.1 et 3.4.2 :

Il existe deux processus de Wiener $\{\widetilde{W}_1(t) : t \geq 0\}$ et $\{\widetilde{W}_2(t) : t \geq 0\}$ tels que, d'après le lemme 3.4.1, nous ayons

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\left(\sum_{i=1}^{\nu(t)} Z_i - \mathbb{E}[Z_1]t\right) - \left(\sigma\widetilde{W}_1\left(t + \widetilde{W}_2(t)\right) + \mathbb{E}[Y]\widetilde{W}_2(t)\right)\right| > A_{15}\log T + y\right] \leq B_{15}\exp\left(-C_{15}y\right), \tag{3.53}$$

pour tout réel $y > 0$, où A_{15}, B_{15} et C_{15} sont des constantes strictement positives. Les inégalités (3.52) et (3.53) conduisent à l'inégalité suivante

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\sqrt{\mathrm{Var}[Z_1] + \mathbb{E}^2[Z_1]}\widetilde{W}(t) - \left(\sigma\widetilde{W}_1\left(t + \widetilde{W}_2(t)\right) + \mathbb{E}[Y]\widetilde{W}_2(t)\right)\right| > A_{16}\log T + x\right] \leq B_{16}\exp\left(-C_{16}x\right), \tag{3.54}$$

où A_{16}, B_{16} et C_{16} sont des constantes strictement positives. Les inégalités (3.54) et (3.51) combinées avec le théorème A.1 de De Acosta [28] impliquent le lemme 3.4.3. $\square$

Pour démontrer l'inégalité (3.41), il suffit de remarquer

$$\begin{aligned}\Pi_c(t) - \lambda\mathbb{E}[Y]t - \sqrt{\lambda\mathbb{E}[Y^2]}W(t) = &\Big[\Pi_c(t) - \lambda\mathbb{E}[Y]t \\ &- \left(\sigma W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t)\right)\Big] \\ &< +\Big[\left(\sigma W_1\left(\lambda t + \sqrt{\lambda}W_2(t)\right) + \mathbb{E}[Y]\sqrt{\lambda}W_2(t)\right) \\ &- \sqrt{\lambda\mathbb{E}[Y^2]}W(t)\Big] \\ = &\ \mathcal{D}_1 + \mathcal{D}_2.\end{aligned}$$

Ensuite nous appliquons le lemme 3.4.1 à la quantité $\mathcal{D}_1$ et le lemme 3.4.3 à la quantité $\mathcal{D}_2$, ce qui conduit à l'inégalité (3.41). L'inégalité (3.42) est une conséquence directe de l'inégalité (3.41).

Pour démontrer l'inégalité (3.43), nous procédons de façon similaire. L'inégalité (3.48) appliquée à la quantité $\mathcal{D}_1$ et le lemme 3.4.3 appliqué à la quantité $\mathcal{D}_2$ donnent l'inégalité (3.43). L'inégalité (3.44) découle de l'inégalité (3.49). □

3.5. Liens entre le processus de Poisson composé et le processus empirique composé

3.5.1. Construction d'un processus de Poisson

Le résultat suivant se rapporte à la définition de la mesure empirique composée ou pondérée par des variables aléatoires pour les processus de Poisson composés. En fait, la définition la plus simple des processus de Poisson composés utilise les mesures empiriques pondérées par une suite de variables aléatoires $\{Y_i : i \geq 1\}$ indépendantes et identiquement distribuées. Cette suite de variables $\{Y_i : i \geq 1\}$ est indépendante de la suite de variables $\{X_i : i \geq 1\}$ qui définit la mesure empirique. Un processus de Poisson composé se construit de la manière suivante.

Soit un espace métrique séparable $\mathcal{E}$, muni d'une σ-algèbre $\mathcal{B}$ de sous-ensembles boréliens. Notons par μ une mesure borélienne positive de Borel *bornée* sur $\mathcal{E}$. Considérons une suite $\{X_i : i \geq 1\}$ de variables aléatoires réelles indépendantes et identiquement distribuées dans l'espace $\mathcal{E}$, de même loi que la variable aléatoire $X = X_1$, c'est-à-dire

$$\mathbb{P}\left[X \in B\right] = \mu_0(B) = \frac{\mu(B)}{\mu(\mathcal{E})} \text{ pour } B \in \mathcal{B}. \tag{3.55}$$

Considérons également une seconde suite $\{Y_i : i \geq 1\}$ de variables aléatoires réelles indépendantes et identiquement distribuées et de fonction de répartition F_Y continue sur $\mathbb{R}$. Les suites de variables $\{X_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont supposées indépendantes. Définissons maintenant une variable aléatoire notée Π, indépendante des deux suites de variables $\{X_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$. De plus, la variable Π suit une loi de Poisson de paramètre $\lambda > 0$, c'est-à-dire

$$\mathbb{P}[\Pi = n] = \frac{\lambda^n}{n!}\exp(-\lambda), \text{ pour } n \in \mathbb{N}. \tag{3.56}$$

Le processus de Poisson composé sur l'espace $\mathcal{E}$ de moyenne $\nu_c = \lambda\mathbb{E}[Y]\mu_0$ est alors défini comme la mesure aléatoire sur l'espace $\mathcal{E}$

$$N_c = \sum_{i=0}^{\Pi} Y_i \delta_{X_i}, \tag{3.57}$$

où on utilise la convention que $\sum_{\emptyset}(.) = 0$.

La caractéristique la plus importante de la mesure aléatoire N_c est que, quelle que soit la suite $B_1, B_2, \ldots$ de parties boréliennes disjointes de l'espace $\mathcal{E}$, les variables aléatoires $N_c(B_1), N_c(B_2), \ldots$ sont *indépendantes*. De plus, pour tout $B \in \mathcal{B}$, la variable aléatoire $N_c(B)$ suit une loi de Poisson composée de moyenne $\nu_c(B)$.

Théorème 3.5.1. *Soient N_c^1 et N_c^2 deux processus de Poisson composés indépendants sur l'espace $\mathcal{E}$, de moyenne respective $\nu_{1,c}$ et $\nu_{2,c}$. Alors la somme $N_c^1 + N_c^2$ est encore un processus de Poisson composé de moyenne $\nu_{1,c} + \nu_{2,c}$.*

Démonstration. Pour la démonstration de ce théorème, on pourra consulter le livre de Karlin et Taylor [80]. □

De ce fait, il en découle que les mesures directrices des processus de Poisson composés sont bornées et définissent ainsi un processus de Poisson composé de moyenne

$$\nu_c = \sum_{i=-\infty}^{+\infty} \nu_{i,c}. \tag{3.58}$$

Ce processus peut se construire formellement en superposant (c'est-à-dire en additionnant) une suite dénombrable de processus de Poisson composés indépendants et chacun de moyenne $\nu_{i,c}$ où i varie de $-\infty$ à $+\infty$. L'intérêt particulier que l'on porte à cette remarque est le cas où la mesure ν_c est la mesure de Lebesgue sur $\mathbb{R}$ pondérée par une suite $\{Y_i : i \geq 1\}$ de variables aléatoires et $\nu_{i,c}$ la mesure uniforme sur $(i, i+1]$ pondérée par la variable aléatoire Y_i. Dans ce cas, on obtient le processus de Poisson composé standard sur $\mathbb{R}$. Si nous notons ce processus par $N_{0,c}$, il est habituel de définir $N_{0,c}$ sur $\mathbb{R}^+$ par la version continue à droite de sa fonction de répartition

$$\Pi_c(t) = N_{0,c}(0, t] \text{ pour } t \geq 0. \tag{3.59}$$

On peut désormais énoncer le théorème suivant

Théorème 3.5.2. *Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées. Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0,1]$. Les suites de variables $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. Soit $\{\Pi_\lambda(t) = \Pi(\lambda t) : t \geq 0\}$ un processus de Poisson homogène, continu à droite et de paramètre $\lambda > 0$. Soit*

$$\left\{ \Pi_{\lambda,c}(t) = \sum_{i=1}^{\Pi(\lambda t)} Y_i : t \geq 0 \right\}$$

un processus de Poisson composé. Alors, pour tout réel $\lambda > 0$, nous avons l'égalité en distribution suivante

$$\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\} \stackrel{d}{=} \{\Pi_{\lambda,c}(t) : 0 \leq t \leq 1 | \Pi(\lambda) = n\},$$

où la fonction empirique composée $\mathbb{U}_{n,c}(t)$ est définie par

$$\mathbb{U}_{n,c}(t) = \frac{1}{n} \sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \leq t\}} \quad \textit{pour } t \in [0,1].$$

Démonstration. La démonstration du théorème 3.5.2 est évidente à partir des résultats énoncés précédemment. □

Remarque 3.5.1. Il existe un résultat analogue pour la fonction de répartition empirique uniforme $\mathbb{U}_n(t)$ qui s'énonce de la façon suivante.

Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0,1]$. Soit $\{\Pi_\lambda(t) : 0 \leq t \leq 1\}$ un processus de Poisson homogène, continu à droite et de paramètre $\lambda > 0$. Alors, on a l'égalité en distribution suivante

$$\{n\mathbb{U}_n(t) : 0 \leq t \leq 1\} \stackrel{\mathrm{d}}{=} \{\Pi_\lambda(t) : 0 \leq t \leq 1 | \Pi(\lambda) = n\},$$

où la fonction de répartition empirique uniforme $\mathbb{U}_n(t)$ est définie par

$$\mathbb{U}_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{U_i \leq t\}} \quad \text{pour } t \in [0,1].$$

Remarque 3.5.2. Comme le choix du paramètre $\lambda > 0$ dans le théorème 3.5.2 est arbitraire, il est préférable pour la suite, de réécrire ce théorème en choisissant $\lambda = n$. Alors, pour tout réel $t \in [0,1]$, nous avons alors l'égalité suivante

$$\mathbb{E}[n\mathbb{U}_{n,c}(t)] = \mathbb{E}[\Pi_{n,c}(t)] = \mathbb{E}[\Pi(nt)]\mathbb{E}[Y] = \mathbb{E}[Y]nt.$$

3.5.2. Un théorème d'approximation pour le processus empirique composé

Le théorème 3.5.2 montre, dans un sens, que le processus $\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\}$ est proche du processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$. Ce qui nous permet, parfois, de remplacer le processus $\{\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\}$ par le processus $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$. Le processus $\{\Pi_{n,c}(t) : t \geq 0\}$ est plus facile à manipuler parce qu'il a des accroissements indépendants (cette propriété a été démontrée dans le théorème 3.2.1) ce que n'a pas le processus empirique composé $\{\alpha_{n,c}(t) = \sqrt{n}\,(\mathbb{U}_{n,c}(t) - \mathbb{E}[\mathbb{U}_{n,c}(t)]) : 0 \leq t \leq 1\}$.

Introduisons quelques notations. Soit un réel $a \in (0,1)$. Considérons l'ensemble $\mathcal{M}[0,a]$ de mesures boréliennes positives bornées sur l'intervalle $[0,a]$. Chaque mesure $\mu \in \mathcal{M}[0,a]$ est caractérisée par sa fonction de répartition $F_\mu(x) = \mu\left([0,x]\right)$, pour $x \in \mathbb{R}$. Nous notons $I_{RC}[0,a]$ l'ensemble de ces fonctions de répartition. Remarquons qu'il existe une correspondance entre $\mu \in \mathcal{M}[0,a]$ et $F_\mu \in I_{RC}[0,a]$. En particulier, la connaissance de $F_\mu(x)$ pour tout $x \in [0,a]$ est suffisante pour déterminer la mesure μ. De plus, la fonction F appartient à $I_{RC}[0,a]$ si et seulement si la fonction F est croissante, positive, bornée, continue à droite et telle que

$$F(x) = 0 \quad \text{pour } x < 0, \quad F(x) = F(a) \quad \text{pour } x \geq a.$$

Nous définissons une métrique (la distance de Lévy) sur $\mathcal{M}[0,a]$ ou de façon équivalente sur $I_{RC}[0,a]$, par

$$d_L\left(\mu,\nu\right) = d_L\left(F_\mu,F_\nu\right) = \inf\{\varepsilon > 0 : F_\nu(x-\varepsilon) \leq F_\mu(x) \leq: F_\nu(x+\varepsilon) \quad \forall x \in \mathbb{R}\}, \quad (3.60)$$

où les mesures $\mu,\nu \in \mathcal{M}[0,a]$. La topologie définie sur la mesure $\mu \in \mathcal{M}[0,a]$ par d_L est aussi appelée la topologie faible (voir le livre de Billingsley [5] pour de plus amples détails). Notons par $C[0,a]$ l'ensemble des fonctions continues sur l'intervalle $[0,a]$. Nous avons alors l'équivalence, pour $\mu \in \mathcal{M}[0,a]$ et $\mu_n \in \mathcal{M}[0,a]$, $n = 1,2,\ldots$,

$$\lim_{n\to\infty} d_L\left(\mu_n,\mu\right) = 0 \Leftrightarrow \lim_{n\to\infty} \int_0^a f d\mu_n = \int_0^a f d\mu \quad \forall f \in C[0,a]. \quad (3.61)$$

Soit $\mathcal{B}_a$ la σ-algèbre de sous-ensemble boréliens induits sur l'espace $I_{RC}[0,a]$ (muni de la topologie de la convergence en loi). Notons que pour tout réel $a \in (0,1)$ et pour tout sous-ensemble $B \in \mathcal{B}_a$, les probabibilités

$$\mathbb{P}\big[\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq a\} \in B\big] \quad \text{et} \quad \mathbb{P}\big[\{\Pi_{n,c}(t) : 0 \leq t \leq a\} \in B\big],$$

sont bien définies. Ces préliminaires établis, nous pouvons énoncer le théorème 3.5.3. Ce théorème permet d'obtenir une “approximation de Poisson”, qui est appropriée pour obtenir des lois limites pour le processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$.

Théorème 3.5.3. *Pour tout réel $a \in (0,1)$, il existe une constante $\mathcal{C}_a$ telle que $0 < \mathcal{C}_a < \infty$ et telle que, pour tout sous-ensemble $B \in \mathcal{B}_a$, nous ayons*

$$\mathbb{P}\big[\{n\mathbb{U}_{n,c}(t) : 0 \le t \le a\} \in B\big] \le \mathcal{C}_a \, \mathbb{P}\big[\{\Pi_{n,c}(t) : 0 \le t \le a\} \in B\big]. \tag{3.62}$$

Remarque 3.5.3. Il existe une version de ce théorème pour la fonction de répartition empirique uniforme $\mathbb{U}_n(t)$ qui s'énonce de la façon suivante.

Pour tout réel $a \in (0,1)$, il existe une constante C_a telle que $0 < C_a < \infty$ et telle que, pour tout sous-ensemble $B \in \mathcal{B}_a$, nous ayons

$$\mathbb{P}\big[\{n\mathbb{U}_n(t) : 0 \le t \le a\} \in B\big] \le C_a \, \mathbb{P}\big[\{\Pi(nt) : 0 \le t \le a\} \in B\big]. \tag{3.63}$$

Il faut noter que ce résultat a souvent été utilisé dans une série d'articles de Deheuvels et de Deheuvels et Mason (voir, e.g., le lemme 2.7 de Deheuvels [29] et les lemmes 2.1 et 3.1 de Deheuvels et Mason [35]).

Démonstration. Soit un réel $\lambda > 0$. Soit une variable aléatoire Z qui suit une loi de Poisson de paramètre $\lambda > 0$. Alors, par définition de la loi de Poisson, nous avons

$$\mathbb{P}[Z = k] = \frac{\lambda^k \exp(-\lambda)}{k!} \quad \text{pour } k \in \mathbb{N}.$$

Il est facile de vérifier que

$$M(\lambda) = \max_{k \in N} \mathbb{P}(Z = k) = \mathbb{P}(Z = \lfloor \lambda \rfloor) = \frac{\lambda^{\lfloor \lambda \rfloor} \exp(-\lambda)}{\lfloor \lambda \rfloor !}, \tag{3.64}$$

où $\lfloor u \rfloor \le u < \lfloor u \rfloor + 1$ représente la partie entière de u. (Pour ce calcul, on peut consulter également le livre de Johnson et Kotz [76] p.92). Par des calculs très simples et en utilisant la formule de Stirling ($n! = (1 + o(1))(n/e)^n \sqrt{2\pi n}$ lorsque $n \to \infty$), nous obtenons

$$M(\lambda) = \frac{1 + o(1)}{\sqrt{2\pi\lambda}}, \quad \text{lorsque } \lambda \to \infty. \tag{3.65}$$

Maintenant sous les conditions du théorème 3.5.3, posons $R = \Pi(na)$ et $\overline{R} = \Pi(n) - \Pi(na)$. Remarquons que les variables aléatoires R et $\overline{R}$ sont indépendantes et suivent une loi de Poisson de paramètres respectivement na et $n(1-a)$. Comme la variable aléatoire $R + \overline{R} = \Pi(n)$ suit une loi de Poisson de paramètre n, d'après (3.64) et (3.65), il s'ensuit, lorsque $n \to \infty$, que

$$\begin{aligned} \sup_{k \in \mathbb{N}} \frac{\mathbb{P}(\overline{R} = k)}{\mathbb{P}(R + \overline{R} = n)} &= \frac{M\big(n(1-a)\big)}{M(n)} \\ &= \left(\frac{1 + o(1)}{\sqrt{2\pi n(1-a)}} \right) \Bigg/ \left(\frac{1 + o(1)}{\sqrt{2\pi n}} \right) \to \frac{1}{\sqrt{1-a}}, \quad \text{d'après (3.65).} \end{aligned} \tag{3.66}$$

D'après (3.66), nous pouvons alors définir une constante $0 < \mathcal{C}_a < \infty$ en posant

$$\mathcal{C}_a = \sup_{n\geq 1}\left\{\frac{M\big(n(1-a)\big)}{M(n)}\right\} \in \left[\frac{1}{\sqrt{1-a}}, \infty\right). \tag{3.67}$$

Introduisons les événements

$$\mathcal{E}_1 = \left\{\left\{\Pi_{n,c}(t) : 0 \leq t \leq a\right\} \in B\right\} \text{ et } \mathcal{E}_2 = \left\{\left\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq a\right\} \in B\right\}. \tag{3.68}$$

D'après le théorème 3.5.2, comme le processus $\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\}$ est égal en distribution au processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$ sachant $\Pi(n) = n$, alors nous avons

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_2] &= \mathbb{P}\left[\mathcal{E}_1 | R + \overline{R} = n\right] \\
&= \frac{\mathbb{P}\left[\mathcal{E}_1 \cap \{R + \overline{R} = n\}\right]}{\mathbb{P}\left[\{R + \overline{R} = n\}\right]} \\
&= \sum_{j=0}^{n} \frac{\mathbb{P}\left[\mathcal{E}_1 \cap \{R = n - j\} \cap \{\overline{R} = j\}\right]}{\mathbb{P}\left[R + \overline{R} = n\right]}.
\end{aligned}$$

De plus, comme les événements $\{\mathcal{E}_1 \cap \{R = n-j\}\}$ et $\{\overline{R} = j\}$ sont indépendants, nous avons

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_2] &= \sum_{j=0}^{n} \mathbb{P}\left[\mathcal{E}_1 \cap \{R = n - j\}\right] \times \frac{\mathbb{P}\left[\overline{R} = j\right]}{\mathbb{P}\left[R + \overline{R} = n\right]} \\
&\leq \mathbb{P}[\mathcal{E}_1] \sup_{j\in\mathbb{N}} \frac{\mathbb{P}\left[\overline{R} = j\right]}{\mathbb{P}\left[R + \overline{R} = n\right]} \\
&\leq \mathcal{C}_a \mathbb{P}[\mathcal{E}_1].
\end{aligned}$$

Ceci achève la démonstration du théorème 3.5.3. □

De ce théorème, nous pouvons énoncer une version uniforme pour le processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ défini par

$$\alpha_{n,c}(t) = \sqrt{n}\left(\mathbb{U}_{n,c}(t) - \mathbb{E}[\mathbb{U}_{n,c}(t)]\right) \quad \text{pour } 0 \leq t \leq 1.$$

Corollaire 3.5.1. *Pour tout réel $a \in (0,1)$, il existe une constante réelle $\mathcal{C}_a$ telle que $0 < \mathcal{C}_a < \infty$ et telle que pour tout réel $\mu > 0$, nous ayons*

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \pm\alpha_{n,c}(t) \geq \mu\right] \leq \mathcal{C}_a \mathbb{P}\left[\sup_{0\leq t\leq a} \frac{\pm\left(\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]\right)}{\sqrt{n}} \geq \mu\right]. \tag{3.69}$$

Démonstration. La démonstration de ce corollaire est semblable à celle du théorème 3.5.3. Par souci du détail, nous en donnons les grandes lignes. Soient les deux événements suivants

$$A_1 = \left\{ \sup_{0\leq t\leq a} \pm\alpha_{n,c}(t) \geq \mu \right\} \text{ et } A_2 = \left\{ \sup_{0\leq t\leq a} \frac{\pm\left(\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]\right)}{\sqrt{n}} \geq \mu \right\}.$$

Comme nous avons l'égalité en distribution suivante

$$\{\alpha_{n,c}(t) : 0 \leq t \leq 1\} \overset{\mathrm{d}}{=} \left\{ \frac{\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]}{\sqrt{n}} : 0 \leq t \leq 1 \middle| \Pi(n) = n \right\}$$

et en utilisant l'indépendance des variables aléatoires $R = \Pi(na)$ et $\overline{R} = \Pi(n) - \Pi(na)$, variables introduites dans la démonstration du théorème 3.5.3, nous obtenons successivement

$$\begin{aligned}
\mathbb{P}[A_1] &= \mathbb{P}[A_2|\Pi(n) = n] \\
&= \frac{\mathbb{P}\left[A_2 \cap \{\Pi(n) = n\}\right]}{\mathbb{P}\left[\Pi(n) = n\right]} \\
&= \frac{\displaystyle\sum_{k=0}^{n} \mathbb{P}\left[A_2 \cap \{R = k\} \cap \{\overline{R} = n-k\}\right]}{\mathbb{P}\left[\Pi(n) = n\right]}.
\end{aligned}$$

Comme les événements $\{A_2 \cap \{R = k\}\}$ et $\{\overline{R} = n - k\}$ sont indépendants, nous avons

$$\begin{aligned}
\mathbb{P}[A_1] &= \sum_{k=0}^{n} \frac{\mathbb{P}\big[\overline{R} = n-k\big]}{\mathbb{P}\big[\Pi(n) = n\big]} \mathbb{P}\big[A_2 \cap \{R = k\}\big] \\
&\leq \sup_{k\in\mathbb{N}} \frac{\mathbb{P}\big[\overline{R} = n-k\big]}{\mathbb{P}\big[\Pi(n) = n\big]} \mathbb{P}[A_2].
\end{aligned}$$

En utilisant à nouveau l'approximation (3.66) et la notation (3.67) nous obtenons

$$\mathbb{P}[A_1] \leq \mathcal{C}_a \, \mathbb{P}[A_2].$$

Ceci achève la démonstration du corollaire 3.5.1. □

Remarque 3.5.4. Une version du corollaire 3.5.1 pour le processus empirique multivarié $\{\alpha_n(t) : t \in [0,1]^d\}$ est établie dans l'article de Ruymgaart et Wellner [118]. Cette version s'énonce ainsi.

Soit la classe

$$\mathcal{R} = \{R(x,y) : R(x,y) \subset [0,1]^d\}$$

de rectangles semi-ouverts dans $[0,1]^d$. Soit $\{X_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F. Soit $F_n(t)$ la fonction de répartition empirique construite sur les n premières variables aléatoires X_i. Soit $\{\alpha_n(t) = \sqrt{n}\,(F_n(t) - F(t)) : t \in [0,1]^d\}$ le processus empirique multivarié. Pour tout rectangle $R \in \mathcal{R}$ tel que $F(R) < 1$ et pour tout réel $\mu \geq 0$, nous avons

$$\mathbb{P}\left[\sup_{S \subset R} |\alpha_n(S)| \geq \mu\right] \leq C\left(F\left(R^C\right)\right)\mathbb{P}\left[\sup_{S \subset R}\left|\frac{\Pi_n(S) - nF(S)}{\sqrt{n}}\right| \geq \mu\right],$$

où le supremum est pris sur tous les rectangles S de $\mathcal{R}$ tels que $S \subset R$, $C(\theta)$ tend vers 1 lorsque θ tend vers 1, et $C(\theta) \leq 2$ pour $\theta \geq \frac{1}{2}$.

Chapitre 4

Le comportement des oscillations du processus empirique composé

4.1. Introduction et historique

Soit D l'espace des fonctions (càdlàg) définies sur l'intervalle $[0,1]$ qui sont continues à droite et qui admettent une limite à gauche pour tout point de l'intervalle $[0,1]$. Alors

$$(D, ||.||) \text{ est un espace métrique complet non séparable.}$$

On désigne par $\mathcal{D}_{||.||}$ la σ-tribu de Borel des parties de D, par $\mathcal{D}^B_{||.||}$ la σ-tribu des parties de D engendrée par les boules ouvertes et par $\mathcal{D}$ la σ-tribu engendrée par les parties de dimension finie de D. On en déduit que

$$\mathcal{D} = \mathcal{D}^B_{||.||} \subset \mathcal{D}_{||.||}.$$

Pour de plus amples détails, on pourra consulter le chapitre 2 du livre de Shorack et Wellner [125]. Soit $\mathbb{X}$ un processus défini sur l'espace $(D, \mathcal{D})$. Dans ce chapitre, C désigne un intervalle de type $(s,t]$ dans l'intervalle $[0,1]$ et $|C|$ représente la longueur de l'intervalle, i.e., $|C| = t - s$. Nous notons par $\mathbb{X}(C) = \mathbb{X}(t) - \mathbb{X}(s)$ l'accroissement du processus $\mathbb{X}$. Le module d'oscillation $\omega_{\mathbb{X}}$ du processus $\mathbb{X}$ est défini par

$$\begin{aligned} \omega_{\mathbb{X}}(a) &= \sup_{|C| \le a} |\mathbb{X}(C)| = \sup_{0 \le t-s \le a} |\mathbb{X}(s,t]| = \sup_{0 \le t-s \le a} |\mathbb{X}(t) - \mathbb{X}(s)| \\ &= \sup_{0 \le h \le a} \sup_{0 \le t \le 1-h} |\mathbb{X}(t+h) - \mathbb{X}(t)| . \end{aligned} \tag{4.1}$$

Comme les grands intervalles contrôlent le comportement du module d'oscillation $\omega_{\mathbb{X}}$, nous définissons également

$$\overline{\omega}_{\mathbb{X}}(a) = \sup_{|C| = a} |\mathbb{X}(C)| = \sup_{0 \le t \le 1-a} |\mathbb{X}(t+a) - \mathbb{X}(t)| .$$

Nous allons nous intéresser uniquement dans ce paragraphe au pont brownien $\mathbb{B}$ et plus particulièrement à ses modules d'oscillation notés

$$\omega = \omega_{\mathbb{B}} \quad \text{et} \quad \overline{\omega} = \overline{\omega}_{\mathbb{B}}.$$

Comme le mouvement brownien ou processus de Wiener $\mathbb{W}$ a des accroissements indépendants, il est plus facile d'étudier le module d'oscillation du mouvement brownien $\mathbb{W}$ et d'utiliser ensuite le fait que nous avons l'égalité en distribution suivante

$$\mathbb{B} \overset{\mathrm{d}}{=} \mathbb{W} - \mathcal{I}\mathbb{W}(1), \tag{4.2}$$

où $\mathcal{I}$ désigne l'application identité.

4.1.1. Le module de continuité du pont brownien

Comme la trajectoire du pont brownien $\mathbb{B}$ est continue presque partout, alors nous avons

$$\omega(a) \to 0 \text{ presque sûrement (p.s.) lorsque } a \to 0.$$

En 1937, Lévy [89] a précisé la vitesse de convergence du pont brownien $\mathbb{B}$ vers 0, dans le théorème suivant.

Théorème 4.1.1. *Nous avons*

$$\lim_{a\downarrow 0} \frac{\omega(a)}{\sqrt{2a\log(1/a)}} = 1 \quad p.s. \tag{4.3}$$

Remarque 4.1.1. En fait, Lévy a établi une vitesse de convergence pour le mouvement brownien $\mathbb{W}$. Le théorème s'énonce ainsi.

$$\lim_{a\downarrow 0} \frac{\omega_{\mathbb{W}}(a)}{\sqrt{2a\log(1/a)}} = 1 \quad \text{p.s.} \tag{4.4}$$

La transition de (4.4) à (4.3) via (4.2) est triviale.

En ce qui concerne le module d'oscillation $\overline{\omega}$ du pont brownien $\mathbb{B}$, nous avons un théorème analogue au théorème 4.1.1 qui s'énonce de la façon suivante.

Théorème 4.1.2. *Nous avons*

$$\lim_{a\downarrow 0} \frac{\overline{\omega}(a)}{\sqrt{2\log(1/a)}} = 1 \quad p.s. \tag{4.5}$$

Pour la démonstration des théorèmes 4.1.1 et 4.1.2, nous renvoyons au livre de Shorack et Wellner [125] p.537 et suivantes.

Remarque 4.1.2. Le pont brownien $\mathbb{B}$ est la limite naturelle du processus empirique uniforme $\{\alpha_n(t) : 0 \le t \le 1\}$ défini par

$$\alpha_n(t) = \sqrt{n}(\mathbb{U}_n(t) - \mathbb{E}[\mathbb{U}_n(t)]) \qquad \text{pour } t \in [0,1],$$

où $\mathbb{U}_n(t)$ est la fonction de répartition empirique uniforme définie par

$$\mathbb{U}_n(t) = \sum_{i=1}^{n} \mathbf{1}_{\{U_i \le t\}} \qquad \text{pour } t \in [0,1].$$

4.1.2. Le module d'oscillation du processus empirique uniforme

Dans ce paragraphe, nous allons nous intéresser au module d'oscillation ω_n du processus empirique uniforme $\{\alpha_n(t) : 0 \le t \le 1\}$. Posons

$$\begin{aligned}\omega_n = \omega_{\alpha_n}(a) &= \sup_{|C|\le a} |\alpha_n(C)| = \sup_{0\le t-s\le a} |\alpha_n(s,t]| \\ &= \sup_{0\le t-s\le a} |\alpha_n(t)-\alpha_n(s)| = \sup_{0\le h\le a}\ \sup_{0\le t\le 1-h} |\alpha_n(t+h)-\alpha_n(t)| \,.\end{aligned} \tag{4.6}$$

Par analogie avec le théorème 4.1.1 de Lévy, nous aimerions montrer que

$$\lim_{n\to\infty} \frac{\omega_n(a_n)}{\sqrt{2a_n \log(1/a_n)}} = 1 \quad \text{p.s.} \quad \text{pour } a_n \downarrow 0.$$

Toutefois, ce résultat ne peut être établi que sous des conditions additionnelles partant sur a_n.

Remarque 4.1.3. La vitesse avec laquelle la suite de réels $(a_n)_{n\ge 1}$ tend vers 0 est ici un facteur crucial. En effet, celle-ci peut à la fois influencer la valeur de la limite presque sûrement et la forme de la constante appropriée à la normalisation.

Suite à cette remarque, nous allons supposer que la suite $(a_n)_{n\ge1}$ vérifie les conditions suivantes.
(S.1) (i) $a_n \downarrow 0$ et (ii) $na_n \uparrow \infty$ lorsque $n\to\infty$;
(S.2) $\log(1/a_n)/\log\log n \to \infty$ lorsque $n \to \infty$;
(S.3) $\log(1/a_n)/na_n \to 0$ lorsque $n\to\infty$.

Remarque 4.1.4. Ces conditions sont connues sous le nom de conditions de Csörgő-Révész-Stute [CRS] (voir par exemple Csörgő-Révész [27], Stute [131]).

En 1982, Stute [131] établit un résultat analogue à celui de Lévy pour le module d'oscillation ω_n du processus empirique uniforme $\{\alpha_n(t) : 0 \le t \le 1\}$.

Théorème 4.1.3. *Soit une suite $(a_n)_{n\ge1}$ qui vérifie les conditions (S.1–2–3). Alors, nous avons*

$$\lim_{n\to\infty} \frac{\omega_n(a_n)}{\sqrt{2a_n \log(1/a_n)}} = 1 \quad p.s.$$

Démonstration. La démonstration du théorème 4.1.3 se trouve dans l'article de Stute [131]. On peut également la trouver dans le livre de Shorack et Wellner [125], p.552 et suivantes. □

Remarque 4.1.5. La condition (S.1) est une condition de lissage. La condition (S.3) impose à la suite $(a_n)_{n\ge1}$ de ne pas être trop petite ; en effet si

$$a_n = \frac{c_n \log n}{n},$$

alors la condition $(S.3)$ est réalisée si la suite c_n tend vers ∞ lorsque $n \to \infty$. Par contre, si $c_n \in (0, \infty)$ alors la condition $(S.3)$ n'est plus vérifiée. La condition $(S.2)$ impose à la suite $(a_n)_{n\geq 1}$ de ne pas être trop grande ; en effet si

$$a_n = \frac{1}{(\log n)^{c_n}},$$

alors la condition $(S.2)$ est réalisée si la suite c_n tend vers ∞ lorsque $n \to \infty$. Si $c_n \in (0, \infty)$ alors la condition $(S.2)$ n'est plus vérifiée. En résumé, le théorème 4.1.3 de Stute [131] traite les cas de lissage qui englobent les valeurs de la suite $(a_n)_{n\geq 1}$ aussi petites que $a_n = (c_n \log n)/n$ avec $c_n \to \infty$ aux valeurs de la suite $(a_n)_{n\geq 1}$ aussi grandes que $1/(\log n)^{c_n}$ avec $c_n \to \infty$.

Remarque 4.1.6. Le théorème 4.1.3 sert à Stute [131] pour établir une loi du logarithme pour estimer la densité par la méthode du noyau de convolution. L'estimateur à noyau pour la densité univariée a été introduit par Rosenblatt [115] puis par Parzen [109]. Il se définit ainsi

$$f_n(t) = a_n^{-1} \int K\left(\frac{t-x}{a_n}\right) F_n(dx), \quad t \in \mathbb{R},$$

où $(a_n)_{n\geq 1}$ est une suite positive qui tend vers 0 lorsque $n \to \infty$ et K est une fonction à variation bornée sur $\mathbb{R}$, positive et telle que $\int_{\mathbb{R}} K(u)du = 1$.
La loi du logarithme pour l'estimateur à noyau pour la densité établie par Stute [131] s'énonce ainsi.

Théorème 4.1.4. *Soit* $\{X_i : i \geq 1\}$ *une suite de variables aléatoires indépendantes et identiquement distribuées sur* $\mathbb{R}$ *de fonction de répartition* $F(x) = \mathbb{P}(X_1 \leq x)$*, où* $x \in \mathbb{R}$ *et de densité* $f(x) = F'(x)$ *supposée continue et positive sur un intervalle compact* $[A, B]$ *avec* $A < B$*. Soit* $(a_n)_{n\geq 1}$ *une suite de réels qui vérifie les conditions* $(S.1-2-3)$*. Soit* K *un noyau à variation bornée sur* $\mathbb{R}$ *tel que* $K(u) = 0$ *si* $u \notin [r, s)$ *et* $\int_{\mathbb{R}} K(u)du = 1$*. Alors, pour tous réels* C *et* D *tels que* $A < C < D < B$*, nous avons*

$$\lim_{n\to\infty} \sqrt{\frac{na_n}{2\log(1/a_n)}} \sup_{C\leq t\leq D} \left| \frac{f_n(t) - \mathbb{E}[f_n(t)]}{\sqrt{f(t)}} \right| = \sqrt{\int_{\mathbb{R}} K^2(u)du} \quad p.s. \tag{4.7}$$

Démonstration. La démonstration du théorème 4.1.4 se trouve dans l'article [131] de Stute. Elle s'appuie essentiellement sur deux idées : une intégration par parties et l'utilisation du théorème 4.1.3. □

Remarque 4.1.7. L'intégration par parties a déjà été utilisée par Nadaraya [103].

Remarque 4.1.8. En 1981, Hall [62] a établi la loi du logarithme itéré pour l'estimateur à noyau pour la densité.

Remarque 4.1.9. En 1992, Deheuvels and Mason [35] établissent également une loi du logarithme pour l' estimateur à noyau pour la densité. La nouveauté réside dans la démonstration du théorème 4.1.5 et par une extension de (4.7).

Théorème 4.1.5. *Soit* $(a_n)_{n\geq 1}$ *une suite de réels qui vérifie les conditions* (S.1–2–3). *Soit* K *un noyau à variation bornée sur* $\mathbb{R}$ *tel que* $K(u) = 0$ *pour* $|u| \geq M$*, où* $0 < M < \infty$ *et* $\int_{\mathbb{R}} K(u)du = 1$*. Alors, pour tous réels* C *et* D *tels que* $A < C < D < B$*, nous avons*

$$\lim_{n\to\infty} \pm\sqrt{\frac{na_n}{2\log(1/a_n)}} \sup_{C\leq t\leq D} \left(\frac{f_n(t) - \mathbb{E}[f_n(t)]}{\sqrt{f(t)}} \right) = \sqrt{\int_{\mathbb{R}} K^2(u)du} \quad p.s.$$

Démonstration. La démonstration du théorème 4.1.5 fait appel à des lois limites fonctionnelles locales. Elle se trouve dans l'article de Deheuvels et Mason [35]. □

4.2. Notations et hypothèses

Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0, 1]$. Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. Nous supposons maintenant que $\mathbb{E}[Y] = 1$, $\mathrm{Var}[Y] < \infty$ et que $|Y - 1| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable. De plus, les suites de variables $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes.

Remarque 4.2.1. Comme $\mathbb{E}[Y] = 1$ et $\mathrm{Var}[Y] < \infty$, ceci implique que

$$0 < \mathbb{E}[Y^2] < \infty.$$

Pour tout entier $n \geq 1$, nous notons

$$\mathbb{U}_{n,c}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i\leq t\}}, \qquad \text{pour } t \in [0, 1], \tag{4.8}$$

la fonction composée empirique, continue à droite, et définie en fonction des n premières variables aléatoires des suites de variables $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$. On rappelle que, ici $\mathbf{1}_A$ représente la fonction indicatrice de l'événement A et la lettre "c" est utilisée pour rappeler "composé".

Pour tout entier $n \geq 1$, nous introduisons

$$\alpha_{n,c}(t) = \sqrt{n}\left(\mathbb{U}_{n,c}(t) - \mathbb{E}[\mathbb{U}_{n,c}(t)]\right) \qquad \text{pour } t \in [0, 1], \tag{4.9}$$

le processus empirique composé. Par la suite, pour tout entier $n \geq 1$ et pour tout réel $t \notin [0, 1]$, nous posons $\alpha_{n,c}(t) = 0$.

Le module d'oscillation $\omega_{n,c}$ du processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ sur l'intervalle $[0, a]$ est défini par

$$\begin{aligned} \omega_{n,c}(a) &= \omega_{\alpha_{n,c}}(a) = \sup_{|C|\leq a} |\omega_{n,c}(C)| = \sup_{0\leq t-s\leq a} |\alpha_{n,c}(s, t]| \\ &= \sup_{0\leq t-s\leq a} |\alpha_{n,c}(t) - \alpha_{n,c}(s)| = \sup_{0\leq h\leq a}\ \sup_{0\leq t\leq 1-h} |\alpha_{n,c}(t + h) - \alpha_{n,c}(t)|. \end{aligned}$$

4.3. Résultat sur le comportement des oscillations du processus empirique composé

Le théorème 4.3.1 donne un résultat décrivant le comportement limite du module d'oscillation $\omega_{n,c}$ du processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ défini en (4.9).

Théorème 4.3.1. *Supposons que les hypothèses du paragraphe* 4.2 *sont satisfaites. Soit* $(a_n)_{n\geq 1}$ *une suite de constantes à valeurs dans* $(0,1)$ *qui vérifie les conditions de Csörgő-Révész-Stute* [CRS], *c'est-à-dire*
(S.1) (i) $a_n \downarrow 0$ *et (ii)* $na_n \uparrow \infty$ *lorsque* $n \to \infty$ *;*
(S.2) $\log(1/a_n)/\log\log n \to \infty$ *lorsque* $n \to \infty$ *;*
(S.3) $\log(1/a_n)/na_n \to 0$ *lorsque* $n \to \infty$.
Alors, nous avons

$$\lim_{n\to\infty} \frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} = 1 \quad p.s. \tag{4.10}$$

Pour démontrer le théorème 4.3.1 nous allons utiliser une technique dite de "poissonisation", consistant à approcher le processus empirique composé $\{\alpha_{n,c}(t) : 0 \leq t \leq 1\}$ par un processus de Poisson composé.

4.4. Résultats préliminaires : deux inégalités

Pour démontrer le théorème 4.3.1, nous utiliserons les deux inégalités qui vont suivre. La première inégalité se déduit de l'approximation du processus empirique composé par un processus de Poisson composé et d'une inégalité de type exponentiel pour le processus de Poisson composé établies au chapitre précédent. La seconde inégalité est un exemple d'inégalité maximale ; ces inégalités jouent un rôle fondamental pour établir des résultats asymptotiques presque sûrs ; elles sont attachées au nom de Kolmogorov.

Inégalité 4.4.1. *Soit un réel* a *tel que* $0 < a \leq \delta \leq \frac{1}{2}$. *Alors, pour tout réel* $\mu \geq 0$, *nous avons*

$$\mathbb{P}\left[\omega_{n,c}(a) \geq \mu\sqrt{a}\,\right] \leq \frac{1280}{a\delta^3} \exp\left(-(1-\delta)^3 \frac{\mu^2}{32\mathbb{E}[Y^2]} \psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right), \tag{4.11}$$

où la fonction ψ *est définie par*

$$\psi(u) = \frac{1}{u^2}\{(1+u)\log(1+u) - u\} \qquad pour \quad u > 0. \tag{4.12}$$

Remarque 4.4.1. Pour le processus empirique uniforme $\{\alpha_n(t) : 0 \leq t \leq 1\}$ nous avons une inégalité semblable qui s'énonce de la façon suivante.

Soit un réel a tel que $0 < a \le \delta \le \frac{1}{2}$. Alors pour tout réel $\mu \ge 0$, nous avons

$$\mathbb{P}\left[\omega_n(a) \ge \mu\sqrt{a}\right] \le \frac{20}{a\delta^3}\exp\left(-(1-\delta)^4\frac{\mu^2}{2}\psi\left(\frac{\mu}{\sqrt{an}}\right)\right),$$

où ω_n est le module d'oscillation du processus empirique uniforme $\{\alpha_n(t) : 0 \le t \le 1\}$ défini en (4.6). Cette inégalité est dûe à Mason, Shorack et Wellner [95]. Pour la démonstration de cette inégalité, on pourra consulter l'article de Mason, Shorack et Wellner [95] ou encore le livre de Shorack et Wellner [125].

Démonstration. La démonstration de l'inégalité (4.11) se fait en plusieurs étapes. Nous allons d'abord rappeler le théorème 3.3.3 et le corollaire 3.5.1 du chapitre précédent.

Théorème 4.4.1. *Soit un réel $a > 0$. Soit $\{\Pi_{n,c}(t) : 0 \le t \le 1\}$ un processus de Poisson composé défini dans le chapitre précédent. Supposons que les variables aléatoires Y_i vérifient, pour $i \ge 1$, $|Y_i - \mathbb{E}[Y_i]| \le \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable. Alors, pour tout réel $\mu \ge 0$, nous avons*

$$\mathbb{P}\left[\sup_{0\le t\le a}\frac{|\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]|}{\sqrt{n}} \ge \mu\right] \le 32\exp\left(-\frac{\mu^2}{32a\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4a\sqrt{n}\mathbb{E}[Y^2]}\right)\right), \quad (4.13)$$

où la fonction ψ est définie en (4.12).

Corollaire 4.4.1. *Soit $\{\Pi_{n,c}(t) : 0 \le t \le 1\}$ un processus de Poisson composé. Pour tout réel $a \in (0,1)$, il existe une constante réelle $\mathcal{C}_a$ telle que $0 < \mathcal{C}_a < \infty$ et telle que pour tout réel $\mu \ge 0$, nous ayons*

$$\mathbb{P}\left[\sup_{0\le t\le a} \pm\alpha_{n,c}(t) \ge \mu\right] \le \mathcal{C}_a\mathbb{P}\left[\sup_{0\le t\le a}\frac{\pm\left(\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]\right)}{\sqrt{n}} \ge \mu\right]. \quad (4.14)$$

Remarque 4.4.2. Nous allons préciser la constante $\mathcal{C}_a$ dans le cas qui nous intérèsse. Rappelons que nous avons défini la constante $\mathcal{C}_a$ par

$$\mathcal{C}_a = \sup_{n\ge 1}\left\{\frac{M(n(1-a))}{M(n)}\right\},$$

où

$$M(\lambda) = \frac{\lambda^{\lfloor\lambda\rfloor}}{\lfloor\lambda\rfloor!}\exp(-\lambda),$$

où $\lfloor u\rfloor$ désigne la partie entière de u. Comme $n! = (n/e)^n(2\pi n)^{1/2}\exp(\theta_n/n)$, avec $0 < \theta_n < 1/12$ pour $n \ge 1$, alors nous obtenons que

$$\frac{M(n(1-a))}{M(n)} \le \frac{n^{1/2}\exp(1/12)}{\left(\frac{1}{2}n - 1\right)^{1/2}} \le 2 \quad \text{pour } n \ge 5.$$

Désormais la constante $\mathcal{C}_a$ sera prise égale à 2, sans perte de généralité pour $n \ge 5$.

Première étape. D'après le corollaire 4.4.1, pour tout couple de réels $a \in (0, 1/2]$ et $\mu \geq 0$, nous avons

$$\begin{aligned}\mathbb{P}\left[\omega_{n,c}^{+}(a) \geq \mu\sqrt{a}\,\right] &= \mathbb{P}\left[\sup_{0\leq t-s\leq a} \alpha_{n,c}(s,t] \geq \mu\sqrt{a}\right] \\ &\leq 2\,\mathbb{P}\left[\sup_{0\leq t-s\leq a} \frac{\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]]}{\sqrt{n}} \geq \mu\sqrt{a}\right].\end{aligned}$$

Soit N un entier non nul. Remarquons que, pour tous réels s et t tels que $t - s \leq a$, nous avons

$$\begin{aligned}\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]] &= \Pi_{n,c}(t) - \Pi_{n,c}(s) - n(t-s) \\ &\leq \Pi_{n,c}(t) - \Pi_{n,c}\left(\frac{j}{N}\right) - n\left(t - \frac{j}{N}\right) + \frac{n}{N},\end{aligned}$$

si j/N est le point du maillage le plus proche du point s et qui ne le dépasse pas. Parfois nous utiliserons la notation suivante : $\Pi'_{n,c}(t) = \Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]$. Pour tout réel $a \in (0, 1/2]$, nous avons

$$\sup_{0\leq t-s\leq a} \Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]] \leq \max_{0\leq j\leq N-1} \sup_{0\leq r\leq a+1/N} \left\{\Pi'_{n,c}\left(\frac{j}{N}, \frac{j}{N} + r\right] + \frac{n}{N}\right\}.$$

Comme le processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$ a des accroissements stationnaires (propriété établie dans le théorème 3.2.1 du chapitre précédent), nous avons

$$\mathbb{P}\left[\sup_{0\leq t-s\leq a} \frac{\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]]}{\sqrt{n}} \geq \mu\sqrt{a}\right] \leq N\mathbb{P}\left[\sup_{0\leq r\leq a+1/N} \frac{\Pi'_{n,c}(r)}{\sqrt{n}} \geq \mu\sqrt{a} - \frac{\sqrt{n}}{N}\right]. \tag{4.15}$$

Ainsi, pour tous réels $a \in (0, 1/2]$ et $\mu \geq 0$, nous en déduisons que

$$\mathbb{P}\left[\omega_{n}^{+}(a) \geq \mu\sqrt{a}\,\right] \leq 2\,N\,\mathbb{P}\left[\sup_{0\leq r\leq a+1/N} \frac{\Pi_{n,c}(r) - \mathbb{E}[\Pi_{n,c}(r)]}{\sqrt{n}} \geq \mu\sqrt{a} - \frac{\sqrt{n}}{N}\right].$$

En ce qui concerne le terme $\omega_{n,c}^{-}(a)$, nous pouvons faire une démonstration analogue. En effet, nous avons

$$\begin{aligned}-\left(\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]]\right) &= n(t-s) - (\Pi_{n,c}(t) - \Pi_{n,c}(s)) \\ &\leq n\left(\frac{k+1}{N} - s\right) - \left(\Pi_{n,c}\left(\frac{k}{N} \vee s\right) - \Pi_{n,c}(s)\right).\end{aligned}$$

Cette dernière inégalité conduit à

$$\mathbb{P}\left[\omega_{n,c}^{-}(a) \geq \mu\sqrt{a}\,\right] \leq 2N\mathbb{P}\left[\sup_{0\leq r\leq a} \frac{-\left(\Pi_{n,c}(r) - \mathbb{E}[\Pi_{n,c}(r)]\right)}{\sqrt{n}} \geq \mu\sqrt{a} - \frac{\sqrt{n}}{N}\right]. \tag{4.16}$$

Deuxième étape. En utilisant le théorème 4.4.1, nous obtenons que

$$\begin{aligned}
\mathbb{P}\left[\omega_{n,c}^{+}(a) \geq \mu\sqrt{a}\right] &\leq 2\,N\,\mathbb{P}\left[\sup_{0\leq r\leq a+1/N} \frac{\Pi_{n,c}(r)-\mathbb{E}[\Pi_{n,c}(r)]}{\sqrt{n}} \geq \mu\sqrt{a}-\frac{\sqrt{n}}{N}\right] \\
&\leq 64\,N\exp\left(-\frac{\left(\mu\sqrt{a}-\sqrt{n}/N\right)^2}{32\mathbb{E}[Y^2]\,(a+1/N)}\psi\left(\frac{\mathcal{M}\left(\mu\sqrt{a}-\sqrt{n}/N\right)}{4\mathbb{E}[Y^2]\,(a+1/N)\,\sqrt{n}}\right)\right) \\
&\leq 64\,N\exp\left(-\frac{\mu^2}{32\mathbb{E}[Y^2]}\,\frac{a}{(a+1/N)}\left(1-\frac{\sqrt{n}}{\mu N\sqrt{a}}\right)^2\right. \\
&\qquad\left.\times\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right),
\end{aligned}$$

puisque la fonction ψ est décroissante. Ce résultat a été établi dans la proposition 3.3.1 du chapitre précédent.

Troisième étape. Maintenant nous choisissons N comme étant le plus petit entier non nul pour lequel nous avons

$$\frac{1}{N} < a\delta^3, \text{ ce qui implique que } \frac{2}{N} \geq \frac{1}{N-1} \geq a\delta^3. \tag{4.17}$$

Remarquons aussi que si $\mu \geq \delta^2\sqrt{an}$, alors l'inégalité (4.17) implique que

$$\mu \geq \delta^2\sqrt{an} \geq \frac{\sqrt{n}}{N\sqrt{a}\delta}. \tag{4.18}$$

En utilisant l'inégalité (4.17) puis l'inégalité (4.18), nous avons

$$\frac{a}{a+1/N} = \frac{1}{1+\delta^3} \geq 1-\delta \tag{4.19}$$

et

$$\left(1-\frac{\sqrt{n}}{\mu N\sqrt{a}}\right) \geq (1-\delta)\,. \tag{4.20}$$

D'après les inégalités (4.17), (4.19) et (4.20), et si $\mu \geq \delta^2\sqrt{an}$, alors nous obtenons que

$$\mathbb{P}\left[\omega_{n,c}^{+}(a) \geq \mu\sqrt{a}\right] \leq \frac{128}{a\delta^3}\exp\left(-(1-\delta)^3\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right).$$

Quatrième étape. En ce qui concerne le terme $\omega_{n,c}^{-}(a)$, la démonstration est analogue.

D'après l'inégalité (4.16), le corollaire 4.4.1 et le théorème 4.4.1, nous avons successivement

$$
\begin{aligned}
\mathbb{P}\left[\omega^{-}_{n,c}(a) \geq \mu\sqrt{a}\,\right] &\leq 2\,N\,\mathbb{P}\left[\sup_{0\leq r\leq a} \frac{-\left(\Pi_{n,c}(r) - \mathbb{E}[\Pi_{n,c}(r)]\right)}{\sqrt{n}} \geq \mu\sqrt{a} - \frac{\sqrt{n}}{N}\right] \\
&\leq 64\,N\,\exp\left(-\frac{\left(\mu\sqrt{a} - \sqrt{n}/N\right)^2}{32\mathbb{E}[Y^2]a}\psi\left(-\frac{\mathcal{M}\left(\mu\sqrt{a} - \sqrt{n}/N\right)}{4\mathbb{E}[Y^2]a\sqrt{n}}\right)\right) \\
&\leq 64\,N\,\exp\Bigg(-\frac{\mu^2}{32\mathbb{E}[Y^2]}\left(1 - \frac{\sqrt{n}}{\mu N\sqrt{a}}\right)^2 \\
&\qquad\qquad \times\psi\left(-\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\left(1 - \frac{\sqrt{n}}{N\mu\sqrt{a}}\right)\right)\Bigg) \\
&\leq \frac{128}{a\delta^3}\exp\left(-\left(1-\delta\right)^2\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(-\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\left(1-\delta\right)\right)\right) \\
&\leq \frac{128}{a\delta^3}\exp\left(-\left(1-\delta\right)^2\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right),
\end{aligned}
$$

puisque la fonction ψ est décroissante.
Cinquième étape. En regroupant le cas positif et le cas négatif et si $\mu \geq \delta^2\sqrt{an}$, alors pour tous réels $a \in (0, 1/2]$ et $\mu \geq 0$, nous obtenons que

$$\mathbb{P}\left[\omega_{n,c}(a) \geq \mu\sqrt{a}\,\right] \leq \frac{128}{a\delta^3}\exp\left(-\left(1-\delta\right)^3\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right).$$

Sixième étape. Nous allons démontrer l'inégalité dans le cas où $\mu \leq \delta^2\sqrt{an}$. D'abord d'après le corollaire 4.4.1, pour tout réel $a \in (0, 1/2]$, nous avons

$$\mathbb{P}\left[\omega_{n,c}(a) \geq \mu\sqrt{a}\,\right] \leq 4\,\mathbb{P}\left[\sup_{0\leq t-s\leq a} \frac{|\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]]|}{\sqrt{n}} \geq \mu\sqrt{a}\right].$$

Soit N un entier non nul. Remarquons que, pour tous réels s et t tels que $0 \leq t - s \leq a$, nous avons

$$
\begin{aligned}
\sup_{0\leq t-s\leq a} |\Pi_{n,c}(s,t] - \mathbb{E}[\Pi_{n,c}(s,t]]| \leq \max_{1\leq j\leq N}\sup_{0\leq r\leq a}\left|\Pi'_{n,c}\left(\frac{j}{N}, \frac{j}{N} + r\right]\right| \\
+ 2\max_{1\leq j\leq N}\sup_{0\leq r\leq 1/N}\left|\Pi'_{n,c}\left(\frac{j}{N} - r, \frac{j}{N}\right]\right|.
\end{aligned}
\tag{4.21}
$$

Le facteur 2 sur le second terme est nécessaire dans le cas où aucun point de la forme j/N se trouve entre les points s et t.
Septième étape. Maintenant nous choisissons N comme étant le plus petit entier non nul tel que nous ayons

$$\frac{1}{N} \leq \frac{a\delta^2}{4}, \qquad \text{ce qui implique que} \qquad N < \frac{4}{a\delta^2} + 1 \leq \frac{5}{a\delta^2}. \tag{4.22}$$

En utilisant l'inégalité (4.21) et la stationnarité des accroissements du processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$ nous obtenons que

$$\begin{aligned}
\mathbb{P}\left[\omega_{n,c}(a) \geq \mu\sqrt{a}\right] &\leq 4\sum_{j=1}^{N}\mathbb{P}\left[\sup_{0\leq r\leq a}\frac{|\Pi_{n,c}(r)-\mathbb{E}[\Pi_{n,c}(r)]|}{\sqrt{n}} \geq \mu\sqrt{a}\,\frac{1}{1+\delta}\right] \\
&\quad +4\sum_{j=1}^{N}\mathbb{P}\left[\sup_{0\leq r\leq 1/N}\frac{|\Pi_{n,c}(r)-\mathbb{E}[\Pi_{n,c}(r)]|}{\sqrt{n}} \geq \mu\sqrt{a}\,\frac{\delta}{2(1+\delta)}\right] \\
&= R+S.
\end{aligned}$$

Neuvième étape. En appliquant successivement le corollaire 4.4.1 et le théorème 4.4.1 à la quantité R, nous obtenons que

$$\begin{aligned}
R &\leq 4\,N\,\mathbb{P}\left[\sup_{0\leq r\leq a}\frac{|\Pi_{n,c}(r)-\mathbb{E}[\Pi_{n,c}(r)]|}{\sqrt{n}} \geq \mu\sqrt{a}\,\frac{1}{1+\delta}\right] \\
&\leq 128\,N\,\exp\left(-\frac{\mu^2 a}{32a\mathbb{E}[Y^2](1+\delta)^2}\psi\left(\frac{\mathcal{M}\mu\sqrt{a}}{4\mathbb{E}[Y^2]a\sqrt{n}\,(1+\delta)}\right)\right) \\
&\leq \frac{640}{a\delta^2}\,\exp\left(-\frac{\mu^2}{32\mathbb{E}[Y^2](1+\delta)^2}\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}\,(1+\delta)}\right)\right) \\
&\leq \frac{640}{a\delta^2}\,\exp\left(-(1-\delta)^2\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{4\mathbb{E}[Y^2]\sqrt{an}}\right)\right) \\
&\leq \frac{640}{a\delta^2}\,\exp\left(-(1-\delta)^3\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\right),
\end{aligned}$$

en utilisant l'inégalité (4.22), le fait que $\dfrac{1}{(1+\delta)^2} \geq (1-\delta)^2$, la décroissance de la fonction ψ et la propriété suivante $\psi(t) \geq (1-\delta)$ pour $0 \leq t \leq 3\delta$ où $0 \leq \delta \leq 1$. Cette dernière propriété a été énoncée dans la proposition 3.3.1 du chapitre précédent.

Pour la quantité S, nous procédons de manière analogue. Ainsi nous obtenons que

$$\begin{aligned}
S &\leq 4\,N\,\mathbb{P}\left[\sup_{0\leq r\leq 1/N}\frac{|\Pi_{n,c}(r)-\mathbb{E}[\Pi_{n,c}(r)]|}{\sqrt{n}} \geq \mu\sqrt{a}\,\frac{\delta}{2(1+\delta)}\right] \\
&\leq 128\,N\,\exp\left(-\frac{\mu^2 a\delta^2}{32\mathbb{E}[Y^2](1/N)4(1+\delta)^2}\psi\left(\frac{\mathcal{M}\mu\sqrt{a}\,\delta}{4\mathbb{E}[Y^2](1/N)\,2\,\sqrt{n}\,(1+\delta)}\right)\right) \\
&\leq \frac{640}{a\delta^2}\,\exp\left(-(1-\delta)^2\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu\sqrt{a}\,\delta}{4\mathbb{E}[Y^2](1/N)\,2\,\sqrt{n}\,(1+\delta)}\right)\right), \\
&\quad \text{puisque } \frac{4}{a\delta^2} \leq N \leq \frac{5}{a\delta^2} \text{ et } \frac{1}{(1+\delta)^2} \geq (1-\delta)^2, \\
&\leq \frac{640}{a\delta^2}\,\exp\left(-(1-\delta)^2\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\,(1-\delta)\right) \quad \text{si } \frac{\mathcal{M}\,\mu\sqrt{a}\,\delta\,N}{4\,\mathbb{E}[Y^2]\,2\,\sqrt{n}(1+\delta)} \leq 3\delta.
\end{aligned}$$

Ainsi, nous avons

$$\mathbb{P}\left[\omega_{n,c}(a) \geq \mu\sqrt{a}\,\right] \leq \frac{1280}{a\delta^2}\exp\left(-(1-\delta)^3\,\frac{\mu^2}{32\mathbb{E}[Y^2]}\right) \quad \text{si } \mu \leq \frac{24\mathbb{E}[Y^2]\sqrt{n}(1+\delta)}{\mathcal{M}\sqrt{a}N}.$$

Cette dernière condition sur μ est compatible avec la conditon $\mu \leq \delta^2\sqrt{an}$ pour un δ bien choisi lorsque la constante $\mathcal{M}$ et $\mathbb{E}[Y^2]$ sont fixées. Ainsi, l'inégalité (4.11) est démontrée dans le cas $\mu \leq \delta^2\sqrt{an}$. □

Inégalité 4.4.2. *Soit un réel $r > 0$. Soient $(a_m)_{m\geq 1}$ et $(\mu_m)_{m\geq 1}$ deux suites de réels qui vérifient les conditions suivantes*
(i) $a_m \downarrow$ et $\mu_m \uparrow$;
(ii) $ma_m \uparrow$;
(iii) $\mu_m/\sqrt{ma_m} \leq d$ où d est une constante strictement positive ;
(iv) $\mu_m \geq \sqrt{2\log(1/a_m)}$.
Soit un réel $\varepsilon > 0$ fixé. Pour $n_k = \lfloor (1+\theta)^k \rfloor$ nous pouvons choisir un réel $\theta = \theta_{\varepsilon,d}$ si petit que pour tout entier $k \geq k_{\varepsilon,d}$, nous avons

$$\mathbb{P}\left[\max_{n_{k-1}\leq m\leq n_k} \frac{\omega_{m,c}(a_m)}{\mu_m\sqrt{\mathbb{E}[Y^2]a_m}} \geq r + 2\varepsilon\right] \leq 2\mathbb{P}\left[\frac{\omega_{n_{k+1},c}((1+\theta)^2\,a_{n_k})}{(1+\theta)\sqrt{\mathbb{E}[Y^2]a_{n_k}}} \geq \frac{r+\varepsilon}{1+\theta}\mu_{n_{k-1}}\right]. \tag{4.23}$$

Remarque 4.4.3. Pour le processus empirique uniforme $\{\alpha_n(t) : 0 \leq t \leq 1\}$, nous avons une inégalité semblable qui a été démontrée par Stute [131] et qui s'énonce ainsi.
Soit un réel $r > 0$. Soient $(a_m)_{m\geq 1}$ et $(\mu_m)_{m\geq 1}$ deux suites de réels qui vérifient les conditions de l'inégalité (4.4.2). Soit un réel $\varepsilon > 0$ fixé. Pour $n_k = \lfloor (1+\theta)^k \rfloor$ nous pouvons choisir un réel $\theta = \theta_{\varepsilon,d}$ si petit que pour tout entier $k \geq k_{\varepsilon,d}$, nous avons

$$\mathbb{P}\left[\max_{n_{k-1}\leq m\leq n_k} \frac{\omega_m(a_m)}{\mu_m\sqrt{a_m}} \geq r + 2\varepsilon\right] \leq 2\mathbb{P}\left[\frac{\omega_{n_{k+1}}((1+\theta)^2\,a_{n_k})}{(1+\theta)\sqrt{a_{n_k}}} \geq \frac{r+\varepsilon}{1+\theta}\mu_{n_{k-1}}\right]. \tag{4.24}$$

La démonstration de cette inégalité se trouve dans l'article de Stute [131] ou encore dans le livre de Shorack et Wellner [125].

Démonstration. Soit un réel $\varepsilon > 0$ fixé. Pour plus de facilité d'écriture, posons

$$\underline{n} = n_{k-1}, \quad n = n_k \quad \overline{n} = n_{k+1}.$$

Pour $\underline{n} \leq m \leq n$, définissons

$$A_{m,c} = \left[\sup_{t-s\leq a_m} \pm\alpha_{m,c}(s,t] \geq (r+2\varepsilon)\,\mu_m\sqrt{\mathbb{E}[Y^2]a_m}\right].$$

Notons

$$M = \overline{n} - m \quad \text{et} \quad \mathbb{V}_{M,c}(t) = \frac{1}{\sqrt{M}}\sum_{i=m+1}^{\overline{n}} \left(Y_i\mathbf{1}_{\{U_i\leq t\}} - t\right) \quad \text{pour } t \in [0,1].$$

Définissons également, pour $\underline{n} \leq m \leq n$,

$$B_{m,c} = \left[\sup_{t-s \leq a_m} |\mathbb{V}_{M,c}(s,t]| < K\sqrt{\theta}\mu_m\sqrt{\mathbb{E}[Y^2]a_m} \right],$$

où $K = K_{\varepsilon,d}$ est une constante très grande et $\theta = \theta_{\varepsilon,d}$ est une constante très petite qui seront précisées ci-dessous. Remarquons que

$$\alpha_{\overline{n}}(s,t] = \sqrt{m/\overline{n}}\,\alpha_m(s,t] + \sqrt{M/\overline{n}}\,\mathbb{V}_{M,c}(s,t]. \tag{4.25}$$

Nous allons d'abord montrer que nous pouvons choisir les constantes $K = K_{\varepsilon,d}$ et $\theta = \theta_{\varepsilon,d}$ telles que

$$\inf_{\underline{n} \leq m \leq n} \mathbb{P}\left[B_{m,c}\right] > \frac{1}{2}. \tag{4.26}$$

Nous avons successivement

$$\begin{aligned}
\mathbb{P}\left[B^c_{m,c}\right] \;&\leq\; \frac{10240}{a_m}\exp\left(-\frac{K^2\theta\mu_m^2}{256}\psi\left(\frac{\mathcal{M}K\sqrt{\theta}\,\mu_m}{4\sqrt{\mathbb{E}[Y^2]Ma_m}}\right)\right)\\
&\quad\text{d'après l'inégalité (4.11) en prenant } \delta = 1/2,\\
&\leq\; \frac{10240}{a_m}\exp\left(-\frac{K^2\theta}{128}\log\left(\frac{1}{a_m}\right)\psi\left(\frac{\sqrt{2}\,\mathcal{M}K\mu_m}{4\sqrt{\mathbb{E}[Y^2]ma_m}}\right)\right)\\
&\quad\text{d'après la condition (iv) et puique la fonction } \psi \text{ est décroissante}\\
&\quad\text{et } M = m(\overline{n}/m - 1) \geq m(\overline{n}/n - 1) \geq m\theta/2 \text{ pour } k \geq k_\theta,\\
&\leq\; \frac{10240}{a_m}\exp\left(-\frac{K^2\theta}{128}\log\left(\frac{1}{a_m}\right)\psi\left(\frac{\sqrt{2}\,\mathcal{M}Kd}{4\sqrt{\mathbb{E}[Y^2]}}\right)\right)\\
&\quad\text{d'après la condition (iii).}
\end{aligned}$$

Comme $\psi(\lambda) \sim \dfrac{2\log(\lambda)}{\lambda}$ lorsque $\lambda \to \infty$ (propriété établie dans la proposition 3.3.1 du chapitre précédent), alors lorsque $K \to \infty$ nous avons

$$\frac{K^2\theta}{128}\log\left(\frac{1}{a_m}\right)\psi\left(\frac{\sqrt{2}\,\mathcal{M}Kd}{4\sqrt{\mathbb{E}[Y^2]}}\right) \sim \frac{K\theta\sqrt{\mathbb{E}[Y^2]}}{16\sqrt{2}\mathcal{M}d}\log\left(\frac{1}{a_m}\right)\log\left(\frac{\mathcal{M}Kd}{\sqrt{\mathbb{E}[Y^2]}}\right).$$

Par conséquent, nous en déduisons que

$$\begin{aligned}
\mathbb{P}\left[B^c_{m,c}\right] \;&\leq\; \frac{10240}{a_m}\exp\left(-\frac{K\theta\sqrt{\mathbb{E}[Y^2]}}{32\mathcal{M}d}\log\left(\frac{\mathcal{M}Kd}{\sqrt{\mathbb{E}[Y^2]}}\right)\log\left(\frac{1}{a_m}\right)\right)\\
&\quad\text{pour } K \text{ suffisamment grand},\\
&\leq\; \frac{10240}{a_m}\exp\left(-G\log\left(\frac{1}{a_m}\right)\right) = 10240\,a_m^{G-1} \quad \text{d'après (4.27)},\\
&\leq\; 10240\,a_1^{G-1} \quad \text{puisque la suite } (a_m)_{m\geq 1} \text{ est décroissante},\\
&\leq\; \frac{1}{2} \text{ si } G \text{ est choisi suffisamment grand},
\end{aligned}$$

où

$$K = K_{\varepsilon,d} = \frac{D}{\varepsilon}, \quad \theta = \theta_{\varepsilon,d} = \frac{\varepsilon^2}{2D}, \quad D = D_{\varepsilon,d} = \frac{\varepsilon\sqrt{\mathbb{E}[Y^2]}}{\mathcal{M}d} \exp\left(\frac{64G\mathcal{M}d}{\varepsilon\sqrt{\mathbb{E}[Y^2]}}\right). \tag{4.27}$$

Nous avons ainsi démontré l'inégalité (4.26) et précisé les constantes K et θ. Démontrons maintenant que

$$\mathbb{P}\left[\frac{\omega_{n_{k+1},c}((1+\theta)^2 a_{n_k})}{(1+\theta)\sqrt{\mathbb{E}[Y^2]a_{n_k}}} \geq \frac{r+\varepsilon}{1+\theta}\mu_{n_{k-1}}\right] \geq \frac{1}{2}\mathbb{P}\left[\max_{n_{k-1}\leq m\leq n_k} \frac{\omega_{m,c}(a_m)}{\mu_m\sqrt{\mathbb{E}[Y^2]a_m}} \geq r+2\varepsilon\right].$$

Remarquons d'abord que pour $k \geq k_\theta$ et pour tout $\underline{n} \leq m \leq n$, nous avons

$$a_m \leq (1+\theta)^2(\underline{n}\,/n)\,a_m \leq (1+\theta)^2(ma_m/n) \leq (1+\theta)^2 a_n, \tag{4.28}$$

en utilisant la croissance de la suite $(ma_m)_{m\geq 1}$. Ainsi, pour $\underline{n} \leq m \leq n$, pour l'événement $\{A_{m,c} \cap B_{m,c}\}$, nous avons successivement

$$\begin{aligned}
\sup_{t-s\leq(1+\theta)^2 a_n} \frac{\pm\alpha_{\overline{n},c}(s,t]}{\sqrt{\mathbb{E}[Y^2]a_n}} &\geq \sup_{t-s\leq a_m} \frac{\pm\alpha_{\overline{n},c}(s,t]}{\sqrt{\mathbb{E}[Y^2]a_m}} \\
&\quad \text{d'après (4.28) et la décroissance de la suite } (a_m)_{m\geq 1}, \\
&\geq \sqrt{\frac{m}{\overline{n}}} \sup_{t-s\leq a_m} \frac{\pm\alpha_{m,c}(s,t]}{\sqrt{\mathbb{E}[Y^2]a_m}} - \sqrt{\frac{M}{\overline{n}}} \sup_{t-s\leq a_m} \frac{|\mathbb{V}_{M,c}(s,t]|}{\sqrt{\mathbb{E}[Y^2]a_m}} \\
&\quad \text{d'après (4.25)}, \\
&\geq \sqrt{\frac{\underline{n}}{\overline{n}}}(r+2\varepsilon)\mu_m - \sqrt{\frac{\overline{n}-\underline{n}}{\overline{n}}}K\sqrt{\theta}\,\mu_m \\
&\quad \text{d'après ce que nous venons de démontrer.}
\end{aligned}$$

Par définition des entiers $\underline{n}$ et $\overline{n}$, lorsque $k \to \infty$, nous avons

$$\sqrt{\frac{\underline{n}}{\overline{n}}}(r+2\varepsilon)\mu_m - \sqrt{\frac{\overline{n}-\underline{n}}{\overline{n}}}K\sqrt{\theta}\,\mu_m \sim \frac{\mu_m}{(1+\theta)}\left[(r+2\varepsilon) - \sqrt{(1+\theta)^2-1}\,K\sqrt{\theta}\right].$$

Par conséquent, nous avons

$$\begin{aligned}
\sup_{t-s\leq(1+\theta)^2 a_n} \frac{\pm\alpha_{\overline{n},c}(s,t]}{\sqrt{\mathbb{E}[Y^2]a_n}} &\geq \mu_m(r+\varepsilon) \quad \text{puisque d'après (4.27) nous avons } K\theta = \frac{\varepsilon}{2}, \\
&\geq \mu_{\underline{n}}(r+\varepsilon) \text{ d'après la croissance de la suite } (\mu_m)_{m\geq 1}.
\end{aligned}$$

Ainsi, nous obtenons que

$$\bigcup_{m=\underline{n}}^{n} (A_{m,c} \cap B_{m,c}) \subset D_{\overline{n},c} = \left[\sup_{t-s\leq(1+\theta)^2 a_n} \frac{|\alpha_{\overline{n},c}(s,t]|}{\sqrt{\mathbb{E}[Y^2]a_n}} \geq \mu_{\underline{n}}(r+\varepsilon)\right].$$

Par conséquent, nous avons

$$\begin{aligned}
\mathbb{P}\left[D_{\overline{n},c}\right] &\geq \sum_{m=\underline{n}}^{n} \mathbb{P}\left[(A_{m,c}\cap B_{m,c}) \setminus \bigcup_{k=\underline{n}}^{m-1} (A_{k,c}\cap B_{k,c})\right] \\
&\geq \sum_{m=\underline{n}}^{n} \mathbb{P}\left[(A_{m,c}\cap B_{m,c}) \setminus \bigcup_{k=\underline{n}}^{m-1} A_{k,c}\right] \\
&= \sum_{m=\underline{n}}^{n} \mathbb{P}\left[\left(A_{m,c} \setminus \bigcup_{k=\underline{n}}^{m-1} A_{k,c}\right)\cap B_{m,c}\right] \\
&= \sum_{m=\underline{n}}^{n} \mathbb{P}\left[A_{m,c} \setminus \bigcup_{k=\underline{n}}^{m-1} A_{k,c}\right]\mathbb{P}\left[B_{m,c}\right] \\
&\qquad \text{par indépendance des événements,} \\
&\geq \left[\inf_{\underline{n}\leq m\leq n} \mathbb{P}\left[B_{m,c}\right]\right]\mathbb{P}\left[\bigcup_{k=\underline{n}}^{n} A_{k,c}\right] \\
&\geq \frac{1}{2}\mathbb{P}\left[\bigcup_{k=\underline{n}}^{n} A_{k,c}\right] \qquad \text{d'après (4.26) .}
\end{aligned}$$

Ce qui achève la démonstration de l'inégalité (4.23). □

4.5. Démonstration du théorème 4.3.1

Pour démontrer le théorème 4.3.1, nous procéderons en deux étapes, en montrant que, pour tout $\varepsilon > 0$
(i) la limite supérieure est presque sûrement inférieure à $1+\varepsilon$,
(ii) la limite inférieure est presque sûrement supérieure à $1-\varepsilon$.
Nous pourrons conclure en faisant tendre ε vers zéro. Nous associons, à chaque ε un ensemble presque sûr et, en considérant une suite de ε qui tend vers zéro, nous utilisons le fait qu'une réunion dénombrables d'ensembles négligeables est négligeable.

Première étape. D'abord nous allons montrer que

$$\forall\varepsilon > 0, \quad \sum_{n=1}^{+\infty} \mathbb{P}\left(\frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n\log\left(1/a_n\right)}} \geq 1+\varepsilon\right) < \infty.$$

Ensuite le premier lemme de Borel-Cantelli nous permettra de conclure que, presque sûrement, nous avons

$$\limsup_{n\to\infty} \frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n\log\left(1/a_n\right)}} < 1+\varepsilon.$$

La démonstration qui va suivre s'inspire de celle écrite dans l'article de Mason, Shorack et Wellner [95]. Soit un réel $\varepsilon > 0$ fixé. Définissons l'événement

$$A_{n,c} = \left[\frac{\omega_{n,c}(a_n)}{\sqrt{\mathbb{E}[Y^2]a_n}} \geq (r + 2\epsilon)\sqrt{2\log\left(\frac{1}{a_n}\right)} \right], \tag{4.29}$$

où le réel r sera précisé plus tard. Nous cherchons à montrer que

$$\sum_{n=1}^{\infty} \mathbb{P}\left[A_{n,c}\right] < \infty.$$

En fait ceci est une conséquence de l'inégalité maximale à condition de montrer que

$$\sum_{k=1}^{n} \mathbb{P}\left[D_{k,c}\right] < \infty, \tag{4.30}$$

où l'événement $D_{k,c}$ est défini par

$$D_{k,c} = \left[\frac{\omega_{n_{k+1},c}\left((1+d)^2 a_{n_k}\right)}{(1+d)\sqrt{\mathbb{E}[Y^2]a_{n_k}}} \geq \frac{(r+\varepsilon)}{1+d} \sqrt{2\log\left(\frac{1}{a_{n_{k-1}}}\right)} \right]$$

avec

$$n_k = \lfloor (1+d)^k \rfloor \quad \text{et} \quad d = d_\varepsilon \text{ un nombre suffisamment petit.}$$

D'après l'inégalité (4.11) nous avons

$$\begin{aligned} \mathbb{P}\left[D_{k,c}\right] &\leq \frac{1280}{(1+d)^2 a_{n_k} \delta^3} \exp\left(-(1-\delta)^3 \gamma_k \left(\frac{r+\varepsilon}{4(1+d)}\right)^2 \log\left(\frac{1}{a_{n_{k-1}}}\right)\right) \\ &\leq \mathcal{C} a_{n_{k-1}}^{[(1-\delta)^3 (r+\varepsilon)^2 \gamma_k / 16(1+d)^2] - 1} \text{ d'après } (H.1), \end{aligned} \tag{4.31}$$

où

$$\gamma_k = \psi\left(\frac{\mathcal{M}(r+\varepsilon)}{4(1+d)^2 \sqrt{\mathbb{E}[Y^2]a_{n_k} n_{k+1}}} \sqrt{2\log\left(\frac{1}{a_{n_{k-1}}}\right)} \right). \tag{4.32}$$

Supposons désormais que les hypothèses du théorème 4.3.1 sont vérifiées. Prenons $r = 1$ dans (4.29), la définition de l'événement $A_{n,c}$. D'après (4.32) et les conditions $(S.1)$ et $(S.3)$, nous remarquons que

$$\gamma_k \to 1 \quad \text{lorsque } k \to \infty,$$

puisque la fonction $\psi(\lambda)$ tend vers 1 lorsque λ tend vers 0. Ainsi, d'après (4.31) nous avons pour $\delta = \delta_\varepsilon$ suffisamment petit et $K = K_{\varepsilon,\delta}$ suffisamment grand

$$\mathbb{P}\left[D_{k+1}\right] \leq \mathcal{C} a_{n_k}^{\varepsilon} \quad \text{pour } k \geq K. \tag{4.33}$$

La pire situation pour la convergence du membre droit de l'inégalité (4.33) est lorsque les termes de la suite $(a_n)_{n\geq 1}$ sont aussi grands que possible. Or d'après l'hypothèse $(S.2)$, nous savons que pour k suffisamment grand nous avons

$$\begin{aligned} a_{n_k} \quad &\leq \quad \frac{1}{(\log n_k)^C} \quad \text{pour toute constante } C \in (0,\infty), \\ &\sim \frac{1}{(k\log(1+d))^C} \\ &\sim \frac{1}{(k\log(1+d))^{2/\varepsilon}} \quad \text{en spécifiant maintenant la constante } C = 2/\varepsilon. \end{aligned} \tag{4.34}$$

Ainsi les inégalités (4.33) et (4.34) montrent que

$$\mathbb{P}\left[D_{k,c}\right] \leq \frac{\mathcal{K}_\varepsilon}{k^2}, \quad \text{où } \mathcal{K}_\varepsilon \text{ est une constante strictement positive.}$$

Ainsi l'inégalité (4.30) est vérifiée puisque

$$\sum_{k=1}^{\infty} \frac{\mathcal{K}_\varepsilon}{k^2} < \infty.$$

Nous en déduisons par une application du premier lemme de Borel-Cantelli, sous les hypothèses du théorème 4.3.1, que nous avons

$$\limsup_{n\to\infty} \frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} \leq 1 \text{ p.s.}$$

Deuxième étape. Maintenant nous allons montrer que

$$\forall \varepsilon > 0, \quad \sum_{n=1}^{\infty} \mathbb{P}\left[\frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} \leq 1-\varepsilon\right] < \infty.$$

Le premier lemme de Borel-Cantelli nous permettra de conlure que, presque sûrement, nous avons

$$\liminf_{n\to\infty} \frac{\omega_{n,c}(a_n)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} > 1-\varepsilon.$$

Nous commençons la démonstration en introduisant une suite de points t_i tels que

$$t_i = (i-1)a_n, \quad \text{pour } 1 \leq i \leq M_n,$$

où

$$M_n = \left\lfloor \frac{1}{a_n} \right\rfloor. \tag{4.35}$$

Nous définissons également

$$B^{\pm}_{n,c} = \left[\max_{1 \le i \le M_n} \frac{\pm \alpha_{n,c}((i-1)a_n, ia_n]}{\sqrt{a_n}} \le (1-\varepsilon) r \sqrt{2\mathbb{E}[Y^2] \log\left(\frac{1}{a_n}\right)} \right].$$

Rappelons le théorème 3.5.2 du chapitre précédent qui va nous être très utile pour continuer la démonstration du théorème 4.3.1.

Théorème 4.5.1. *Soit* $\{\Pi(nt) : 0 \le t \le 1\}$ *un processus de Poisson homogène, continu à droite et de paramètre* 1. *Soit* $\{\Pi_{n,c}(t) : 0 \le t \le 1\}$ *un processus de Poisson composé. Alors, nous avons l'égalité en distribution suivante*

$$\{n\mathbb{U}_{n,c}(t) : 0 \le t \le 1\} \stackrel{d}{=} \{\Pi_{n,c}(t) : 0 \le t \le 1 | \Pi(n) = n\},$$

où la fonction $\mathbb{U}_{n,c}(t)$ *est la fonction empirique composée définie en* (4.8).

Pour des facilités d'écriture, posons

$$\Pi'_{n,c}(t) = \Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)] \quad \text{pour } 0 \le t \le 1.$$

Donc, d'après le théorème 4.5.1 nous en déduisons que

$$\begin{aligned} \mathbb{P}[B^{\pm}_{n,c}] &= \mathbb{P}\left[\max_{1 \le i \le M_n} \frac{\pm \Pi'_{n,c}((i-1)a_n, ia_n]}{\sqrt{na_n}} \le (1-\varepsilon) r \sqrt{2\mathbb{E}[Y^2] \log\left(\frac{1}{a_n}\right)} \middle| \Pi(n) = n \right] \\ &= \frac{\mathbb{P}\left[\frac{\pm \Pi'_{n,c}(a_n)}{\sqrt{na_n}} \le (1-\varepsilon) r \sqrt{2\mathbb{E}[Y^2] \log(1/a_n)} \right]^{M_n}}{\mathbb{P}[\Pi(n) = n]}, \end{aligned}$$

en utilisant l'inégalité classique

$$\mathbb{P}(E|F) \le \frac{\mathbb{P}(E)}{\mathbb{P}(F)}$$

et le fait que le processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \le t \le 1\}$ a des accroissements indépendants (cette propriété a été démontrée dans le théorème 3.2.1 du chapitre précédent).

Nous rappelons que pour un processus de Poisson de paramètre 1, nous avons

$$\mathbb{P}[\Pi(n) = n] = \left(\frac{n^n}{n!}\right) \exp(-n).$$

En utilisant la formule de Stirling, nous en déduisons que

$$\mathbb{P}[\Pi(n) = n] \sim \frac{1}{\sqrt{2\pi n}} \quad \text{lorsque } n \to \infty.$$

Et par conséquent nous pouvons affirmer que

$$\mathbb{P}[\Pi(n) = n] \geq \frac{1}{3\sqrt{n}}.$$

Rappelons ici le théorème 3.3.2 du chapitre précédent qui va nous permettre de continuer la démonstration.

Théorème 4.5.2. *Soit un réel $b > 0$. Soit $\Pi_c(b)$ une variable de Poisson composée de paramètre $b > 0$. Supposons que les variables Y_i vérifient $|Y_i - \mathbb{E}[Y_i]| \leq \mathcal{M}$ pour $i \geq 1$. Alors, pour tout réel $\mu \geq 0$, nous avons*

$$\mathbb{P}\left[|\Pi_c(b) - \mathbb{E}[\Pi_c(b)]| \geq \mu\right] \leq 2\exp\left(-\frac{\mu^2}{2b\mathbb{E}[Y^2]}\psi\left(\frac{\mathcal{M}\mu}{b\mathbb{E}[Y^2]}\right)\right),$$

où la fonction ψ est définie en (4.12).

En utilisant le théorème 4.5.2, nous en déduisons, pour n suffisamment grand que

$$\begin{aligned}
\mathbb{P}[B_{n,c}^{\pm}] &\leq 3\sqrt{n}\Bigg[1 - \exp\Bigg(-(1-\varepsilon)^2 r^2 \log(1/a_n) \\
&\qquad \times\psi\left(\frac{\mathcal{M}(1-\varepsilon)r\sqrt{2\log(1/a_n)}}{\sqrt{\mathbb{E}[Y^2]a_n n}}\right)\Bigg)\Bigg]^{M_n} \\
&\leq 3\sqrt{n}\left[1 - \exp\left(-(1-\varepsilon)^2 r^2 \log(1/a_n)\right)\right]^{M_n},
\end{aligned}$$

d'après l'hypothèse (H.3) et le fait que la fonction $\psi(\lambda)$ tend vers 1 lorsque λ tend vers 0 (propriété établie dans la proposition 3.1 du chapitre précédent). Ainsi, nous obtenons successivement que

$$\begin{aligned}
\mathbb{P}[B_{n,c}^{\pm}] &\leq 3\sqrt{n}\left(1 - a_n^{(1-\varepsilon)^2 r^2}\right)^{M_n} \\
&\leq 3\sqrt{n}\exp\left(-M_n a_n^{(1-\varepsilon)^2 r^2}\right) \\
&\sim 3\sqrt{n}\exp\left(-a_n^{(1-\varepsilon)^2 r^2 - 1}\right) \quad \text{d'après (4.35).}
\end{aligned}$$

Ainsi en prenant $r = 1$, nous avons

$$\begin{aligned}
\mathbb{P}[B_{n,c}^{\pm}] &\leq \frac{3\sqrt{n}}{\exp(a_n^{-\varepsilon})} \\
&= \frac{3\sqrt{n}}{\exp((\log n)^{\varepsilon C_n})} \quad \text{où } C_n \to \infty \text{ d'après } (H.2).
\end{aligned}$$

Or nous savons que

$$\sum_{n=1}^{\infty} \frac{3\sqrt{n}}{\exp((\log n)^{\varepsilon C_n})} < \infty.$$

Ainsi le lemme de Borel-Cantelli implique que

$$\mathbb{P}[B^{\pm}_{n,c}\text{i.s.}] = 0,$$

où l'abréviation "i.s." signifie infiniment souvent. Par conséquent, sous les hypothèses du théorème 4.3.1, nous avons

$$\liminf_{n\to\infty} \max_{1\leq i\leq M_n} \frac{\pm\alpha_{n,c}((i-1)a_n, ia_n]}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} \geq 1 \quad \text{p.s.}$$

où

$$M_n = \left\lfloor \frac{1}{a_n} \right\rfloor.$$

Sous les hypothèses du théorème 4.3.1, nous avons également

$$\liminf_{n\to\infty} \sup_{|C|=a_n} \frac{\pm\alpha_{n,c}(C)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} \geq 1 \quad \text{p.s.}$$

Ceci complète la deuxième étape et achève la démonstration du théorème 4.3.1.□

Remarque 4.5.1. Notons que la démonstration de la borne inférieure du théorème 4.3.1 peut être modifiée pour montrer, sous les hypothèses du théorème 4.3.1, que

$$\liminf_{n\to\infty} \sup_{|C|=a_n, C\subset[C_0,D_0]} \frac{\pm\alpha_{n,c}(C)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} \geq 1 \quad \text{p.s.} \quad \text{pour chaque } 0 \leq C_0 \leq D_0 \leq 1;$$

il suffit simplement de modifier la constante M_n en remplaçant $\left\lfloor \frac{1}{a_n} \right\rfloor$ par $\left\lfloor \frac{D_0 - C_0}{a_n} \right\rfloor$.

Chapitre 5

Lois fonctionnelles du L.I. pour les accroissements du processus $\alpha_{n,c}$

5.1. Introduction

Dans ce paragraphe, nous allons rappeler brièvement les lois fonctionnelles du logarithme itéré pour le processus empirique uniforme de queue et pour les accroissements d'ordre $a_n > 0$ du processus empirique uniforme. Les limites de ces lois fonctionnelles dépendent principalement, comme les limites des oscillations du processus empirique, des conditions que vérifie la suite $(a_n)_{n\geq 1}$. A cet effet, nous allons rappeler des résultats limites connus sur le comportement des oscillations du processus empirique uniforme.

Une suite de constantes $(a_n)_{n\geq 1}$ satisfait les conditions de Csörgő-Révész-Stute [CRS] (voir par exemple Csörgő-Révész [27], Stute [131]) si elle vérifie les trois conditions suivantes

(S.1) (i) $0 < a_n < 1$ pour $n \geq 1$, (ii) $a_n \downarrow 0$ et (iii) $na_n \uparrow \infty$ lorsque $n \to \infty$;
(S.2) $\log(1/a_n)/\log\log n \to \infty$ lorsque $n \to \infty$;
(S.3) $na_n/\log n \to \infty$ lorsque $n \to \infty$.

Introduisons également les trois conditions suivantes qui nous seront utiles par la suite. Dans ce qui suit, $\mathcal{C}$ désigne une constante.

(S.4) $na_n/\log\log n \to \infty$ lorsque $n \to \infty$;
(S.5) $na_n/\log\log n \to \mathcal{C} \in (0,\infty)$ lorsque $n \to \infty$;
(S.6) $na_n/\log n \to \mathcal{C} \in (0,\infty)$ lorsque $n \to \infty$.

Sous les conditions [CRS] (S.1–2–3), Stute [131] a montré que

$$\lim_{n\to\infty} \sup_{0\leq t\leq 1-a_n} \frac{\pm\left(\alpha_n(t+a_n)-\alpha_n(t)\right)}{\sqrt{2a_n\log(1/a_n)}} = 1 \quad \text{p.s.}$$

Dans le cas borné c'est-à-dire lorsque $na_n/\log n \to \mathcal{C} \in (0,\infty)$, Mason, Shorack et Wellner [95] ont prouvé que

$$\lim_{n\to\infty} \sup_{0\leq t\leq 1-a_n} \frac{\pm\left(\alpha_n(t+a_n)-\alpha_n(t)\right)}{\sqrt{2a_n\log(1/a_n)}} = \pm\sqrt{\frac{\mathcal{C}}{2}}\left(\delta^{\pm}_{\mathcal{C}}-1\right) \quad \text{p.s.,}$$

où $0 \leq \delta_{\mathcal{C}}^- < 1 < \delta_{\mathcal{C}}^+ < \infty$ sont les racines de l'équation $h(\delta) = 1/\mathcal{C}$, avec la convention que $\delta_{\mathcal{C}}^- = 0$ pour $0 < \mathcal{C} < 1$ et où la fonction $h(\cdot)$ est définie par

$$h(u) = \begin{cases} u \log u - u + 1 & \text{pour } u > 0, \\ 1 & \text{pour } u = 0, \\ \infty & \text{pour } u < 0. \end{cases} \tag{5.1}$$

Sur cet exemple, nous notons que les limites varient en fonction des conditions que vérifie la suite $(a_n)_{n\geq 1}$ et plus particulièrement en fonction de la limite de $na_n/\log n$.

Afin de rappeler les lois fonctionnelles du logarithme itéré valables dans le cas qui nous intéresse, nous avons besoin d'introduire quelques notations.

Notons $I_{RC}(0,1)$ l'ensemble des fonctions de répartition continues à droite, positives, de mesures de Radon bornées de support inclus dans l'intervalle $[0,1]$. Pour toute fonction $f \in I_{RC}(0,1)$ et pour tout réel t, posons

$$f(t\pm) = \lim_{\varepsilon \downarrow 0} f(t \pm \varepsilon).$$

De plus, pour toute fonction $f \in I_{RC}(0,1)$ et pour tout réel t, posons

$$f(t) = \int_0^t f'(s)\, ds + f_S(t),$$

où la fonction $f_S \in I_{RC}(0,1)$ est la fonction de répartition de la composante singulière dans la décomposition de Lebesgue de df et la fonction f' est la dérivée au sens de Lebesgue de la partie absolument continue de cette décomposition.

Nous noterons $\mathbb{S}_0$ l'ensemble de Strassen (voir par exemple l'article de Strassen [130]) qui est composé des fonctions f absolument continues sur l'intervalle $[0,1]$ et telles que

$$f(0) = 0 \quad \text{et} \quad \int_0^1 \left[f'(s)\right]^2 ds \leq 1,$$

où la fonction f' désigne la dérivée au sens de Lebesgue de la fonction f.

Notons $B(0,1)$ l'espace des fonctions bornées définies sur l'intervalle $[0,1]$ et à valeurs réelles. La norme de la convergence uniforme d'une fonction $f \in B(0,1)$ est notée par

$$\|f\| = \sup_{0\leq t\leq 1} |f(t)|.$$

Pour tout $\varepsilon > 0$ et tout ensemble $C \subseteq B(0,1)$, on note C^ε l'ensemble des fonctions $f \in B(0,1)$ telles qu'il existe une fonction $g \in C$ vérifiant $\|f - g\| < \varepsilon$.

Enfin, nous allons rappeler la définition du processus empirique uniforme de queue. Ce dernier (voir par exemple l'article de Mason [94]) est défini pour tout entier $n \geq 1$ par

$$\frac{\alpha_n\left(a_n t\right)}{\sqrt{a_n}}, \quad \text{pour} \quad 0 \leq t \leq 1,$$

où $\{\alpha_n(t) : 0 \leq t \leq 1\}$ est le processus empirique uniforme, et $(a_n)_{n\geq 1}$ est une suite de constantes positives qui vérifie $a_n \to 0$ et $na_n \to \infty$ lorsque $n \to \infty$.

Nous pouvons maintenant énoncer les théorèmes qui établissent des lois fonctionnelles du logarithme itéré pour le processus empirique uniforme de queue. Le premier théorème est dû à Mason en 1988.

Théorème 5.1.1. *Sous les hypothèses* $(S.1)$ *et* $(S.4)$*, la suite de fonctions*

$$\left\{ \frac{\alpha_n(a_n t)}{\sqrt{2a_n \log\log n}} : 0 \leq t \leq 1 \right\} \tag{5.2}$$

est presque sûrement relativement compacte dans $B(0,1)$ *muni de la topologie de la convergence uniforme. De plus, l'ensemble des points limites de cette suite est égal à l'ensemble de Strassen* $\mathbb{S}_0$.

Démonstration. Pour la démonstration de ce théorème, on peut consulter l'article de Mason [94]. □

Remarque 5.1.1. Cette loi fonctionnelle du logarithme itéré pour le processus empirique uniforme de queue permet de donner une nouvelle démonstration de la loi limite suivante, déjà établie par Kiefer [82].
Sous les hypothèses $(S.1)$ et $(S.4)$, nous avons

$$\limsup_{n\to\infty} \frac{\pm\alpha_n(a_n)}{\sqrt{2a_n \log\log n}} = 1 \quad \text{p.s.}$$

Remarque 5.1.2. Ce théorème peut être également appliqué à l'étude du comportement asymptotique presque sûr de sommes de valeurs extrêmes. Pour cette étude, nous renvoyons à l'article de Deheuvels et Mason [32].

Le second théorème est dû à Deheuvels et Mason en 1990.

Théorème 5.1.2. *Sous les hypothèses* $(S.1)(i)$ *et* $(S.5)$*, la suite de fonctions*

$$\left\{ \frac{n\mathbb{U}_n(a_n t)}{\sqrt{\log\log n}} : 0 \leq t \leq 1 \right\} \tag{5.3}$$

est presque sûrement relativement compacte dans $I_{RC}(0,1)$ *muni de la topologie de la convergence uniforme. De plus, l'ensemble des points limites de cette suite est égal à* $\Delta_{\mathcal{C}}$*, où* $\Delta_{\mathcal{C}}$ *désigne l'ensemble des fonctions absolument continues dans* $I_{RC}(0,1)$ *et telles que*

$$\mathcal{C} \int_0^1 h\left(\frac{f'(s)}{\mathcal{C}}\right) ds \leq 1,$$

où la fonction $h(\cdot)$ *est définie en* (5.1).

Démonstration. Pour la démonstration de ce théorème, on peut consulter l'article de Deheuvels et Mason [33]. □

Remarque 5.1.3. Cette loi fonctionnelle non-standard du logarithme itéré pour le processus empirique de queue peut être utilisée pour obtenir des lois limites comme

$$\limsup_{n\to\infty} \frac{\pm\alpha_n(a_n)}{\sqrt{2a_n \log\log n}} = \pm\sqrt{\frac{\mathcal{C}}{2}}\left(\delta_{\mathcal{C}}^{\pm} - 1\right) \quad \text{p.s.},$$

sous l'hypothèse $(S.5)$. Cette loi limite a d'abord été établie par Kiefer [82]. Pour d'autres utilisations de cette loi fonctionnelle du logarithme itéré, nous renvoyons à l'article de Deheuvels et Mason [34].

Ces lois fonctionnelles du logarithme itéré existent également pour les accroissements η_n de la fonction de répartition empirique uniforme $\mathbb{U}_n$ ou pour les accroissements ξ_n du processus empirique uniforme α_n définis respectivement par

$$\eta_n\left(a_n, t; s\right) = \frac{n}{\log n}\left(\mathbb{U}_n(t + a_n s) - \mathbb{U}_n(t)\right),$$

pour $0 \leq s \leq 1$ et $0 \leq t \leq 1 - a_n$, et

$$\xi_n(a_n, t; s) = \alpha_n(t + a_n s) - \alpha_n(t),$$

pour $0 \leq s \leq 1$ et $0 \leq t \leq 1$.

Théorème 5.1.3. *Sous l'hypothèse* $(S.6)$*, pour tout* $\varepsilon > 0$*, il existe presque sûrement un entier* N'_ε *tel que, pour tout entier* $n \geq N'_\varepsilon$*, nous ayons*

$$\{\eta_n(a_n, t; \cdot) : 0 \leq t \leq 1 - a_n\} \subset \widetilde{\Delta}_{\mathcal{C}}^{\varepsilon}.$$

De plus, pour toute fonction $f \in \Delta_{\mathcal{C}}$ *et pour tout* $\varepsilon > 0$*, il existe presque sûrement un entier* $N''_{\varepsilon,f}$ *tel que, pour tout entier* $n \geq N''_\varepsilon$*, il existe un réel* $\tilde{t}$*, avec* $0 \leq \tilde{t} \leq 1 - a_n$*, tel que nous ayons*

$$||\eta_n(a_n, \tilde{t}; \cdot) - f|| < \varepsilon.$$

Démonstration. La démonstration de ce théorème se trouve dans l'article de Deheuvels et Mason [35]. □

Enfin, le dernier théorème établit une loi fonctionnelle du logarithme itéré pour les accroissements du processus empirique uniforme.

Théorème 5.1.4. *Sous les conditions [CRS]* $(S.1$–2–$3)$*, pour tout* $\varepsilon > 0$*, il existe presque sûrement un entier* n'_ε *tel que, pour tout entier* $n \geq n'_\varepsilon$*, nous ayons*

$$\left\{\frac{\xi_n(a_n, t; \cdot)}{\sqrt{2a_n \log(1/a_n)}} : 0 \leq t \leq 1 - a_n\right\} \subset \mathbb{S}_0^{\varepsilon}. \tag{5.4}$$

De plus, pour toute fonction $f \in \mathbb{S}_0$ *et pour tout* $\varepsilon > 0$*, il existe presque sûrement un entier* $n''_{\varepsilon,f}$ *tel que, pour tout entier* $n \geq n''_{\varepsilon,f}$*, il existe un réel* t*, avec* $0 \leq t = t_{n,\varepsilon,f} \leq 1 - a_n$*, tel que nous ayons*

$$\left\|\frac{\xi_n(a_n, t; \cdot)}{\sqrt{2a_n \log(1/a_n)}} - f\right\| < \varepsilon. \tag{5.5}$$

Démonstration. La démonstration de ce théorème se trouve également dans l'article de Deheuvels et Mason [35]. □

Remarque 5.1.4. Cette loi fonctionnelle du logarithme itéré pour les accroissements du processus empirique uniforme permet d'établir la loi du logarithme itéré pour les estimateurs non paramétriques comme l'estimateur à noyau pour la densité, résultat déjà mentionné dans le chapitre précédent ou encore pour les estimateurs des k-plus proches voisins. Énonçons ce dernier résultat.

Soit $\{X_i : i \geq 1\}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées de fonction de répartition $F(t) = \mathbb{P}[X_1 \leq t]$ et de densité $f(t) = F'(t)$ supposées continues et positives sur un intervalle compact $[A, B]$ (où $A < B$). Soit $\{\lambda_n : n \geq 1\}$ une suite de réels telle que pour tout entier $n \geq 1$, $0 < \lambda_n < 1$. L'estimateur de densité par la méthode des k points les plus proches, introduit par Fix et Hodges [56] et étudié par la suite par Loftsgaarden et Quesenberry [90], estime la densité f associée à la suite $\{X_i : i \geq 1\}$ par

$$\widehat{f_n}(t) = \lambda_n / \inf\left\{a > 0 : F_n\left(t + \frac{a}{2}\right) - F_n\left(t - \frac{a}{2}\right) \geq \lambda_n\right\}. \tag{5.6}$$

Notons, pour $n \geq 1$, M_n la fonction définie par

$$M_n(t) = \lambda_n / \inf\left\{a > 0 : F\left(t + \frac{a}{2}\right) - F\left(t - \frac{a}{2}\right) \geq \lambda_n\right\}. \tag{5.7}$$

Nous énonçons une loi du logarithme itéré pour l'estimateur $\widehat{f_n}$ défini en (5.6) de la densité.

Théorème 5.1.5. *Soit $\{a_n : n \geq 1\}$ une suite de constantes qui vérifie les conditions [CRS] (S.1–2–3). Alors, pour tous réels C et D tels que $A < C < D < B$, nous avons*

$$\lim_{n\to\infty} \pm\sqrt{\frac{na_n}{2\log(1/a_n)}} \sup_{C \leq t \leq D} \left(\frac{\widehat{f_n}(t) - M_n(t)}{f(t)}\right) = 1 \quad p.s. \tag{5.8}$$

Démonstration. La démonstration de ce théorème se trouve dans l'article de Deheuvels et Mason [35]. □

5.2. Résultat principal

Nous allons maintenant établir une version de la loi fonctionnelle du logarithme itéré pour les accroissements $\xi_{n,c}$ du processus empirique composé $\alpha_{n,c}$ définis par

$$\xi_{n,c}(a, t; s) = \alpha_{n,c}(t + as) - \alpha_{n,c}(t), \tag{5.9}$$

où $0 < a < 1$, $0 \leq t \leq 1$ et $0 \leq s \leq 1$.

Théorème 5.2.1. *Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $[0,1]$. Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. Nous supposons que $\mathbb{E}[Y] = 1$, $\mathrm{Var}[Y] < \infty$ et que la variable aléatoire Y vérifie $|Y - 1| \leq \mathcal{M}$ p.s. pour une constante $\mathcal{M}$ convenable. De plus, les suites de variables $\{U_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont supposées indépendantes. Alors, sous les conditions [CRS] (S.1–2–3), pour tout $\varepsilon > 0$, il existe presque sûrement un entier n'_ε tel que, pour tout entier $n \geq n'_\varepsilon$, nous ayons*

$$\left\{ \frac{\xi_{n,c}(a_n, t; \cdot)}{\sqrt{2\mathbb{E}[Y^2] a_n \log(1/a_n)}} : 0 \leq t \leq 1 - a_n \right\} \subset \mathbb{S}_0^\varepsilon. \tag{5.10}$$

De plus, pour toute fonction $f \in \mathbb{S}_0$ et pour tout $\varepsilon > 0$, il existe presque sûrement un entier $n''_{\varepsilon,f}$ tel que, pour tout entier $n \geq n''_{\varepsilon,f}$, il existe un réel t, avec $0 \leq t = t_{n,\varepsilon,f} \leq 1 - a_n$, tel que nous ayons

$$\left\| \frac{\xi_{n,c}(a_n, t; \cdot)}{\sqrt{2\mathbb{E}[Y^2] a_n \log(1/a_n)}} - f \right\| < \varepsilon. \tag{5.11}$$

5.3. Démonstration du théorème 5.2.1

Sauf mention du contraire, nous supposons que les conditions [CRS] (S.1–2–3.) sont vérifiées.

Pour tout réel $a \in (0,1)$ posons

$$L_{n,c}(a, t, s) = \frac{[\Pi_{n,c}(t + as) - \Pi_{n,c}(t) - nas]}{\sqrt{n}}, \tag{5.12}$$

pour $0 \leq t \leq 1$ et $0 \leq s \leq 1$ et où $\{\Pi_{n,c}(t) : t \geq 0\}$ désigne un processus de Poisson composé. Pour une définition de ce processus, nous renvoyons au chapitre 3.

Lemme 5.3.1. *Il existe une constante $\mathcal{C}_0$ (dépendant de la suite $(a_n)_{n\geq 1}$ seulement) avec $0 < \mathcal{C}_0 < \infty$, et telle qu'on ait la propriété suivante. Pour tout choix de $\{t_1, \cdots, t_m\} \subseteq \{ka_n : 0 \leq k \leq a_n^{-1} - 1\}$ avec $0 < ma_n \leq \frac{1}{2}$ et pour tout choix de sous-ensembles boréliens $B_1, \ldots, B_m$ de $B(0,1)$ muni de la topologie de la convergence uniforme sur l'intervalle $[0,1]$, si*

$$A_1 = \{\xi_{n,c}(a_n, t_i; .) \in B_i, i = 1, \ldots, m\}$$

et

$$A_2 = \{L_{n,c}(a_n, t_i; .) \in B_i, i = 1, \ldots, m\},$$

alors

$$\mathbb{P}[A_1] \leq \mathcal{C}_0 \, \mathbb{P}[A_2].$$

Démonstration. Posons
$$I = \bigcup_{i=1}^{m} (t_i, t_i + a_n].$$
Notons que la mesure de Lebesgue de l'ensemble I, notée $|I|$, est égale à $ma_n \leq \frac{1}{2}$. Posons
$$\overline{I} = [0,1] - I.$$
D'après le théorème 3.5.2 du chapitre 3, le processus $\{n\mathbb{U}_{n,c}(t) : 0 \leq t \leq 1\}$ est égal en distribution au processus de Poisson composé $\{\Pi_{n,c}(t) : 0 \leq t \leq 1\}$ sachant que $\Pi(n) = n$. Par conséquent, nous avons
$$\begin{aligned}
\mathbb{P}[A_1] &= \mathbb{P}[A_2 | \Pi(n) = n] \\
&= \frac{\mathbb{P}[A_2 \cap \Pi(n) = n]}{\mathbb{P}[\Pi(n) = n]} \\
&= \sum_{j=0}^{n} \frac{\mathbb{P}\left[A_2 \cap \{\Pi(nI) = j\} \cap \{\Pi(n\overline{I}) = n - j\}\right]}{\mathbb{P}[\Pi(n) = n]}.
\end{aligned}$$
Comme les événements $\{A_2 \cap \{\Pi(nI) = j\}\}$ et $\{\Pi(n\overline{I}) = n - j\}$ sont indépendants, alors nous obtenons que
$$\begin{aligned}
\mathbb{P}[A_1] &= \sum_{j=0}^{n} \mathbb{P}[A_2 \cap \{\Pi(nI) = j\}] \times \frac{\mathbb{P}\left[\Pi(n\overline{I}) = n - j\right]}{\mathbb{P}[\Pi(n) = n]} \\
&\leq \frac{\mathbb{P}[A_2]}{\mathbb{P}[\Pi(n) = n]} \max_{0 \leq j \leq n} \mathbb{P}\left[\Pi(n\overline{I}) = n - j\right] \\
&\leq \frac{\mathbb{P}[A_2]}{\mathbb{P}[\Pi(n) = n]} \mathbb{P}\left[\Pi(n\overline{I}) = \lfloor n|\overline{I}| \rfloor\right],
\end{aligned}$$
où $\lfloor u \rfloor \leq u \leq \lfloor u \rfloor + 1$ représente la partie entière de u. Nous avons utilisé ici le fait que la variable aléatoire $\Pi(n\overline{I})$ suit une loi de Poisson de paramètre $n\overline{I}$ et que (voir, e.g., le livre de Johnson et Kotz [76], p.92) $\mathbb{P}[\Pi(n\Lambda) = j] \leq \mathbb{P}[\Pi(n\Lambda) = \lfloor n|\Lambda| \rfloor]$, pour tout choix de j et de $\Lambda \subset [0, \infty)$. En posant
$$\frac{\mathbb{P}\left[\Pi(n\overline{I}) = \lfloor n|\overline{I}| \rfloor\right]}{\mathbb{P}[\Pi(n) = n]} = \mathcal{C}_0,$$
nous obtenons le résultat annoncé. De plus, en utilisant la formule de Stirling ($n! = (n/e)^n (2\pi n)^{1/2} \exp(\theta_n/n)$, où $0 < \theta_n < 1/12$ pour $n \geq 1$), nous obtenons que
$$0 < \frac{\mathbb{P}\left[\Pi(n\overline{I}) = \lfloor n|\overline{I}| \rfloor\right]}{\mathbb{P}[\Pi(n) = n]} \leq \frac{n^{1/2} e^{1/2}}{(\frac{1}{2}n - 1)^{1/2}} \leq 2, \quad \text{pour } n \geq 5.$$
Ce qui permet de démontrer que $0 < \mathcal{C}_0 < \infty$ et de conclure. □

Le lemme suivant (connu sous le nom de *théorème de Schilder* ; voir par exemple l'article de Schilder [119]) donne un résultat de grandes déviations qui est un outil dont nous aurons besoin par la suite. Pour toute fonction $f \in B(0,1)$, posons

$$J(f) = \begin{cases} \displaystyle\int_0^1 [f'(s)]^2 \, ds, & \text{si la fonction } f \text{ est absolument continue sur } [0,1], \\ & \text{où } f' \text{ est la dérivée au sens de Lebesgue,} \\ \infty, & \text{sinon,} \end{cases}$$

et, pour tout ensemble $B \subset B(0,1)$, posons

$$J(B) = \inf_{f \in B} J(f).$$

Lemme 5.3.2. *Soit* $\{W(t) : t \geq 0\}$ *un processus de Wiener standard. Pour tout réel* $\delta > 0$ *et pour tout réel* $t \in [0,1]$, *posons*

$$W_\delta(t) = 2^{-1/2}\delta^{-1}W(\gamma t). \tag{5.13}$$

Nous avons alors la propriété suivante.
(i) *Pour tout sous-ensemble fermé* F *de* $B(0,1)$ *muni de la topologie de la convergence uniforme sur l'intervalle* $[0,1]$,

$$\limsup_{\delta\to\infty} \delta^{-1} \log \mathbb{P}\left[W_\delta \in F\right] \leq -J(F). \tag{5.14}$$

(ii) *Pour tout sous-ensemble ouvert* G *de* $B(0,1)$ *muni de la topologie de la convergence uniforme sur l'intervalle* $[0,1]$,

$$\liminf_{\delta\to\infty} \delta^{-1} \log \mathbb{P}\left[W_\delta \in G\right] \geq -J(G). \tag{5.15}$$

Démonstration. Pour la démonstration du lemme 5.3.2, nous renvoyons à l'article de Ventsel [134]. □

Le lemme suivant démontre la deuxième partie [*i.e.* l'inégalité (5.11)] du théorème 5.2.1 pour le processus $\xi_{n,c}$.

Lemme 5.3.3. *Sous les conditions* $[CRS]$ $(S.1–2–3)$, *pour toute fonction* $f \in \mathbb{S}_0$ *et pour tout* $\varepsilon > 0$, *il existe presque sûrement un entier* $n''_{\varepsilon,f}$ *tel que, pour tout entier* $n \geq n''_{\varepsilon,f}$, *il existe un réel* $t = t_{n,\varepsilon,f} \in [0, 1 - a_n]$ *tel que nous ayons*

$$\left\| \frac{\xi_{n,c}(a_n, t; .)}{\sqrt{2\mathbb{E}[Y^2]a_n \log(1/a_n)}} - f \right\| < \varepsilon.$$

Démonstration. La démonstration reprend certaines composantes de l'article de Deheuvels et Mason [35] en les adaptant au contexte élargi de nos hypothèses. La démonstration de ce lemme se fait en plusieurs étapes.
Première étape. Posons, pour $\varepsilon > 0$ et $f \in B(0,1)$,

$$N_\varepsilon(f) = \{g \in B(0,1) : \|f-g\| < \varepsilon\}.$$

D'après le lemme 5.3.1 appliqué à la suite de points $t_i = ia_n, i = 1, \cdots, m_n = \lfloor 1/(2a_n) \rfloor$ et $B_i = B(0,1) - N_\varepsilon(f)$, pour $i = 1, \cdots, m_n$, nous obtenons que

$$\begin{aligned}
\mathbb{P}_n &= \mathbb{P}\left[\bigcap_{i=1}^{m_n}\left(\frac{\xi_{n,c}(a_n,t_i;.)}{\sqrt{2\mathbb{E}[Y^2]a_n\log(1/a_n)}} \notin N_\varepsilon(f)\right)\right] \\
&\le \mathcal{C}_0\mathbb{P}\left[\bigcap_{i=1}^{m_n}\left(\frac{L_{n,c}(a_n,t_i;.)}{\sqrt{2\mathbb{E}[Y^2]a_n\log(1/a_n)}} \notin N_\varepsilon(f)\right)\right] \\
&= \mathcal{C}_0\left(1-\mathbb{P}\left[\frac{L_{n,c}(a_n,0;.)}{\sqrt{2\mathbb{E}[Y^2]a_n\log(1/a_n)}} \in N_\varepsilon(f)\right]\right)^{m_n} \\
&= \mathcal{C}_0\big(1-\mathbb{P}_{1,n}\big)^{m_n}.
\end{aligned} \tag{5.16}$$

Deuxième étape. Maintenant, nous allons évaluer $\mathbb{P}_{1,n}$. Pour cela, nous allons appliquer le principe d'invariance de Komlós, Major et Tusnády [83] au processus de Poisson composé. Ceci nous permet de construire sur un même espace de probabilité un processus de Poisson composé $\{\Pi_c(t) : t \ge 0\}$ et un processus de Wiener standard $\{W(t) : t \ge 0\}$. Nous avons alors le théorème suivant.

Théorème 5.3.1. *Soit $\{\Pi_c(t) : t \ge 0\}$ un processus de Poisson composé défini par*

$$\Pi_c(t) = \sum_{i=1}^{\Pi(t)} Y_i,$$

où $\{Y_i : i \ge 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que $Y = Y_1$ et $\{\Pi(t) : t \ge 0\}$ un processus de Poisson homogène, continu à droite et de paramètre $\lambda > 0$. La suite $\{Y_i : i \ge 1\}$ et le processus $\{\Pi(t) : t \ge 0\}$ sont supposés indépendants. Supposons de plus qu'il existe un t_0 tel que

$$\mathbb{E}[\exp(tY)] < \infty \text{ pour } |t| \le t_0.$$

Alors sur un espace de probabilité éventuellement élargi, il existe un processus de Wiener $\{W(t) : t \ge 0\}$ tel que, pour tout $T \ge 1$,

$$\mathbb{P}\left[\sup_{0\le t\le T}\left|\frac{\Pi_c(t)-\mathbb{E}[\Pi_c(t)]}{\sqrt{\mathbb{E}[Y^2]}} - W(t)\right| \ge C_1\log T + y\right] \le C_2\exp(-C_3 y), \tag{5.17}$$

pour tout réel $y > 0$, où C_1, C_2 et C_3 sont des constantes strictement positives.

Démonstration. Pour la démonstration de ce théorème, nous renvoyons au théorème 3.4.3 du chapitre 3. □

D'après la définition (5.12) et les inégalités (5.16) et (5.17), nous obtenons que

$$\begin{aligned}\mathbb{P}_{1,n} &= \mathbb{P}\left[\frac{L_{n,c}(a_n,0;\cdot)}{\sqrt{2\mathbb{E}[Y^2]a_n\log(1/a_n)}}\in N_\varepsilon(f)\right]\\ &= \mathbb{P}\left[\frac{\Pi_c(na_n\cdot)-na_n\cdot}{\sqrt{2\mathbb{E}[Y^2]na_n\log(1/a_n)}}\in N_\varepsilon(f)\right],\end{aligned}$$

avec $\Pi_c(0)=0$. Ainsi, nous avons

$$\begin{aligned}\mathbb{P}_{1,n} &\geq \mathbb{P}\left[\frac{W(na_n\cdot)}{\sqrt{2na_n\log(1/a_n)}}\in N_{\varepsilon/2}(f)\right] \\ &\quad -\mathbb{P}\left[\sup_{0\leq s\leq 1}\frac{\left|\Pi_c(na_ns)-na_ns-\sqrt{\mathbb{E}[Y^2]}W(na_ns)\right|}{\sqrt{2\mathbb{E}[Y^2]na_n\log(1/a_n)}}\geq\varepsilon/2\right]\\ &= \mathbb{P}_{2,n}-\mathbb{P}_{3,n}^{\varepsilon/2}.\end{aligned}\tag{5.18}$$

Troisième étape. Intéressons nous maintenant à la quantité $\mathbb{P}_{2,n}$. Remarquons l'égalité en distribution

$$\frac{W(na_n\cdot)}{\sqrt{2na_n\log(1/a_n)}}\stackrel{\mathrm{d}}{=}2^{-1/2}\big(\log(1/a_n)\big)^{-1}W\big(\log(1/a_n)\cdot\big),$$

sous les conditions $(S.2-3)$. Par conséquent, nous obtenons que

$$\mathbb{P}_{2,n}=\mathbb{P}\left[W_{\log(1/a_n)}(\cdot)\in N_{\varepsilon/2}(f)\right],$$

où le processus $W_\gamma(\cdot)$ est défini en (5.13). Comme $f\in\mathbb{S}_0$, nous avons $J(N_{\varepsilon/2}(f))<1$. Alors, d'après l'inégalité (5.15), il s'ensuit que pour tout réel $\rho\in(J(N_{\varepsilon/2}(f)),1)$, nous avons

$$\mathbb{P}_{2,n}\geq\exp(-\rho\log(1/a_n))=a_n^\rho,\qquad\text{lorsque } n\to\infty.\tag{5.19}$$

Quatrième étape. Maintenant étudions la quantité $\mathbb{P}_{3,n}^{\varepsilon/2}$. La condition $(S.1)$ implique que

$$\frac{\log(na_n)}{\sqrt{2na_n\log(1/a_n)}}\to 0\qquad\text{lorsque } n\to\infty.$$

Ainsi, en appliquant l'inégalité (5.17), tenant compte du fait de nos hypothèses impliquent que, à partir d'un certain rang,

$$C_1\log(na_n)>\frac{\varepsilon}{4}\sqrt{2na_n\log(1/a_n)},$$

nous pouvons écrire que, pour tout n assez grand

$$\begin{aligned}
\mathbb{P}_{3,n}^{\varepsilon/2} &\leq \mathbb{P}\left[\sup_{0\leq t\leq na_n}\left|\frac{\Pi_c(t)-t-\sqrt{\mathbb{E}[Y^2]}W(t)}{\sqrt{\mathbb{E}[Y^2]}}\right|\right.\\
&\qquad\qquad\left.\geq C_1\log(na_n)+\frac{\varepsilon}{4}\sqrt{2na_n\log(1/a_n)}\right]\\
&\leq C_2\exp\left(-C_3\left(\frac{\varepsilon}{4}\right)\sqrt{2na_n\log(1/a_n)}\right).
\end{aligned}$$

De même, comme, à partir d'un certain rang,

$$C_3\left(\frac{\varepsilon}{4}\right)\sqrt{2na_n\log(1/a_n)}>\rho\log(1/a_n),$$

nous avons, pour tout n assez grand

$$\mathbb{P}_{3,n}^{\varepsilon/2}\leq\frac{1}{2}\exp\left(-\rho\log(1/a_n)\right)=\frac{1}{2}a_n^{\rho}. \tag{5.20}$$

Cinquième étape. D'après les inégalités (5.18), (5.19) et (5.20), lorsque $n\to\infty$, alors nous obtenons que

$$\begin{aligned}
\mathbb{P}_{1,n} &\geq \mathbb{P}_{2,n}-\mathbb{P}_{3,n}^{\varepsilon/2}\geq a_n^{\rho}-\frac{1}{2}a_n^{\rho}\\
\mathbb{P}_{1,n} &\geq \frac{1}{2}a_n^{\rho}.
\end{aligned}$$

Sixième étape. Ainsi, d'après l'inégalité (5.16), nous obtenons successivement

$$\begin{aligned}
\mathbb{P}_n &\leq \mathcal{C}_0\exp\left(-\frac{1}{2}m_na_n^{\rho}\right),\ \text{d'après ci-dessus,}\\
\mathbb{P}_n &\leq \mathcal{C}_0\exp\left(-\frac{1}{2}\left(\frac{1}{4a_n}\right)a_n^{\rho}\right),\ \text{puisque } m_n=\lfloor 1/(2a_n)\rfloor,\\
\mathbb{P}_n &\leq \mathcal{C}_0\exp\left(-\frac{1}{8}\,a_n^{\rho-1}\right).
\end{aligned} \tag{5.21}$$

Comme $\rho-1<0$ et d'après la condition $(S.2)$, le membre droit de l'inégalité (5.21) est finalement inférieur ou égal à

$$\mathcal{C}_0\exp\left(-\frac{1}{8}\left(\log n\right)^r\right)$$

pour un réel $r>1$ arbitraire. Il vient alors que $\sum_n\mathbb{P}_n<\infty$, ce qui d'après le lemme de Borel-Cantelli implique l'inégalité (5.11) du théorème 5.2.1, ce qui achève par conséquent la démonstration du lemme 5.3.3. $\square$

Maintenant démontrons la première partie [*ie* l'inclusion (5.10)] du théorème 5.2.1. Fixons un réel $\gamma > 0$ et un réel $\varepsilon > 0$. Posons

$$\nu_k = [(1+\gamma)^k],$$

pour tout entier $k \geq 1$ et

$$b_n = \sqrt{2\mathbb{E}[Y^2] a_n \log(1/a_n)}.$$

Considérons les deux événements suivants

$$C_k(\varepsilon,\gamma) = \Big\{ (n/\nu_{k+1})^{1/2}\, b_{\nu_{k+1}}^{-1} \xi_{n,c}(a_{\nu_{k+1}}, t; \cdot) \notin \mathbb{S}_0^{\varepsilon} \quad \text{pour certains } 0 \leq t \leq 1 - a_{\nu_{k+1}} \text{ et } \nu_k < n \leq \nu_{k+1} \Big\},$$

et

$$D_k(\varepsilon,\gamma) = \left\{ b_{\nu_{k+1}}^{-1} \xi_{\nu_{k+1},c}\left(a_{\nu_{k+1}}, t; \cdot\right) \notin \mathbb{S}_0^{\varepsilon} \text{ pour certains } 0 \leq t \leq 1 - a_{\nu_{k+1}} \right\}.$$

Lemme 5.3.4. *Pour tout $\varepsilon > 0$ et pour tout réel $\gamma > 0$, il existe un entier $K = K_{\varepsilon,\gamma}$ tel que pour tout entier $k \geq K$, nous ayons*

$$\mathbb{P}\left[C_k(\varepsilon,\gamma)\right] \leq \mathcal{C}_0\, \mathbb{P}\left[D_k(\varepsilon/2,\gamma)\right]. \tag{5.22}$$

Démonstration. Soit $\{r_i : i \geq 1\}$ une suite de rationnels dans l'intervalle $[0, 1 - a_{\nu_{k+1}}]$. Pour tout entier $i \geq 1$ et pour tout entier n tel que $\nu_k < n \leq \nu_{k+1}$ introduisons les trois événements suivants

$$\begin{aligned} E_{k,i,n}(\varepsilon) &= \left\{ (n/\nu_{k+1})^{1/2}\, b_{\nu_{k+1}}^{-1} \xi_{n,c}(a_{\nu_{k+1}}, r_i; \cdot) \notin \mathbb{S}_0^{\varepsilon} \right\}, \\ E_{k,n}(\varepsilon) &= \bigcup_{i \geq 1} E_{k,i,n}(\varepsilon) \\ \text{et} & \\ F_{k,i,n}(\varepsilon) &= \left\{ b_{\nu_{k+1}}^{-1} \left\| \xi_{\nu_{k+1},c}(a_{\nu_{k+1}}, r_i; \cdot) - (n/\nu_{k+1})^{1/2}\, \xi_{n,c}(a_{\nu_{k+1}}, r_i; \cdot) \right\| < \varepsilon \right\}. \end{aligned}$$

Observons que pour tout $\varepsilon_1 > 0$ et tout $\varepsilon_2 > 0$, les suites d'événements $\{E_{k,i,n}(\varepsilon_1) : i \geq 1\}$ et $\{F_{k,i,n}(\varepsilon_2) : i \geq 1\}$ sont indépendantes. Notons $\overline{A}$ le complémentaire de l'événement A. Nous avons

$$\mathbb{P}\left[C_k(\varepsilon,\gamma)\right] = \sum_{q=\nu_k+1}^{\nu_{k+1}} \sum_{i=1}^{\infty} \mathbb{P}\left[E_{k,i,q}(\varepsilon) \cap \bigcap_{j=1}^{i-1} \overline{E}_{k,j,q}(\varepsilon) \bigcap_{r=\nu_k+1}^{q-1} \overline{E}_{r,q}(\varepsilon) \right].$$

Ainsi nous avons

$$\begin{aligned} &\inf_{\nu_k < n \leq \nu_{k+1}} \inf_{m \geq 1} \mathbb{P}\left[F_{k,m,n}(\varepsilon/2)\right]\, \mathbb{P}\left[C_k(\varepsilon,\gamma)\right] \\ &\quad \leq \sum_{q=\nu_k+1}^{\nu_{k+1}} \sum_{i=1}^{\infty} \mathbb{P}\left[E_{k,i,q}(\varepsilon) \cap F_{k,i,q}(\varepsilon/2) \cap \bigcap_{j=1}^{i-1} \overline{E}_{k,j,q}(\varepsilon) \bigcap_{r=\nu_k+1}^{q-1} \overline{E}_{r,q}(\varepsilon) \right] \\ &\quad \leq \mathbb{P}\left[\bigcup_{i=1}^{\infty} E_{k,i,\nu_{k+1}}(\varepsilon/2) \right] = \mathbb{P}\left[D_k(\varepsilon/2,\gamma)\right]. \end{aligned} \tag{5.23}$$

Nous allons maintenant montrer qu'il existe un entier k_0 tel que pour tout entier $k \geq k_0$, nous ayons

$$\inf_{\nu_k < n \leq \nu_{k+1}} \inf_{m \geq 1} \mathbb{P}\left[F_{k,m,n}(\varepsilon/2)\right] > \frac{1}{2}.$$

Pour cela, nous allons montrer qu'il existe un entier K_0 tel que pour tout entier $k \geq K_0$, nous ayons

$$\mathbb{P}\left[\overline{F}_{k,m,n}(\varepsilon/2)\right] \leq \varepsilon.$$

Remarque 5.3.1. Deheuvels et Mason [35] utilisent pour établir cette inégalité le fait que

$$\left\{\frac{n\left(\mathbb{U}_n(t)-t\right)}{1-t} : 0 \leq t < 1\right\}$$

est une martingale par rapport à la σ-algèbre $\mathcal{F}_n(t)$ définie par

$$\mathcal{F}_n(t) = \sigma[\mathbf{1}_{\{U_i \leq s\}} : 1 \leq i \leq n, 0 \leq s \leq t] \quad \text{pour } t \in [0,1).$$

Dans le cas qui nous intéresse, nous ne pouvons pas utiliser un argument semblable puisque les variables aléaoires Y_i ne vérifient pas la condition suivante

$$\sum_{i=1}^{n} Y_i = n,$$

qui permettrait d'énoncer la propriété suivante qui serait

$$\left\{\frac{n\left(\mathbb{U}_{n,c}(t)-t\right)}{1-t} : 0 \leq t < 1\right\}$$

est une martingale par rapport à la σ-algèbre $\mathcal{F}_{n,c}(t)$ définie par

$$\mathcal{F}_{n,c}(t) = \sigma[Y_i\mathbf{1}_{\{U_i \leq s\}} : 1 \leq i \leq n, 0 \leq s \leq t] \quad \text{pour } t \in [0,1).$$

Nous allons donc procéder autrement. Nous allons rappeler deux résultats déjà établis dans le chapitre 3.

Inégalité 5.3.1. *Pour tout réel $a \in (0,1)$, il existe une constante $\mathcal{C}_a$ telle que $0 < \mathcal{C}_a < \infty$ et telle que, pour tout réel $\mu > 0$, nous ayons*

$$\mathbb{P}\left[\sup_{0 \leq t \leq a} \pm\sqrt{n}\alpha_{n,c}(t) \geq \mu\right] \leq \mathcal{C}_a \mathbb{P}\left[\sup_{0 \leq t \leq a} \pm\left(\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]\right) \geq \mu\right]. \tag{5.24}$$

Démonstration. On peut consulter la démonstration de cette inégalité dans le chapitre 3 corollaire 3.5.1. □

Propriété 5.3.1. *Le processus de Poisson composé centré $\{\Pi_c(t) - \mathbb{E}[\Pi_c(t)] : t \geq 0\}$ est une martingale relativement à la filtration $\mathcal{A}_s = \sigma\left\{\Pi_c(u) : u \leq s\right\}, s \geq 0$.*

Démonstration. La démonstation de cette propriété se trouve dans le chapitre 3 théorème 3.2.2. □

Ainsi, pour tout entier $n \geq 1$ et pour tout couple de réels $0 < a < 1$ et $\beta > 0$, nous avons

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \sqrt{n}\,|\alpha_{n,c}(t)| \geq \beta\right] \leq 2\,\mathcal{C}_a \mathbb{P}\left[\sup_{0\leq t\leq a} |\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)]| \geq \beta\right].$$

D'après la propriété 5.3.1, le processus de Poisson composé centré $\{\Pi_{n,c}(t) - \mathbb{E}[\Pi_{n,c}(t)] : t \geq 0\}$ est une martingale par rapport à la σ-algèbre $\mathcal{A}_s$ définie dans la propriété 5.3.1. Nous en déduisons que

$$\mathbb{P}\left[\sup_{0\leq t\leq a} \sqrt{n}\,|\alpha_{n,c}(t)| \geq \beta\right] \leq 2\,\mathcal{C}_a \frac{\mathbb{E}\left[|\Pi_{n,c}(a) - \mathbb{E}[\Pi_{n,c}(a)|^2\right]}{\beta^2} = 2\,\mathcal{C}_a \frac{na\mathbb{E}[Y^2]}{\beta^2},$$

puisque d'après le théorème 3.2.1 du chapitre 3, nous avons

$$\mathrm{Var}[\Pi_c(t)] = \lambda t \mathbb{E}[Y^2].$$

À partir de cette dernière inégalité, pour tout couple d'entiers m et n tels que $m \geq 1$ et $\nu_k < n \leq \nu_{k+1}$, nous obtenons que

$$\begin{aligned}
\mathbb{P}\left[\overline{F}_{k,m,n}(\varepsilon/2)\right] &= \mathbb{P}\left[\nu_{k+1}^{-1/2} \sup_{0\leq u\leq a_{\nu_{k+1}}} (\nu_{k+1} - n)|\mathbb{U}_{\nu_{k+1}-n,c}(u) - u| \geq (\varepsilon/2) b_{\nu_{k+1}}\right] \\
&\leq \frac{4\mathbb{E}[Y^2](\nu_{k+1} - \nu_k) a_{\nu_{k+1}}}{\varepsilon^2 \nu_{k+1} b^2_{\nu_{k+1}}} \\
&\leq \frac{2}{\varepsilon^2 \log(1/a_{\nu_{k+1}})} \to 0 \quad \text{lorsque } k \to \infty, \text{ d'après la condition } (S.1).
\end{aligned}$$

Ainsi, nous avons

$$\mathbb{P}\left[\overline{F}_{k,m,n}(\varepsilon/2)\right] \to 0 \quad \text{lorsque } k \to \infty.$$

Ainsi, nous en déduisons qu'il existe un entier k_0 tel que pour tout entier $k \geq k_0$ nous ayons

$$\inf_{\nu_k < n \leq \nu_{k+1}} \inf_{m\geq 1} \mathbb{P}\left[F_{k,m,n}(\varepsilon/2)\right] > \frac{1}{2}.$$

En utilisant cette borne dans (5.23), on en déduit l'inégalité (5.22). □

Lemme 5.3.5. *Nous avons*

$$\lim_{\gamma\downarrow 0}\Bigg(\limsup_{k\to\infty} \max_{\nu_k<n\leq\nu_{k+1}} \sup_{0\leq t\leq 1-a_{\nu_{k+1}}} \left|(n/\nu_{k+1})^{1/2}\, b^{-1}_{\nu_{k+1}} - b_n^{-1}\right| \times \left\|\xi_{n,c}(a_{\nu_{k+1}}, t; \cdot)\right\|\Bigg) = 0 \quad p.s. \tag{5.25}$$

Démonstration. Le membre de gauche de (5.25) est inférieur ou égal à

$$\lim_{\gamma\downarrow 0}\limsup_{k\to\infty}\left(\max_{\nu_k<n\le\nu_{k+1}}\left|b_n b_{\nu_{k+1}}^{-1}-\big(\nu_{k+1}/n\big)^{1/2}\right|\right)\times\limsup_{n\to\infty} b_n^{-1}\omega_{n,c}(a_n), \tag{5.26}$$

où $\omega_{n,c}(a_n)$ est défini, pour $0<a_n<1$ et pour $n\ge 1$, par

$$\omega_{n,c}(a_n)=\sup_{0\le t\le 1-a_n}\big|\xi_{n,c}(a_n,t;\cdot)\big|.$$

Remarque 5.3.2. Le module d'oscillation $\omega_{n,c}$ du processus empirique composé a été défini dans le chapitre 4.

Sous les conditions [CRS] (S.1–2–3), nous pouvons appliquer le théorème 4.4.1 du chapitre précédent et nous avons

$$\lim_{n\to\infty} b_n^{-1}\omega_{n,c}(a_n)=1 \quad \text{p.s.} \tag{5.27}$$

De plus, la condition (S.1) implique que

$$a_n\downarrow 0 \quad \text{et} \quad na_n\uparrow\infty \quad \text{lorsque } n\to\infty.$$

Par conséquent la suite $(b_n)_{n\ge1}$ est décroissante à partir d'un certain rang et est telle que, pour tout entier n tel que $\nu_k<n\le\nu_{k+1}$,

$$0\le b_n/b_{\nu_{k+1}}-1\le b_{\nu_k}/b_{\nu_{k+1}}-1\le\big(1+o(1)\big)\big(\nu_{k+1}/\nu_k\big)-1\to\gamma \quad \text{lorsque } k\to\infty.$$

Ainsi, en combinant les résultats (5.26) et (5.27), nous obtenons (5.25). □

Lemme 5.3.6. *Nous avons*

$$\lim_{\gamma\downarrow 0}\left(\limsup_{k\to\infty}\max_{\nu_k<n\le\nu_{k+1}}\sup_{0\le t\le 1-a_{\nu_{k+1}}} b_n^{-1}\big\|\xi_{n,c}(a_{\nu_{k+1}},t;\cdot)-\xi_{n,c}(a_n,t;\cdot)\big\|\right)=0 \quad p.s. \tag{5.28}$$

Démonstration. Comme la suite $(a_n)_{n\ge1}$ tend vers 0 en décroissant lorsque $n\to\infty$ (d'après la condition (S.1) (ii)), pour tout entier n tel que $\nu_k<n\le\nu_{k+1}$, nous avons

$$a_n-a_{\nu_{k+1}}=a_n\big(1-a_{\nu_{k+1}}/a_n\big)\le a_n\big(1-a_{\nu_{k+1}}/a_{\nu_k}\big). \tag{5.29}$$

Comme la suite $(na_n)_{n\ge1}$ est croissante lorsque $n\to\infty$ (d'après la condition (S.1) (iii)), nous avons

$$a_{\nu_{k+1}}/a_{\nu_k}\ge\nu_k/\nu_{k+1}\ge 1-\gamma \quad \text{lorsque } k\to\infty. \tag{5.30}$$

Ainsi, pour tout entier n tel que $\nu_k<n\le\nu_{k+1}$, en rassemblant (5.29) et (5.30), nous en déduisons que

$$a_n-a_{\nu_{k+1}}\le a_n\big(1-a_{\nu_{k+1}}/a_{\nu_k}\big)\le\gamma a_n \quad \text{lorsque } k\to\infty.$$

Ainsi, le membre de gauche de (5.28) est inférieur ou égal à

$$\lim_{\gamma\downarrow 0}\left(\lim_{n\to\infty} b_n^{-1}\omega_{n,c}(\gamma a_n)\right). \tag{5.31}$$

D'après (5.27) et (5.31) il est facile d'obtenir le résultat (5.28). Ce qui achève la démonstration du lemme 5.3.6. □

Avec les lemmes 5.3.4 et 5.3.6, nous allons montrer dans ce qui suit que la démonstration de l'inclusion (5.10) du théorème 5.2.1 se ramène à montrer que, pour tout $\varepsilon > 0$ et tout $\gamma > 0$

$$\sum_{k=1}^{\infty} \mathbb{P}\left[D_k(\varepsilon,\gamma)\right] < \infty. \tag{5.32}$$

A cette fin, nous allons utiliser l'inégalité suivante.

Lemme 5.3.7. *Pour tout réel θ tel que $0 < \theta < 1$, pour tout $\varepsilon > 0$ et pour tout entier $n \geq 1$, nous avons*

$$\begin{aligned}\mathbb{P}\Big[b_n^{-1}\xi_{n,c}(a_n,t;\cdot) \notin \mathbb{S}_0^{\varepsilon} \textit{ pour certains } 0 \leq t \leq 1-a_n\Big] \\ \leq (\theta a_n)^{-1}\,\mathbb{P}\left[b_n^{-1}\xi_{n,c}(a_n,0;\cdot) \notin \mathbb{S}_0^{\varepsilon/2}\right] + \mathbb{P}\left[b_n^{-1}\omega_{n,c}(\theta a_n) > \varepsilon/4\right].\end{aligned}$$

Démonstration. Pour t tel que $\theta i a_n \leq t \leq \theta(i+1)a_n$ et pour $i \geq 0,1,\ldots,[(a_n^{-1}-1)/\theta]-1 = \mu_n - 1$, ou pour t tel que $\mu_n a_n \leq t \leq 1-a_n$ et pour $i=\mu_n$, nous avons uniformément pour $0 \leq s \leq 1$,

$$\begin{aligned}\left|\xi_{n,c}(a_n,t;s) - \xi_{n,c}(a_n,i\theta a_n;s)\right| &\leq \left|\alpha_{n,c}(t+sa_n) - \alpha_{n,c}(i\theta a_n + sa_n)\right| \\ &\quad + \left|\alpha_{n,c}(t) - \alpha_{n,c}(i\theta a_n)\right| \\ &\leq 2\omega_{n,c}(\theta h_n).\end{aligned}$$

Le reste de la démonstration du lemme 5.3.7 est évident. □

Maintenant nous pouvons montrer que, pour tout $\varepsilon > 0$ et tout $\gamma > 0$,

$$\sum_{k=1}^{\infty} \mathbb{P}\left[D_k(\varepsilon,\gamma)\right] < \infty.$$

Pour cela, nous allons procéder en plusieur étapes.
Première étape. Nous utilisons le lemme 5.3.7 pour obtenir que, pour tout réel $\theta > 0$ suffisamment petit,

$$\begin{aligned}\mathbb{P}\left[D_k(\varepsilon,\gamma)\right] &\leq \left(\theta a_{\nu_{k+1}}\right)^{-1}\mathbb{P}\left[b_{\nu_{k+1}}^{-1}\xi_{\nu_{k+1},c}(a_{\nu_{k+1}},0;.) \notin \mathbb{S}_0^{\varepsilon/2}\right] \\ &\quad + \mathbb{P}\left[b_{\nu_{k+1}}^{-1}\omega_{\nu_{k+1},c}(\theta a_{\nu_{k+1}}) \geq \varepsilon/4\right] \\ &\leq Q_{1,k} + Q_{2,k}.\end{aligned}$$

Il ne reste plus qu'à montrer que

$$\sum_{k=1}^{\infty} Q_{1,k} < \infty \quad \text{et} \quad \sum_{k=1}^{\infty} Q_{2,k} < \infty.$$

Deuxième étape. Pour montrer que

$$\sum_{k=1}^{\infty} Q_{1,k} < \infty,$$

nous allons également procéder en plusieurs étapes. D'abord, nous utilisons le lemme 5.3.1 pour obtenir l'inégalité suivante

$$\begin{aligned}(\theta h_{\nu_{k+1}})Q_{1,k} &= \mathbb{P}\left[b^{-1}_{\nu_{k+1}}\xi_{\nu_{k+1}}(a_{\nu_{k+1}},0;.) \notin \mathbb{S}_0^{\varepsilon/2}\right]\\ &\leq \mathcal{C}_0\mathbb{P}\left[b^{-1}_{\nu_{k+1}}L(a_{\nu_{k+1}},0;.) \notin \mathbb{S}_0^{\varepsilon/2}\right]\\ &= \mathcal{C}_0 Q_{3,k}.\end{aligned}$$

Ensuite nous reprenons les arguments de la démonstration du lemme 5.3.3 pour obtenir une inégalité analogue à (5.18) et alors nous obtenons que

$$Q_{3,k} \leq \mathbb{P}\Big(W_{\log(1/a_{\nu_{k+1}})} \notin \mathbb{S}_0^{\varepsilon/4}\Big) + \mathbb{P}^{\varepsilon/4}_{3,\nu_{k+1}} = Q_{4,k} + Q_{5,k}. \tag{5.33}$$

Étudions d'abord la quantité $Q_{4,k}$. Comme $B(0,1) - \mathbb{S}_0^{\varepsilon/4}$ est fermé (pour la topologie de la convergence uniforme) et satisfait à la condition

$$J\left(B(0,1) - \mathbb{S}_0^{\varepsilon/4}\right) > 1,$$

d'après l'inégalité de grandes déviations (5.14), pour tout réel $\rho \in \left(J\left(B(0,1) - \mathbb{S}_0^{\varepsilon/4}\right), 1\right)$, nous en déduisons que

$$Q_{4,k} \leq \exp\big(-\rho\log(1/a_{\nu_{k+1}})\big) = a^{\rho}_{\nu_{k+1}} \quad \text{lorsque } k \to \infty. \tag{5.34}$$

Pour la quantité $Q_{5,k}$, les mêmes arguments que ceux utilisés pour (5.19) et (5.20) entraînent que

$$Q_{5,k} \leq h^{\rho}_{\nu_{k+1}} \quad \text{lorsque } k \to \infty. \tag{5.35}$$

Ainsi, en combinant les inégalités (5.33), (5.34) et (5.35), nous obtenons que

$$\begin{aligned}Q_{1,k} &\leq \frac{\mathcal{C}_0}{\theta a_{\nu_{k+1}}} Q_{3,k}\\ &\leq \frac{2\mathcal{C}_0}{\theta} a^{\rho-1}_{\nu_{k+1}}\\ &= \frac{2\mathcal{C}_0}{\theta}\exp\left(-\big(\rho-1\big)\log\big(1/a_{\nu_{k+1}}\big)\right).\end{aligned} \tag{5.36}$$

Comme la condition (S.2) implique que

$$\log(1/a_{\nu_{k+1}})/\log k \to \infty \quad \text{lorsque } k \to \infty,$$

d'après l'inégalité (5.36), nous en déduisons que

$$Q_{1,k} \leq \frac{1}{k^2} \quad \text{lorsque } k \to \infty$$

et par conséquent que

$$\sum_{k=1}^{\infty} Q_{1,k} < \infty. \tag{5.37}$$

Troisième étape. Étudions maintenant la quantité $Q_{2,k}$. Pour cela, nous allons utiliser l'inégalité 4.3.1 du chapitre précédent. Posons

$$\delta = 1/2, \quad a = \theta a_{\nu_{k+1}}, \quad n = \nu_{k+1}, \quad \mu = \frac{\varepsilon}{4\sqrt{\theta}}\sqrt{2\mathbb{E}[Y^2]\log(1/a_{\nu_{k+1}})}.$$

Nous obtenons alors

$$Q_{2,k} \leq \frac{10240}{\theta a_{\nu_{k+1}}} \exp\left(-\frac{\varepsilon^2 \log(1/a_{\nu_{k+1}})}{2048\,\theta}\psi\left(\frac{\mathcal{M}\varepsilon\sqrt{2\log(1/a_{\nu_{k+1}})}}{4\theta\sqrt{\mathbb{E}[Y^2]\nu_{k+1}a_{\nu_{k+1}}}}\right)\right). \tag{5.38}$$

Comme la condition $(S.3)$ implique que

$$na_n/\log(1/a_n) \to \infty \quad \text{lorsque } n \to \infty,$$

et d'autre part, d'après la proposition 3.3.1 du chapitre 3, nous savons que

$$\psi(x) \to 1 \quad \text{lorsque} \quad x \downarrow 0,$$

nous en déduisons que le membre de droite de (5.38) est inférieur ou égal à

$$\frac{10240}{\theta} \exp\left(-\left(\frac{\varepsilon^2}{4096\,\theta} - 1\right)\log\left(\frac{1}{a_{\nu_{k+1}}}\right)\right).$$

Ainsi, le même argument utilisé pour établir l'inégalité (5.37) montre que, pour tout $0 < \theta < \varepsilon/128$, nous avons

$$\sum_{k=1}^{\infty} Q_{2,k} < \infty. \tag{5.39}$$

Quatrième étape. En combinant (5.37) et (5.39), nous remarquons que (5.32) reste vraie. D'après le lemme de Borel-Cantelli et les lemmes 5.3.4 et 5.3.6, il vient que pour tout $\varepsilon > 0$, dès que le réel $\gamma > 0$ est suffisamment petit, l'événement

$$\bigcup_{\nu_k < n \leq \nu_{k+1}} \left\{b_n^{-1}\xi_{n,c}(a_n, t; \cdot) : 0 \leq t \leq 1 - a_{\nu_{k+1}}\right\} \not\subset \mathbb{S}_0^{\varepsilon}$$

arrive un nombre fini de fois avec une probabilité égale à 1. Le lemme 5.3.8 montre que la même chose est vraie pour l'événement

$$\bigcup_{\nu_k < n \leq \nu_{k+1}} \left\{b_n^{-1}\xi_{n,c}(a_n, t; \cdot) : 1 - a_{\nu_{k+1}} < t \leq 1 - a_n\right\} \not\subset \mathbb{S}_0^{\varepsilon},$$

ce qui achève la démonstration de l'inclusion (5.10) du théorème 5.2.1.

Lemme 5.3.8. *Nous avons*

$$\lim_{\gamma\downarrow 0}\Bigg(\limsup_{k\to\infty}\sup_{\nu_k<n\le\nu_{k+1}}\sup_{1-a_{\nu_{k+1}}\le t\le 1-a_n} b_n^{-1} \times \left\|\xi_{n,c}(a_n,t;\cdot)-\xi_{n,c}(a_n,1-a_{\nu_{k+1}};\cdot)\right\|\Bigg)=0 \qquad p.s. \tag{5.40}$$

Démonstration. Avec les mêmes arguments que ceux utilisés pour la démonstration du lemme 5.3.6, nous remarquons que le membre droit de (5.40) est inférieur ou égal à

$$\lim_{\gamma\downarrow 0}\left(2\lim_{n\to\infty} b_n^{-1}\omega_{n,c}(\gamma a_n)\right),$$

qui vaut 0 presque sûrement d'après (5.27) et (5.31). □

Chapitre 6

Lois du L.I. pour le processus empirique composé indexé par des ensembles

6.1. Notations et définitions

Dans ce paragraphe, nous allons généraliser la définition du processus empirique composé défini en un point à celle du processus empirique composé indexé par des ensembles et plus particulièrement par des intervalles qui ne sont en fait qu'une famille particulière d'ensembles. Pour cela introduisons quelques notations.

Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur $[0,1]^d$ où $d \geq 1$. Soit $\mathbb{B}$ la classe de sous-ensembles de Borel de $[0,1]^d$. Notons $\lambda(.)$ la mesure de Lebesgue sur $\mathbb{R}^d$, et $\lambda_0(.)$ la restriction de la mesure $\lambda(.)$ sur $[0,1]^d$. Introduisons la *mesure empirique uniforme indexée par la classe* $\mathbb{B}$, définie par

$$\lambda_n(B) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{U_i \in B\}}, \quad \text{où } B \in \mathbb{B} \text{ et } n \geq 1. \tag{6.1}$$

En calquant sur le modèle précédent, définissons la *mesure empirique composée indexée par la classe* $\mathbb{B}$ par

$$\lambda_{n,c}(B) = \frac{1}{n}\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \in B\}}, \quad \text{où } B \in \mathbb{B}, \tag{6.2}$$

et $\{Y_i : i \geq 1\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. De plus, les suites de variables aléatoires $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes.

Pour toute sous-classe $\mathbb{D}$ de la classe $\mathbb{B}$, introduisons le *processus empirique uniforme indexé par* la sous-classe $\mathbb{D}$ et le *processus empirique composé indexé par* la sous-classe $\mathbb{D}$, définis respectivement par

$$\alpha_n(D) = \sqrt{n}\left(\lambda_n(D) - \lambda_0(D)\right), \quad \text{où } D \in \mathbb{D} \tag{6.3}$$

et

$$\alpha_{n,c}(D) = \sqrt{n}\left(\lambda_{n,c}(D) - \mathbb{E}[\lambda_{n,c}(D)]\right), \quad \text{où } D \in \mathbb{D}. \tag{6.4}$$

Soit la sous-classe $\mathbb{D}$ définie ainsi. Soient $\mathbf{t} \in [0,1]^d$ et $\mathbb{C}$ une classe de sous-ensembles de Borel de la forme $[a,b]^d$ où $a < b$ et $b - a = 1$. Soit $\mathbf{a} = (a, \ldots, a) \in \mathbb{R}^d$ et posons

$$\mathbb{D} = \{\mathbf{t} + C : C \in \mathbb{C}\}. \tag{6.5}$$

Nous allons donner deux définitions qui seront utiles par la suite.

Définition 6.1.1. *La classe* $\mathbb{C}$ *est* engendrée par une famille dénombrable *si il existe une sous-classe dénombrable* $\mathbb{G}$ *de la classe* $\mathbb{C}$ *telle que, pour tout* $D \in \mathbb{C}$*, il existe une suite* $\{D_n : n \geq 1\}$ *de* $\mathbb{G}$ *qui vérifie, pour tout* $x \in \mathbb{R}^d$*,*

$$\lim_{n\to\infty} \mathbf{1}_{\{x \in D_n\}} = \mathbf{1}_{\{x \in D\}}. \tag{6.6}$$

Remarque 6.1.1. Par la suite, nous dirons qu'une classe $\mathbb{G}$ de ce type est une famille génératrice dénombrable de $\mathbb{C}$.

Définition 6.1.2. *Une classe* $\mathbb{D}$ *est une classe de Donsker pour la mesure* λ_0 *si les mesures empiriques normalisées* $\{\alpha_n(D) : D \in \mathbb{D}\}$ *définies en* (6.3) *convergent en loi [dans le sens de Dudley [48], section 1, Kuelbs et Dudley [75], p.406 et Gaenssler [57], p.46,65 et 113] vers une mesure gaussienne* $\{G_\lambda(D) : D \in \mathbb{D}\}$*.*

Remarque 6.1.2. Nous renvoyons au théorème B, p.113 de Gaenssler [57] et aux articles de Giné et Zinn [59], Talagrand [132] et Alexander [1] pour de plus amples détails concernant les classes de Donsker.

De plus, nous supposons que
(C.1) $\mathbf{t} + C \subseteq [0,1]^d$ pour tout $C \in \mathbb{C}$ et, pour tout $h > 0$ suffisamment petit, $\mathbf{t} + h^{1/d}[a,b]^d \subseteq [0,1]^d$.
(C.2) $\mathbb{C}$ est *engendrée par une famille dénombrable.*
(C.3) $\mathbb{C} - \{\mathbf{a}\}$ est une *classe de Donsker pour* λ_0.
(C.4) Pour tous $1/2 \leq h_1 \leq h_2 \leq 1$, $h_1\mathbb{C} \subseteq h_2\mathbb{C} \subseteq \mathbb{C}$.
(C.5) Pour tout sous-ensemble fini $\{C_1, \cdots, C_M\} \subseteq \mathbb{C}$, avec $\lambda(C_i) > 0$, la classe $\mathbb{C}$ peut être agrandie, si cela est nécessaire, pour inclure un ensemble fini d'éléments disjoints $\{D_1, \cdots, D_M\} \subseteq \mathbb{C}$, avec $\lambda(D_i) > 0$, tel que pour chaque C_i il existe un sous-ensemble $J \subseteq \{1, \cdots, N\}$ pour lequel $\cup_{j\in J} D_j = C_i$.

Pour tout $C \in \mathbb{C}$, notons que

$$\lim_{h \uparrow 1} d_\lambda(h\,C, C) = 0, \tag{6.7}$$

où d_λ est la *pseudométrique* définie sur la classe $\mathbb{C}$ par

$$d_\lambda(A,B) = \lambda(A\Delta B), \qquad \text{pour tous } A \text{ et } B \in \mathbb{C}, \tag{6.8}$$

et où $A\Delta B = (A - B) \cup (B - A)$ représente la *différence symétrique* entre A et B.

Posons

$$\log^+ u = \log(\max(u,e)) \quad \text{ et } \log_2 u = \log^+\left(\log^+ u\right).$$

Soit une suite $\{k(n) : n \geq 1\}$ qui vérifie les conditions suivantes.

(K.1) (i) $0 < k(n) \leq n$, (ii) $k(n) \uparrow \infty$ et (iii) $n^{-1}k(n) \downarrow 0$ lorsque $n \to \infty$;

(K.2) $k(n)/\log_2 n \to \infty$ lorsque $n \to \infty$;

(K.3) $k(n)/\log_2 n \to \mathcal{C} \in (0,\infty)$ lorsque $n \to \infty$.

Introduisons enfin le *processus empirique composé local au point* $\mathbf{t}$ *indexé par la classe* $\mathbb{C}$ défini par

$$\begin{aligned}\Theta_{n,c}(C) &= \Theta_{n,c}\left(C, \frac{k(n)}{n}\right) = \Theta_{n,c}\left(C, \mathbf{t}, \frac{k(n)}{n}\right) \\ &= \left(\frac{k(n)}{n}\mathbb{E}[Y^2]\right)^{-1/2} \alpha_{n,c}\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right), \quad \text{où } C \in \mathbb{C} \qquad (6.9)\end{aligned}$$

et le processus $\alpha_{n,c}$ est défini en (6.4).

Remarque 6.1.3. La définition (6.9) du processus $\Theta_{n,c}$ s'appuie sur celle du processus empirique local Θ_n au point $\mathbf{t}$ indexé par la classe $\mathbb{C}$. Ce processus Θ_n a été introduit par Deheuvels et Mason [36] p.1620. Il se définit par

$$\begin{aligned}\Theta_n(C) &= \Theta_n\left(C, \frac{k(n)}{n}\right) = \Theta_n\left(C, \mathbf{t}, \frac{k(n)}{n}\right) = \left(\frac{k(n)}{n}\right)^{-1/2} \alpha_n\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right) \\ &= (k(n))^{-1/2}\left\{\sum_{i=1}^{n} \mathbf{1}\left(U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right) - k(n)\lambda_0(C)\right\}, \quad \text{où } C \in \mathbb{C}.\end{aligned} \qquad (6.10)$$

Remarque 6.1.4. Nous pouvons particulariser la classe $\mathbb{C}$ et la sous-classe $\mathbb{D}$ afin de définir le processus $\Theta_{n,c}$ indexé par des intervalles. Soient un réel $t \in [0,1]$ et

$$\mathbb{C}_1 = \{(y,z] : -1/2 \leq y \leq z \leq 1/2\} \qquad (6.11)$$

la classe des sous-intervalles de l'intervalle $[-1/2, 1/2]$. Par conséquent la sous-classe $\mathbb{D}$ devient

$$\mathbb{D}_1 = \{t + I : I \in \mathbb{C}_1\}. \qquad (6.12)$$

Le processus $\Theta_{n,c}$ indexé par des intervalles se définit alors par

$$\begin{aligned}\Theta_{n,c}((y,z]) &= \Theta_{n,c}\left((y,z], \frac{k(n)}{n}\right) = \Theta_{n,c}\left((y,z], t, \frac{k(n)}{n}\right) \\ &= \left(\frac{k(n)}{n}\mathbb{E}[Y^2]\right)^{-1/2} \alpha_{n,c}\left(t + \frac{k(n)}{n}(y,z]\right), \quad \text{où } (y,z] \in \mathbb{C}_1, \qquad (6.13)\end{aligned}$$

et où le processus $\alpha_{n,c}$ est défini en (6.4).

6.2. Résultats antérieurs concernant le processus empirique local

Les propriétés limite du processus empirique local Θ_n défini en (6.10) ont déjà été étudiées dans le cas particulier où $d = 1$, $\mathbf{t} = 0$ et où $\mathbb{C}$ représente la classe des intervalles de la forme $[0, s]$ pour $0 \leq s \leq 1$. Lorsque $t = 0$, le processus est appelé le *processus empirique de queue*. Nous allons commencer par rappeler très brièvement les principales lois fortes qui ont été obtenues dans ce cas.

Pour $s \in [0, 1]$, posons

$$\xi_n(s) = \Theta_n\left([0, s], 0, \frac{k(n)}{n}\right), \quad \eta_n(s) = \lambda_n\left(\left[0, \frac{k(n)}{n}\, s\right]\right), \tag{6.14}$$

où le processus Θ_n est défini en (6.10) et la mesure empirique λ_n en (6.1).

Rappelons que B$(0, 1)$ désigne l'espace des fonctions bornées définies sur l'intervalle $[0, 1]$ et à valeurs réelles muni de la norme de la convergence uniforme et que la fonction h est définie par

$$h(u) = \begin{cases} u \log u - u + 1 & \text{pour } u > 0, \\ 1 & \text{pour } u = 0, \\ \infty & \text{pour } u < 0. \end{cases} \tag{6.15}$$

Rappelons deux résultats déjà cités dans le chapitre précédent.

Théorème 6.2.1 (Mason (1988)). *Sous les hypothèses* $(K.1)$ *et* $(K.2)$*, la suite de fonctions*

$$\left\{ \frac{\xi_n(t)}{\sqrt{2 \log \log n}} : 0 \leq t \leq 1 \right\}$$

est presque sûrement relativement compacte dans $B(0, 1)$ *muni de la topologie de la convergence uniforme. De plus, l'ensemble des points limites de cette suite est égal à l'ensemble de Strassen* $\mathbb{S}_0$ *défini dans le chapitre précédent.*

Démonstration. Pour la démonstration de ce théorème, on pourra consulter l'article de Mason [94]. □

Théorème 6.2.2 (Deheuvels et Mason (1990)). *Sous les hypothèses* $(K.1)(i)$ *et* $(K.3)$*, la suite de fonctions*

$$\left\{ \frac{n\eta_n(t)}{\sqrt{\log \log n}} : 0 \leq t \leq 1 \right\}$$

est presque sûrement relativement compacte dans $B(0, 1)$ *muni de la topologie de la convergence uniforme. De plus, l'ensemble des points limites de cette suite est égal à l'ensemble* *composé des fonctions f absolument continues sur l'intervalle $[0,1]$ telles que $f(0)=0$ et qui ont une dérivée f' au sens de Lebesgue qui vérifie*

$$\mathcal{C}\int_0^1 h\left(\frac{f'(s)}{\mathcal{C}}\right) ds \leq 1, \tag{6.16}$$

où la fonction $h(\cdot)$ est définie en (6.15)

Démonstration. Pour la démonstration de ce théorème, on pourra consulter l'article de Deheuvels et Mason [33]. □

Des résultats tels que ceux donnés dans les théorèmes 6.2.1 et 6.2.2 sont très utiles pour décrire le comportement limite des *statistiques de queue*. Ces dernières se basent sur les valeurs extrêmes supérieures (respectivement inférieures) d'un échantillon univarié. Pour de plus amples détails nous renvoyons à l'article de Deheuvels et Mason [34].

Enfin citons un dernier résultat qui est la base du résultat principal de ce chapitre. Ce théorème est une généralisation de la loi fonctionnelle du logarithme itéré énoncée dans le théorème 6.2.1. Pour énoncer ce résultat, introduisons la classe $B(\mathbb{C})$ qui est composée des fonctions définies et bornées sur $\mathbb{C}$, où $\mathbb{C}$ est muni de la topologie de la convergence uniforme.

Théorème 6.2.3. *Sous les hypothèses (K.1–2) et (C.1–2–3–4–5), la suite de fonctions*

$$\left\{\frac{\Theta_n(C)}{\sqrt{2\log_2 n}} : C\in\mathbb{C}\right\}$$

est presque sûrement relativement compacte dans $B(\mathbb{C})$ muni de la topologie de la convergence uniforme. De plus, l'ensemble limite est égal à

$$\mathbb{S}(\mathbb{C}) = \left\{f\in B(\mathbb{C}) : f(C) = \int_C \phi(\mathbf{s})\,d\lambda(\mathbf{s}) \text{ avec } \int_{\mathbb{R}^d}[\phi(\mathbf{s})]^2\,d\lambda(\mathbf{s})\leq 1\right\}. \tag{6.17}$$

Démonstration. La démonstration de ce théorème se trouve dans l'article de Deheuvels et Mason [36]. □

Appliquons ce théorème à une classe particulière, la classe $\mathbb{C}_1$ définie par

$$\mathbb{C}_1 = \{(y,z] : -1/2\leq y\leq z\leq 1/2\} \tag{6.18}$$

la classe des sous-intervalles de l'intervalle $[-1/2,1/2]$. Par conséquent la sous-classe $\mathbb{D}$ devient

$$\mathbb{D}_1 = \{t+I : I\in\mathbb{C}_1\}.$$

Nous obtenons ainsi le corollaire suivant.

Corollaire 6.2.1. *Sous les hypothèses (K.1–2) et (C.1–2–3–4–5), la suite de fonctions*

$$\left\{ \frac{\Theta_n\left((y,z]\right)}{\sqrt{2\log_2 n}} : (y,z] \in \mathbb{C}_1 \right\}$$

est presque sûrement relativement compacte dans $B(\mathbb{C}_1)$ muni de la topologie de la convergence uniforme. De plus, l'ensemble limite est égal à

$$\mathbb{S}(\mathbb{C}_1) = \left\{ f \in B(\mathbb{C}_1) : f((y,z]) = \int_y^z \phi(s)\, ds \ \textit{avec} \ \int_{\mathbb{R}} [\phi(s)]^2\, ds \leq 1 \right\}. \tag{6.19}$$

Ce corollaire sera utilisé dans le paragraphe “Applications” pour établir une loi du logarithme itéré pour la convergence forte en un point de l'estimateur de la densité f_n qui améliore les résultats établis par Hall [62].

6.3. Résultat principal

L'objet de ce chapitre est d'établir une approche unifiée de l'étude du comportement presque sûre des statistiques univariées locales ou de queue telles que les statistiques locales qui étudient par exemple les estimateurs de régression non paramètriques dans le cas univarié. Ces derniers seront étudiés dans le chapitre 7.

Rappelons que $\mathbb{C}$ désigne une classe de sous-ensembles de Borel de la forme $[a,b]^d$ où $a < b$ et $b - a = 1$.

De plus, introduisons les hypothèses suivantes.

(H.1) La suite $\{Y_i : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées est telle que

$$\mathbb{E}[Y_i] = 1 \quad \text{pour tout } i \geq 1.$$

(H.2) La suite $\{Y_i : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées est telle que

$$|Y_i - \mathbb{E}[Y_i]| \leq \mathcal{M} \quad \text{pour tout } i \geq 1.$$

Théorème 6.3.1. *Sous les hypothèses (K.1–2), (C.1–2–3–4–5) et (H.1–2), la suite de fonctions*

$$\left\{ \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} : C \in \mathbb{C} \right\}$$

est presque sûrement relativement compacte dans $B(\mathbb{C})$ muni de la topologie de la convergence uniforme. De plus, l'ensemble limite est égal à

$$\mathbb{S}(\mathbb{C}) = \left\{ f \in B(\mathbb{C}) : f(C) = \int_C \phi(\mathbf{s})\, d\lambda(\mathbf{s}) \ \textit{avec} \ \int_{\mathbb{R}^d} [\phi(\mathbf{s})]^2\, d\lambda(\mathbf{s}) \leq 1 \right\}. \tag{6.20}$$

Appliquons le théorème 6.3.1 à la classe $\mathbb{C}_1$ définie en (6.18). Le corollaire suivant en découle.

Corollaire 6.3.1. *Sous les hypothèses (K.1–2), (C.1–2–3–4–5) et (H.1–2), la suite de fonctions*

$$\left\{ \frac{\Theta_{n,c}((y,z])}{\sqrt{2\mathbb{E}[Y^2]\log_2 n}} : (y,z] \in \mathbb{C}_1 \right\}$$

est presque sûrement relativement compacte dans $B(\mathbb{C}_1)$ *muni de la toplogie de la convergence uniforme. De plus, l'ensemble limite est égal à*

$$\mathbb{S}(\mathbb{C}_1) = \left\{ f \in B(\mathbb{C}_1) : f((y,z]) = \int_y^z \phi(s)\, ds \ \textit{avec} \ \int_{\mathbb{R}} [\phi(s)]^2\, ds \leq 1 \right\}. \tag{6.21}$$

Dans le paragraphe 6.4, nous donnons une preuve du théorème 6.3.1. Les méthodes utilisées s'appuient sur celles utilisées par Deheuvels et Mason [36] et sont essentiellement basées sur la théorie générale des processus empiriques et n'utilisent pas de principes d'invariance forte. Nous présenterons une utilisation du corollaire 6.3.1 dans le paragraphe "Applications".

6.4. Démonstration du résultat principal

Pour démontrer le théorème 6.3.1, nous avons besoin d'établir quelques résultats préliminaires.

6.4.1. Résultats préliminaires

Soit la suite $\{U_i : i \geq 1\}$ de variables aléatoires indépendantes, uniformément distribuées sur $[0,1]^d$ et définies sur un même espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$. Nous commençons par énoncer des propriétés utiles qui découlent des hypothèses (C.1–2–3–4–5).

Notons $B_0(\mathbb{C})$ l'espace linéaire composé de l'espace des fonctions bornées, définies sur $\mathbb{C}$, à valeurs réelles et uniformément continues par rapport à la pseudométrique d_λ définie en (6.8).

Remarque 6.4.1. Notons que $C \to \lambda(C)$ appartient à $B_0(\mathbb{C})$, sous la condition que pour deux ensembles mesurables C_1 et C_2, nous ayons

$$|\lambda(C_1) - \lambda(C_2)| \leq d_\lambda(C_1, C_2).$$

Notons $B_1(\mathbb{C})$ l'espace linéaire composé de l'espace des fonctions ϕ définies sur $\mathbb{C}$, à valeurs réelles et de la forme

$$\phi(C) = \phi_1(C) + \sum_{i=1}^{k} a_i \mathbf{1}_{\{x_i \in C\}}, \tag{6.22}$$

où $a_i \in \mathbb{R}$, $x_i \in \mathbb{R}^d$, $1 \leq i \leq k$, $k \in \mathbb{N}$, et $\phi_1 \in B_0(\mathbb{C})$.

Remarque 6.4.2. Notons que

$$B_0(\mathbb{C}) \subset B_1(\mathbb{C}).$$

Munissons ces deux espaces de la norme de la convergence uniforme définie par

$$||\phi||_{\mathbb{C}} = \sup_{C \in \mathbb{C}} |\phi(C)|. \tag{6.23}$$

Pour tout $C \in \mathbb{C}$, considérons la fonction $\pi_C : B_1(\mathbb{C}) \rightarrow \mathbb{R}$ définie par $\pi_C(\phi) = \phi(C)$. Notons $\mathbb{A} = \sigma(\pi_C : C \in \mathbb{C})$ la σ-algèbre de $B_1(\mathbb{C})$ engendrée par $\{\pi_C : C \in \mathbb{C}\}$, et $\mathbb{B}_U$ la σ-algèbre engendrée par les boules ouvertes par rapport à la norme $||.||_{\mathbb{C}}$.

Propriété 6.4.1. *Sous l'hypothèse* $(C.2)$, *nous avons*

$$\mathbb{B}_U \subseteq \mathbb{A}. \tag{6.24}$$

De plus, si $\mathbb{G}$ *est une sous classe dénombrable de* $\mathbb{C}$, *alors, pour tout* $\phi \in B(\mathbb{C})$ *et* $\varepsilon > 0$, *nous avons*

$$\{f \in B(\mathbb{C}) : ||\phi - f||_{\mathbb{C}} \leq \varepsilon\} = \bigcap_{D \in \mathbb{G}} \{f \in B(\mathbb{C}) : |\phi(D) - f(D)| \leq \varepsilon\}. \tag{6.25}$$

De plus, pour tout $n \geq 1$, *l'ensemble des fonctions dont les applications qui à* $C \in \mathbb{C}$ *associent* $\lambda_n(\mathbf{a} + C)$ *[resp.* $\Theta_n(C)$*] est* $(\mathcal{A}, \mathbb{A})$ *mesurable, et chaque chemin échantillon de* $\{\lambda_n(\mathbf{a} + C), C \in \mathbb{C}\}$ *[resp.* $\{\Theta_n(C), C \in \mathbb{C}\}$*] est uniquement déterminé par ses valeurs sur* $\mathbb{G}$.

Démonstration. Pour la démonstration de cette propriété, on pourra, par exemple, consulter le lemme 20, p.108 de Gaenssler [57]. □

Pour tout $n \geq 1$ et pour tout $A \in \mathbb{B}$, soit

$$\alpha_{n,c}(A) = \sqrt{n}\left(\lambda_{n,c}(A) - \mathbb{E}\left[\lambda_{n,c}(A)\right]\right).$$

Pour tout $\varepsilon > 0$, posons

$$\theta_{n,c}(\varepsilon) = \sup\left\{|\alpha_{n,c}(A) - \alpha_{n,c}(B)| : A, B \in \mathbb{C} - \mathbf{a}, d_\lambda(A, B) < \varepsilon\right\}, \tag{6.26}$$

où d_λ est la pseudométrique définie en (6.8).

Remarque 6.4.3. Sous l'hypothèse (H.1), notons que

$$\mathbb{E}[\lambda_{n,c}(A)] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Y_i \mathbf{1}_{\{U_i \in A\}}\right] = \mathbb{E}[Y_1]\mathbb{E}[\mathbf{1}_{\{U_1 \in A\}}] = \lambda_0(A),$$

puisque les suites $\{U_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$ sont supposées indépendantes et identiquement distribuées.

Pour tout réel $h > 0$ et pour tout $C \in \mathbb{B}$, posons

$$\Theta_{n,c}(C,h) = \frac{1}{\sqrt{\mathbb{E}[Y^2]h}}\, \alpha_{n,c}(\mathbf{t} + h^{1/d}C) \tag{6.27}$$

et, pour tout $\varepsilon > 0$, posons également

$$\omega_{n,c}(\varepsilon,h) = \sup\left\{|\Theta_{n,c}(C,h) - \Theta_{n,c}(D,h)| : C, D \in \mathbb{C}\,, d_\lambda(C,D) < \varepsilon\right\}. \tag{6.28}$$

Propriété 6.4.2. *Supposons que les hypothèses (H.1–2) sont vérifiées. Les hypothèses (C.2) et (C.3) impliquent que*

$$\textit{la classe } \mathbb{C} \textit{ est totalement bornée par rapport à } d_\lambda, \tag{6.29}$$

et pour tout $\varepsilon > 0$ et $\gamma > 0$,

$$\lim_{\varepsilon \downarrow 0}\left(\limsup_{n\to\infty} \mathbb{P}\big[\theta_{n,c}(\varepsilon) > \gamma\big]\right) = 0. \tag{6.30}$$

Démonstration. Pour démontrer la propriété (6.29), on pourra consulter, par exemple, le théorème 1.2, p.903 de Dudley [48] et plus particulièrement l'article [47] de Dudley, ou encore le théorème B, p.113 de Gaenssler [57]. Pour établir le résultat (6.30), on utilisera de nouveau le théorème 1.2, p.903 de Dudley [48], l'hypothèse (H.2) et l'indépendance des deux suites de variables aléatoires $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ entre elles. □

6.4.2. Résultats sur le module d'oscillation du processus empirique composé local

Dans ce paragraphe, nous allons établir une borne supérieure uniforme pour les oscillations du processus $\{\Theta_{n,c}(C) : C \in \mathbb{C}\}$ par rapport à la pseudométrique d_λ définie en (6.8). Ce résultat, établi dans le lemme 6.4.7, sera un outil essentiel dans la démonstration du théorème 6.3.1.

Lemme 6.4.1. *Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur $[0,1]^d$. Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. Alors, pour m tel que $1 \leq m < n$, nous avons l'égalité en distribution*

$$\left\{ \sum_{i=1}^{m} Y_i \mathbf{1}_{\{U_i \in \mathbf{t}+h^{1/d}C\}} \middle| U_i \in \mathbf{t} + h^{1/d}[a,b]^d, i = 1,\ldots,m, \right.$$
$$\left. U_i \notin \mathbf{t} + h^{1/d}[a,b]^d, i = m+1,\ldots,n : C \in \mathbb{C} \right\} \stackrel{d}{=} \left\{ \sum_{i=1}^{m} Y_i \mathbf{1}_{\{U_i \in C - \mathbf{a}\}} : C \in \mathbb{C} \right\}.$$

Démonstration. Évident. □

Lemme 6.4.2. *Sous les hypothèses (K.1) et (H.1–2) et si la classe $\mathbb{C}$ vérifie (6.30), alors, pour tout $\eta > 0$, nous avons*

$$\lim_{\varepsilon \downarrow 0} \left\{ \limsup_{n\to\infty} \mathbb{P}\left[\omega_{n,c}\left(\varepsilon, \frac{k(n)}{n}\right) > \eta\right]\right\} = 0. \tag{6.31}$$

Démonstration. Pour tout $\varepsilon > 0$ et $\eta > 0$, posons

$$\delta_c(\varepsilon, \eta) = \limsup_{n\to\infty} \mathbb{P}\left[\theta_{n,c}(\varepsilon) > \eta/4\right].$$

Soit Z une variable aléatoire qui suit une loi normale standard $N(0,1)$. Pour tout $\varepsilon \in (0, 1/2)$, définissons $z(\varepsilon)$ par

$$\mathbb{P}\left[|Z| > z(\varepsilon)\right] = \varepsilon. \tag{6.32}$$

D'après le résultat (6.30), pour chaque $\eta > 0$, nous avons

$$\lim_{\varepsilon\downarrow 0} \delta_c(\varepsilon, \eta) = 0. \tag{6.33}$$

Comme

$$\mathbb{P}\left[|Z| > z\right] = (1 + o(1))\sqrt{\frac{2}{\pi}}\,\frac{1}{z}\exp\left(-\frac{1}{2}z^2\right) \qquad \text{lorsque } z \to \infty,$$

un calcul simple montre que

$$\lim_{\varepsilon\downarrow 0} \varepsilon z(\varepsilon) = 0. \tag{6.34}$$

Maintenant, pour tout $\varepsilon \in (0, 1/2)$ tel que $\eta\mathbb{E}[Y^2] - \varepsilon z(\varepsilon) > \eta/2$, nous allons établir que

$$\limsup_{n\to\infty} \mathbb{P}\left[\omega_{n,c}\left(\varepsilon, \frac{k(n)}{n}\right) > \eta\right] \le \delta_c(\varepsilon, \eta) + \varepsilon. \tag{6.35}$$

Cette dernière inégalité combinée avec (6.33) et (6.34) implique alors le résultat (6.31). Il reste donc à démontrer (6.35). Pour cela, introduisons

$$N_n = \#\left\{U_1, \ldots, U_n \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d}[a, b]^d\right\}.$$

La variable aléatoire N_n suit une loi binomiale $\mathrm{B}\,(n,\, k(n)/n)$ qui vérifie, sous l'hypothèse (K.1), [remarquons que (K.1) implique que $n(k(n)/n)(1 - k(n)/n) = (1 + o(1))k(n) \to \infty$ lorsque $n \to \infty$],

$$\frac{N_n - k(n)}{\sqrt{k(n)}} \xrightarrow{\mathrm{d}} N(0,1) \qquad \text{lorsque } n \to \infty, \tag{6.36}$$

où "$\xrightarrow{\mathrm{d}}$" signifie converge en loi. Nous avons alors l'inégalité suivante

$$\begin{aligned}
&\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,\frac{k(n)}{n}\right)>\eta\right]\\
&\quad\leq \sum_{m:|m-k(n)|\leq z(\varepsilon)\sqrt{k(n)}} \mathbb{P}\left[\omega_{n,c}\left(\varepsilon,\frac{k(n)}{n}\right)>\eta, N_n=m\right]\\
&\qquad+\mathbb{P}\left[|N_n-k(n)|>z(\varepsilon)\sqrt{k(n)}\right]\\
&\quad\leq \sum_{m:|m-k(n)|\leq z(\varepsilon)\sqrt{k(n)}} \mathbb{P}_{1,n}+\mathbb{P}_{2,n}.
\end{aligned}$$

D'après (6.36) et (6.32), nous en déduisons que

$$\limsup_{n\to\infty}\mathbb{P}_{2,n}=\varepsilon. \tag{6.37}$$

Étudions maintenant la quantité $\mathbb{P}_{1,n}$. D'après la définition de la probabilité conditionnelle $\mathbb{P}[A\cap B]=\mathbb{P}[A|B]\,\mathbb{P}[B]$, nous avons

$$\begin{aligned}
&\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,\frac{k(n)}{n}\right)>\eta, N_n=m\right]\\
&\quad=C_n^m\left(1-\frac{k(n)}{n}\right)^{n-m}\left(\frac{k(n)}{n}\right)^m\\
&\qquad\times\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,\frac{k(n)}{n}\right)>\eta\,\middle|\,U_i\in\mathbf{t}+\left(\frac{k(n)}{n}\right)^{1/d}[a,b]^d, i=1,\dots,m,\right.\\
&\qquad\qquad\left.U_i\notin\mathbf{t}+\left(\frac{k(n)}{n}\right)^{1/d}[a,b]^d, i=m+1,\dots,n\right],
\end{aligned}$$

puisque la variable N_n suit une loi binomiale $B(n,k(n)/n)$. Les deux inégalités suivantes

$$\lambda\left(C\Delta D\right)<\varepsilon \quad \text{et} \quad |m-k(n)|\leq z\left(\varepsilon\right)\sqrt{k(n)}$$

impliquent que

$$|\lambda(C)-\lambda(D)|\times|m-k(n)|\leq\varepsilon z(\varepsilon)\sqrt{k(n)}. \tag{6.38}$$

Ainsi, d'après le lemme 6.4.1 et l'inégalité (6.38), nous obtenons

$$\begin{aligned}
\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,\frac{k(n)}{n}\right)>\eta, N_n=m\right]\leq & C_n^m\left(1-\frac{k(n)}{n}\right)^{n-m}\left(\frac{k(n)}{n}\right)^m\\
&\mathbb{P}\left[\theta_{n,c}\left(\varepsilon\right)>\sqrt{\frac{k(n)}{m}}\left(\eta\mathbb{E}[Y^2]-\varepsilon z(\varepsilon)\right)\right].
\end{aligned}$$

Remarquons également que

$$\sqrt{\frac{k(n)}{m}} > \frac{1}{2} \quad \text{uniformément sur } |m-k(n)| \le z(\varepsilon)\sqrt{k(n)} \quad \text{lorsque } n \to \infty.$$

Comme nous avons choisi ε tel que $\eta\mathbb{E}[Y^2] - \varepsilon z(\varepsilon) > \eta/2$, nous avons, uniformément sur $|m-k(n)| \le z(\varepsilon)\sqrt{k(n)}$, que

$$\mathbb{P}\left[\omega_{n,c}\left(\varepsilon, \frac{k(n)}{n}\right) > \eta, N_n = m\right] \le \mathbb{P}\left[N_n = m\right]\mathbb{P}\left[\theta_{n,c}\left(\varepsilon\right) > \eta/4\right].$$

Ainsi nous en déduisons que

$$\begin{aligned}\sum_{m:|m-k(n)|\le z(\varepsilon)\sqrt{k(n)}} &\mathbb{P}\left[\omega_{n,c}\left(\varepsilon, \frac{k(n)}{n}\right) > \eta, N_n = m\right] \\ &\le \mathbb{P}\left[\theta_{n,c}\left(\varepsilon\right) > \eta/4\right] \sum_{m:|m-k(n)|\le z(\varepsilon)\sqrt{k(n)}} \mathbb{P}\left[N_n = m\right] \\ &\le \mathbb{P}\left[\theta_{n,c}\left(\varepsilon\right) > \eta/4\right].\end{aligned}$$

Le passage à la "$\limsup_{n\to\infty}$" sur cette inégalité et (6.37) impliquent le résultat (6.35). La démonstration du lemme 6.4.2 est maintenant complète. □

Soit $\{s_i : i \ge 1\}$ une suite de signes aléatoires indépendants et identiquement distribués. De plus, nous supposerons que la suite $\{s_i : i \ge 1\}$ est indépendante des suites de variables aléatoires $\{U_i : i \ge 1\}$ et $\{Y_i : i \ge 1\}$ et est telle que $\mathbb{P}[s_n = 1] = \mathbb{P}[s_n = -1] = \frac{1}{2}$. Posons, pour tout $C \in \mathbb{C}$,

$$S_{n,c}(C) = \sum_{i=1}^{n} s_i Y_i \mathbf{1}_{\{U_i \in \mathbf{t}+C\}}. \tag{6.39}$$

Remarque 6.4.4. Comme les suites $\{s_i : i \ge 1\}$, $\{Y_i : i \ge 1\}$ et $\{U_i : i \ge 1\}$ sont supposées indépendantes, notons que

$$\mathbb{E}[S_{n,c}(C)] = \sum_{i=1}^{n} \mathbb{E}[s_i]\mathbb{E}[Y_i]\mathbb{E}[\mathbf{1}_{\{U_i \in \mathbf{t}+C\}}] = 0,$$

car pour tout $i \ge 1$, nous avons

$$\mathbb{E}[s_i] = 1 \times \frac{1}{2} - 1 \times \frac{1}{2} = 0.$$

Lemme 6.4.3. *Sous les hypothèses (H.1–2), pour tout $\eta > 0$, l'hypothèse (C.3) implique que*

$$\lim_{\varepsilon\downarrow 0}\left\{\limsup_{n\to\infty}\mathbb{P}\left[\sup\left\{\frac{|S_{n,c}(C) - S_{n,c}(D)|}{\sqrt{\mathbb{E}[Y^2]n}} : d_\lambda(C,D) < \varepsilon\right\} > \eta\right]\right\} = 0. \tag{6.40}$$

Démonstration. Pour de plus amples détails sur la démonstration de ce lemme, on pourra consulter le théorème 2.14 p.941 de l'article de Giné et Zinn [58]. □

Pour tout $\varepsilon > 0$ et pour tout $C \in \mathbb{C}$, posons

$$\Delta_{n,c}(C,\varepsilon) = \sup\{|S_{n,c}(C) - S_{n,c}(D)| : D \in \mathbb{C}, d_\lambda(C,D) < \varepsilon\}. \tag{6.41}$$

Lemme 6.4.4. *Sous l'hypothèse (H.2), pour tout* $\varepsilon > 0$ *il existe un entier* $N = N(\varepsilon) > 0$ *tel que, pour tout entier* $k = 1, 2, \cdots, \lfloor n/N \rfloor$, *et pour tous réels* $\lambda > 0$ *et* $b > 0$, *nous ayons*

$$\mathbb{P}\left[\Delta_{kN,c}(C,\varepsilon) \geq 2\lambda b\right] \leq \exp\left(-\lambda^2 + \frac{\lambda^2}{2}\phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, k, \lambda, b, \varepsilon\right)\right), \tag{6.42}$$

où

$$\phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, k, \lambda, b, \varepsilon\right) = \left\{\frac{\sqrt{17\varepsilon^2\mathbb{E}[Y^2]kN}}{\lambda b} + \frac{1}{b^2}\exp\left(\frac{2N\left(\mathcal{M}+1\right)\lambda}{b}\right)17\varepsilon^2\mathbb{E}[Y^2]kN\right\}. \tag{6.43}$$

Démonstration. La démonstration de ce lemme se trouve dans le paragraphe "Appendice". □

Remarque 6.4.5. Pour tout entier $n \geq N$ et pour tous réels $\lambda > 0$ et $b > 0$, l'inégalité (6.42) conduit à

$$\mathbb{P}\left[\Delta_{n,c}(C,\varepsilon) \geq 2\lambda b + N\right] \leq \exp\left(-\lambda^2 + \frac{\lambda^2}{2}\phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, \frac{n}{N}, \lambda, b, \varepsilon\right)\right), \tag{6.44}$$

où nous utilisons le fait que $\Delta_{n,c}(C,\varepsilon) \leq \Delta_{kN,c}(C,\varepsilon) + N$, où $k = \lfloor n/N \rfloor$ et $\lfloor u \rfloor$ représente la partie entière de u.

Pour tout réel $h > 0$ et pour tout $C \in \mathbb{C}$, posons

$$T_{n,c}(C,h) = \sum_{i=1}^{n} Y_i \left\{\mathbf{1}_{\{U_i \in \mathbf{t} + h^{1/d}C\}} - \mathbf{1}_{\{U_i' \in \mathbf{t} + h^{1/d}C\}}\right\}, \tag{6.45}$$

où $\{U_i' : i \geq 1\}$ est une suite de variables aléatoires indépendantes et uniformément distribuées sur $[0,1]^d$. De plus, les suites de variables aléatoires $\{U_i' : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. Posons également

$$D_{n,c}(C,h,\varepsilon) = \sup\left\{|T_{n,c}(C,h) - T_{n,c}(D,h)| : D \in \mathbb{C}, d_\lambda(C,D) < \varepsilon\right\}. \tag{6.46}$$

Lemme 6.4.5. *Supposons que les hypothèses (C.1–2–3–4), (K.1)(ii), (K.1)(iii) et (K.2) sont vérifiées. Alors pour tout* $0 < \varepsilon < 1$ *et* $C \in \mathbb{C}$, *il existe une constante* $K \leq 832\mathcal{M}^2 + 128$ *telle que nous ayons*

$$\limsup_{n\to\infty} \frac{D_{n,c}\left(C, \frac{k(n)}{n}, \frac{\varepsilon}{2}\right)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}} \leq K\varepsilon \quad p.s. \tag{6.47}$$

Démonstration. Posons $n_r = \lfloor \rho^r \rfloor$, où $r \geq 0$ et ρ est tel que $1 < \rho < 2$. D'après les hypothèses (K.1)(ii), (K.1)(iii), (C.4), le résultat (6.7), et l'inégalité suivante

$$\lambda\big((hC)\Delta D\big) \leq \lambda\big((hC)\Delta C\big) + \lambda(C\Delta D),$$

pour ρ tel que $1 < \rho < 2$ qui dépend de $\varepsilon > 0$, nous obtenons que

$$\max_{n_r < n \leq n_{r+1}} \frac{D_{n,c}(C, k(n)/n, \varepsilon/2)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}} \leq 2 \max_{n_r < n \leq n_{r+1}} \frac{D_{n,c}(C, k(n_r)/n_r, \varepsilon)}{\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}}.$$

Maintenant, d'après une généralisation de l'inégalité de Lévy [voir le paragraphe "Appendice", lemme 6.6.1], pour tout $\zeta > 0$ et $\varepsilon > 0$, nous en déduisons que

$$\begin{aligned}
&\mathbb{P}\left[\max_{n_r < n \leq n_{r+1}} \frac{D_{n,c}(C, k(n_r)/n_r, \varepsilon)}{\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}} > 8\zeta\varepsilon\right] \\
&\quad \leq 2\mathbb{P}\left[\frac{D_{n_{r+1},c}(C, k(n_r)/n_r, \varepsilon)}{\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}} > 8\zeta\varepsilon\right] = 2\mathbb{P}(r).
\end{aligned} \tag{6.48}$$

Pour $r \geq 1$, posons

$$\overline{N}_r = \#\left\{U_1, \ldots, U_{n_{r+1}} \in \mathbf{t} + \left(\frac{k(n_r)}{n_r}\right)^{1/d}[a,b]^d\right\}$$

et

$$\overline{N'}_r = \#\left\{U'_1, \ldots, U'_{n_{r+1}} \in \mathbf{t} + \left(\frac{k(n_r)}{n_r}\right)^{1/d}[a,b]^d\right\}.$$

Notons que les variables aléatoires $\overline{N}_r$ et $\overline{N'}_r$ sont indépendantes et suivent une loi binomiale $B(n_{r+1}, k(n_r)/n_r)$. Pour tout réel $h > 0$ et pour tout $C \in \mathbb{C}$, posons

$$\overline{T}_{n,c}(C,h) = \sum_{i=1}^{n} s_i Y_i \left(\mathbf{1}_{\{U_i \in \mathbf{t} + h^{1/d}C\}} - \mathbf{1}_{\{U'_i \in \mathbf{t} + h^{1/d}C\}}\right),$$

où $\{s_i : i \geq 1\}$ est une suite de signes aléatoires définie en (6.39), indépendante des suites $\{U_i : i \geq 1\}$, $\{U'_i : i \geq 1\}$ et $\{Y_i : i \geq 1\}$. Comme nous avons l'égalité en distribution

$$\left\{\overline{T}_{n,c}(C,h) : C \in \mathbb{C}\right\} \stackrel{d}{=} \{T_{n,c}(C,h) : C \in \mathbb{C}\},$$

nous obtenons alors l'égalité en distribution suivante

$$\left\{\overline{D}_{n,c}(C,h,\varepsilon) : C \in \mathbb{C}\right\} \stackrel{d}{=} \{D_{n,c}(C,h,\varepsilon) : C \in \mathbb{C}\},$$

où

$$\overline{D}_{n,c}(C,h,\varepsilon) = \sup\left\{|\overline{T}_{n,c}(C,h) - \overline{T}_{n,c}(D,h)| : D \in \mathbb{C}, d_\lambda(C,D) < \varepsilon\right\}. \tag{6.49}$$

Pour $\{i_1, \dots, i_m, j_1, \dots, j_{m'}\} \subseteq \{1, \dots, n_{r+1}\}$, posons

$$\begin{aligned}\mathbb{P}'_r(\varepsilon) = \mathbb{P}\Big[&\overline{D}_{n_{r+1},c}\,(C, k(n_r)/n_r, \varepsilon) \\ &> 8\zeta\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}\,\Big|A\big(i_1, \dots, i_m, j_1, \dots, j_{m'}\big)\Big],\end{aligned} \tag{6.50}$$

où $A\big(i_1, \dots, i_m, j_1, \dots, j_{m'}\big)$ est l'événement suivant

$$\begin{array}{lcll} U_i & \in & \mathbf{t} + \left(\dfrac{k(n_r)}{n_r}\right)^{1/d} [a,b]^d & \text{pour } i \in \{i_1, \dots, i_m\}, \\ U_i & \notin & \mathbf{t} + \left(\dfrac{k(n_r)}{n_r}\right)^{1/d} [a,b]^d & \text{pour } i \in \{1, \dots, n_{r+1}\} - \{i_1, \dots, i_m\}, \\ U'_j & \in & \mathbf{t} + \left(\dfrac{k(n_r)}{n_r}\right)^{1/d} [a,b]^d & \text{pour } j \in \{j_1, \dots, j_{m'}\}, \\ U'_j & \notin & \mathbf{t} + \left(\dfrac{k(n_r)}{n_r}\right)^{1/d} [a,b]^d & \text{pour } j \in \{1, \dots, n_{r+1}\} - \{j_1, \dots, j_{m'}\}. \end{array}$$

De ceci, nous en déduisons que

$$\begin{aligned}\mathbb{P}'_r(\varepsilon) &= \mathbb{P}\Bigg[\sup\Bigg\{\Bigg|\sum_{k=1}^{m} s_{i_k}Y_{i_k}\left(\mathbf{1}_{\{U_{i_k}\in C\}} - \mathbf{1}_{\{U_{i_k}\in D\}}\right) - \sum_{l=1}^{m'} s_{j_l}Y_{j_l}\left(\mathbf{1}_{\{U'_{j_l}\in C\}} - \mathbf{1}_{\{U'_{j_l}\in D\}}\right)\Bigg| \\ &\qquad > 8\,\zeta\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r} : D \in \mathbb{C}\,, d_\lambda(C,D) < \varepsilon\Bigg\}\Bigg] \\ &\le \mathbb{P}\left[\Delta_{m,c}(C,\varepsilon) > 4\,\zeta\,\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}\right] \\ &\quad + \mathbb{P}\left[\Delta_{m',c}(C,\varepsilon) > 4\,\zeta\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2 n_r}\right] = \mathbb{P}(r,m) + \mathbb{P}(r,m').\end{aligned} \tag{6.51}$$

D'après (6.48), (6.49), (6.50) et (6.51), il s'ensuit que

$$\begin{aligned}\mathbb{P}(r) \le &\sum_{k(n_r)\le m,m'\le 3k(n_r)} \left(\mathbb{P}(r,m) + \mathbb{P}\,(r,m')\right)\mathbb{P}\left[\,\overline{N}_r = m\right]\mathbb{P}\left[\,\overline{N}'_r = m'\,\right] \\ &+ 2\mathbb{P}\left[\,\overline{N}'_r > 3k(n_r) \text{ ou } \overline{N}_r < k(n_r)\right] = \mathbb{P}_1(r) + \mathbb{P}_2(r).\end{aligned} \tag{6.52}$$

Étudions d'abord $\mathbb{P}_2(r)$. Rappelons que la variable aléatoire $\overline{N}_r$ suit une loi binomiale $B\,(n_{r+1}, k(n_r)/n_r)$. À ce propos, énonçons une inégalité classique.

Inégalité 6.4.1. *Soit Z une variable aléatoire qui suit une loi binomiale $B(n,p)$. Alors, pour tout $\gamma > 0$, et pour tout p tel que $0 < p \le 1/2$, nous avons*

$$\mathbb{P}\left[\,|Z - np| \ge \gamma\sqrt{n}\,\right] \le 2\exp\left(-\frac{\gamma^2}{2p}\psi\left(\frac{\gamma}{p\sqrt{n}}\right)\right), \tag{6.53}$$

où la fonction ψ est définie par

$$\psi(u) = \frac{2}{u^2}\left\{(1+u)\log(1+u) - u\right\} \qquad \textit{pour } u > 0. \tag{6.54}$$

Démonstration. La démonstration de cette inégalité se trouve, par exemple, dans le livre de Shorack et Wellner [125], inégalité 1, p.440. $\square$

Dans (6.53), remplaçons n par n_{r+1}, p par $k(n_r)/n_r$ et

$$\gamma \text{ par } k(n_r)/\sqrt{n_{r+1}} \min\left\{\big(3 - (n_{r+1}/n_r)\big), \big((n_{r+1}/n_r) - 1\big)\right\} = k(n_r)/\sqrt{n_{r+1}}(\rho - 1),$$

puisque par définition $\rho = n_{r+1}/n_r \in (1,2)$. Par conséquent, nous obtenons que

$$\frac{\gamma^2}{2p} = \frac{(\rho-1)^2}{2\rho}\,k(n_r) \quad \text{ et } \quad \frac{\gamma}{p\sqrt{n}} = \frac{\rho - 1}{\rho}.$$

Ainsi nous en déduisons que

$$\mathbb{P}_2(r) \leq 4\exp\left(-\frac{(\rho-1)^2}{2\rho}\,k(n_r)\psi\left(\frac{\rho-1}{\rho}\right)\right) = 4\exp\left(-c_1 k(n_r)\right), \tag{6.55}$$

où l'on pose

$$c_1 = \frac{(\rho-1)^2}{2\rho}\psi\left(\frac{\rho-1}{\rho}\right).$$

L'hypothèse (K.2) implique que, à partir d'un certain rang,

$$k(n_r) \geq (3/c_1)\log_2 n_r \geq 2\log r,$$

ce qui, d'après (6.55), implique que

$$\mathbb{P}_2(r) \leq \frac{4}{r^2}.$$

Ainsi nous en déduisons que

$$\sum_r \mathbb{P}_2(r) < \infty. \tag{6.56}$$

Maintenons intéressons nous à la quantité $\mathbb{P}_1(r)$. Soit $N = N(\varepsilon)$ défini comme en (6.44). Comme $\sqrt{k(n_r)\log_2(n_r)} \to \infty$ lorsque $r \to \infty$, nous avons, pour r assez grand et m tel que $k(n_r) \leq m \leq 3k(n_r)$,

$$\begin{aligned}\mathbb{P}(r,m) &= \mathbb{P}\left[\Delta_{m,c}(C,\varepsilon) > 4\,\zeta\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2(n_r)}\right] \\ &\leq \mathbb{P}\left[\Delta_{m,c}(C,\varepsilon) > 2\,\zeta\varepsilon\sqrt{\mathbb{E}[Y^2]k(n_r)\log_2(n_r)} + N\right].\end{aligned} \tag{6.57}$$

En remplaçant λ par $\lambda_r = \sqrt{\zeta \log_2 n_r}$, k par m/N et b par $b_r = \varepsilon\sqrt{\zeta k(n_r)}$ dans (6.43), pour r assez grand et m tel que $k(n_r) \leq m \leq 3k(n_r)$, nous obtenons que

$$\phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, \frac{m}{N}, \lambda_r, b_r, \varepsilon\right) \leq \left\{\frac{\sqrt{51\mathbb{E}[Y^2]}}{\zeta\sqrt{\log_2 n_r}} + \frac{51\mathbb{E}[Y^2]}{\zeta} \exp\left(\frac{2N(\mathcal{M}+1)}{\varepsilon}\sqrt{\frac{\log_2 n_r}{k(n_r)}}\right)\right\}. \tag{6.58}$$

D'après l'hypothèse (K.2), nous obtenons que

$$\phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, \frac{m}{N}, \lambda_r, b_r, \varepsilon\right) \leq \frac{52\mathbb{E}[Y^2]}{\zeta} \quad \text{lorsque } r \to \infty.$$

Maintenant choissisons ζ tel que $\zeta = 52\mathbb{E}[Y^2] + 8$. En combinant (6.44), (6.57) et (6.58), pour r assez grand et m tel que $k(n_r) \leq m \leq 3k(n_r)$, nous en déduisons que

$$\begin{aligned}\mathbb{P}[r,m] &\leq \exp\left(\lambda_r^2\left(-1 + \phi\left(\mathbb{E}[Y^2], \mathcal{M}, N, \frac{m}{N}, \lambda_r, b_r, \varepsilon\right)\right)\right) \\ &\leq \exp\left((26\mathbb{E}[Y^2] - \zeta)\log_2 n_r\right) = \exp\left(-4\log_2 n_r\right) \\ &\leq \frac{1}{r^3}.\end{aligned} \tag{6.59}$$

D'après (6.52) et (6.59), nous avons alors

$$\mathbb{P}_1(r) \leq \frac{2}{r^2} \quad \text{pour } r \text{ assez grand.}$$

Cette inégalité implique alors que

$$\sum_r \mathbb{P}_1(r) < \infty.$$

De plus, d'après (6.48), (6.52) et (6.56), nous concluons grâce au lemme de Borel-Cantelli que l'inégalité (6.47) est vérifiée en prenant $K = 16\,\zeta = 832\mathcal{M}^2 + 128$. Ceci complète la preuve du lemme 6.4.5. □

Nous allons utiliser une procédure de symétrisation qui figure dans le lemme 2.1 de l'article de Ledoux et Talagrand [88] (on peut également consulter le lemme 3.16 de l'article de Kuelbs et Dudley [75]).

Lemme 6.4.6. *Soit* $\mathbf{X}$ *un espace vectoriel topologique de Hausdorff. Soient* $\mathbf{B}$ *les ensembles correspondant de la* σ*-algèbre de Borel. Soit* $||\cdot||$ *une semi-norme* $\mathbf{B}$*-mesurable sur* $\mathbf{X}$*. Soient* $\{Z_n : n \geq 1\}$ *et* $\{Z'_n : n \geq 1\}$ *deux suites indépendantes et identiquement distribuées de variables aléatoires à valeurs dans* $\mathbf{X}$ *telles que la suite* $\{Z_n - Z'_n : n \geq 1\}$ *est presque sûrement bornée (resp. convergente vers* 0*) et* $\{Z'_n : n \geq 1\}$ *est bornée (resp. convergente vers* 0*) en probabilité. Alors* $\{Z_n : n \geq 1\}$ *est presque sûrement bornée (resp. convergente vers* 0*). De plus, si, pour tous réels* M *et* A*, nous avons*

$$\limsup_{n\to\infty} ||Z_n - Z'_n|| \leq M \quad p.s. \tag{6.60}$$

et

$$\limsup_{n\to\infty} \mathbb{P}\,[\,||Z'_n|| > A] < 1, \tag{6.61}$$

alors nous avons

$$\limsup_{n\to\infty} ||Z_n|| \leq 2M + A \quad p.s. \tag{6.62}$$

Démonstration. On pourra consulter à ce propos le lemme 2.3 p.1631 de l'article de Deheuvels et Mason [36]. □

Soit $\Theta_{n,c}(C,h)$ défini en (6.27). Pour $C \in \mathbb{C}$, $h > 0$, $\varepsilon > 0$ et $n \geq 1$, définissons

$$W_{n,c}(C,h,\varepsilon) = \sup\{|\Theta_{n,c}(C,h) - \Theta_{n,c}(D,h)| : D \in \mathbb{C}\,, d_\lambda(C,D) < \varepsilon\}. \tag{6.63}$$

Lemme 6.4.7. *Supposons que les conditions (C.1–2–3–4), (K.1)(ii), (K.1)(iii) et (K.2) soient vérifiées. Alors il existe un $\varepsilon_0 > 0$ tel que, pour tout $0 < \varepsilon < \varepsilon_0$, nous ayons*

$$\limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{C}} \frac{W_{n,c}\left(C, \dfrac{k(n)}{n}, \dfrac{\varepsilon}{2}\right)}{\sqrt{\log_2 n}} \right\} \leq 2K\varepsilon \quad p.s. \tag{6.64}$$

Démonstration. Nous allons appliquer le lemme 6.4.6 aux deux suites définies respectivement par

$$Z_{n,c} = \frac{1}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}} \sum_{i=1}^{n} \left\{ Y_i \mathbf{1}_{\left\{U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right\}} - k(n)\lambda_0(C) \right\}$$

et

$$Z'_{n,c} = \frac{1}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}} \sum_{i=1}^{n} \left\{ Y_i \mathbf{1}_{\left\{U'_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right\}} - k(n)\lambda_0(C) \right\}.$$

Remarquons que les variables aléatoires $Z_{n,c}$ et $Z'_{n,c}$ sont à valeurs dans $B(\mathbb{C})$. Pour tout $\varepsilon > 0$ et $\phi \in B(\mathbb{C})$, posons

$$||\phi||_\varepsilon = \sup\{|\phi(C) - \phi(D)| : C \in \mathbb{C}\,, D \in \mathbb{C}\,, d_\lambda(C,D) < \varepsilon\}. \tag{6.65}$$

Selon la propriété 6.4.1, les hypothèses $(C.1)$, $(C.2)$ et$(C.4)$ impliquent que $||.||_\varepsilon$ est une semi-norme mesurable sur $B(\mathbb{C})$. D'après (6.45), (6.46) et de (6.65), nous obtenons que

$$\limsup_{n\to\infty} ||Z_{n,c} - Z'_{n,c}||_\varepsilon = \limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{C}} \frac{D_{n,c}\left(C, k(n)/n, \varepsilon\right)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}} \right\}. \tag{6.66}$$

En utilisant la propriété 6.4.1 puis le lemme 6.4.5, nous remarquons que si $\mathbb{G}$ est une sous-classe dénombrable de $\mathbb{C}$, alors nous obtenons, presque sûrement,

$$\begin{aligned}\limsup_{n\to\infty}\left\{\sup_{C\in\mathbb{C}}\frac{D_{n,c}\left(C,k(n)/n,\varepsilon\right)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}}\right\}&=\limsup_{n\to\infty}\left\{\sup_{C\in\mathbb{G}}\frac{D_{n,c}\left(C,k(n)/n,\varepsilon\right)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}}\right\}\\&=\sup_{C\in\mathbb{G}}\left\{\limsup_{n\to\infty}\frac{D_{n,c}\left(C,k(n)/n,\varepsilon\right)}{\sqrt{\mathbb{E}[Y^2]k(n)\log_2 n}}\right\}\leq K\varepsilon.\end{aligned}\tag{6.67}$$

Ainsi, en combinant (6.66) et (6.67), nous obtenons que

$$\limsup_{n\to\infty}||Z_{n,c}-Z'_{n,c}||_\varepsilon\leq K\varepsilon\quad\text{p.s.}\tag{6.68}$$

D'autre part, d'après (6.27) et (6.28), nous remarquons que

$$||Z_{n,c}||_\varepsilon=\frac{\omega_{n,c}\left(\varepsilon,k(n)/n\right)}{\sqrt{k(n)\log_2 n}}.\tag{6.69}$$

Soient $A>0$ et $\eta>0$ choisis arbitrairement. Le lemme 6.4.2 appliqué à (6.69), implique que, en choisissant, indépendamment de A, un $\varepsilon(\eta)>0$ suffisamment petit, pour tout ε tel que $0\leq\varepsilon\leq\varepsilon(\eta)$, nous obtenons alors

$$\begin{aligned}\limsup_{n\to\infty}\mathbb{P}\left[\,||Z_{n,c}||_\varepsilon>A\right]&=\limsup_{n\to\infty}\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,k(n)/n\right)>A\sqrt{k(n)\log_2 n}\right]\\&\leq\limsup_{n\to\infty}\mathbb{P}\left[\omega_{n,c}\left(\varepsilon,k(n)/n\right)>\eta\,\right]<1.\end{aligned}\tag{6.70}$$

D'après (6.68), (6.70) et le lemme 6.4.6, pour tout ε tel que $0<\varepsilon\leq\varepsilon_0=\varepsilon(\eta)$, il s'ensuit que,

$$\limsup_{n\to\infty}\left\{\sup_{C\in\mathbb{C}}\frac{W_{n,c}\left(C,k(n)/n,\varepsilon/2\right)}{\sqrt{\log_2 n}}\right\}=\limsup_{n\to\infty}||Z_{n,c}||_\varepsilon\leq 2K\varepsilon+A\qquad\text{p.s.}\tag{6.71}$$

Comme A est arbitraire, nous obtenons directement le résultat (6.64) en appliquant (6.71) à une suite $A_n\downarrow 0$. Ceci complète la preuve du lemme 6.4.7. □

6.4.3. Démonstration du théorème 6.3.1

Commençons par établir un cas particulier du théorème 6.3.1 en considérant le comportement limite

$$\left\{\frac{\Theta_{n,c}(C_i)}{\sqrt{2\log_2 n}}:1\leq i\leq M\right\},$$

où l'entier $M\geq 1$ et les intervalles $C_1,\cdots,C_M\in\mathbb{C}$ sont fixés, et $\lambda(C_i)>0$ pour $i=1,\cdots,M$. Ce résultat est donné dans le lemme 6.4.13 ci-dessous. Avant d'établir ce lemme, nous avons besoin des résultats préliminaires qui vont suivre. Nous supposons, maintenant, que les hypothèses du théorème 6.3.1 sont vérifiées.

Choisissons $D_1, \ldots, D_N$ comme dans l'hypothèse (C.5). Soit

$$X_{n,c}(j) = \frac{\Theta_{n,c}(D_j)}{\sqrt{2\nu_j \log_2 n}} = \sqrt{\frac{n}{2\mathbb{E}[Y^2]\nu_j k(n) \log_2 n}}\alpha_{n,c}\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} D_j\right). \tag{6.72}$$

Posons alors

$$\mathbb{X}_{n,c} = (X_{n,c}(1), \ldots, X_{n,c}(N)) \in \mathbb{R}^N.$$

Nous allons décrire le comportement limite du vecteur $\mathbb{X}_{n,c}$ dans les lemmes 6.4.11 et 6.4.12. Les démonstrations de ces deux lemmes nécessitent deux résultats techniques établis dans les lemmes 6.4.8 et 6.4.10.

Soit $\{\Pi_c(t) : t \geq 0\}$ un processus de Poisson composé continu à droite tel que $\mathbb{E}[\Pi_c(t)] = \mathbb{E}[Y]\mathbb{E}[\Pi(t)] = \mathbb{E}[\Pi(t)]$, puisque l'hypothèse (H.1) est vérifiée. Pour $j = 1, \ldots, N$, soit

$$Y_{n,c}(j) = \sqrt{\frac{1}{2\nu_j \mathbb{E}[Y^2] k(n) \log_2 n}}\left(\Pi_c\left(k(n)\sum_{i=1}^{j}\nu_i\right) - \Pi_c\left(k(n)\sum_{i=1}^{j-1}\nu_i\right) - k(n)\nu_j\right). \tag{6.73}$$

Posons alors

$$\mathbb{Y}_{n,c} = (Y_{n,c}(1), \ldots, Y_{n,c}(N)) \in \mathbb{R}^N.$$

Lemme 6.4.8. *Lorsque* $n \geq 5$ *et*

$$\frac{k(n)}{n}\sum_{i=1}^{N}\nu_i \leq \frac{1}{2},$$

nous avons, pour tout sous-ensemble de Borel A *de* $\mathbb{R}^N$*,*

$$\mathbb{P}\left[\mathbb{X}_{n,c} \in A\right] \leq 2\,\mathbb{P}\left[\mathbb{Y}_{n,c} \in A\right].$$

Démonstration. En utilisant le fait que la distribution de $\mathbb{Y}_{n,c}$, sachant que $\Pi(n) = n$, est égale à la distribution de $\mathbb{X}_{n,c}$, la démonstration est similaire à celle du théorème 5.3 du chapitre 3. □

Lemme 6.4.9. *Soit* $N \geq 1$ *fixé. Soient* $L = L(n)$ *et* $n \geq 1$ *des entiers positifs. Soit* $\mathbf{R} = (R_1, \ldots, R_N)$ *un vecteur de loi multinomiale de la forme*

$$\mathbb{P}\left[R_1 = r_1, \ldots, R_N = r_N\right] = \frac{L!}{r_1! \ldots r_N! \left(L - \sum_{j=1}^{N} r_j\right)!} p_1^{r_1} \ldots p_N^{r_N} \left(1 - \sum_{j=1}^{N} p_j\right)^{L - \sum_{j=1}^{N} r_j},$$

pour les entiers $r_1 \geq 0, \ldots, r_N \geq 0$ *et* $\sum_{j=1}^{N} r_j \leq L$*. De plus, soit*

$$p_j = \frac{k(n)}{n}\nu_j \quad \text{et} \quad r_j = Lp_j + \delta_j\sqrt{2np_j \log_2 n} \quad \text{pour } j = 1, \ldots, N,$$

et supposons que n/L et $\delta_j = \delta_{n,j}$ pour $j = 1, \ldots, N$ sont bornés lorsque $n \to \infty$. Alors, pour tout $\varepsilon > 0$, il existe une constante $C > 0$ telle que nous ayons, pour n suffisamment grand,

$$\mathbb{P}\left[R_1 = r_1, \ldots, R_N = r_N\right] \geq Ck(n)^{-N/2} \exp\left(-\left(1-\varepsilon\right)\frac{n}{L}\left(\sum_{j=1}^{N}\delta_j^2\right)\log_2 n\right). \qquad (6.74)$$

Démonstration. Nous commençons la démonstration de ce lemme dans le cas d'une variable binomiale. Dans ce cas, le lemme 6.4.9 devient :
Soit R une variable binomiale de paramètres $L = L(n)$ et p. De plus, posons

$$p = \frac{k(n)}{n}\nu \quad \text{ et } \quad r = Lp + \delta_n\sqrt{2np\log_2 n}.$$

De plus, supposons que n/L et δ_n sont bornés lorsque $n \to \infty$. Alors, pour tout $\varepsilon > 0$, il existe une constante $C_1 > 0$ telle que, pour n suffisamment grand, nous ayons

$$\mathbb{P}\left[R = r\right] \geq Ck(n)^{-1/2} \exp\left(-\left(1-\varepsilon\right)\frac{n}{L}\delta^2\log_2 n\right). \qquad (6.75)$$

Démonstration. Commmençons par rappeler quelques résultats dûs essentiellement à Höglund [69].

Théorème 6.4.1. *Supposons que*

(F.1) la distribution de la variable aléatoire X est non dégénérée, i.e. $\mathbb{P}[X = x] < 1$ pour $x \in \mathbb{R}$;

(F.2) $t_0 = \sup\{t : \psi(t) = \mathbb{E}[\exp(tX)] < \infty\} > 0$;

(F.3) $\mathbb{E}[X] = \mu \in \mathbb{R}$ existe ;

(F.4) $t_1 = \inf\{t : \psi(t) = \mathbb{E}[\exp(tX)] < \infty\} < 0$.

Posons $\mathrm{Var}[X] = \sigma^2 < \infty$ *(avec $\sigma > 0$). Alors, uniformément pour tout $1 \leq x = O(\sqrt{n})$, lorsque $n \to \infty$, nous avons*

$$\mathbb{P}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq x\right] = \left[1 - \Phi(x)\right]\exp\left(\frac{x^2}{2} - n\Psi\left(\mu + \frac{\sigma x}{\sqrt{n}}\right) + O\left(\frac{x}{\sqrt{n}}\right)\right),$$

où la fonction Ψ est définie par

$$\Psi\left(u\right) = \sup_{t \in I_0}\left\{ut - \log\left(\psi(u)\right)\right\}, \qquad (6.76)$$

où $I_0 = \{t : \psi(t) < \infty\}$.

Démonstration. Voir le théorème A.(i) p.107 de l'article de Höglund [69]. □

Propriété 6.4.3. *La fonction τ qui à x associe $[1 - \Phi(\sqrt{x})]\exp(x/2)$ est strictement monotone, vérifie $\tau(0) = 1/2$ et*

$$1 - \frac{1}{x} < \tau(x)\sqrt{2\pi x} < 1 \quad \text{pour } x > 0. \qquad (6.77)$$

Démonstration. Voir l'article de Höglund [69], p.115 pour la première partie de cette propriété. Pour les inégalités (6.77), nous renvoyons au livre de Feller [54], p.166. □

Corollaire 6.4.1. *Supposons que*

(F.1) la distribution de la variable aléatoire X *est non dégénérée, i.e.* $\mathbb{P}[X = x] < 1$ *pour* $x \in \mathbb{R}$ *;*

(F.2) $t_0 = \sup t : \psi(t) = \mathbb{E}[\exp(tX)] < \infty > 0$ *;*

(F.3) $\mathbb{E}[X] = \mu \in \mathbb{R}$ *existe ;*

(F.4) $t_1 = \inf \{t : \psi(t) = \mathbb{E}[\exp(tX)] < \infty\} < 0$.

Posons $\mathrm{Var}[X] = \sigma^2 < \infty$ *(avec* $\sigma > 0$*). Alors, uniformément pour tout* $1 \leq x = O(\sqrt{n})$*, lorsque* $n \to \infty$*, nous avons*

$$\mathbb{P}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq x\right] \geq \frac{\mathcal{C}}{\sqrt{n}} \exp\left(-n\Psi\left(\mu + \frac{\sigma x}{\sqrt{n}}\right)\right),$$

où $\mathcal{C}$ *est une constante convenable et où la fonction* Ψ *est définie en* (6.76).

Démonstration. On se sert du théorème 6.4.1 et des inégalités (6.77). □

Nous avons maintenant tous les outils à notre disposition pour démontrer l'inégalité 6.74 dans le cas d'une variable binomiale R de paramètres L et p. Ainsi, nous obtenons que

$$\begin{aligned}
\mathbb{P}[R = r] &= \mathbb{P}\left[\frac{R - Lp}{\sqrt{Lp(1-p)}} = \frac{r - Lp}{\sqrt{Lp(1-p)}}\right] \\
&\geq \frac{\mathcal{C}}{\sqrt{k(n)}} \exp(-\Psi(r)), \quad \text{en appliquant le corollaire 6.4.1 avec } n = 1.
\end{aligned}$$

D'autre part, d'après l'article de Deheuvels et Steinebach [38], nous savons que lorsque $\alpha \to 0$

$$\Psi(\mu + \alpha) = \frac{\alpha^2}{2\sigma^2} + o(\alpha^2). \tag{6.78}$$

Ainsi, en utilisant (6.78) nous obtenons que

$$\begin{aligned}
\mathbb{P}[R = r] &\geq \frac{C}{\sqrt{k(n)}} \exp(-\Psi(r)) \\
&\geq \frac{\mathcal{C}}{\sqrt{k(n)}} \exp(-\Psi(Lp + (r - Lp))) \\
&\geq \frac{\mathcal{C}}{\sqrt{k(n)}} \exp\left(-(1-\varepsilon)\frac{n}{L}\delta^2 \log_2 n\right),
\end{aligned}$$

ce qui achève la démonstration du lemme 6.4.9 dans le cas d'une variable binomiale R de paramètres L et p. □

Pour un vecteur $\mathbf{R}$ de loi multinomiale, la démonstration se généralise facilement et utilise sensiblement les mêmes arguments. $\square$

Le lemme 6.4.9 est une étape intermédiaire pour établir le lemme technique 6.4.10 mentionné ci-dessus. Rappelons que nous avons l'égalité en distribution suivante

$$n\mathbb{U}_{n,c}(x) \overset{\mathrm{d}}{=} \sum_{i=1}^{B(n,x)} Y_i \quad \text{pour } x \in [0,1],$$

où $B(n,x)$ est une variable binomiale de paramètres n et x. Cette égalité peut se généraliser dans le cas du processus empirique composé local au point $\mathbf{t}$ indexé par la classe $\mathbb{C}$. Nous en déduisons alors l'égalité en distribution suivante

$$\sum_{i=1}^{n} Y_i \mathbf{1}\left(U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right) \overset{\mathrm{d}}{=} \sum_{i=1}^{\mathbf{R}} Y_i,$$

où le vecteur $\mathbf{R}$ suit une loi multinomiale de paramètres.

Lemme 6.4.10. *Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable aléatoire $Y = Y_1$. De plus, supposons que $\mathbb{E}[Y] = 1$ et $\mathrm{Var}[Y] = \sigma^2 < \infty$. Soit $N \geq 1$ fixé. Soient $L = L(n)$ et $n \geq 1$ des entiers positifs. Soit $\mathbf{R} = (R_1, \ldots, R_N)$ un vecteur de loi multinomiale de la forme*

$$\mathbb{P}\left[R_1 = r_1, \ldots, R_N = r_N\right] = \frac{L!}{r_1! \ldots r_N! \left(L - \sum_{j=1}^{N} r_j\right)!} p_1^{r_1} \ldots p_N^{r_N} \left(1 - \sum_{j=1}^{N} p_j\right)^{L - \sum_{j=1}^{N} r_j},$$

pour les entiers $r_1 \geq 0, \ldots, r_N \geq 0$ et $\sum_{j=1}^{N} r_j \leq L$. De plus, soit

$$p_j = \frac{k(n)}{n}\nu_j \quad \text{et} \quad r_j = Lp_j + \delta_j \sqrt{2np_j \log_2 n} \quad \text{pour } j = 1, \ldots, N,$$

et supposons que $n/L(n)$ et $\delta_j = \delta_{n,j}$ pour $j = 1, \ldots, N$ sont bornés lorsque $n \to \infty$. Alors, pour tous $\varepsilon > 0$ et $\varepsilon' > 0$, il existe une constante $C > 0$ telle que, pour n suffisamment grand, et pour tout réel $\beta \geq 0$ nous ayons

$$\mathbb{P}\left[\sum_{i=1}^{\mathbf{R}} Y_i \geq \beta\right] \geq Ck(n)^{-(N+1)/2} \exp\left(-(1-\varepsilon)\frac{n}{L}\left(\sum_{j=1}^{N} \delta_j^2\right) \log_2 n\right) \exp\left(-(1-\varepsilon')\frac{\beta^2}{L}\right). \tag{6.79}$$

Démonstration. Ici, nous présentons la démonstration de ce lemme uniquement lorsque le vecteur aléatoire $\mathbf{R}$ est réduit à une variable aléatoire binomiale. La démonstration du cas où le vecteur aléatoire $\mathbf{R}$ suit une loi multinomiale utilise les mêmes arguments que celle du cas d'une variable binomiale que nous allons faire maintenant. La démonstration du cas multinomial s'achève en utilisant le lemme 6.4.9.

Soit R une variable aléatoire binomiale de paramètres $L = L(n)$ et p. De plus, posons

$$p = \frac{k(n)}{n}\nu \quad \text{et} \quad r = Lp + \delta_n\sqrt{2np\log_2 n}.$$

Supposons que n/L et δ_n sont bornés lorsque $n \to \infty$. Pour tout réel $\beta \geq 0$, nous avons

$$\mathbb{P}\left[\sum_{i=1}^{R} Y_i \geq \beta\right] \;\geq\; \mathbb{P}\left[R = r\right]\mathbb{P}\left[\sum_{i=1}^{r} Y_i \geq \beta\right].$$

D'une part, remarquons que

$$\mathbb{P}\left[\sum_{i=1}^{r} Y_i \geq \beta\right] \;=\; \mathbb{P}\left[\frac{\sum\limits_{i=1}^{r} Y_i - r}{\sigma\sqrt{r}} \geq \frac{\beta - r}{\sigma\sqrt{r}}\right]$$

En appliquant le corollaire 6.4.1, nous en déduisons qu'il existe une constante $\mathcal{C}_2$ telle que, uniformément pour tout r, lorsque $n \to \infty$, nous ayons

$$\mathbb{P}\left[\frac{\sum_{i=1}^{r} Y_i - r}{\sigma\sqrt{r}} \geq \frac{\beta - r}{\sigma\sqrt{r}}\right] \geq \frac{\mathcal{C}_2}{\sqrt{k(n)}}\exp\left(-r\Psi\left(\frac{\beta}{r}\right)\right),$$

où la fonction Ψ est définie en (6.76). De plus, d'après l'article de Deheuvels et Steinebach [38], nous savons que lorsque $\alpha \to 0$

$$\Psi\left(\mu + \alpha\right) = \frac{\alpha^2}{2\sigma^2} + o\left(\alpha^2\right). \tag{6.80}$$

Ainsi, nous obtenons que pour tout $\varepsilon' > 0$, il existe une constante $\mathcal{C}_2$ telle que, lorsque $n \to \infty$, nous ayons

$$\mathbb{P}\left[\frac{\sum_{i=1}^{r} Y_i - r}{\sigma\sqrt{r}} \geq \frac{\beta - r}{\sigma\sqrt{r}}\right] \geq \frac{\mathcal{C}_2}{\sqrt{k(n)}}\exp\left(-(1-\varepsilon')\frac{\beta^2}{2r}\right).$$

Nous en déduisons qu'il existe une constante $\mathcal{C}_2$ telle que, lorsque $n \to \infty$, nous ayons

$$\mathbb{P}\left[\sum_{i=1}^{R} Y_i \beta\right] \geq \frac{\mathcal{C}}{k(n)}\exp\left(-(1-\varepsilon)\frac{n}{L}\delta^2\log_2 n\right)\exp\left(-(1-\varepsilon')\frac{\beta^2}{2r}\right),$$

en utilisant le lemme 6.4.9 dans le cas d'une variable R binomiale. □

Lemme 6.4.11. *Supposons que les hypothèses (C.1–2–3–4–5), (K.1–2) et (H.1–2) sont vérifiées. Alors la suite $\{\mathbb{X}_{n,c} : n \geq 1\}$ est presque sûrement relativement compacte dans $\mathbb{R}^N$. De plus, l'ensemble limite est inclus dans*

$$\mathbb{L}_N = \left\{\mathbb{x} = (x_1, \ldots, x_N) : \sum_{j=1}^{N} x_j^2 \leq 1\right\}. \tag{6.81}$$

Démonstration. La démonstration de ce lemme se décompose en quatre étapes.

Première étape. Fixons $\gamma \in (0,1)$ et introduisons la suite $n(m) = \lfloor (1+\gamma)^m \rfloor$. Considérons les ensembles définis, pour n tel que et $m \geq 0$ par

$$D'_j(n) = \left(\frac{n}{k(n)} \frac{k(n(m))}{n(m)} \right)^{1/d} D_j = \rho_n D_j. \tag{6.82}$$

En conséquence, pour $j = 1, \ldots, N$ et $n \geq 1$, posons

$$X'_{n,c}(j) = \sqrt{\frac{n}{2\nu_j \mathbb{E}[Y^2] k(n) \log_2 n}} \alpha_{n,c} \left(\mathbf{t} + \left(\frac{k(n)}{n} \right)^{1/d} D'_j(n) \right) \tag{6.83}$$

et

$$\mathbb{X}'_{n,c} = \left(X'_{n,c}(1), \ldots, X'_{n,c}(N) \right).$$

Nous allons montrer que, pour tout $\varepsilon > 0$, il existe un réel $\gamma_0 > 0$ tel que, pour tout $0 < \gamma \leq \gamma_0$, nous ayons

$$\limsup_{n \to \infty} \left(\max_{1 \leq j \leq N} |X_{n,c}(j) - X'_{n,c}(j)| \right) < \varepsilon \quad \text{p.s.} \tag{6.84}$$

D'abord, remarquons que, pour tout m assez grand et n tel que $n(m-1) < n \leq n(m)$, nous avons

$$1 - \gamma \leq \frac{n}{n(m)} \leq 1. \tag{6.85}$$

Comme $k(n) \uparrow 0$ et $n^{-1}k(n) \downarrow 0$, (6.85) implique, pour tout m assez grand et n tel que $n(m-1) < n \leq n(m)$ que

$$(1-\gamma)\, k(n(m)) \frac{\log_2 n(m)}{\log_2 n} \leq \frac{n}{n(m)} k(n(m)) \leq k(n) \leq k(n(m)). \tag{6.86}$$

Ainsi, d'après (6.82), (6.85) et (6.86), pour tout m assez grand et n tel que $n(m-1) < n \leq n(m)$, nous avons

$$D'_j(n) = \rho_n\, D_j, \quad \text{où } (1-\gamma)^{1/d} \leq \rho_n \leq 1.$$

D'après (6.7) appliqué à $\{D_j : 1 \leq j \leq N\}$, il s'ensuit que, pour tout $\varepsilon > 0$, nous pouvons choisir $\gamma = \gamma(\varepsilon) > 0$ si petit que, pour tout m assez grand,

$$\max_{n(m-1) < n \leq n(m)} \left(\max_{1 \leq j \leq N} d_\lambda \left(D_j, D'_j(n) \right) \right) < \frac{\varepsilon}{2} \min \left(1, \frac{1}{K} \right), \tag{6.87}$$

où K est définie en (6.64). Pour conclure nous appliquons le lemme 6.4.7, qui sous l'hypothèse (H.2), lorsqu'il est combiné à (6.87), suffit à prouver (6.84).

Deuxième étape. Pour $j = 1, \ldots, N$, $m \geq 1$ et n tel que $n(m-1) < n \leq n(m)$, introduisons

$$\begin{aligned} X''_{n,c}(j) &= \sqrt{\frac{n}{2\nu_j \mathbb{E}[Y^2] k(n(m)) \log_2 n(m)}} \alpha_{n,c} \left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)} \right)^{1/d} D_j \right) \\ &= \sqrt{\frac{k(n) \log_2 n}{k(n(m)) \log_2 n(m)}} X'_{n,c}(j) \\ &\leq \sqrt{\frac{k(n)}{k(n(m))}} X'_{n,c}(j), \end{aligned} \tag{6.88}$$

et

$$\mathbb{X}''_{n,c} = (X''_{n,c}(1), \ldots, X''_{n,c}(N)).$$

D'après (6.83), nous remarquons que

$$X''_{n(m),c}(j) = X'_{n(m),c}(j) = X_{n(m),c}(j). \tag{6.89}$$

Pour n tel que $n(m-1) < n \leq n(m)$, notons également que nous avons l'égalité en distribution suivante

$$X''_{n(m),c}(j) - X''_{n,c}(j) \stackrel{\mathrm{d}}{=} \sqrt{\frac{n(m) - n}{2\nu_j \mathbb{E}[Y^2] k(n(m)) \log_2 n(m)}} \alpha_{n(m)-n,c} \left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)} \right)^{1/d} D_j \right). \tag{6.90}$$

Soient $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$ et $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)$. Introduisons les événements

$$\begin{aligned} C_{m,c}(\boldsymbol{\delta} + \boldsymbol{\eta}) &= \Bigg\{ \bigcap_{1 \leq j \leq N} |X''_{n,c}(j)| \geq \delta_j + \eta_j \\ &\qquad \text{pour } n \text{ tel que } n(m-1) < n \leq n(m) \Bigg\}, \\ D_{m,c}(\boldsymbol{\delta}) &= \left\{ \bigcap_{1 \leq j \leq N} |X''_{n(m),c}(j)| \geq \delta_j \right\} = \left\{ \bigcap_{1 \leq j \leq N} |X_{n(m),c}(j)| \geq \delta_j \right\} \text{ d'après (6.89)} . \end{aligned}$$

Nous allons prouver que si $\delta_j > 0$ et $\eta_j > 0$ ou si $\delta_j = \eta_j = 0$ pour $j = 1, \ldots, N$, alors il existe un entier m_0 tel que, pour tout $m \geq m_0$, nous ayons

$$\mathbb{P}\left[C_{m,c}(\boldsymbol{\delta} + \boldsymbol{\eta})\right] \leq 2\mathbb{P}\left[D_{m,c}(\boldsymbol{\delta})\right]. \tag{6.91}$$

Pour établir l'inégalité (6.91), posons

$$\begin{aligned} E_{n,c}(\boldsymbol{\delta} + \boldsymbol{\eta}) &= \left\{ \bigcap_{1 \leq j \leq N} |X''_{n,c}(j)| \geq \delta_j + \eta_j \right\}, \\ F_{m,n,c}(\boldsymbol{\eta}) &= \left\{ \bigcap_{1 \leq j \leq N} |X''_{n(m),c}(j) - X''_{n,c}(j)| < \eta_j \right\}. \end{aligned} \tag{6.92}$$

Dans la suite de la démonstration, sans perdre en généralité, nous supposons que

$$\min(\delta_1, \ldots, \delta_N) \geq 0 \quad \text{et} \quad \eta = \min(\eta_1, \ldots, \eta_N) > 0.$$

Pour n tel que $n(m-1) < n \leq n(m)$, remarquons que les événements $\{E_{q,c}(\boldsymbol{\delta} + \boldsymbol{\eta}) : n(m-1) < q \leq n\}$ et $F_{m,n,c}(\boldsymbol{\eta})$ sont indépendants. Notons $\overline{E}$ le complémentaire de l'événement E. Il s'ensuit que

$$\mathbb{P}\left[C_{m,n}(\boldsymbol{\delta} + \boldsymbol{\eta})\right] = \sum_{r=n(m-1)+1}^{n(m)} \mathbb{P}\left[E_{r,c}(\boldsymbol{\delta} + \boldsymbol{\eta}) \cap \bigcap_{q=n(m-1)+1}^{r-1} \overline{E}_{q,c}(\boldsymbol{\delta} + \boldsymbol{\eta})\right].$$

Par conséquent, nous avons

$$\begin{aligned} \inf_{n(m-1)<n\leq n(m)} & \mathbb{P}\left[F_{m,n,c}(\boldsymbol{\eta})\, C_{m,n}(\boldsymbol{\delta} + \boldsymbol{\eta})\right] \\ & \leq \sum_{r=n(m-1)+1}^{n(m)} \mathbb{P}\left[E_{r,c}(\boldsymbol{\delta} + \boldsymbol{\eta}) \cap F_{m,r,c}(\boldsymbol{\delta}) \cap \bigcap_{q=n(m-1)+1}^{r-1} \overline{E}_{q,c}(\boldsymbol{\delta} + \boldsymbol{\eta})\right] \\ & \leq \mathbb{P}\left[D_{m,c}(\boldsymbol{\delta})\right]. \end{aligned} \tag{6.93}$$

D'autre part, d'après (6.90) et l'inégalité de Chebyshev, uniformément pour n tel que $n(m-1) < n \leq n(m)$, lorsque $m \to \infty$, nous avons que,

$$\mathbb{P}\left[\,\overline{F}_{m,n,c}(\boldsymbol{\eta})\right] \leq \eta^{-2} \sum_{j=1}^{N} \mathbb{E}\Big[\left|X''_{n(m),c}(j) - X''_{n,c}(j)\right|^2\Big] \leq \frac{N(n(m)-n)\mathbb{E}[Y^2]}{2\,\eta^2 n(m) \log_2 n(m)} \to 0.$$

Ainsi, nous en déduisons qu'il existe un entier m_0 tel que, pour tout $m \geq m_0$,

$$\inf_{n(m-1)<n\leq n(m)} \mathbb{P}\left[F_{m,n,c}(\boldsymbol{\eta})\right] > \frac{1}{2},$$

ce qui, d'après (6.93), implique (6.91).

Troisième étape. Notons $|.|$ la norme euclidienne sur $\mathbb{R}^N$. Pour tout sous-ensemble E de $\mathbb{R}^N$ et $\varepsilon > 0$, définissons

$$E^{\varepsilon} = \{\mathbb{x} \in \mathbb{R}^N : |\mathbb{x} - \mathbb{y}| < \varepsilon \quad \text{pour } \mathbb{y} \in E\}.$$

Fixons $\varepsilon \in (0,1)$. Nous allons d'abord démontrer qu'il existe une constante $K_1(\varepsilon)$ telle que, pour n assez grand, nous ayons

$$\begin{aligned} \mathbb{P}_m(\varepsilon) &= \mathbb{P}\left[\mathbb{X}''_{n,c} \notin \mathbb{L}_N^{\varepsilon/3} \quad \text{pour } n \text{ tel que } n(m-1) < n \leq n(m)\right] \\ &\leq \mathbb{P}\left[|\mathbb{X}''_{n,c}|^2 \geq 1 + \varepsilon \quad \text{pour } n \text{ tel que } n(m-1) < n \leq n(m)\right] \\ &\leq K_1(\varepsilon) \exp\left(-\left(1 + \frac{\varepsilon}{4}\right) \log_2 n(m)\right). \end{aligned} \tag{6.94}$$

Soit $\lceil 16\, N/\varepsilon \rceil$. Remarquons que si

$$x_1^2 + \cdots + x_N^2 = 1 + \varepsilon < 2,$$

alors, pour $|\delta_1| \leq 2/Q, \ldots, |\delta_N| \leq 2/Q$, nous avons

$$(x_1 + \delta_1)^2 + \cdots + (x_N + \delta_N)^2 \geq 1 + \varepsilon - 4 \sum_{i=1}^{N} |\delta_i| \geq 1 + \frac{\varepsilon}{2}. \tag{6.95}$$

Supposons que $|\mathbb{X}''_{n,c}|^2 \geq 1 + \varepsilon$, et posons, pour $j = 1, \ldots, N$,

$$x_j = \left(\frac{\sqrt{1+\varepsilon}}{|\mathbb{X}''_{n,c}|} \right) X_{n,c}(j), \quad m_j = \lfloor Q |x_j| \rfloor \quad \text{et} \quad \delta_j = \begin{cases} \frac{m_j}{Q} - x_j & \text{si } m_j = 0, \\ \frac{m_j - 1}{Q} & \text{sinon.} \end{cases}$$

Comme $|\delta_j| \leq 2/Q$ pour $j = 1, \ldots, N$, il s'ensuit d'après (6.95) que $\mathbb{m} = (m_1, \ldots, m_N)$ varie dans l'ensemble

$$S_N(\varepsilon) = \left\{ \mathbb{m} \in \{0, \ldots, 2\, Q\}^N : \sum_{1 \leq j \leq N : m_j \geq 2} \left(\frac{m_j - 1}{Q} \right)^2 \geq 1 + \frac{\varepsilon}{2} \right\}. \tag{6.96}$$

De plus, en utilisant la définition (6.92) de l'événement $E_{n,c}$, nous avons les inclusions des événements suivantes

$$\begin{aligned} \left\{ |\mathbb{X}''_{n,c}|^2 \geq 1 + \varepsilon \right\} &\subseteq \bigcup_{\mathbb{m} \in S_N(\varepsilon)} E_{n,c}(\mathbb{m}/Q) \\ &\subseteq \bigcup_{\mathbb{m} \in S_N(\varepsilon)} \left\{ \bigcap_{1 \leq j \leq N : m_j \geq 2} \left\{ |X''_{n,c}(j)| \geq \frac{m_j - 1}{Q} \right\} \right\}. \end{aligned} \tag{6.97}$$

Posons $M_j = \max(0, m_j - 1)$ pour $j = 1, \ldots, N$. D'après (6.91) et (6.97) nous obtenons que

$$\begin{aligned} &\mathbb{P}\Big[|\mathbb{X}''_{n,c}|^2 \geq 1 + \varepsilon \text{ pour } n \text{ tel que } n(m-1) < n \leq n(m) \Big] \\ &\quad \leq \sum_{\mathbb{m} \in S_N(\varepsilon)} \mathbb{P}\left[C_{m,c}\left(\frac{m_1}{Q}, \ldots, \frac{m_N}{Q} \right) \right] \\ &\quad \leq 2 \sum_{\mathbb{m} \in S_N(\varepsilon)} \mathbb{P}\left[D_{m,c}\left(\frac{M_1}{Q}, \ldots, \frac{M_N}{Q} \right) \right] \\ &\quad \leq 2 \sum_{\mathbb{m} \in S_N(\varepsilon)} \mathbb{P}\left[\bigcap_{1 \leq j \leq N : M_j \geq 1} \left\{ |X''_{n(m),c}(j)| \geq \frac{M_j}{Q} \right\} \right]. \end{aligned} \tag{6.98}$$

D'après le lemme 6.4.8 et le fait que les variables aléatoires $Y_{n(m),c}$ sont indépendantes pour $j = 1, \ldots, N$, nous avons

$$\mathbb{P}\left[\bigcap_{1\le j\le N: M_j \ge 1} \left\{|X''_{n(m),c}(j)| \ge \frac{M_j}{Q}\right\}\right] \le 2 \prod_{1\le j\le N: M_j\ge 1} \mathbb{P}\left[\left\{|Y_{n(m),c}(j)| \ge \frac{M_j}{Q}\right\}\right]. \quad (6.99)$$

À cette étape de la démonstration, nous allons rappeler une inégalité établie au chapitre 3.

Inégalité 6.4.2. *Soit un réel $b > 0$. Soit $\Pi_c(b)$ une variable de Poisson composée. Sous la condition $(H.2)$, pour tout $\mu > 0$, nous avons*

$$\mathbb{P}\left[\frac{|\Pi_c(b) - \mathbb{E}[\Pi_c(b)]|}{\sqrt{\mathbb{E}[Y^2]b}} \ge \mu\right] \le 2\exp\left(-\frac{\mu^2}{2}\psi\left(\frac{\mathcal{M}\mu}{\sqrt{\mathbb{E}[Y^2]b}}\right)\right), \quad (6.100)$$

où la fonction ψ est définie en (6.54).

Démonstration. La démonstration de cette inégalité se trouve dans le chapitre 3, paragraphe 3, théorème 3.3.2. □

Rappelons également une propriété de la fonction ψ.

Propriété 6.4.4. *Au voisinage de 0, la fonction ψ admet le développement suivant*

$$\psi(u) = 1 - \frac{u}{3} + \frac{u^2}{6} - \frac{u^3}{10} + \cdots + \frac{(-1)^n 2u^n}{(n+2)(n+1)} + \cdots \qquad \text{pour } |u| < 1.$$

Démonstration. Pour la démonstration de cette propriété, on pourra consulter la propriété 15, p.441 du livre de Shorack et Wellner [125]. □

D'après l'inégalité 6.100, la propriété et la définition (6.73), nous obtenons, pour $M_j \ge 1$, la borne supérieure suivante

$$\mathbb{P}\left[|Y_{n,c}(j)| \ge \frac{M_j}{Q}\right] \le 2\exp\left(-(\log_2 n)\left(\frac{M_j}{Q}\right)^2\left(1 - \frac{\mathcal{M}M_j}{Q}\sqrt{\frac{2\log_2 n}{\nu_j \mathbb{E}[Y^2]k(n)}}\right)\right). \quad (6.101)$$

D'après les hypothèses (K.1–2), l'expression ci-dessus est, pour n assez grand, pour $j = 1, \ldots, N$ et pour $1 \le M_j \le Q$, inférieure à

$$2\exp\left(-\left(\frac{1+\varepsilon/4}{1+\varepsilon/2}\right)(\log_2 n)\left(\frac{M_j}{Q}\right)^2\right).$$

Soit $K_0(\varepsilon) = \#S_N(\varepsilon)$. D'après (6.94), (6.96), (6.98), (6.99) et (6.101), nous obtenons, pour n assez grand que,

$$\begin{aligned}\mathbb{P}_m(\varepsilon) &\le 2^{N+2} \sum_{\mathbb{m}\in S_N(\varepsilon)} \exp\left(-\left(\frac{1+\varepsilon/4}{1+\varepsilon/2}\right)(\log_2 n(m)) \sum_{j=1}^N \left(\frac{M_j}{Q}\right)^2\right) \\ &\le 2^{N+2} K_0(\varepsilon) \exp\left(-\left(1+\frac{\varepsilon}{4}\right)\log_2 n(m)\right).\end{aligned} \quad (6.102)$$

En choisissant $K_1(\varepsilon) = 2^{N+2}K_0(\varepsilon)$, nous obtenons (6.94) d'après (6.98) et (6.102).

Quatrième étape. D'après (6.94) nous remarquons que pour tout $\varepsilon > 0$,

$$\sum_{m=1}^{\infty} \mathbb{P}_m(\varepsilon) < \infty,$$

ce qui, d'après le lemme de Borel-Cantelli, implique que

$$\mathbb{P}\left[|\mathbb{X}''_{n,c}|^2 \geq 1+\varepsilon \quad \text{i.s.}\right] = 0. \tag{6.103}$$

D'après (6.88), remarquons que, lorsque l'encadrement (6.86) est vérifié, nous avons

$$(1-\gamma)\,|\mathbb{X}'_{n,c}|^2 \leq |\mathbb{X}''_{n,c}|^2.$$

Ainsi, d'après (6.103), nous obtenons que

$$\mathbb{P}\left[|\mathbb{X}'_{n,c}|^2 \geq \frac{1+\varepsilon}{1-\gamma} \quad \text{i.s.}\right] = 0. \tag{6.104}$$

Comme $\varepsilon > 0$ et $\gamma > 0$ peuvent être choisis arbitrairement petit dans (6.84) et (6.104), nous complétons la preuve du lemme (6.4.11) d'après l'inégalité triangulaire. □

Lemme 6.4.12. *Supposons que les hypothèses (C.1–2–3–4–5), (K.1–2) et (H.1–2) sont vérifiées. Alors, la suite $\{\mathbb{X}_{n,c} : n \geq 1\}$ est presque sûrement relativement compacte dans $\mathbb{R}^N$. De plus, l'ensemble limite contient*

$$\mathbb{L}_N = \left\{ \mathbb{x} = (x_1, \ldots, x_N) : \sum_{j=1}^{N} x_j^2 \leq 1 \right\}.$$

Démonstration. La démonstration se décompose en deux étapes.

Première étape. Soit $\gamma > 0$ fixé et posons $n(m) = \lfloor (1+\gamma)^m \rfloor$ pour $m \geq 0$. Pour $j = 1, \ldots, N$, introduisons

$$Z'_{m,c}(j) = \sqrt{\frac{n(m)}{2\nu_j \mathbb{E}[Y^2] k(n(m)) \log_2 n(m)}}\, \alpha_{n(m-1),c}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right)$$

et

$$\begin{aligned} Z''_{m,c}(j) &= \sqrt{\frac{n(m)}{2\nu_j \mathbb{E}[Y^2] k(n(m)) \log_2 n(m)}}\, \alpha_{n(m),c}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right) \\ &\quad - \sqrt{\frac{n(m-1)}{2\nu_j \mathbb{E}[Y^2] k(n(m)) \log_2 n(m)}}\, \alpha_{n(m-1),c}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right) \qquad (6.105) \\ &= X_{n(m),c}(j) - \sqrt{\frac{n(m-1)}{n(m)}} Z'_{m,c}(j). \end{aligned}$$

Remarquons que les variables aléatoires $Z''_{m,c}(j)$ sont indépendantes pour $m \geq 1$ et nous avons l'égalité en distribution suivante

$$Z''_{m,c}(j) \stackrel{\mathrm{d}}{=} \sqrt{\frac{n(m)-n(m-1)}{2\nu_j \mathbb{E}[Y^2]k(n(m))\log_2 n(m)}} \times \alpha_{n(m)-n(m-1),c}\left(\mathbf{t}+\left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right).$$

En répétant les arguments utilisés dans la troisième et quatrième étapes de la preuve du lemme 6.4.11, nous obtenons directement que, pour tout $j = 1, \ldots, N$,

$$\limsup_{m\to\infty} |Z'_{m,c}(j)| \leq 1 \quad \text{p.s.} \tag{6.106}$$

Soit

$$\mathbb{Z}''_{m,c} = \left(Z''_{m,c}(1), \ldots, Z''_{m,c}(N)\right).$$

Comme

$$\frac{n(m-1)}{n(m)} \to 1 \quad \text{lorsque } m \to \infty,$$

(6.105) et (6.106) impliquent que, pour tout $\varepsilon > 0$, il existe presque sûrement un réel $\gamma_1 = \gamma_1(\varepsilon) > 0$ tel que, pour tout $\gamma \geq \gamma_1$, nous ayons pour m assez grand

$$|\mathbb{X}_{n(m),c} - \mathbb{Z}''_{m,c}| < \varepsilon. \tag{6.107}$$

Deuxième étape. Soit $Q \geq 1$, $(m_1, \ldots, m_N) \in \{1, \ldots, Q\}$ fixés, $(e_1, \ldots, e_N) \in \{-1, 1\}^N$. Introduisons

$$\mathbb{P}''_m = \mathbb{P}\left(e_1\, Z''_{m,c}(1) \in \left(\frac{m_1-1}{Q}, \frac{m_1}{Q}\right], \ldots, e_N\, Z''_{m,c}(N) \in \left(\frac{m_N-1}{Q}, \frac{m_N}{Q}\right]\right). \tag{6.108}$$

Nous allons démontrer dans cette étape que, lorsque

$$\sum_{j=1}^{N}\left(\frac{m_j}{Q}\right)^2 < 1, \tag{6.109}$$

nous pouvons choisir $\gamma > 0$ aussi grand que

$$\sum_m \mathbb{P}''_m = \infty. \tag{6.110}$$

Comme $Q \geq 1$ peut être choisi arbitrairement grand dans (6.108) et $\varepsilon > 0$ arbitrairement petit dans (6.107), le lemme de Borel-Cantelli, l'inégalité (6.109) le résultat (6.110) et le fait que les variables aléatoires $\mathbb{Z}''_m$ sont indépendantes, impliquent alors la conclusion du lemme 6.4.12.

Afin de prouver (6.110), nous allons utiliser le lemme 6.4.10. D'après (6.73), (6.108) et en appliquant le lemme 6.4.10 où l'on pose $n = n(m)$, $\delta_j = e_j(m_j/Q)$ et $L = n(m)-n(m-1)$, il s'ensuit que pour tout $\varepsilon > 0$ et pour m suffisamment grand,

$$\mathbb{P}''_m \geq \mathcal{K}\exp\left(-(1-\varepsilon)\frac{n(m)}{n(m)-n(m-1)}\left\{\sum_{j=1}^{N}\left(\frac{m_j}{Q}\right)^2\right\}\log_2 n(m)\right).$$

Finalement, nous remarquons, d'après cette minoration, si nous choisissons ε suffisamment petit et γ suffisamment grand, que

$$\frac{1-\varepsilon}{1-(1/\gamma)}\sum_{j=1}^{N}\left(\frac{m_j}{Q}\right)^2<1,$$

alors (6.110) est vérifiée. Ceci achève la preuve du lemme. □

Lemme 6.4.13. *La suite de fonctions*

$$\left\{\frac{\Theta_{n,c}(C_1)}{\sqrt{2\log_2 n}},\cdots,\frac{\Theta_{n,c}(C_M)}{\sqrt{2\log_2 n}}:n\geq 3\right\}$$

est presque sûrement relativement compacte dans $\mathbb{R}^M$. De plus, l'ensemble limite est égal à

$$\left\{\left(\int_{C_1}\phi(\mathbf{s})\,d\lambda(\mathbf{s}),\cdots,\int_{C_M}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right):\int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s})\leq 1\right\}. \tag{6.111}$$

Démonstration. D'après les lemmes 6.4.11 et 6.4.12, nous obtenons que la suite de fonctions

$$\left\{\frac{\Theta_{n,c}\left(D_1\right)}{\sqrt{2\log_2 n}},\ldots,\frac{\Theta_{n,c}\left(D_N\right)}{\sqrt{2\log_2 n}}:n\geq 3\right\}$$

est presque sûrement relativement compacte dans $\mathbb{R}^N$. De plus, l'ensemble limite est égal à

$$\left\{(y_1,\ldots,y_N):\sum_{j=1}^{N}\frac{y_j^2}{\lambda(D_j)}\leq 1\right\}. \tag{6.112}$$

Soit

$$\psi\left(t,\mathbf{s}\right)=\sum_{j=1}^{N}\frac{y_j}{\lambda(D_j)}\,\mathbf{1}_{\{\mathbf{s}\in D_j\}}.$$

Remarquons que, si (6.112) est vérifié, alors pour $j=1,\ldots,N$, nous avons

$$y_j=\int_{D_j}\psi\left(\mathbf{s}\right)\,d\lambda\left(\mathbf{s}\right)\quad\text{ et }\quad\int_{\mathbb{R}^d}\psi^2\left(\mathbf{s}\right)\,d\lambda\left(\mathbf{s}\right)\leq 1.$$

Inversement, supposons que $\phi(\mathbf{s})$ est une fonction quelconque mesurable sur $\mathbb{R}^d$, telle que

$$\int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda\left(\mathbf{s}\right)\leq 1,$$

et introduisons

$$y_j=\int_{D_j}\phi(\mathbf{s})\,d\lambda\left(\mathbf{s}\right)\quad\text{ pour } j=1\ldots,N.$$

D'après l'inégalité de Cauchy Schwarz et le fait que les ensembles D_j sont disjoints, nous obtenons que

$$\begin{aligned}
\sum_{j=1}^{N} \frac{y_j^2}{\lambda(D_j)} &= \sum_{j=1}^{N} \frac{1}{\lambda(D_j)} \left\{ \int_{D_j} \phi(\mathbf{s})\, d\lambda(\mathbf{s}) \right\}^2 \\
&\leq \sum_{j=1}^{N} \frac{1}{\lambda(D_j)} \left\{ \int_{D_i} \phi^2(\mathbf{s})\, d\lambda(\mathbf{s}) \right\} \left\{ \int_{D_i} d\lambda(\mathbf{s}) \right\} \\
&\leq \int_{\mathbb{R}^d} \phi^2(\mathbf{s})\, d\lambda(\mathbf{s}) \\
&\leq 1.
\end{aligned}$$

D'après ceci, et étant donné (6.81), il s'ensuit que l'ensemble défini en (6.112) coincide avec l'ensemble suivant

$$\left\{ \left(\int_{D_1} \phi(\mathbf{s})\, d\lambda(\mathbf{s}), \ldots, \int_{D_N} \phi(\mathbf{s})\, d\lambda(\mathbf{s}) \right) : \int_{\mathbb{R}^d} \phi^2(\mathbf{s})\, d\lambda(\mathbf{s}) \leq 1 \right\}. \tag{6.113}$$

Comme chaque C_i, pour $i = 1, \ldots, M$, est, presque partout pour la mesure λ_0, l'union disjointe des ensembles pris parmi $\{D_1, \ldots, D_N\}$, (6.111) s'ensuit directement de (6.113). □

Nous avons maintenant tous les outils nécessaires pour prouver le théorème 6.3.1. Fixons $\varepsilon > 0$ arbitrairement. Cette démonstration reprend les arguments de ceux utilisés dans le théorème 2.1 p.415-416 de l'article de Kuelbs et Dudley [75]. Si nous choisissons $\varepsilon_1 = \frac{1}{2}\min(\varepsilon_0, \varepsilon/8K, \varepsilon^2/4)$ dans le lemme 6.4.7, alors, d'après ce lemme, il existe presque sûrement un entier n_1 tel que, pour tout $n \geq n_1$, nous ayons

$$\sup_{C \in \mathbb{C}} \frac{W_{n,c}(C, k(n)/n, \varepsilon/2)}{\sqrt{2\log_2 n}} \leq \frac{\varepsilon}{2}. \tag{6.114}$$

D'après (6.29) il existe une suite $\{C_i : 1 \leq i \leq M\} \subset \mathbb{C}$ avec la propriété suivante.

Propriété 6.4.5. *Pour tout $C \in \mathbb{C}$, il existe un entier $i \in \{1, \cdots, M\}$, tel que nous ayons $d_\lambda(C, C_i) < \varepsilon_1/2$.*

D'après (6.59) et (6.114), nous obtenons

$$\left| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - \frac{\Theta_{n,c}(C_i)}{\sqrt{2\log_2 n}} \right| \leq \frac{\varepsilon}{2}. \tag{6.115}$$

Soit $1 \leq n'_1 < \cdots < n'_m < \cdots$ une suite telle que, la suite

$$\left\{ \frac{\Theta_{n,c}(C_i)}{\sqrt{2\log_2 n}} \right\}$$

converge vers une limite pour chaque $1 \leq i \leq M$. D'après le lemme 6.4.13, il existe une fonction mesurable ϕ qui vérifie

$$\int_{\mathbb{R}} \phi^2(\mathbf{s})\, d\lambda\,(\mathbf{s}) \leq 1$$

et telle que,

$$\max_{1\leq i\leq M} \left| \frac{\Theta_{n,c}(C_i)}{\sqrt{2\log_2 n}} - \int_{C_i} \phi(\mathbf{s}) d\lambda\,(\mathbf{s}) \right| \leq \frac{\varepsilon}{4}. \tag{6.116}$$

D'autre part, l'inégalité de Cauchy Schwarz entraîne

$$\begin{aligned} \left| \int_C \phi(\mathbf{s})\, d\lambda\,(\mathbf{s}) - \int_{C_i} \phi(\mathbf{s})\, d\lambda\,(\mathbf{s}) \right| &\leq \sqrt{\int_{C\Delta C_i} \phi^2(\mathbf{s}) d\lambda\,(\mathbf{s})}\, \sqrt{d_\lambda(C, C_i)} \\ &\leq \frac{\varepsilon}{4}, \end{aligned} \tag{6.117}$$

uniformément par rapport à $\phi \in \mathbb{S}(\mathbb{C}\,)$ et $d_\lambda(C, C_i) \leq \dfrac{\varepsilon^2}{16}$.
Ainsi, en combinant (6.115), (6.116) et (6.117), nous obtenons que,

$$\sup_{C\in\mathbb{C}} \left| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - \int_C \phi(\mathbf{s})\, d\lambda\,(\mathbf{s}) \right| \leq \varepsilon. \tag{6.118}$$

Nous allons maintenant montrer que, presque sûrement,

$$\limsup_{n\to\infty} \left\{ \inf_{f\in\mathbb{S}(\mathbb{C}\,)} \left(\sup_{C\in\mathbb{C}} \left| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - f(C) \right| \right) \right\} \leq \varepsilon. \tag{6.119}$$

Pour prouver cette inégalité, nous remarquons que si (6.119) n'est pas vérifiée, alors il existe un $\varepsilon' > 0$ et une suite $1 \leq n_1 < n_2 < \cdots$ telle que, nous ayons

$$\inf_{f\in\mathbb{S}(\mathbb{C}\,)} \left(\sup_{C\in\mathbb{C}} \left| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - f(C) \right| \right) \geq \varepsilon + \varepsilon'.$$

D'autre part, le lemme 6.4.13 nous permet d'extraire de $\{n_m\}$ une suite $\{n'_m\}$ telle que

$$\left\{ \frac{\Theta_{n,c}(C_i)}{\sqrt{2\log_2 n}} \right\}$$

converge pour tout $1 \leq i \leq N$. D'après (6.118), nous obtenons une contradiction. Comme $\varepsilon > 0$ peut être choisi aussi petit que l'on veut (6.119), il s'ensuit que, presque sûrement,

$$\lim_{n\to\infty} \left\{ \inf_{f\in\mathbb{S}(\mathbb{C}\,)} \left\| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - f(C) \right\|_{\mathbb{C}} \right\} = 0. \tag{6.120}$$

D'autre part, nous avons la propriété suivante.

Propriété 6.4.6. $\mathbb{S}(\mathbb{C})$ *défini en* (6.20) *est un sous-ensemble compact de* $B(\mathbb{C})$*.*

Démonstration. Afin de prouver que cette propriété est réalisée, il suffit de montrer que $\mathbb{S}(\mathbb{C})$ est totalement borné et complet par rapport à la topologie de la convergence uniforme. Étant donné l'égalité des ensembles définis en (6.112) et (6.113) et d'après la compacité de la boule unité de $\mathbb{R}^M$, la première partie est une conséquence évidente de la propriété (6.29) combinée avec l'existence, pour chaque $\varepsilon > 0$, de $\{C_i : 1 \leq i \leq M\} \in \mathbb{C}$, telle que l'inégalité (6.117) soit réalisée. La seconde partie découle directement du fait que $\mathbb{S}(\mathbb{C})$ est un sous-ensemble fermé de $B(\mathbb{C})$. $\square$

D'après (6.120), il s'ensuit que l'ensemble limite de

$$\left\{ \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} : C \in \mathbb{C} \right\}$$

est presque sûrement inclus dans $\mathbb{S}(\mathbb{C})$. Pour montrer que nous avons l'égalité, nous choisissons une fonction quelconque $\phi \in \mathbb{S}(\mathbb{C})$. D'après le lemme 6.4.13, il existe une suite $1 \leq n_1' < n_2' < \ldots$, qui dépend de $\varepsilon > 0$, suivant que (6.116) et (6.118) ont lieu au voisinage de ∞. De plus, il existe un entier $n = n(\varepsilon, \phi)$ tel que (6.118) ait lieu. En répétant cet argument pour une suite $\varepsilon = \varepsilon_k \downarrow 0$, nous obtenons facilement l'existence d'une suite croissante $1 \leq n_1'' < n_2'' < \ldots$, suivant laquelle

$$\left\| \frac{\Theta_{n,c}(C)}{\sqrt{2\log_2 n}} - \phi \right\|_{\mathbb{C}} \to 0.$$

Ceci implique que ϕ appartient à l'ensemble limite de

$$\left\{ \frac{\Theta_{n,c}}{\sqrt{2\log_2 n}} : n \geq 1 \right\}$$

et complète la démonstration du théorème 6.3.1.

6.5. Applications

6.5.1. Quelques théorèmes locaux pour des échantillons non-uniformes

Le premier théorème donne un exemple de la manière dont on peut appliquer le corollaire 6.2.1 à des suites de variables aléatoires de distribution non uniforme.

Théorème 6.5.1. *Soit* $\{X_n : n \geq 1\}$*, une suite de variables aléatoires indépendantes de fonction de répartition* F *dérivable au point* x *de dérivée* $F'(x) = f(x) > 0$*. Supposons* *que la suite* $\{k(n) : n \geq 1\}$ *vérifie les conditions* $(H.1)$ *et* $(H.2)$. *Alors, pour tout réel* t *fixé, la suite de fonctions de* $s \in [0,1]$ *définie par*

$$\Big\{\Psi_n(s) \;=\; (2\log_2 n)^{-1/2}\Theta_n\Bigg(\left(\frac{k(n)}{n}f(x)\right)^{-1}\left(F\left(x+\frac{k(n)}{n}t\right)-F(x),\right. \\ \left.\left. F\left(x+\frac{k(n)}{n}(t+s)\right)-F(x)\right), F(x), \frac{k(n)}{n}f(x)\Bigg) : n \geq 1\Big\} \tag{6.121}$$

est presque sûrement relativement compacte dans l'ensemble des fonctions bornées sur l'intervalle $[0,1]$, *muni de la topologie de la convergence uniforme. L'ensemble de ses points limites contient les fonctions de la forme*

$$\psi(s) = \int_0^s \phi(u)\,du \text{ avec } \int_0^1 \phi(u)\,du \leq 1. \tag{6.122}$$

Démonstration. Nous nous limitons ici au cas où la fonction de répartition F est continue et strictement croissante sur $\mathbb{R}$. Soit la suite $\{U_n = F_n(X_n) : n \geq 1\}$, où les U_n sont des variables aléatoires indépendantes et uniformément distribuées sur l'intervalle $(0,1)$. En appliquant le corollaire 6.2.1, nous obtenons que, pour tout $A > 0$, la suite de fonctions de (y,z) définie par

$$\overline{\Psi}_n(y,z) = \frac{\Theta_n\left((y,z], F(x), \dfrac{k(n)}{n}f(x)\right)}{\sqrt{2\log_2 n}}, \quad \text{pour} - A \leq y \leq z \leq A \tag{6.123}$$

est presque sûrement relativement compacte par rapport à la topologie de la convergence uniforme. L'ensemble limite contient les fonctions de la forme

$$\overline{\Psi}(y,z) = \int_y^z \phi(t)dt \quad \text{avec} \int_{-\infty}^{\infty} \phi^2(t)dt \leq 1. \tag{6.124}$$

Un argument basé sur l'équicontinuité uniforme de la fonction Ψ définie en (6.124) permet de montrer que la suite $\overline{\Psi}_n(g_n(y), g_n(z))$ est aussi relativement compacte et l'ensemble limite est caractérisé par (6.124), lorsque g_n est une suite de fonctions non décroissantes telle que, pour tout $M > 0$,

$$\lim_{n\to}\left(\sup_{|z|\leq M} |g_n(z) - z|\right) = 0. \tag{6.125}$$

Soit

$$g_n(z) = \left(\frac{k(n)}{n}f(x)\right)^{-1}\left(F\left(x+\frac{k(n)}{n}z\right)-F(x)\right).$$

Comme cette fonction satisfait (6.125), nous pouvons conclure au théorème 6.5.1. $\square$

Notre second résultat donne également un exemple de la manière dont on peut appliquer le corollaire 6.3.1 à des suites de variables aléatoires de distribution non uniforme.

Théorème 6.5.2. *Soit $\{Y_i : i \geq 1\}$, une suite de variables aléatoires indépendantes et identiquement distribuées, de fonction de répartition F_Y dérivable au point x, de dérivée $f_Y(x) > 0$. De plus, la suite $\{Y_i : i \geq 1\}$ vérifie les hypothèses (H.1–2). Soit $\{X_i : i \geq 1\}$, une suite de variables aléatoires indépendantes et identiquement distribuées, de fonction de répartition F_Z dérivable au point x, de dérivée $f_Z(x) > 0$. Supposons que la suite $\{k(n) : n \geq 1\}$ vérifie les hypothèses (K.1–2). Alors, pour tout réel t fixé, la suite de fonctions de $s \in [0,1]$ définie par*

$$\Big\{\Psi_{n,c}(s) \;=\; (2\log_2 n)^{-1/2}$$

$$\Theta_{n,c}\left(\left(\frac{k(n)}{n}f_X(x)\right)^{-1}\left(F_X\left(x+\frac{k(n)}{n}t\right)-F_X(x),\right.\right.$$

$$\left.\left.F_X\left(x+\frac{k(n)}{n}(t+s)\right)-F_X(x)\right), F_X(x), \frac{k(n)}{n}f_X(x)\right) : n \geq 1\Big\}$$

est presque sûrement relativement compacte dans l'ensemble des fonctions bornées sur l'intervalle $[0,1]$, muni de la topologie de la convergence uniforme. L'ensemble de ses points limites contient les fonctions de la forme

$$\psi(s) = \int_0^s \phi(u)\,du \text{ avec } \int_0^1 \phi(u)\,du \leq 1.$$

Démonstration. La démonstration se fait de la même manière que celle du théorème 6.5.1. □

6.5.2. Estimation de la densité

Soit $X, X_1, X_2, \cdots$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs réelles de fonction de répartition F et de densité f. Le problème de l'estimation de f par des techniques non-paramétriques a été largement abordé. En particulier, Parzen [109] et Rosenblatt [115] s'y sont intéressés et ont introduit l'estimateur à noyau qui porte leur nom. Il se définit ainsi

$$\begin{aligned} f_n(x) &= \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right) \\ &= \int_{-\infty}^{\infty} \frac{1}{h_n}K\left(\frac{x-t}{h_n}\right) dF_n(t), \end{aligned} \tag{6.126}$$

où $\{h_n : n \geq 1\}$ est une suite de constantes positives, et le noyau $K(\cdot)$ est une fonction qui vérifie les conditons suivantes.

(K.1) $\int_{\mathbb{R}} K(u)\,du = 1$.
(K.2) $K(\cdot)$ est à variation bornée sur $\mathbb{R}$.
(K.3) Il existe un $M < \infty$ tel que $K(u) = 0$ pour $|u| \leq M/2$.

Théorème 6.5.3. *Supposons que la fonction de répartition F soit dérivable au point $x \in \mathbb{R}$ et admette une dérivée $f(x) > 0$. Supposons de plus que la suite $\{h_n : n \geq 1\}$ vérifie les conditions suivantes*

$$(S.1) \quad (i) \quad h_n \searrow 0; \quad (ii) \quad nh_n \nearrow \infty \quad \textit{lorsque} \quad n \to \infty.$$
$$(S.2) \quad \frac{nh_n}{\log_2 n} \to \infty \quad \textit{lorsque} \quad n \to \infty.$$

Soit un noyau $K(\cdot)$ qui vérifie les propriétés (K.1–2–3). Alors nous avons

$$\limsup_{n\to\infty} \pm \sqrt{\frac{nh_n}{2\log_2 n}} \Big(f_n(x) - \mathbb{E}[f_n(x)] \Big) = \sqrt{f(x) \int_{\mathbb{R}} K^2(u)\,du} \quad p.s. \tag{6.127}$$

Démonstration. Posons pour $x \in \mathbb{R}$,

$$F_n(x) = n^{-1}\#\{X_i \leq x : 1 \leq i \leq n\} \quad \text{et} \quad F(x) = \mathbb{P}(X_1 \leq x). \tag{6.128}$$

Pour la suite de cette démonstration, nous poserons également $\tilde{K}(t) = K(-t)$ pour $t \in \mathbb{R}$. Nous obtenons, après un changement de variable et d'après $(K.3)$,

$$\begin{aligned} f_n(x) - \mathbb{E}[f_n(x)] &= \int_{-\infty}^{\infty} \frac{1}{h_n} K\left(\frac{x-t}{h_n}\right) dF_n(t) - \int_{-\infty}^{\infty} \frac{1}{h_n} K\left(\frac{x-t}{h_n}\right) dF(t) \\ &= \int_{-\infty}^{\infty} \frac{1}{h_n} K\left(\frac{x-t}{h_n}\right) \{dF_n(t) - dF(t)\} \\ &= \frac{1}{h_n} \int_{-M}^{M} \tilde{K}(u)\, d\{F_n(x+h_n u) - F(x+h_n u)\}. \end{aligned} \tag{6.129}$$

En intégrant par parties (6.129), nous obtenons

$$\begin{aligned} &\pm \sqrt{\frac{nh_n}{2\log_2 n}} \{ f_n(x) - \mathbb{E}[f_n(x)] \} \\ &= \pm \int_{-M}^{M} \left\{ \frac{\sqrt{n}\big(F_n(x+h_n u) - F_n(x-h_n M) - F(x+h_n u) + F(x-h_n M)\big)}{\sqrt{2h_n \log_2 n}} \right\} d\{-\tilde{K}(u)\}. \end{aligned} \tag{6.130}$$

Maintenant, une application du théorème 6.5.1 montre que l'ensemble des points limites du membre de droite de (6.130) est presque sûrement égal à

$$\left\{ \pm\sqrt{f(x)} \int_{-M}^{M} \left\{ \int_{-M}^{u} \phi(s)ds \right\} d\{-\tilde{K}(u)\} : \int_{-M}^{M} \phi^2(u)du \leq 1 \right\}$$

Nous concluons en procédant de nouveau à une intégration par parties pour obtenir que l'ensemble limite est égal à

$$\left\{ \pm\sqrt{f(x)} \int_{-M}^{M} \phi(u)\tilde{K}(u)du : \int_{-M}^{M} \phi^2(u)du \leq 1 \right\}.$$

Enfin, nous appliquons l'inégalité de Cauchy Schwarz et nous obtenons la conclusion du théorème 6.5.3. □

Remarque 6.5.1. Ce théorème se généralise en dimension $d \geq 1$. Pour de plus amples détails, nous renvoyons à l'article [36].

6.6. Appendice

Dans cette section, nous allons établir le lemme 6.4.4. Pour cela, nous allons rappeler et établir quelques résultats.

Lemme 6.6.1 (Inégalité de Lévy). *Soit V un espace vectoriel réel et $||\cdot||$ une semi-norme sur V. Soient $\{D_i : i \geq 1\}$ une suite de variables aléatoires à valeurs dans V, $S_{n,} = D_1 + \cdots + D_n$ et $S_{mn} = S_m - (D_{m+1} + \cdots + D_n)$, $1 \leq m \leq n$. Supposons que, pour $1 \leq m \leq n$, $||S_m||$ et $||S_{mn}||$ sont mesurables et que $< ||S_1||, \cdots, ||S_m||, ||S_n|| >$ ont la même distribution jointe sur $\mathbb{R}^{m+1}$ que $< ||S_1||, \cdots, ||S_m||, ||S_{mn}|| >$. Alors, pour tout réel $K > 0$ et pour tout entier $n \geq 1$, nous avons*

$$\mathbb{P}\left[\max_{j\leq n} ||S_j|| \geq K\right] \leq 2\mathbb{P}\left[||S_n|| \geq K\right].$$

Démonstration. La démonstration de ce lemme se trouve dans l'article de Kuelbs et Dudley [75]. □

Le lemme qui suit est très connu dans le cas séparable. Nous renvoyons par exemple aux articles de Kahane [79], p.16, Hoffman-Jorgensen [68], 3.3, Jain et Marcus [74], lemme 3.4.

Lemme 6.6.2. *Supposons que les hypothèses du lemme 6.6.1 sont vérifées et qu'il existe un réel K tel que pour tout entier $i \geq 1$, $||D_i|| \leq K$. De plus, supposons que, pour tout $1 \leq m \leq n$, les deux variables aléatoires $< ||S_1||, \cdots, ||S_m|| >$ de $\mathbb{R}^m$ et $||S_n - S_m||$ de $\mathbb{R}$ sont indépendantes. Alors, pour tout réel $t > K$ et pour tout entier $n \geq 1$, nous avons*

$$\mathbb{P}\left[||S_n|| \geq 3t\right] \leq 4\mathbb{P}\left[||S_n|| \geq t\right]^2.$$

Démonstration. La démonstration de ce lemme se trouve dans l'article de Kuelbs et Dudley [75]. □

Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la variable $Y = Y_1$ et qui vérifie l'hypothèse (H.2). Soit $\{U_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et uniformément distribuées sur l'intervalle

$[0, 1]$. Soit $\{V_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et de même loi que la suite $\{U_i : i \geq 1\}$. Nous allons appliquer les lemmes 6.6.1 et 6.6.2 aux variables $D_i = Y_i \mathbf{1}_{U_i} - Y_i \mathbf{1}_{V_i}$, pour $i \geq 1$, et nous allons remplacer la norme $||.||$ par la norme $||.||_{\mathcal{C}}$ définie dans la proposition 3.2 de Dudley [48].

Lemme 6.6.3. *Pour toute classe de Donsker $\mathbb{P}$ $\in$-Suslin $\mathcal{C}$ et pour tout $\varepsilon > 0$, il existe un entier r et une décomposition $\mathcal{C} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$, où les $\mathcal{C}_j$ sont disjoints, tels que chaque $\mathcal{C}_j = \mathcal{C}(j)$ est $\mathbb{P}$ $\in$-Suslin et pour $A_j \in \mathcal{C}_j$, $N = N(\varepsilon) < \infty$, et $||f||_j = \sup_{C \in \mathcal{C}(j)} |f(C) - f(A_j)|$, nous avons pour chaque $j = 1, \ldots, r$,*

$$\sup_{n \geq N} \mathbb{P}\left[||S_n||_j > \varepsilon \sqrt{\mathbb{E}[Y^2]n} \right] < \frac{1}{80}.$$

Démonstration. La démonstration de ce lemme est identique à celle du lemme 3.3 p.411 de l'article de Kuelbs et Dudley [75]. □

Lemme 6.6.4. *Pour tout $\varepsilon > 0$, pour tout $j = 1, \ldots, r$ et pour $n \geq \max(N(\varepsilon), 2\varepsilon^{-2})$, où $N(\varepsilon)$ et $\mathcal{C}_j$ sont définis dans le lemme 6.6.3, nous avons*

$$\mathbb{E}\left[||S_n||_j^2 \right] \leq 17\varepsilon^2 \, \mathbb{E}[Y^2]n.$$

Démonstration. Posons $u = \varepsilon^2 \, \mathbb{E}[Y^2]n \geq 2(\mathcal{M} + 1)$. Comme $||S_{n,c}||_j$ est mesurable et rappelons que $D_i = Y_i \mathbf{1}_{U_i} - Y_i \mathbf{1}_{V_i}$, les lemmes 6.6.1 et 6.6.2 sont vérifiés en prenant $K = 2(\mathcal{M} + 1)$ et la norme $||.||_j$ dans le cas présent. Alors nous avons

$$\begin{aligned}
\mathbb{E}\left[||S_n||_j^2/9 \right] &= \int_0^\infty \mathbb{P}\left[||S_n||_j > 3\sqrt{t} \right] dt \\
&\leq u + \int_u^\infty \mathbb{P}\left[||S_n||_j > 3\sqrt{t} \right] dt \\
&\leq u + \int_u^\infty 4\mathbb{P}\left[||S_n||_j > \sqrt{t} \right]^2 dt \qquad \text{d'après le lemme 6.6.2} \\
&\leq u + \int_u^\infty \mathbb{P}\left[||S_n||_j > \sqrt{t} \right] dt/20 \qquad \text{d'après le lemme 6.6.3} \\
&\leq u + \int_0^\infty \mathbb{P}\left[||S_n||_j > \sqrt{t} \right] dt/20.
\end{aligned}$$

Ainsi, nous concluons à

$$\mathbb{E}\left[||S_n||_j^2 \right] \leq \frac{180}{11} u < 17\varepsilon^2 \, \mathbb{E}[Y^2]n.$$

□

Lemme 6.6.5. *Soit $N > \max(N(\varepsilon), 2\varepsilon^{-2})$, où $N(\varepsilon)$ est défini dans le lemme 6.6.3 et $\varepsilon > 0$. Soit $T_k = S_{kN}$ et $Z_k = T_k - T_{k-1}$ pour $k = 1, \ldots, N$. Alors, pour tout $j = 1, \ldots, r$*

et $k = 1, \ldots, N$, nous avons

$$
\begin{aligned}
||Z_k||_j &\leq 2\,(\mathcal{M}+1)\,N, && (6.131)\\
\mathbb{E}\left[\,||Z_k||_j^2\,\right] &\leq 17\varepsilon^2\,\mathbb{E}[Y^2]N,, && (6.132)\\
\mathbb{E}\left[\,||T_k||_j\,\right] &\leq \sqrt{17\varepsilon^2\,\mathbb{E}[Y^2]kN},. && (6.133)
\end{aligned}
$$

Pour tout entier $m \geq 1$ et pour tout réel h, nous avons

$$
\mathbb{E}\left[\exp\left(h||T_m||_j\right)\right] \leq \exp\left(h\mathbb{E}[\,||T_m||_j\,] + 2h^2\exp\left(4\,(\mathcal{M}+1)\,Nh\right)\sum_{k=1}^{m}\mathbb{E}\left[\,||Z_k||_j^2\,\right]\right). \tag{6.134}
$$

Pour tous réels $b > 0$ et $\lambda > 0$, nous avons

$$
\begin{aligned}
&\mathbb{P}\big[\,||T_m||_j \geq 2\lambda b\big]\\
&\quad\leq \exp\left(-\lambda^2 + \frac{\lambda^2}{2}\left\{\frac{\mathbb{E}[\,||T_m||_j\,]}{\lambda b} + \frac{1}{b^2}\exp\left(\frac{2\,(\mathcal{M}+1)\,N\lambda}{b}\right)\sum_{k=1}^{m}\mathbb{E}\left[\,||Z_k||_j^2\,\right]\right\}\right)\\
&\quad\leq \exp\left(-\lambda^2 + \frac{\lambda^2}{2}\left\{\frac{\sqrt{17\varepsilon^2\,\mathbb{E}[Y^2]mN}}{\lambda b} + \frac{1}{b^2}\exp\left(\frac{2\,(\mathcal{M}+1)\,N\lambda}{b}\right)17\varepsilon^2\,\mathbb{E}\left[Y^2\right]mN\right\}\right).
\end{aligned} \tag{6.135}
$$

Démonstration. Pour tous entiers i et j, nous avons

$$
||D_i||_j \leq 2\,(\mathcal{M}+1)\,,
$$

ce qui implique (6.131). Le lemme implique (6.132) et (6.133). Pour démontrer (6.134) introduisons $\mathcal{B}_k$ la plus petite σ-algèbre pour laquelle Y_i, U_i et V_i sont mesurables, pour $1 \leq i \leq kN$ et $\mathcal{B}_0$ la σ-algèbre usuelle. Soient $\mathbb{E}_k$ l'espérance conditionnelle par rapport à la σ-algèbre $\mathcal{B}_k$ et $\mathbb{E}_0[X] = \mathbb{E}[X]$ lorsque $\mathbb{E}[|X|] < \infty$. Soit $F_k = (\mathbb{E}_k - \mathbb{E}_{k-1})\,||T_m||_j$. Alors nous avons

$$
\mathbb{E}[F_k] = 0
$$

et

$$
||T_m||_j = \mathbb{E}[\,||T_m||_j\,] + \sum_{k=1}^{m} F_k.
$$

Alors pour tout réel h, nous avons

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(h||T_m||_j\right)\right] &= \mathbb{E}\left[\exp\left(h\mathbb{E}[\,||T_m||_j\,] + \sum_{k=1}^{m} F_k\right)\right]\\
&= \mathbb{E}\left[\exp\left(h\mathbb{E}[\,||T_m||_j\,] + \sum_{k=1}^{m-1} F_k\right)\right]\mathbb{E}_{m-1}\exp\left(hF_m\right),
\end{aligned} \tag{6.136}
$$

puisque, pour tout $k < m$, F_k est $\mathcal{B}_{m-1}$ mesurable. Posons

$$X_k = T_m - Z_k, \quad \text{pour } 1 \leq k \leq m.$$

Nous en déduisons alors que

$$||X_k||_j - ||Z_k||_j \leq ||T_m||_j \leq ||X_k||_j + ||Z_k||_j,$$

ce qui implique que

$$\begin{aligned} \mathbb{E}_k[\,||T_m||_j\,] &\leq \mathbb{E}_k[\,||X_k||_j\,] + \mathbb{E}_k[\,||Z_k||_j\,], \\ \mathbb{E}_{k-1}[\,||T_m||_j\,] &\geq \mathbb{E}_{k-1}[\,||X_k||_j\,] - \mathbb{E}_{k-1}[\,||Z_k||_j\,]. \end{aligned}$$

Par conséquent, nous obtenons

$$F_k \leq \mathbb{E}_k[\,||X_k||_j\,] + \mathbb{E}_k[\,||Z_k||_j\,] - \mathbb{E}_{k-1}[\,||X_k||_j\,] + \mathbb{E}_{k-1}[\,||Z_k||_j\,].$$

Nous avons également

$$\mathbb{E}_k[\,||Z_k||_j\,] = ||Z_k||_j, \quad \mathbb{E}_{k-1}[\,||Z_k||_j\,] = \mathbb{E}[\,||Z_k||_j\,] \quad \text{et} \quad \mathbb{E}_k[\,||X_k||_j\,] = \mathbb{E}_{k-1}[\,||X_k||_j\,].$$

Ainsi, nous avons

$$F_k \leq ||Z_k||_j + \mathbb{E}[\,||Z_k||_j\,].$$

De manière analogue, nous obtenons

$$F_k \geq -||Z_k||_j - \mathbb{E}[\,||Z_k||_j\,].$$

Nous en déduisons alors que

$$|F_k| \leq ||Z_k||_j + \mathbb{E}[\,||Z_k||_j\,]. \tag{6.137}$$

En utilisant (6.136) et en itérant cette inégalité, l'inégalité (6.134) découlera de

$$E_{k-1}\left[\exp\left(hF_k\right)\right] \leq \exp\left(2h^2 \exp\left(4Nh\right) \mathbb{E}\left[\,||Z_k||_j^2\,\right]\right) \quad \text{pour } k = 1\ldots, m. \tag{6.138}$$

Il ne reste donc plus qu'à démontrer (6.138). Remarquons que $\mathbb{E}_{k-1}[F_k] = 0$. Pour $r \geq 2$, d'après (6.137) nous avons

$$\begin{aligned} \mathbb{E}_{k-1}\left[F_k^r\right] &\leq \mathbb{E}_{k-1}\left[\left(||Z_k||_j + \mathbb{E}\left[\,||Z_k||_j\,\right]\right)^r\right] = \mathbb{E}\left[\left(||Z_k||_j + \mathbb{E}\left[\,||Z_k||_j\,\right]\right)^r\right] \\ &\leq 2^r \mathbb{E}\left[\,||Z_k||_j^r\,\right], \end{aligned} \tag{6.139}$$

puisque, pour $s = 0, 1, \ldots, r$ et en posant $f = ||Z_k||_j \geq 0$, nous avons les inégalités suivantes

$$\mathbb{E}[f] \leq \left(\mathbb{E}[f^r]\right)^{1/r} \quad \text{et} \quad \mathbb{E}[f^s] \leq \left(\mathbb{E}[f^r]\right)^{s/r}$$

qui impliquent que

$$\mathbb{E}[f^s]\left(\mathbb{E}[f]\right)^{r-s} \leq \mathbb{E}[f^r].$$

Nous avons

$$\exp(hF_k) = 1 + hF_k + \frac{h^2 F_k^2}{2!} + \frac{h^3 F_k^3}{3!} + \ldots.$$

Comme $\mathbb{E}_{k-1} F_k = 0$, nous obtenons

$$\begin{aligned}
\mathbb{E}_{k-1}\left[\exp(hF_k)\right] &= 1 + \frac{h^2}{2!}\mathbb{E}_{k-1}\left[F_k^2\right] + \frac{h^3}{3!}\mathbb{E}_{k-1}\left[F_k^3\right] + \ldots \\
&\leq 1 + \frac{h^2}{2!} \times 2^2 \mathbb{E}\left[\,||Z_k||_j^2\,\right] + \frac{h^3}{3!} \times 2^3 \mathbb{E}\left[\,||Z_k||_j^3\,\right] + \ldots \qquad \text{d'après (6.139)} \\
&\leq 1 + 2h^2 \mathbb{E}\left[\,||Z_k||_j^2\,\right] S \qquad \text{d'après (6.131)}\ ,
\end{aligned}$$

où

$$S = 1 + \frac{2h}{3} 2(\mathcal{M}+1)N + \frac{2^2h^2}{3.4}\left(2(\mathcal{M}+1)N\right)^2 + \ldots \leq \exp\left(4(\mathcal{M}+1)Nh\right).$$

Ainsi, nous en concluons que

$$\begin{aligned}
\mathbb{E}_{k-1}\left[\exp(hF_k)\right] &\leq 1 + 2h^2 \mathbb{E}\left[\,||Z_k||_j^2\,\right] \exp\left(4(\mathcal{M}+1)Nh\right) \\
&\leq \exp\left(2h^2 \mathbb{E}\left[\,||Z_k||_j^2\,\right] \exp\left(4(\mathcal{M}+1)Nh\right)\right) \qquad \text{puisque } 1+x \leq \exp x,
\end{aligned}$$

ce qui achève la démonstration de l'inégalité (6.138) et par conséquent celle de (6.134). Alors en posant $h = \lambda/2b$ dans l'inégalité (6.134), nous obtenons la première partie de l'inégalité (6.135). Nous appliquons ensuite les inégalités (6.132) et (6.133) pour obtenir la seconde partie de l'inégalité (6.135). □

Chapitre 7

Estimation non-paramétrique de la courbe de régression

7.1. Introduction et préliminaires

7.1.1. Introduction

Soient X un vecteur aléatoire de $\mathbb{R}^p$, Y une variable aléatoire réelle et (X_i, Y_i), $i = 1, \cdots, n$ une suite de n couples aléatoires indépendants ayant même loi que le couple (X, Y). On désigne par r la fonction de régression de la variable aléatoire Y sur la variable aléatoire X définie par

$$r(x) = \mathbb{E}[Y|X = x] = \frac{\int y f_{X,Y}(x, y)dy}{f_X(x)}, \tag{7.1}$$

où $f_{X,Y}$ désigne la fonction de densité jointe et f_X la fonction de densité marginale. On considère le problème de l'estimation de cette fonction, lorsqu'on ne fait a priori aucune hypothèse sur la loi du couple (X, Y) autre que celle de l'existence de la fonction de régression r.

Dans ce chapitre, nous nous proposons de passer en revue des travaux relevant de la statistique mathématique et relatifs au problème d'estimation non-paramétrique de la courbe de régression r définie en (7.1). Il existe dans la littérature plusieurs estimateurs non-paramétriques de la régression r (à ce propos, on pourra, par exemple, consulter l'article de Collomb [21]). Ici nous allons nous intéresser uniquement à la méthode du noyau. Pour être plus précis, un noyau $K(\cdot)$ est une fonction réelle de variable réelle (pas nécessairement de signe constant), qui vérifie, pour une constante réelle $M > 0$ convenable, les conditions suivantes.

(K.1) $\int_{\mathbb{R}} K(u)du = 1$;

(K.2) $K(\cdot)$ est continue sur $(-M/2, M/2]$ et à variation bornée sur $\mathbb{R}$;

(K.3) $K(u) = 0$ pour $|u| \geq M/2$.

Cet estimateur r_n défini, pour tout entier $n \geq 1$, par

$$r_n(x) = \begin{cases} \dfrac{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} Y_i K\left(\dfrac{x-X_i}{h_n}\right)}{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right)} & \text{si } \dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) \neq 0, \\ 0 & \text{si } \dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) = 0. \end{cases} \tag{7.2}$$

a été introduit et étudié par Nadaraya [102] et Watson [135] indépendamment et simultanément. Nadaraya donne dans son article quelques résultats de convergence pour le cas $p = 1$, en signalant que les démonstrations, omises, sont analogues à celles de Parzen [109] sur l'estimateur de densité défini, pour tout $x \in \mathbb{R}$, par

$$f_n(x) = \frac{1}{nh_n^p}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right) \tag{7.3}$$

déjà étudié par Rosenblatt [115]. Watson [135] l'étudie à l'aide d'échantillons empiriques de loi connue. L'intérêt pour l'estimateur r_n de la régression a toujours augmenté ce qui a provoqué de nombreux travaux dans ce domaine. Citons, par exemple les travaux de Collomb [19]-[20], Devroye [39]-[40], Nadaraya [102], Stone [129] et également l'ouvrage de Bosq et Lecoutre [11] dans lequel un chapitre entier est consacré aux propriétés de l'estimateur r_n. En raison de la forme de l'estimateur r_n qui est un rapport, les propriétés de convergence de cet estimateur sont obtenues plus difficilement et avec des conditions plus restrictives que pour l'estimateur f_n de la densité f de la variable aléatoire X.

Les résultats que nous allons rappeler ici sont souvent des résultats asymptotiques, liant la convergence de l'estimateur r_n vers la fonction r, selon divers modes, aux propriétés de la suite réelle $(h_n)_{n\geq 1}$ intervenant dans la définition de l'estimateur r_n. Ces résultats font également intervenir des hypothèses sur la loi du couple (X, Y), hypothèses que nous allons énoncer dans le paragraphe ci-dessous.

Remarque 7.1.1. Il faut noter que le dénominateur de l'estimateur r_n est constitué, via un coefficient, de l'estimateur à noyau de Rosenblatt-Parzen de la densité f de la variable aléatoire X. On peut consulter à cet effet Rosenblatt [115], Parzen [109].

7.1.2. Hypothèses générales

Par souci du détail, nous allons rappeler brièvement la définition de méthodes non paramétriques. La méthode considérée ici c'est-à-dire celle du noyau est qualifiée de non paramétrique parce qu'elle ne se ramène pas à l'estimation d'un nombre fini de paramètres réels associés à la loi de probabilité $\mathbb{P}$ du couple (X, Y). Cette loi est cependant soumise à des hypothèses trés variées comme

– la seule condition $\mathbb{E}|Y|^q < \infty$, $q \geq 1$, lorsque l'estimateur r_n vérifie la propriété de convergence universelle (Devroye et Wagner [41])

$$\forall q \geq 1, \quad \mathbb{E}|Y|^q < \infty \Rightarrow \mathbb{E}|r_n(x) - r(x)|^q \to 0, \tag{7.4}$$

le vecteur aléatoire X est indépendant de $\{(X_i, Y_i), i = 1, \cdots, n\}$,

– lorsqu'on étudie la convergence de l'estimateur à noyau $r_n(x)$ vers la courbe de régression $r(x)$, x fixé (respectivement la convergence uniforme de l'estimateur à noyau r_n vers la fonction de régression r, sur une partie non vide A de $\mathbb{R}^p$), les conditions imposées sont les suivantes.

(C.1) La fonction r est continue au point x (respectivement sur un voisinage de A).

(C.2) La loi marginale $\mathbb{P}_X$ admet par rapport à la mesure de Lebesgue sur $\mathbb{R}^p$ une fonction de densité f minorée par un nombre positif sur un voisinage de x (respectivement sur un voisinage de A).

(C.3) On a $\mathbb{E}[Y^2] < \infty$ et la variance conditionnelle $\mathcal{V}$ définie par

$$\forall x \in \mathbb{R}^p, \quad \mathcal{V}(x) = \mathbb{E}[(Y - r(x))^2 | X = x] \tag{7.5}$$

est bornée sur un voisinage de x (respectivement sur un voisinage de A).

Ces hypothèses sont de simples hypothèses de régularité, analytique sur la fonction de régression r et probabiliste sur $\mathbb{P}_X$ et la loi conditionnelle $\mathbb{P}^X$ de Y par rapport à X. Beaucoup d'auteurs font intervenir des hypothèses plus restrictives comme Y bornée ou au contraire rendent moins restrictives ces mêmes hypothèses, par exemple Collomb [19] pour l'existence de $\mathbb{E}[Y]$. Ici, nous ne discuterons pas de l'aspect de ce problème mais nous étudierons l'évolution de la nature des résultats de convergence.

7.1.3. Position du problème

Afin de situer les résultats énoncés par la suite, nous donnons un résultat négatif qui est une conséquence directe (Bosq [9]) d'un théorème dû à Bickel et Lehman [3].

Sous les conditions (C.1–2–3), il n'existe pas d'estimateur non biaisé de la fonction r, en ce sens que si $\mathcal{D}$ désigne l'ensemble des lois $\mathbb{P}$ vérifiant les conditions (C.1–2–3), il n'existe pas d'estimateur $\hat{r}_n$ vérifiant

$$\forall \mathbb{P} \in \mathcal{D}, \quad \forall x \in \mathbb{R}^p, \quad \mathbb{E}[\hat{r}_n(x)] = r(x).$$

Cependant, notons de suite que les estimateurs non-paramétriques r_n que nous étudions ici sont *asymptotiquement* sans biais. Pour la démonstration de ce résultat, nous renvoyons au théorème 7.2.3.

7.1.4. Remarque d'ordre historique

Le terme de régression trouve son origine dans les travaux de Pearson et Lee [110] qui introduisent la notion de droite de régression à propos d'observations du couple de variables (X, Y)=(taille du père, taille du fils). Cette droite de régression est déterminée à l'aide d'une représentation graphique (diagramme 1 de Pearson et Lee [110]) associée aux observations (X_i, Y_i), $i = 1 \ldots, n$, et qui n'est qu'un régressogramme. Sur ce travail, qui a été fondamental dans l'évolution ultérieure de la statistique, nous remarquons que

- le régressogramme de Pearson et Lee apporte plus d'informations, pour la variable X 'grand' ou 'petit', que la droite de régression qui en est issue.
- ce régressogramme est introduit pour simplifier le calcul (peu automatisé à cette époque) des coefficients de la droite de régression, mais, curieusement, l'étude mathématique du régressogramme, très tardive par rapport à Pearson et Lee repose sur le développement de la statistique non-paramétrique, développement qui est lié à celui des moyens automatiques de calcul.

7.2. Propriétés asymptotiques de l'estimateur à noyau

Soit (X, Y) un vecteur aléatoire de $\mathbb{R}\times\mathbb{R}$ de fonction de densité jointe $f_{X,Y}$ et f_X la fonction de densité marginale de la variable aléatoire X. Dans ce paragrahe, nous supposons que $\mathbb{E}[|Y|] < \infty$. Le noyau $K(\cdot)$ introduit ci-dessus est une fonction de densité quelconque qui, dans le cadre de ce paragraphe, vérifie les conditions suivantes.

(A.1) $\sup_{x\in\mathbb{R}} K(x) < \infty$,

(A.2) $\lim_{|x|\to\infty} |x|K(x) = 0$,

(A.3) $K(x) = K(-x)$,

(A.4) $x^2K(x) \in L_1(\mathbb{R})$.

Introduisons, pour $n \geq 1$ et pour tout réel x, les fonction suivantes

$$m_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h_n}\right), \tag{7.6}$$

$$f_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right), \tag{7.7}$$

$$m(x) = \int_{\mathbb{R}} y f_{X,Y}(x, y)\, dy, \tag{7.8}$$

$$V(x) = \int_{\mathbb{R}} y^2 f_{X,Y}(x, y)\, dy. \tag{7.9}$$

Théorème 7.2.1. *Supposons que $\mathbb{E}[Y^2] < \infty$ et $nh_n \to \infty$ lorsque $n \to \infty$. Alors pour tout réel x pour lequel les fonctions r, f_X, m et V définies respectivement en* (7.6), (7.7), (7.8) *et* (7.9) *sont continues et pour tout réel x, $f_X(x) > 0$, r_n défini en* (7.2) *est un estimateur convergent de la régression r.*

Démonstration. D'après l'article de Parzen [109], f_n est un estimateur convergent de la densité f. Il suffit donc de montrer que la fonction m_n définie (7.6) en est un estimateur convergent de la fonction m définie en (7.8). Pour cela, calculons $\mathbb{E}[m_n(x)]$. Nous avons

$$\begin{aligned}
\mathbb{E}[m_n(x)] &= \frac{1}{h_n}\mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\right] = \frac{1}{h_n}\int_{\mathbb{R}}\int_{\mathbb{R}} yK\left(\frac{x-u}{h_n}\right) f_{X,Y}(u,y)\,du\,dy \\
&= \frac{1}{h_n}\int_{\mathbb{R}} \mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\middle|X=u\right] f_X(u)du \qquad (7.10)\\
&= \frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{x-u}{h_n}\right) r(u) f_X(u)du \\
&= \frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{x-u}{h_n}\right) m(u)du. \qquad (7.11)
\end{aligned}$$

À ce stade de la démonstration, rappelons le théorème 1A. de l'article de Parzen [109].

Théorème 7.2.2. *Supposons que $K(\cdot)$ soit une fonction de Borel qui vérifie les conditions suivantes.*

$$\begin{aligned}
\sup_{u\in\mathbb{R}} |K(u)| &< \infty, \qquad (7.12)\\
\int_{\mathbb{R}} |K(u)|\,du &< \infty, \qquad (7.13)\\
\lim_{u\to\infty} |uK(u)| &= 0. \qquad (7.14)
\end{aligned}$$

Soit $f(u)$ une fonction qui vérifie

$$\int_{\mathbb{R}} |f(u)|\,du < \infty. \qquad (7.15)$$

Soit $(h_n)_{n\geq 1}$ une suite de réels positifs qui vérifie

$$\lim_{n\to\infty} h_n = 0. \qquad (7.16)$$

Définissons

$$f_n(x) = \frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{u}{h_n}\right) f(x-u)\,du. \qquad (7.17)$$

Alors, en tout point x où la fonction f est continue, nous avons

$$\lim_{n\to\infty} f_n(x) = f(x)\int_{\mathbb{R}} K(u)du. \qquad (7.18)$$

Démonstration. Pour la démonstration de ce théorème, on pourra consulter l'article de Parzen [109]. □

Comme la fonction m est intégrable et continue au point x, on peut donc appliquer le théorème 7.2.2 et on obtient alors que

$$\mathbb{E}[m_n(x)] \to m(x) \quad \text{lorsque} \quad n \to \infty. \tag{7.19}$$

Maintenant calculons $\mathrm{Var}[m_n(x)]$. On a alors

$$\begin{aligned}
\mathrm{Var}[m_n(x)] &= \mathbb{E}[m_n^2(x)] - (\mathbb{E}[m_n(x)])^2 \\
&= \frac{1}{nh_n^2}\mathbb{E}\left[Y^2K^2\left(\frac{x-X}{h_n}\right)\right] + \frac{n-1}{nh_n^2}\left(\mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\right]\right)^2 \\
&\quad -\frac{1}{h_n^2}\left(\mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\right]\right)^2 \\
&= \frac{1}{nh_n^2}\mathbb{E}\left[Y^2K^2\left(\frac{x-X}{h_n}\right)\right] - \frac{1}{nh_n^2}\left(\mathbb{E}\left[YK\left(\frac{x-X}{h_n}\right)\right]\right)^2 \\
&= \frac{1}{nh_n^2}\int_{\mathbb{R}} K^2\left(\frac{x-u}{h_n}\right)V(u)du - \frac{1}{n}\left(\mathbb{E}[m_n(x)]\right)^2 \\
&\sim \frac{1}{nh_n}V(x)\int_{\mathbb{R}} K^2(u)du, \qquad (7.20)
\end{aligned}$$

en appliquant de nouveau le théorème 7.2.2.

Les relations (7.19) et (7.20) impliquent que $\mathbb{E}\left[(m_n(x) - m(x))^2\right] \to 0$ lorsque $n \to \infty$. Le théorème 7.2.1 est ainsi prouvé. □

Théorème 7.2.3. *a) Si $|Y| \le \mathcal{C}_1 < \infty$ presque sûrement et $nh_n \to \infty$ lorsque $n \to \infty$, alors on a*

$$\mathbb{E}[r_n(x)] = \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} + O\left(\frac{1}{nh_n}\right) \qquad \textit{lorsque } n \to \infty. \tag{7.21}$$

b) Si $\mathbb{E}[Y^2] < \infty$ et $nh_n^2 \to \infty$ lorsque $n \to \infty$, alors on a

$$\mathbb{E}[r_n(x)] = \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} + O\left(\frac{1}{\sqrt{n}h_n}\right) \qquad \textit{lorsque } n \to \infty. \tag{7.22}$$

Remarque 7.2.1. Les assertions a) et b) impliquent que r_n est un estimateur asymptotiquement sans biais de la courbe de régression r pour tout point x où les fonctions m, f_X et V sont continues.

Démonstration. En utilisant l'identité

$$\frac{1}{u} = \frac{1}{C} - \frac{u-C}{C^2} + \frac{(u-C)^2}{uC^2}, \quad C \neq 0, \quad u \neq 0$$

nous obtenons

$$\begin{aligned}
\mathbb{E}[r_n(x)] &= \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} - \frac{\mathbb{E}\left[m_n(x)(f_n(x) - \mathbb{E}[f_n(x)])\right]}{(\mathbb{E}[f_n(x)])^2} + \mathbb{E}\left[\frac{m_n(x)(f_n(x) - \mathbb{E}[f_n(x)])^2}{f_n(x)\left(\mathbb{E}[f_n(x)]\right)^2}\right] \\
&= \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} + \frac{C_n^1(x) + C_n^2(x)}{(\mathbb{E}[f_n(x)])^2}. \qquad (7.23)
\end{aligned}$$

Nous allons d'abord prouver a). Pour cela, nous avons besoin d'un lemme.

Lemme 7.2.1. *Pour la variance de l'estimateur f_n, nous avons*

$$\operatorname{Var}[f_n(x)] \sim \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(u)\,du \qquad \textit{lorsque } n \to \infty. \tag{7.24}$$

Démonstration. Ce résultat a été démontré dans le cadre de l'article de Parzen [109]. □

En combinant (7.24) et (7.20), on en déduit alors que

$$C_n^1(x) = O\left(\frac{1}{nh_n}\right) \qquad \text{lorsque } n \to \infty. \tag{7.25}$$

Comme la distribution de la variable aléatoire Y est bornée, on obtient que

$$\left|C_n^2(x)\right| \leq \mathcal{C}_1 \mathbb{E}\left[(f_n(x) - \mathbb{E}[f_n(x)])^2\right] \sim \frac{1}{nh_n}\, \mathcal{C}_1\, f(x) \int_{\mathbb{R}} K^2(u)du \qquad \text{lorsque } n \to \infty. \tag{7.26}$$

Les relations (7.23), (7.25) et (7.26) conduisent à l'assertion a).

Pour démontrer l'assertion b), remarquons d'abord que la relation (7.25) reste valable dans le cas présent. Intéressons nous maintenant à la quantité $C_n^2(x)$. On a

$$\begin{aligned} \left|C_n^2(x)\right| &\leq \mathbb{E}\left[\max\{|Y_1|, \cdots, |Y_n|\} \left|f_n(x) - \mathbb{E}[f_n(x)]\right|^2\right] \\ &\leq \left(\sum_{i=1}^n \mathbb{E}[Y_i^2]\right)^{1/2} \left(\mathbb{E}\left[(f_n(x) - \mathbb{E}[f_n(x)])^4\right]\right)^{1/2} \\ &\leq \sqrt{n}\left(\mathbb{E}[Y^2]\right)^{1/2} \left(\mathbb{E}\left[(f_n(x) - \mathbb{E}[f_n(x)])^4\right]\right)^{1/2}. \end{aligned}$$

En utilisant le lemme 1.2.3 p.35 du livre de Nadaraya [105], nous obtenons que

$$\mathbb{E}\left[(f_n(x) - \mathbb{E}[f_n(x)])^4\right] = O\left(\frac{1}{n^2 h_n^2}\right).$$

Par conséquent, on en déduit que

$$C_n^2(x) = O\left(\frac{1}{\sqrt{n} h_n}\right). \tag{7.27}$$

Les relations (7.23), (7.25) et (7.27) conduisent à l'assertion b). □

Théorème 7.2.4. *Supposons que $\mathbb{E}[|Y|^{2+\delta}] < \infty$ pour $\delta \in (1/2, 1]$. De plus, supposons que la fonction V est continue, que les fonctions m et f_X admettent des dérivées bornées jusqu'à l'ordre 2. Alors, si $n^\delta h_n^{2+\delta} \to \infty$ et $nh_n^5 \to 0$ lorsque $n \to \infty$, nous obtenons que*

$$\sqrt{nh_n}\,(r_n(x) - r(x))$$

suit asymptotiquement une loi normale $\mathcal{N}(0, \sigma(x))$ où

$$\sigma^2(x) = \frac{D\,(Y|X = x)}{f_X(x)} \times \int_{\mathbb{R}} K^2(u)du, \tag{7.28}$$

où la fonction D est définie par

$$D\,(Y|X=x) = \frac{V(x)}{f_X(x)} - \left(\frac{m(x)}{f_X(x)}\right)^2.$$

Démonstration. Nous avons

$$r_n(x) - \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} = \frac{\zeta_n(x)}{f_n(x)} \times \frac{1}{\mathbb{E}[f_n(x)]}, \tag{7.29}$$

où

$$\zeta_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n}(\alpha_n Y_i - \beta_n)\,K\left(\frac{x-X_i}{h_n}\right),$$

et où l'on a posé

$$\alpha_n = \mathbb{E}[f_n(x)] \to f_X(x) \quad \text{lorsque } n\to\infty, \quad \text{d'après l'article de Parzen [109]}$$

et

$$\beta_n = \mathbb{E}[m_n(x)] \to m(x) \quad \text{lorsque } n\to\infty \quad \text{d'après (7.19).}$$

De plus, par un calcul, on remarque que

$$\mathbb{E}[\zeta_n(x)] = 0 \tag{7.30}$$

D'après (7.30), il en découle que

$$\begin{aligned} n\,h_n\,D\,\zeta_n(x) =&\ n\,h_n\left(\mathbb{E}[\zeta_n^2(x)] - \mathbb{E}^2[\zeta_n(x)]\right) = n\,h_n\,\mathbb{E}[\zeta_n^2(x)] \\ =&\ \frac{\alpha_n^2}{h_n}\int_{\mathbb{R}} K^2\left(\frac{x-u}{h_n}\right)V(u)f_X(u)\,du - 2\,\frac{\alpha_n\beta_n}{h_n}\int_{\mathbb{R}} K^2\left(\frac{x-u}{h_n}\right)m(u)\,du \\ &+ \frac{\beta_n^2}{h_n}\int_{\mathbb{R}} K^2\left(\frac{x-u}{h_n}\right)f_X(u)\,du. \end{aligned}$$

En appliquant de nouveau le théorème 7.2.2, nous en déduisons que

$$n\,h_n\,D\,\zeta_n(x) \to \nu(x) = D\,(Y|X=x)\,f_X^3(x)\int_{\mathbb{R}} K^2(u)\,du \quad \text{lorsque } n\to\infty. \tag{7.31}$$

Nous allons montrer que la variable aléatoire

$$\frac{\zeta_n(x)}{\sqrt{D\zeta_n(x)}}$$

suit asymptotiquement une loi normale $\mathcal{N}(0,1)$. Pour cela, il suffit de montrer que

$$n\,\mathbb{P}\left[\left|\frac{z_n}{\sqrt{Dz_n}}\right| \geq \varepsilon\sqrt{n}\right] \to 0 \quad \text{lorsque } n\to\infty \tag{7.32}$$

où

$$z_n = \frac{1}{h_n} \left(\alpha_n Y - \beta_n \right) K \left(\frac{x - X}{h_n} \right).$$

Pour vérifier (7.32), il suffit de montrer que

$$L_n = \frac{\mathbb{E}\left[\, |z_n|^{2+\delta} \, \right]}{n^{\delta/2} (D z_n)^{1+\delta/2}} \to 0. \tag{7.33}$$

La condition (7.33) est remplie. En effet, nous avons

$$\begin{aligned}
\mathbb{E}\left[\, |z_n|^{2+\delta} \, \right] &\leq \frac{\alpha_n^{2+\delta}}{h_n^{2+\delta}} \int_{\mathbb{R}} \int_{\mathbb{R}} |y|^{2+\delta} K^{2+\delta} \left(\frac{x-u}{h_n} \right) f_{X,Y}(u,y) \, du \, dy \\
&\quad + \frac{\beta_n^{2+\delta}}{h_n^{2+\delta}} \int_{\mathbb{R}} \int_{\mathbb{R}} K^{2+\delta} \left(\frac{x-u}{h_n} \right) f_{X,Y}(u,y) \, du \, dy \\
&\leq \frac{C_2}{h_n^{2+\delta}} \, \mathbb{E}\left[|Y|^{2+\delta} \right]
\end{aligned}$$

et

$$\left(h_n \, D z_n \right)^{1+\delta/2} \to \left(\nu(x) \right)^{1+\delta/2}.$$

Ainsi, nous obtenons

$$L_n = O\left(\frac{1}{h_n \, (n h_n)^{\delta/2}} \right).$$

Il s'ensuit, d'après (7.31) et de l'assertion précédente que

$$\sqrt{\frac{n \, h_n}{\nu(x)}} \, \zeta_n(x)$$

suit asymptotiquement une loi normale $\mathcal{N}(0,1)$. D'où, en utilisant (7.29) et en utilisant le fait que $f_n(x)\mathbb{E}[f_n(x)]$ converge en probabilité vers $f_X^2(x)$, nous obtenons que

$$\mathbb{P}\left[\frac{\sqrt{n \, h_n}}{\sigma(x)} \left(r_n(x) - \frac{\mathbb{E}\left[m_n(x) \right]}{\mathbb{E}[f_n(x)]} \right) < \lambda \right] \to \phi\left(\lambda \right). \tag{7.34}$$

Maintenant, nous utilisons (7.10) pour obtenir que

$$\begin{aligned}
\mathbb{E}\left[m_n(x) \right] &= m(x) + \frac{h_n^2}{2} \int_{\mathbb{R}} m''(x) \left(x + \theta u h_n \right) u^2 K(u) \, du \\
&= m(x) + O\left(h_n^2 \right), \quad 0 < \theta < 1.
\end{aligned} \tag{7.35}$$

De façon analogue, on peut démontrer que

$$\mathbb{E}\left[f_n(x) \right] = f_X(x) + O\left(h_n^2 \right). \tag{7.36}$$

Ainsi, (7.34), (7.35) et (7.36) impliquent la validité du théorème 7.2.4. □

Théorème 7.2.5. *Supposons que $|Y| \leq \mathcal{C}_2 < \infty$ presque sûrement, la fonction V soit continue, les fonctions m et f_X admettent des dérivées d'ordre 1 bornées.*

a) Si

$$\frac{1}{nh_n} \to 0 \quad \textit{lorsque } n \to 0,$$

alors lorsque $n \to \infty$, nous avons

$$r_n(x) - r(x) = \frac{1}{f_X(x)} \left[\eta_n(x) - \mathbb{E}[\eta_n(x)]\right] + O_{\mathbb{P}}\left(h_n + \frac{1}{nh_n}\right), \tag{7.37}$$

où

$$\eta_n(x) = \frac{1}{n}\sum_{i=1}^{n} \eta_n^{(i)}(x), \quad \textit{et} \quad \eta_n^{(i)}(x) = \frac{1}{h_n}\left(Y_i - r(x)\right) K\left(\frac{x - X_i}{h_n}\right). \tag{7.38}$$

b) Si

$$n\, h_n^3 \to 0 \quad \textit{et} \quad \frac{1}{nh_n} \to 0 \quad \textit{lorsque } n \to 0,$$

alors nous avons

$$\sqrt{n\, h_n}\,(r_n(x) - r(x))$$

suit asymptotiquement une loi normale $\mathcal{N}(0, \sigma(x))$, où $\sigma(x)$ est défini en (7.28).

Démonstration. Pour tout réel x, posons

$$\overline{r}_n(x) = r_n(x) - r(x).$$

D'après la définition (7.38) de la fonction $\eta_n(x)$, nous avons

$$\eta_n(x) = m_n(x) - r(x) f_n(x).$$

D'autre part, nous avons

$$\begin{aligned}
\overline{r}_n(x) - \frac{\eta_n(x)}{f_X(x)} &= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - \frac{\eta_n(x)}{\overline{r}_n(x)}\right) \\
&= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - \frac{m_n(x) - r(x)f_n(x)}{r_n(x) - r(x)}\right) \\
&= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - \frac{m_n(x) - \frac{r(x)m_n(x)}{r_n(x)}}{r_n(x) - r(x)}\right) \\
&= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - \frac{m_n(x)}{r_n(x)} \times \frac{r_n(x) - r(x)}{r_n(x) - r(x)}\right) \\
&= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - \frac{m_n(x)}{r_n(x)}\right) \\
&= \frac{\overline{r}_n(x)}{f_X(x)}\left(f_X(x) - f_n(x)\right).
\end{aligned}$$

Ainsi, nous avons

$$\mathbb{E}\left[\left|\overline{r}_n(x) - \frac{\eta_n(x)}{f_X(x)}\right|\right] = \mathbb{E}\left[\left|\frac{\overline{r}_n(x)(f_n(x) - f_X(x))}{f_X(x)}\right|\right].$$

De plus, comme $ab \leq 2^{-1}(a^2 + b^2)$, nous avons alors

$$\mathbb{E}\left[\left|\overline{r}_n(x) - \frac{\eta_n(x)}{f_X(x)}\right|\right] \leq \frac{1}{2f_X(x)} \left(\mathbb{E}\left[(r_n(x) - r(x))^2\right] + \mathbb{E}\left[(f_n(x) - f_X(x))^2\right]\right). \quad (7.39)$$

L'identité

$$r_n(x) = \frac{m_n(x)}{f_X(x)} - \frac{m_n(x)}{f_X(x)f_n(x)}\left(f_n(x) - f_X(x)\right)$$

implique que

$$\mathbb{E}\left[(r_n(x) - r(x))^2\right] = \frac{1}{f_X^2(x)}\mathbb{E}\left[\left(m_n(x) - m(x) - \frac{m_n(x)}{f_n(x)}\left(f_n(x) - f_X(x)\right)\right)^2\right].$$

Ainsi en considérant l'inégalité

$$\frac{m_n(x)}{f_n(x)} \leq \mathcal{C}_2 \quad \text{p.s.}$$

(cette dernière s'ensuit du fait que la variable aléatoire Y est bornée presque sûrement), nous obtenons

$$\mathbb{E}\left[(r_n(x) - r(x))^2\right] \leq \frac{\mathcal{C}_3}{f_X^2(x)} \left(\mathbb{E}\left[(m_n(x) - m(x))^2\right] + \mathbb{E}\left[(f_n(x) - f_X(x))^2\right]\right). \quad (7.40)$$

Estimons d'abord la quantité

$$\mathbb{E}\left[(m_n(x) - m(x))^2\right].$$

La continuité et l'intégrabilité de la fonction V impliquent que

$$Dm_n(x) \leq \frac{1}{nh_n}\int_{\mathbb{R}} K^2\left(\frac{x-u}{h_n}\right)V(u)\,du = O\left(\frac{1}{nh_n}\right).$$

Ensuite, la condition sur la fonction m conduit à

$$\mathbb{E}\left[m_n(x)\right] = m(x) + O\left(h_n\right).$$

Par conséquent, nous en déduisons que

$$\mathbb{E}\left[(m_n(x) - m(x))^2\right] = O\left(\frac{1}{nh_n} + h_n^2\right). \quad (7.41)$$

De façon analogue, on peut montrer que

$$\mathbb{E}\left[(f_n(x) - f_X(x))^2\right] = O\left(\frac{1}{nh_n} + h_n^2\right). \quad (7.42)$$

Ainsi d'après (7.41) et (7.42), nous obtenons que

$$\mathbb{E}\left[(r_n(x) - r(x))^2\right] = O\left(\frac{1}{nh_n} + h_n^2\right). \tag{7.43}$$

D'après (7.42), (7.43) et d'après (7.39) nous avons que

$$r_n(x) - r(x) = \frac{\eta_n(x)}{f_X(x)} + O_{\mathbb{P}}\left(\frac{1}{nh_n} + h_n^2\right).$$

Enfin, pour compléter la preuve du théorème, il suffit seulement de remarquer que la borne

$$\mathbb{E}\left[\eta_n(x)\right] = O\left(h_n\right).$$

L'assertion a) est ainsi prouvée.

L'assertion b) découle directement de (7.37) puisque

$$\frac{\eta_n(x) - \mathbb{E}[\eta_n(x)]}{\sqrt{D\eta_n(x)}}$$

suit asymptotiquement une loi normale $\mathcal{N}(0, \sigma(x)\, f_X(x))$. □

Remarque 7.2.2. Les théorèmes 7.2.4 et 7.2.5 nous permettent de construire un intervalle de confiance pour la courbe de régression r. À ce propos, l'écart-type $\sigma(x)$ pourrait être remplacé par un estimateur convergent. Par exemple, on peut choisir pour cet estimateur la statistique suivante

$$\sigma_n^2(x) = nh_n \times \frac{\displaystyle\sum_{i=1}^{n}(Y_i - r_n(x))^2 K\left(\frac{x - X_i}{h_n}\right)}{\left(\displaystyle\sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)\right)^2} \times \int_{\mathbb{R}} K^2(u)\, du.$$

Sous les conditions des théorèmes 7.2.4 et 7.2.5, nous avons

$$\mathbb{P}\left[|r_n(x) - r(x)| < \frac{\lambda\sigma_n(x)}{\sqrt{nh_n}}\right] \to 2\phi(\lambda) - 1 \quad \text{lorsque } n \to \infty.$$

Remarque 7.2.3. La façon la plus simple pour caractériser l'erreur de l'estimation de la courbe de régression r inconnue est d'utiliser l'espérance mathématique de la déviation quadratique

$$U_n = \mathbb{E}\left[\int_{\mathbb{R}} \eta_n^2(x)\, h(x)\, dx\right] = \mathbb{E}\left[\int_{\mathbb{R}} (r_n(x) - r(x))^2 f_n^2(x)\, h(x)\, dx\right],$$

où h est la fonction de poids que nous supposerons bornée et intégrable. La quantité U_n définit la fonction risque de l'estimateur r_n.

Théorème 7.2.6. *Supposons que les dérivées d'ordre deux des fonctions m et f_X soient continues, bornées et de carré intégrable avec la fonction de poids h. De plus, si*

$$\sup_{x\in\mathbb{R}} |V(x)| < \infty, \quad r(x)\, h(x) \quad et \quad r^2(x)\, h(x) \quad \in L_1(\mathbb{R})$$

alors nous avons

$$U_n \sim \frac{1}{nh_n} \mathcal{A}_1(f_{X,Y}, K) + h_n^4 \mathcal{A}_2(f_{X,Y}, K), \tag{7.44}$$

où

$$\begin{aligned}
\mathcal{A}_1(f_{X,Y}, K) &= \int_{\mathbb{R}} D\,(Y|X=x)\, f_X(x)\, h(x)\, dx \int_{\mathbb{R}} K^2(u)\, du, \\
\mathcal{A}_2(f_{X,Y}, K) &= \frac{1}{4} \int_{\mathbb{R}} (m''(x) - f_X''(x) r(x))^2\, h(x)\, dx \left(\int_{\mathbb{R}} u^2\, K(u)\, du \right)^2.
\end{aligned}$$

Démonstration. Le théorème de Fubini conduit à

$$\begin{aligned}
U_n &= \frac{1}{nh_n^2} \int_{\mathbb{R}} \mathbb{E}\left[(Y - r(x))^2 \right] K^2 \left(\frac{x - X}{h_n} \right) h(x)\, dx \\
&\quad + \left(1 - \frac{1}{n} \right) \frac{1}{h_n^2} \int_{\mathbb{R}} \left(\mathbb{E}\left[Y - r(x) \right] K \left(\frac{x - X}{h_n} \right) \right)^2 h(x)\, dx \\
&= \mathcal{D}_n + \mathcal{E}_n.
\end{aligned}$$

Étudions d'abord le comportement de $\mathcal{D}_n$. Comme on peut permuter sous le signe $\int$, on peut alors écrire que

$$\begin{aligned}
\mathcal{D}_n &\sim \frac{1}{nh_n} \left(\int_{\mathbb{R}} V(x)\, h(x)\, dx - 2 \int_{\mathbb{R}} m(x)\, r(x)\, h(x)\, dx + \int_{\mathbb{R}} r^2(x)\, f_X(x)\, h(x)\, dx \right) \\
&\qquad \times \int_{\mathbb{R}} K^2(u)\, du \\
&\sim \frac{1}{nh_n} \int_{\mathbb{R}} D\,(Y|X=x)\;\, f_X(x)\, h(x)\, dx \int_{\mathbb{R}} K^2(u)\, du.
\end{aligned}$$

Maintenant nous allons estimer $\mathcal{E}_n$. En utilisant la formule de Taylor avec reste intégral, nous avons

$$\begin{aligned}
\mathcal{E}_n &= h_n^4 \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1-t) \left\{ m''\,(x + tu\, h_n) - r(x) f_X''\,(x + tu\, h_n) \right\}^2 u^2 K(u)\, du\, dt \right)^2 h(x)\, dx \\
&= h_n^4 \times \mathcal{A}_n.
\end{aligned}$$

D'après le théorème de Lebesgue, nous en déduisons que

$$\mathcal{A}_n \to \mathcal{A}_2(f_{X,Y}, K) \quad \text{lorsque } n \to \infty.$$

Ainsi, nous en concluons que

$$\mathcal{E}_n \sim h_n^4 \times \mathcal{A}_2(f_{X,Y}, K).$$

Les comportements asymptotiques de $\mathcal{D}_n$ et $\mathcal{E}_n$ impliquent le théorème 7.2.6.

Maintenant, nous allons déterminer la valeur asymptotique $h_n = h_n^0$ qui maximise le membre de droite de (7.44). La valeur optimale de h_n est égale à

$$h_n^0 = \left(4 \, \frac{\mathcal{A}_2(f_{X,Y}, K)}{\mathcal{A}_1(f_{X,Y}, K)}\right)^{-1/5} n^{-1/5}.$$

Pour ce choix de h_n, nous avons alors

$$U_n \sim \left(4^{1/5} + \frac{1}{4^{4/5}}\right) \mathcal{A}_2^{1/5}(f_{X,Y}, K) \, \mathcal{A}_1^{4/5}(f_{X,Y}, K) \times \frac{1}{n^{4/5}}.$$

$\square$

Remarque 7.2.4. Le théorème 7.2.6 peut se généraliser. Pour cela, introduisons les notations suivantes.

Soit W_S l'ensemble des fonctions $\phi(x)$ qui admettent des dérivées jusqu'à l'ordre s et dont la dérivée d'ordre s, $\phi^{(s)}(x)$, est une fonction continue, monotone par morceaux, bornée et appartenant à $L_2(\mathbb{R})$.

Soit H_S (où $S \geq 2$ est un nombre pair) la classe des fonctions qui vérifient les conditions de régularité suivantes

$$\begin{aligned}
&K(x) = K(-x), \quad \int_{\mathbb{R}} K(u)\, du = 1, \quad \sup_{x \in \mathbb{R}} |K(x)| \leq A < \infty, \\
&\int_{\mathbb{R}} u^i K(u)\, du = 0, \quad \text{pour } i = 1, \ldots, s-1, \quad \int_{\mathbb{R}} u^s K(u)\, du \neq 0, \\
&\int_{\mathbb{R}} u^s \, |K(u)| \, du < \infty.
\end{aligned}$$

En particulier, si les fonctions $m, f_X \in W_S$ et si le noyau $K \in H_S$ alors la valeur optimale de h_n est égale à

$$h_n^0 = \mathcal{C}_4(f_{X,Y}, K)\, n^{-1/(1+2s)}$$

et pour $h_n = h_n^0$, nous avons

$$U_n \sim \mathcal{C}_5(f_{X,Y}, K) \times \frac{1}{n^{2s/(1+2s)}},$$

où $\mathcal{C}_4(f_{X,Y}, K)$ et $\mathcal{C}_5(f_{X,Y}, K)$ sont des fonctions qui dépendent de la fonction $f_{X,Y}$ et du noyau K.

7.3. Propriétés de convergence ponctuelle de l'estimateur à noyau

On s'intéresse ici uniquement aux résultats concernant les propriétés de l'estimateur r_n en un point fixé, ou, éventuellement, en un nombre fini de points.
Rosenblatt [116] et Konakov [85] étudient l'estimateur $r_n(x)$, dans le cas $p = 1$, en l'écrivant sous la forme

$$r_n(x) = \frac{\int_{\mathbb{R}} y g_n(x,y)dy}{\int_{\mathbb{R}} g_n(x,y)dy}, \tag{7.45}$$

où

$$\begin{aligned} g_n(x,y) &= \frac{1}{nh_n^2}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}, \frac{y-Y_i}{h_n}\right), \\ k \text{ noyau de } \mathbb{R}^2 \text{ tel que } &\int_{\mathbb{R}} k(z,y)\,dy = K(z). \end{aligned} \tag{7.46}$$

Ils utilisent les propriétés de g_n en tant qu'estimateur de la densité jointe $f_{X,Y}$ de la loi du couple aléatoire (X, Y), densité dont on suppose alors l'existence et la continuité. Ils obtiennent ainsi divers résultats de convergence ponctuelle, en imposant à la suite $(h_n)_{n\geq 1}$ la condition suivante

$$nh_n^{s+1} \to \infty, \text{ lorsque } n \to \infty.$$

Des démonstrations plus directes (voir Collomb [16]) permettent d'obtenir ces mêmes propriétés sans supposer l'existence de la densité jointe $f_{X,Y}$ et avec sur $(h_n)_{n\geq 1}$ l'hypothèse nécessaire et suffisante (cas $p \geq 1$)

$$nh_n^s \to \infty, \text{ lorsque } n \to \infty.$$

Le résultat s'énonce ainsi.

Théorème 7.3.1. *Supposons que $\mathbb{E}[Y^2] < \infty$. Si l'estimateur r_n est associé à un noyau de Parzen-Rosenblatt, alors en tout point de continuité x des fonctions f, r et v tel que $f(x) \neq 0$, on a*

$$h_n \to 0, \quad nh_n^s \to \infty \quad \Leftrightarrow \quad r_n(x) \to r(x) \quad \mathbb{P}, \qquad \textit{lorsque } n \to \infty,$$

où $\mathbb{P}$ désigne la convergence en probabilité.

Remarque 7.3.1. Un noyau $K(\cdot)$ de Parzen-Rosenblatt est une application de $\mathbb{R}^s$ dans $\mathbb{R}$, bornée, intégrable par rapport à la mesure de Lebesgue, d'intégrale 1 et qui vérifie

$$\lim_{||x||\to\infty} ||x||^s K(x) = 0.$$

Démonstration. Pour la démonstration du théorème 7.3.1, on pourra consulter la thèse de Collomb [16]. □

La première étude directe de l'estimateur $r_n(x)$, pour x fixé, semble être celle de Schuster [120] qui obtient la propriété de convergence asymptotique normale, en corrigeant le résultat de même type donné par Nadaraya [102]. Il montre également que $r_n(x)$ et $r_n(x')$, $x \neq x'$, sont asymptotiquement indépendants au sens de Paul Lévy. Les théorèmes 7.2.4 et 7.2.5 énoncés dans le paragraphe précédent généralisent ces résultats.

Noda [107] étudie les propriétés de convergence ponctuelle et donne une majoration asymptotique de l'erreur quadratique $\mathbb{E}[r_n(x) - r(x)]^2$.

Collomb [18] donne une évaluation asymptotique du biais et de la variance de l'estimateur $r_n(x)$ pour x fixé, en fonction de $h_n, r'(x), r''(x), f_X(x), f'_X(x)$ et $\mathcal{V}(x)$ définie en (7.5). Nous avons alors le théorème (biais et variance asymptotiques) suivant.

Théorème 7.3.2. *Sous les conditions suivantes :*
la courbe de régression r est bornée ;
les fonctions r et f_X sont deux fois différentiables, de dérivées partielles continues et bornées ;
$\mathbb{E}[Y^2] < \infty$ *;*
la variance conditionnelle $\mathcal{V}$ est continue et bornée au voisinage de x ;
K est un noyau de Parzen-Rosenblatt positif pair et tel que $\int_{\mathbb{R}^s} ||u||^3 K(u)\, du < \infty$.
Alors, si $h_n \to 0$ lorsque $n \to \infty$ et s'il existe $\alpha > 0$ tel que $n^{1-\alpha} h_n^s \to \infty$ lorsque $n \to \infty$, en tout point x où la fonction $f_X(x) \neq 0$, on a les développements asymptotiques suivants

$$\mathbb{E}[r_n(x)] - r(x) = \frac{h_n^2}{2} \sum_{i=1}^{s} \sum_{j=1}^{s} b_{i,j}(x) \int_{\mathbb{R}^s} u_i u_j K(u) du + O\left(\frac{1}{nh_n^s}\right)^2 \qquad \textit{lorsque } n \to \infty,$$

où l'on a posé

$$b_{i,j}(x) = \frac{\partial^2 r(x)}{\partial x_i \partial x_j} + \frac{\partial r(x)}{\partial x_i} \frac{\partial \log f_X(x)}{\partial x_j} + \frac{\partial r(x)}{\partial x_j} \frac{\partial \log f_X(x)}{\partial x_i}.$$

et

$$\mathbb{E}[r_n(x) - \mathbb{E}[r_n(x)]]^2 = \frac{1}{nh_n^s} \frac{\mathcal{V}(x)}{f_X(x)} \int_{\mathbb{R}^s} K^2(u) du + o\left(\frac{1}{nh_n^s}\right) \qquad \textit{lorsque } n \to \infty.$$

Par conséquent l'erreur quadratique s'écrit

$$\Delta_n(x) = \frac{h_n^4}{4} \left(\sum_{i=1}^{s} \sum_{j=1}^{s} b_{i,j} \int_{\mathbb{R}^s} u_i u_j K(u) du \right)^2 + \frac{1}{nh_n^s} \frac{\mathcal{V}(x)}{f_X(x)} \int_{\mathbb{R}^s} K^2(u) du + o\left(h_n^4 + \frac{1}{nh_n^s}\right).$$

Démonstration. On peut consulter l'article de Collomb [18]. □

Remarque 7.3.2. Ces formules permettent notamment de montrer que

$$\min_{h_n \in \mathbb{R}^+} \mathbb{E}[r_n(x) - r(x)]^2 \sim \frac{\mathcal{C}}{n^{4/(p+4)}} \qquad \text{lorsque } n \to \infty,$$

où $\mathcal{C}$ est une constante qui dépend uniquement du noyau K et de la loi du couple (X, Y).

Remarque 7.3.3. Rosenblatt [116] a obtenu l'évaluation du biais et de la variance asymptotique pour $s = 1$.

7.4. Propriétés de convergence en norme

L'étude de propriétés de convergence de l'estimateur fonctionnel r_n en fonction de la vitesse avec laquelle $h_n \to 0$ lorsque $n \to \infty$. Les résultats sont analogues à ceux de Bleuez et Bosq [6] relatifs à un estimateur de la densité. Seules les hypothèses sur la loi $\mathbb{P}$ sont plus complexes.

7.4.1. Notations

Introduisons la quantité $d_G(r_n, r)$ définie par

$$d_G(r_n, r) = \sup_{x \in G} |r_n(x) - r(x)|,$$

où G est un ensemble borné de $\mathbb{R}^s$ contenant un pavé $[a, b]^s$. Soit K un noyau de $\mathbb{R}^s$, i.e. un élément de

$$\mathcal{K} = \left\{ K \in L_1(\mathbb{R}^s), |K| \leq M < \infty \text{ et } \lim_{|u| \to \infty} |u|^s K(u) = 0 \right\}.$$

Pour $\varepsilon > 0$ fixé, nous noterons

$$G_\varepsilon = \bigcup_{x \in G} B(x, \varepsilon)$$

et

$$\mathcal{L} = \mathcal{L}(G, \varepsilon)$$

l'ensemble des lois $\mathbb{P}$ telles que f_X et r soient continues sur G_ε, f_X étant minorée sur G par un nombre $\alpha > 0$, les moments conditionnels existant sur G_ε et vérifiant

$$(\exists M > 0), \quad (\forall x \in G_\varepsilon), \quad (p \geq 2), \quad \mathbb{E}\left[\, |Y - r(x)|^p \,|X = x\right] \leq M^p p!.$$

7.4.2. Résultats

Nadaraya [104] et Schuster et Yakowitz [121] dans le cas $s = 1$, Devroye [39] pour le cas $s \geq 1$, obtiennent le résultat suivant

$$\frac{nh_n^{2s}}{\log n} \to \infty \quad \text{lorsque } n \to \infty \quad \Rightarrow \quad d_G(r_n, r) \to 0 \text{ p.s.,} \quad \text{lorsque } n \to \infty.$$

Le théorème 7.4.1 est dû à Collomb [16] ou [19] et généralise les résultats précédents, en particulier il établit une condition nécessaire et suffisante de convergence uniforme presque sûre sur un ensemble borné.

Théorème 7.4.1. *Soit r_n l'estimateur associé à un noyau lipschitzien positif à support borné. Alors, nous avons*

$$\forall P \in \mathcal{D}(G, K), \quad d_G(r_n, r) \to 0 \ \textit{p.s.,} \quad \textit{lorsque } n \to \infty \quad \Rightarrow \quad h_n \to 0, \quad \frac{nh_n^s}{\log n} \to \infty.$$

Inversement, nous avons

$$h_n \to 0, \quad \frac{nh_n^s}{\log n} \to \infty \quad \Rightarrow \quad \forall P \in \mathcal{D}(G,K), \quad d_G(r_n, r) \to 0 \ \textit{p.s.}, \qquad \textit{lorsque } n \to \infty.$$

Démonstration. On pourra consulter à cet effet la thèse [16] ou l'article de Collomb [19]. □

Le corollaire 7.4.1 donne une condition nécessaire et suffisante de convergence uniforme presque complète.

Corollaire 7.4.1. *Dans les conditions du théorème précédent, si* $\mathbb{E}[Y^2] < \infty$*, alors nous avons*

$$h_n \to 0, \quad \frac{nh_n^s}{\log n} \to \infty \quad \Leftrightarrow \quad \forall P \in \mathcal{D}(G,K), \quad d_G(r_n, r) \to 0 \ \textit{p.co.}, \qquad \textit{lorsque } n \to \infty,$$

où p.co. désigne la convergence uniforme presque complète.

Démonstration. On pourra consulter à cet effet l'article de Collomb [19]. □

Remarque 7.4.1. Révész obtient, dans le cas $s = 1$ un résultat de convergence en loi pour $d_G(r_n, r)$. La loi $\mathbb{P}$ est alors assujettie à l'hypothèse supplémentaire $\mathcal{V}(x) = \sigma^2$, pour tout $x \in \mathbb{R}$.

Maintenant nous allons étuder les résultats relatifs à la convergence L^p définie par

$$L_n = \int |r_n(x) - r(x)|^p \mathbb{P}_X(dx), \quad p \geq 1.$$

Dans le cas où $p = 1$, c'est-à-dire dans le cadre de la convergence L^1, nous avons le théorème suivant.

Théorème 7.4.2. *Si l'estimateur* r_n *est associé à un noyau positif à support compact et minoré par un nombre strictement positif. On suppose que la loi* $\mathbb{P}_X$ *de la variable* X *est abolument continue par rapport à la mesure de Lebesgue de* $\mathbb{R}^s$ *et que* $|Y| \leq \mathcal{C}_4$ *preque sûrement. Alors nous avons*

$$h_n \to 0, \quad nh_n^s \to \infty \quad \Rightarrow \quad L_n \to 0 \quad \mathbb{P}, \qquad \textit{lorsque } n \to \infty$$

et

$$h_n \to 0, \quad (\forall a > 0), \quad \sum_{i=1}^{\infty} \exp\left(-anh_n^s\right) < \infty \quad \Rightarrow \quad L_n \to 0 \ \textit{p.s.}, \qquad \textit{lorsque } n \to \infty.$$

Démonstration. On pourra consulter l'article de Devroye et Wagner [42]. □

Dans le cas $p \geq 1$, nous avons la propriété de convergence universelle définie en (7.4), i.e.,

Théorème 7.4.3. *Si r_n est associé à un noyau positif à support compact et minoré par un nombre strictement positif, si $\mathbb{E}[|Y|^p] < \infty$, alors nous avons*

$$h_n \to 0, \quad nh_n^s \to \infty \Rightarrow \mathbb{E}[|r_n(X) - r(X)|^p] \to 0 \qquad \textit{lorsque } n \to \infty.$$

Démonstration. On pourra consulter l'article de Devroye et Wagner [41]. □

Remarque 7.4.2. Cette condition suffisante de convergence L^p a été obtenue également par Spielgelman et Sacks [128].

Remarque 7.4.3. La plupart des résultats de convergence en loi, pour diverses normes de $r_n - r$, sont obtenus par application de résultats analogues à ceux de Révész [112] (voir également l'article de Komlós, Major et Tusnády [83]) sur l'approximation de processus empiriques multidimensionnels et reposent sur une technique de démonstration introduite, dans le cas $s = 1$, par Bickel et Rosenblatt [4] à propos de l'estimateur de la densité f_n défini en (7.3). Rüschendorf [117] passe en revue de nombreux travaux du même type.

7.5. Lois limites

Collomb généralise par "studentisation" le résultat de Schuster et obtient ainsi des résultats à propos de la convergence asymptotiquement normale indépendante conjointement en un nombre fini de points pour l'estimateur r_n énoncés ci-dessous.

Théorème 7.5.1. *Sous les conditions suivantes :*

-i- K est un noyau de Parzen-Rosenblatt positif ;

-ii- $\mathbb{E}[|Y|^p] < \infty$ pour tout entier p ;

les fonctions r, f_X et $\mathcal{V}$ sont continues au point x, les fonctions f_X et $\mathcal{V}$ sont strictement positives en $x_1, \ldots, x_s$, points distincts de $\mathbb{R}^s$ et la fonction $\mathcal{V}$ est bornée au voisinage de ces points.
Si $h_n \to 0$ lorsque $n \to \infty$ et s'il existe $\alpha > 0$ tel que $n^{1-\alpha} h_n^s \to \infty$ alors, pour tout t fixé dans $\mathbb{R}^s$, nous avons

$$\lim_{n\to\infty} \mathbb{P}\left[\sqrt{nh_n^s}\, \frac{r_n(x_i) - \mathbb{E}[r_n(x_i)]}{\sqrt{\frac{\mathcal{V}(x_i)}{f(x_i)} \int_{\mathbb{R}} K^2(u)\, du}} \le t_i, i = 1 \ldots, s\right] = \prod_{i=1}^{s} N(t_i),$$

où N est la fonction de répartition de la loi $\mathcal{N}(0,1)$.

Démonstration. Pour démontrer ce théorème, on pourra consulter l'article de Collomb [17]. □

Proposition 7.5.1. *Soit $W_n(x)$ la variable aléatoire réelle définie pour tout x fixé dans $\mathbb{R}^s$ par*

$$W_n(x) = \begin{cases} \dfrac{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n}(Y_i - r_n(x))^2 K\left(\dfrac{x-X_i}{h_n}\right)}{\left(\displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right)\right)^2} \int_{\mathbb{R}} K^2(u)\,du & si \ \displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) \neq 0, \\ 0 & si \ \displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) = 0. \end{cases}$$

Si la courbe de régression r est lipschitzienne d'ordre α au point $x = x_l, l = 1, \ldots, m$ sur un voisinage de x, le support de K est borné et la suite $(h_n)_{n\geq 1}$ vérifie

$$nh_n^{p+2\alpha} \to 0 \text{ lorsque } n \to \infty,$$

alors, pour tout t fixé dans $\mathbb{R}^s$, nous avons

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{r_n(x_i) - r(x_i)}{W_n(x_i)} \leq t_i, i = 1 \ldots, s\right] = \prod_{i=1}^{s} N(t_i)$$

Démonstration. Pour la démonstration de ce théorème, on pourra consulter l'article de Collomb [17]. □

Remarque 7.5.1. Ce résultat permet de définir pour la courbe de régression r un intervalle de confiance, de sécurité asymptotique donnée.

Le théorème 7.5.1 montre que le processus

$$\sqrt{nh_n^s}\, \frac{r_n(x) - \mathbb{E}[r_n(x)]}{\sqrt{\frac{\mathcal{V}(x)}{f(x)} \int K^2(u)\,du}}, \quad x \in \mathbb{R}^s$$

est asymptotiquement gaussien. Nous allons préciser ce résultat par l'étude de la loi limite de l'écart maximal entre l'estimateur r_n et la fonction r, permettant ainsi la construction de tests et de régions de confiance pour la courbe de régression r. Pour cela introduisons quelques notations et hypothèses.
Supposons que

(H.1) $f_{X,Y}(x,y) = 0$ pour $|y| \geq A$, où A est une constante, c'est-à-dire que la variable aléatoire Y est bornée presque sûrement.

(H.2) La fonction de densité marginale $f_X(x) = \int f_{X,Y}(x,y)dy$ est bornée sur $\mathbb{R}$.

(H.3) Les fonctions f_X, m et V sont continues sur l'intervalle $[a, b]$.

(H.4) $\min_{a\leq x\leq b} f_X(x) > 0$, $\min_{a\leq x\leq b} \beta(x) > 0$, où $\beta(x) = D(Y|X = x)$.

(H.5) Le noyau $K(\cdot)$ est une fonction de densité qui satisfait les 4 conditions suivantes, le noyau $K(\cdot)$ est à variation bornée, $K(-x) = K(x)$, $K(|x_2|) \leq K(|x_1|)$ pour $|x_2| > |x_1|$ et $\int |u|K(u)\,du < \infty$.

Divisons l'intervalle $[a, b]$ en s_n sous-intervalles $\Delta_1, \Delta_2, \ldots, \Delta_{s_n}$ de même longueur τ_n. Posons

$$\zeta_n(t_j) = \frac{\eta_n(t_j) - \mathbb{E}[\eta_n(t_j)]}{\sqrt{D\eta_n(t_j)}}, \qquad \text{où } j = 1, \ldots, s_n \text{ et } \eta_n \text{ est définie en (7.38)}$$

et

$$\varepsilon_n = \max_{1 \leq i < j \leq s_n} |\mathbb{E}\left[\zeta_n(t_j)\zeta_n(t_i)\right]| \,.$$

Soit la fonction $m_n(x)$, définie pour $x \in [a, b]$, par

$$m_n(x) = r_n(t_i), \qquad \text{où } t_i \text{ sont les centres des sous-intervalles } \Delta_i, \text{ où } i = 1, \ldots, s_n. \tag{7.47}$$

Théorème 7.5.2. *Supposons que les fonctions f_X, m et V admettent des dérivées d'ordre 1 bornées. De plus, les conditions suivantes doivent être vérifiées lorsque $n \to \infty$*

$$\begin{aligned}
\varepsilon_n \log s_n &\to 0, \\
nh_n^3 \log s_n &\to 0, \\
\frac{(\log s_n)^m}{nh_n} &\to 0, \qquad \textit{pour un entier } m \geq 0 \textit{ fixé}, \\
\frac{nh_n \log s_n}{s_n^2} &\to 0.
\end{aligned}$$

Alors, nous avons

$$\lim_{n\to\infty} \mathbb{P}\left[\sqrt{nh_n} \max_{a \leq x \leq b} \left|\frac{m_n(x) - r(x)}{\sigma(x)}\right| < K_0 \left(l_{s_n} + \frac{\lambda}{l_{s_n}}\right)\right] = \exp\left(-2\exp\left(-\lambda\right)\right),$$

où l_{s_n} est la racine de l'équation

$$\frac{1}{s_n} = \frac{1}{\sqrt{2\pi}} \int_{l_{s_n}}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx, \quad l_{s_n} = O\left(\sqrt{\log s_n}\right),$$

m_n est définie en (7.47) et $\sigma(x)$ en (7.28).

Démonstration. Pour la démonstration de ce théorème, on pourra consulter le livre de Nadaraya [105]. □

Énonçons maintenant une loi du logarithme itéré dans le cas où $s = 1$. Ce résultat a été établi par Härdle [64]. Pour énoncer ce théorème, introduisons les notations utilisées par Härdle dans [64]. Soit h une fonction de nombres entiers naturels n qui vérifie les conditions suivantes

$$(i) \quad h(n) \to 0 \quad \text{et} \quad (ii) \quad nh(n) \to \infty \quad \text{lorsque} \quad n \to \infty. \tag{7.48}$$

Définissons

$$\widehat{r}_{n,h}(x) = \frac{\dfrac{1}{nh}\displaystyle\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right)}{\dfrac{1}{nh}\displaystyle\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)}, \quad x \in \mathbb{R} \tag{7.49}$$

$$S^2(x) = \mathbb{E}[Y^2|X=x], \quad \mathrm{Var}[Y|X=x] = S^2(x) - r^2(x), \quad x \in \mathbb{R}. \tag{7.50}$$

Supposons que les fonctions f_X et r soient deux fois dérivables et la fonction S^2 est continue. De plus, supposons que le noyau $K(\cdot)$ soit une fonction continue à support compact sur l'intervalle $(-1,1)$ et que $\int_{-1}^{1} u\,K(u)\,du = 0$. Introduisons également les conditions suivantes

$$nh^5/\log\log n \to 0 \quad \text{lorsque} \quad n \to \infty, \tag{7.51}$$

$$\sum_{n=3}^{\infty}(h/\log\log n)\mathbb{E}[Y^2\mathbf{1}_{\{|Y|>a_n\}}] < \infty, \tag{7.52}$$

où $(a_n)_{n\geq 1}$ est une suite de constantes positives tendant vers l'infini, telle que

$$\begin{gathered} a_n = o\left(\frac{(nh^{-1}\log\log n)^{1/2}}{(\log n)^2}\right). \\ \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left\{\sup_{m\in\Gamma_{n,\varepsilon}} |h(m)/h(n) - 1|\right\} = 0, \end{gathered} \tag{7.53}$$

où $\Gamma_{n,\varepsilon} = \{m : |m-n| \leq \varepsilon n\}$.

Théorème 7.5.3. *Sous les conditions énoncées ci-dessus et si*

$$\sum_{n=1}^{\infty}\frac{1}{n^2 h} < \infty,$$

alors nous avons

$$\begin{aligned} \limsup_{n\to\infty} \pm \sqrt{\frac{nh}{2\log\log n}}\,(\widehat{r}_{n,h}(x) - r(x)) \\ = \sqrt{\frac{\mathrm{Var}[Y|X=x]}{f_X(x)}\int_{\mathbb{R}} K^2(u)\,du} \quad \textit{p.s.} \end{aligned} \tag{7.54}$$

Démonstration. La démonstration de ce théorème se trouve dans l'article de Härdle [64]. □

7.6. Vitesse de convergence

Le problème délicat de la vitesse de convergence a été traité par Schuster et Yakowitz [121], Mack et Silvermann [92], Härdle et Luckhaus [66]. Ce problème est resté ouvert pendant de nombreuses années. Récemment Einmahl et Mason [51] ont apporté de nouveaux résultats qui précisent les résultats antérieurs de Konakov et Piterbarg [86], et de Härdle, Janssen et Serfling [65].

Pour établir une vitesse exacte de convergence uniforme pour l'estimateur r_n, nous allons d'abord énoncer un théorème qui concerne uniquement le numérateur de l'estimateur r_n, c'est-à-dire pour $m_n(x)$.

Théorème 7.6.1. *Soit $I = [a, b]$ un intervalle de $\mathbb{R}$. Soit J un intervalle de $\mathbb{R}$ tel que $J = [a - \eta, b + \eta]$, pour des $\eta > 0$. Supposons que*

(G.1) la densité marginale f_X de la variable aléatoire X soit continue et strictement positive sur J ;

(G.2) la densité jointe $f_{X,Y}$ du couple (X, Y) soit continue sur $J \times \mathbb{R}$.

De plus, supposons que la suite $\{h_n : n \geq 1\}$ vérifie les trois conditions suivantes

(S.1) $h_n \downarrow 0$ et $nh_n \uparrow \infty$ lorsque $n \to \infty$;

(S.2) $\log(1/h_n)/\log\log n \to \infty$ lorsque $n \to \infty$;

(S.3) $nh_n/\log n \to \infty$ lorsque $n \to \infty$.

Alors pour tout noyau $K(\cdot)$ satisfaisant les conditions (K.1–2–3), nous avons

$$\begin{aligned} \lim_{n\to\infty} \sqrt{\frac{nh_n}{2\log(1/h_n)}} \sup_{x\in I} |m_n(x) - \mathbb{E}[m_n(x)]| \\ = \sqrt{\sup_{x\in I} \mathbb{E}[Y^2|X = x] f_X(x) \int_{\mathbb{R}} K^2(u)\, du} \quad p.s. \end{aligned} \tag{7.55}$$

Remarque 7.6.1. Ce théorème est une version particulière du théorème 1, p.5 de l'article de Einmahl et Mason [51].

On peut maintenant énoncer le corollaire qui établit la vitesse exacte de convergence uniforme pour l'estimateur non paramétrique r_n.

Corollaire 7.6.1. *Supposons que les densités f_X et $f_{X,Y}$ satisfont les conditions $(G.1)$ et $(G.2)$ respectivement. De plus, supposons que la suite $\{h_n : n \geq 1\}$ vérifie les conditions (S.1–2–3). Enfin, supposons que la variable aléatoire Y soit bornée. Alors pour tout noyau $K(\cdot)$ satisfaisant les conditions (K.1–2–3), nous avons*

$$\begin{aligned} \lim_{n\to\infty} \sqrt{\frac{nh_n}{2\log(1/h_n)}} \sup_{x\in I} \left| r_n(x) - \frac{\mathbb{E}[m_n(x)]}{\mathbb{E}[f_n(x)]} \right| \\ = \sqrt{\sup_{x\in I} \frac{\mathrm{Var}[Y|X = x]}{f_X(x)} \int_{\mathbb{R}} K^2(u)\, du} \quad p.s. \end{aligned} \tag{7.56}$$

De plus, si la variable aléatoire Y n'est pas nécessairement bornée, mais vérifie pour $p > 2$, la condition suivante

$$\sup_{x \in J} \mathbb{E}\left[|Y|^p | X = x\right] < \infty$$

et si la suite $\{h_n : n \geq 1\}$ vérifie les conditions (S.1–2) et la condition suivante

$$h_n^{-1} \leq \left(n / \log(1/h_n)\right)^{1-2/p}$$

alors (7.56) *reste valable.*

Remarque 7.6.2. Suite à une erreur typographique (confusion entre les membres de droite de (1.15) et (1.16)), on note que l'énoncé du corollaire 3, p.7 de l'article de Einmahl et Mason [51] est inexact. Il faut remplacer le membre de droite de (1.16), c'est-à-dire

$$||K||_2 / \inf_{z \in I} \sqrt{g_X(z)}$$

par

$$||K||_2 \sup_{z \in I} \frac{\sigma_I(z)}{\sqrt{g_X(z)}},$$

où $\sigma_I^2(z) = \mathrm{Var}[Y|X = z]$.

Démonstration. La démonstration de ce corollaire figure p.30 dans l'article de Einmahl et Mason [51]. □

7.7. Nouveaux résultats : une loi du logarithme itéré en un point pour l'estimateur à noyau

Afin d'établir les résultats principaux, nous allons d'abord rappeler les hypothèses sous lesquelles nous allons travailler et introduire quelques notations.

7.7.1. Hypothèses et notations

Afin d'assurer l'existence de la courbe de régression r, nous supposons que $\mathbb{E}[|Y|] < \infty$. De plus, nous supposons également que la loi de la variable aléatoire X admette une densité f_X. En tout point x de $\mathbb{R}$ où la densité marginale f_X est non nulle, nous rappelons que la courbe de régression, est définie par

$$r(x) = \mathbb{E}[Y|X = x] = \frac{m(x)}{f_X(x)}, \tag{7.57}$$

où la fonction $m(x)$ est définie par

$$m(x) = \int_{\mathbb{R}} y f_{XY}(x, y) dy,$$

et où f_{XY} est la densité jointe.

Nous allons considérer dans ce paragraphe, deux types d'estimateurs à noyau pour la courbe de régression $r(x)$ définie en (7.57). Le premier estimateur est dû à Nadaraya-Watson (voir Nadaraya [102], Watson [135]) et se définit, pour tout entier $n \geq 1$ et pour tout réel $x \in \mathbb{R}$ par

$$r_n(x) = \begin{cases} \dfrac{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} Y_i K\left(\dfrac{x-X_i}{h_n}\right)}{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right)}, & \text{si } \displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) \neq 0 \\ 0, & \text{si } \displaystyle\sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h_n}\right) = 0. \end{cases} \tag{7.58}$$

Dans ce paragraphe, nous envisageons également des applications où la densité marginale f_X de la variable aléatoire X est connue par le statisticien. Il est alors approprié de remplacer l'estimateur f_n de la densité dans le dénominateur de l'estimateur à noyau r_n de la régression par la densité marginale f_X. Ce changement conduit alors à l'estimateur suivant

$$\widetilde{m}_n(x) = \begin{cases} \dfrac{\dfrac{1}{nh_n}\displaystyle\sum_{i=1}^{n} Y_i K\left(\dfrac{x-X_i}{h_n}\right)}{f_X(x)}, & \text{si } f_X(x) \neq 0 \\ 0, & \text{si } f_X(x) = 0, \end{cases} \tag{7.59}$$

étudié principalement par Johnston dans les articles [77] et [78].

Supposons que le noyau K soit une fonction qui vérifie les 3 conditions suivantes.

(K.1) $\int_{\mathbb{R}} K(u)\, du = 1$.

(K.2) $K(\cdot)$ est à variation bornée sur $\mathbb{R}$.

(K.3) Il existe un réel $M < \infty$ tel que $K(u) = 0$ pour $|u| \geq M/2$.

Enfin, pour des commodités d'écriture, posons pour tout réel $x \in \mathbb{R}$,

$$S(x) = \mathbb{E}[Y^2|X = x] \quad \text{et} \quad V(x) = S(x) - r^2(x),$$

où $r(\cdot)$ est la courbe de régression définie en (7.57).

7.7.2. Principaux résultats

Nous allons commencer par établir une loi du logarithme itéré en un point pour l'estimateur à noyau de la régression dans le cas où la densité marginale f_X de la variable X est connue par le statisticien.

Théorème 7.7.1. *Supposons que*

(G.1) la densité marginale f_X soit continue et strictement positive sur $\mathbb{R}$,

(G.2) les fonctions r et S soient continues sur $\mathbb{R}$,

(G.3) il existe une constante $\mathcal{M} < \infty$ telle que $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s.

Supposons de plus que la suite $\{h_n : n \geq 1\}$ vérifie

$$(S.1) \quad (i) \quad h_n \searrow 0; \quad (ii) \quad nh_n \nearrow \infty \quad \textit{lorsque} \quad n \to \infty.$$
$$(S.2) \quad \frac{nh_n}{\log_2 n} \to \infty \quad \textit{lorsque} \quad n \to \infty.$$

Soit un noyau $K(\cdot)$ qui vérifie les propriétés $(K.1)$, $(K.2)$ et $(K.3)$. Alors, nous avons

$$\limsup_{n\to\infty} \pm\sqrt{\frac{nh_n}{2\log_2 n}}\,(\widetilde{m}_n(x) - r(x)) = \sqrt{\frac{\mathbb{E}[Y^2|X=x]}{f_X(x)}\int_{\mathbb{R}} K^2(u)\,du} \ \textit{p.s.} \tag{7.60}$$

Le deuxième théorème établit une loi du logarithme itéré en un point pour l'estimateur à noyau de Nadaraya-Watson de la courbe de régression.

Théorème 7.7.2. *Supposons que*

(G.1) la densité marginale f_X soit continue et strictement positive sur $\mathbb{R}$,

(G.2) la densité jointe f_{XY} soit, pour y fixé, continue et strictement positive sur $\mathbb{R}$,

(G.3) les fonctions r et S soient continues sur $\mathbb{R}$,

(G.4) il existe une constante $\mathcal{M} < \infty$ telle que $|Y - \mathbb{E}[Y]| \leq \mathcal{M}$ p.s.

Supposons de plus que la suite $\{h_n : n \geq 1\}$ vérifie

$$(S.1) \quad (i) \quad h_n \searrow 0; \quad (ii) \quad nh_n \nearrow \infty \quad \textit{lorsque} \quad n \to \infty.$$
$$(S.2) \quad \frac{nh_n}{\log_2 n} \to \infty \quad \textit{lorsque} \quad n \to \infty.$$

Soit un noyau $K(\cdot)$ qui vérifie les propriétés $(K.1)$, $(K.2)$ et $(K.3)$. Alors, nous avons

$$\limsup_{n\to\infty} \pm\sqrt{\frac{nh_n}{2\log_2 n}}\,(r_n(x) - r(x)) = \sqrt{\frac{\mathrm{Var}[Y|X=x]}{f_X(x)}\int_{\mathbb{R}} K^2(u)\,du} \ \textit{p.s.} \tag{7.61}$$

Remarque 7.7.1. Notons que la seule différence entre le théorème 7.7.1 et le théorème 7.7.2 est le facteur multiplicatif différent. Johnston [77] a montré que l'estimateur $\widetilde{m}_n(x)$ a une variance asymptotique proportionnelle à $S(x)$ tandis que l'estimateur à noyau $r_n(x)$ a une variance asymptotique proportionnelle à $V(x)$. Comme nous avons $S(x) \geq V(x)$, nous espérons alors obtenir des intervalles de confiance asymptotiquement plus proche pour l'estimateur $r_n(x)$ de Nadaraya-Watson que pour l'estimateur $\widetilde{m}_n(x)$.

7.7.3. Démonstration des résultats

Pour établir les théorèmes 7.7.1 et 7.7.2 ci-dessus, nous allons avoir besoin d'un lemme technique.

Lemme 7.7.1. *Il existe un espace de probabilités* $(\Omega, \mathbb{P}, \mathcal{A})$ *sur lequel il existe une suite de variables aléatoires* $\{Y_i^* : i \geq 1\}$ *indépendantes et identiquement distribuées de même loi que la suite de variables aléatoires* $\{Y_i : i \geq 1\}$ *telle que*

-i- $\{Y_i^* : i \geq 1\}$ *et* $\{X_i : i \geq 1\}$ *soient indépendantes et*

-ii- il existe une suite $\{\varepsilon_i : i \geq 1\}$ *qui tend vers 0 lorsque* $i \to \infty$ *telle que* $|Y_i - Y_i^*| \leq \varepsilon_i$.

Démonstration. Soit $\{X_i : i \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de densité f_X. Soit $\{Y_i : i \geq 1\}$ une suite de variables aléatoires bornées, indépendantes et identiquement distribuées de densité f_Y. Soit le couple (X, Y) de variables aléatoires.

Supposons que la densité jointe $f_{X,Y}(x, y)$ du couple (X, Y) notée pour plus de commodité $f(x, y)$ soit continue en les 2 variables sur $\mathbb{R} \times \mathbb{R}$. Soit $\{h_n : n \geq 1\}$ une suite qui vérifie les conditions (S.1–2). On a alors pour tout réel x vérifiant $|x - x_0| \leq Ch_n$, où C est une constante positive convenable et pour tout réel y tel que $|y| \leq \mathcal{M}$, où $\mathcal{M}_1$ est également une constante positive

$$|f(x, y) - f(x_0, y)| \to 0 \quad \text{uniformément en } |y| \leq \mathcal{M}.$$

On peut alors construire sur le même espace de probabilités $(\Omega, \mathbb{P}, \mathcal{A})$ une variable aléatoire réelle Y^* de densité $f(x_0, y)$, où $|y| \leq \mathcal{M}_1$ et indépendante de la variable aléatoire X. De cette variable aléatoire Y^*, on peut ainsi définir, sur le même espace de probabilités $(\Omega, \mathbb{P}, \mathcal{A})$, une suite de variables aléatoires $\{Y_i^* : i \geq 1\}$ telles que ces variables aléatoires aient la même densité que la variable Y^*.

D'autre part, on a par définition, pour tout x tel que $|x - x_0| \leq Ch_n$ et pour tout réel y tel que $|y| \leq \mathcal{M}_1$,

$$\mathbb{P}[Y \leq y | X = x] = G(y, x).$$

Définissons maintenant la fonction de quantile associée à la fonction G

$$\inf\{y : G(y, x) \geq t\} = G^{-1}(t, x), \quad \text{pour } t \in (0, 1).$$

Définissons la fonction G^* associée à la variable Y^*

$$G^*(y) = \mathbb{P}[Y^* \leq y],$$

puisque la variable aléatoire Y^* est indépendante de la variable X. Ainsi, on peut relier la variable aléatoire Y et la variable aléatoire Y^* par l'équation suivante

$$Y = G^{-1}(G^*(Y^*); X), \tag{7.62}$$

ce qui équivaut à

$$G(Y, X) = G^*(Y^*) \quad \text{pour } |X - x_0| \leq Ch_n.$$

On montre également par cette construction

$$\mathbb{P}[Y \neq Y^*] \leq \sup_{\substack{|x-x_0| \leq Ch_n \\ |y| \leq \mathcal{M}_1}} |f(x,y) - f(x_0,y)|.$$

□

Remarque 7.7.2. Ce lemme construit une suite de variables aléatoires $\{Y_i^* : i \geq 1\}$ indépendante de la suite de variables aléatoires $\{X_i : i \geq 1\}$. Cette suite de variables $\{Y_i^* : i \geq 1\}$ est introduite pour pouvoir utiliser les résultats des chapitres précédents et en particulier ceux du chapitre 6. En effet, il faut noter que dans la définition du processus empirique composé $\{\alpha_{n,c}(x) : 0 \leq x \leq 1\}$, les suites $\{Y_i : i \geq 1\}$ et $\{U_i : i \geq 1\}$ sont supposées indépendantes. C'est pour cette raison que nous avons besoin de construire une suite de variables aléatoire $\{Y_i^* : i \geq 1\}$ qui est indépendante de la suite $\{X_i : i \geq 1\}$ contrairement à la suite $\{Y_i : i \geq 1\}$ qui elle, est dépendante de la suite $\{X_i : i \geq 1\}$.

Nous allons d'abord démontrer le théorème 7.7.1.

Première étape. D'abord nous allons montrer que nous pouvons centrer l'estimateur $\widetilde{m}_n(x)$ autour de son espérance $\mathbb{E}[\widetilde{m}_n(x)]$. Cela vient du fait que nous avons, lorsque $n \to \infty$

$$\mathbb{E}[\widetilde{m}_n(x)] = \frac{1}{h_n f_X(x)} \int_{\mathbb{R}} K\left(\frac{x-u}{h_n}\right) r(u) f_X(u)\, du = r(x) + O\left(h_n^2\right). \tag{7.63}$$

Ce résultat s'établit en utilisant la régularité de la courbe de régression $r(\cdot)$ ainsi que celle de la densité marginale $f_X(\cdot)$ et des hypothéses mentionnées ci-dessus sur le noyau $K(\cdot)$. On pourra consulter, pour de plus amples détails sur la démonstration de (7.63) les articles de Parzen [109] et de Rosenblatt [116]. L'hypothèse $(S.1)(i)$ implique alors que le biais $\mathbb{E}[\widetilde{m}_n(x)] - r(x)$ tend vers 0 lorsque $n \to \infty$.

Deuxième étape. Cette première étape implique qu'il reste à montrer que

$$\limsup_{n\to\infty} \pm \frac{1}{\sqrt{2nh_n \log_2 n}} \left(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)]\right) = \sqrt{S(x) f_X(x) \int_{\mathbb{R}} K^2(u)\, du} \quad \text{p.s.,} \tag{7.64}$$

où la fonction $\widehat{m}_n(\cdot)$ est définie, pour tout $x \in \mathbb{R}$, par

$$\widehat{m}_n(x) = \sum_{i=1}^n Y_i K\left(\frac{x-X_i}{h_n}\right). \tag{7.65}$$

Pour montrer (7.64), nous allons étudier la quantité $(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)])$ en faisant apparaître la suite de variables aléatoires $\{Y_i^* : i \geq 1\}$ définies dans le lemme 7.7.1. Remarquons que, pour tout entier $n \geq 1$ et pour tout $x \in \mathbb{R}$, nous avons

$$\begin{aligned}
\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)] &= \sum_{i=1}^n (Y_i - Y_i^*)\, K\left(\frac{x-X_i}{h_n}\right) - \mathbb{E}\left[\sum_{i=1}^n (Y_i - Y_i^*)\, K\left(\frac{x-X_i}{h_n}\right)\right] \\
&\quad + \sum_{i=1}^n Y_i^* K\left(\frac{x-X_i}{h_n}\right) - \mathbb{E}\left[\sum_{i=1}^n Y_i^* K\left(\frac{x-X_i}{h_n}\right)\right] \\
&= \mathcal{I}_1 + \mathcal{I}_2.
\end{aligned}$$

Troisième étape. Étudions la quantité $\mathcal{J}_2$ qui se définit, pour tout $x \in \mathbb{R}$, par

$$\mathcal{J}_2 = \widehat{m}_n^*(x) - \mathbb{E}[\widehat{m}_n^*(x)],$$

où la fonction $\widehat{m}_n^*(x)$ se définit, pour tout $x \in \mathbb{R}$, par

$$\widehat{m}_n^*(x) = \sum_{i=1}^n Y_i^* K\left(\frac{x - X_i}{h_n}\right).$$

D'abord, remarquons que

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^n Y_i^* K\left(\frac{x - X_i}{h_n}\right)\right] &= n\mathbb{E}\left[Y_1^* K\left(\frac{x - X_1}{h_n}\right)\right] \\
&\quad \text{puisque les 2 suites } \{Y_i^* : i \geq 1\} \text{ et } \{X_i : i \geq 1\} \\
&\quad \text{sont identiquement distribuées,} \\
&= n\mathbb{E}\left[Y_1^*\right]\mathbb{E}\left[K\left(\frac{x - X_1}{h_n}\right)\right] \\
&\quad \text{puisque les suites } \{Y_i^* : i \geq 1\} \text{ et } \{X_i : i \geq 1\} \\
&\quad \text{sont supposées indépendantes,} \\
&= n\mathbb{E}\left[K\left(\frac{x - X_1}{h_n}\right)\right] \\
&\quad \text{puisque } \mathbb{E}[Y_1^*] = 1, \\
&= n\int_{\mathbb{R}} K\left(\frac{x - u}{h_n}\right) dF_X(u).
\end{aligned}$$

D'autre part, pour tout $x \in \mathbb{R}$, nous avons

$$\widehat{m}_n^*(x) = \sum_{i=1}^n Y_i^* K\left(\frac{x - X_i}{h_n}\right) = n\int_{\mathbb{R}} K\left(\frac{x - u}{h_n}\right) dF_{n,c}(u),$$

où la fonction $F_{n,c}$ est définie, pour tout entier $n \geq 1$ et pour tout $x \in \mathbb{R}$, par

$$F_{n,c}(x) = \frac{1}{n}\sum_{i=1}^n Y_i^* \mathbf{1}_{\{X_i \leq x\}}.$$

Ainsi, nous en déduisons que

$$\widehat{m}_n^*(x) - \mathbb{E}\left[\widehat{m}_n^*(x)\right] = n\int_{\mathbb{R}} K\left(\frac{x - u}{h_n}\right)\left[dF_{n,c}(u) - dF_X(u)\right].$$

En posant $K(-u) = \widetilde{K}(u)$, on obtient que

$$\begin{aligned}
\widehat{m}_n^*(x) - \mathbb{E}\left[\widehat{m}_n^*(x)\right] &= n\int_{-M}^{M} \widetilde{K}(u)\Big[dF_{n,c}(x+h_nu) - dF_X(x+h_nu) \\
&\qquad -dF_{n,c}(x-h_nM) + dF_X(x-h_nM)\Big] \\
&= n\Big[\widetilde{K}(u)\big\{F_{n,c}(x+h_nu) - F_X(x+h_nu) \\
&\qquad -\left(F_{n,c}(x-h_nM) - F_X(x-h_nM)\right)\big\}\Big]_{-M}^{M} \\
&\quad -n\int_{-M}^{M} \big\{F_{n,c}(x+h_nu) - F_X(x+h_nu) \\
&\qquad -\left(F_{n,c}(x-h_nM) - F_X(x-h_nM)\right)\big\}\, d\left(\widetilde{K}(u)\right) \\
&\quad \text{en faisant une intégration par parties,} \\
&= n\int_{-M}^{M} \big\{F_{n,c}(x+h_nu) - F_X(x+h_nu) \\
&\qquad -\left(F_{n,c}(x-h_nM) - F_X(x-h_nM)\right)\big\}\, d\left(-\widetilde{K}(u)\right).
\end{aligned}$$

On a alors

$$\begin{aligned}
&\pm\frac{1}{\sqrt{2nh_n\log_2 n}}\left(\widehat{m}_n^*(x) - \mathbb{E}\left[\widehat{m}_n^*(x)\right]\right) \\
&= \pm\int_{-M}^{M}\left\{\frac{\sqrt{n}\left[F_{n,c}(x+h_nu) - F_{n,c}(x-h_nu) - F_X(x+h_nu) + F_X(x-h_nu)\right]}{\sqrt{2h_n\log_2 n}}\right\} d\left(-\widetilde{K}(u)\right).
\end{aligned} \tag{7.66}$$

À cette étape de la démonstration, nous allons rappeler un théorème établi au chapitre 6.

Théorème 7.7.3. *Soit* $\{Y_n^* : n \geq 1\}$ *une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que la varialbe aléatoire* $Y^* = Y_1^*$, *de fonction de répartition* F_Y *dérivable au point* x *et de dérivée* $f_Y(x)$. *De plus, nous supposons que* $\mathbb{E}[Y^*] = 1$, $\mathrm{Var}[Y^*] = \sigma^2 < \infty$ *et que* $|Y^* - 1| \leq \mathcal{M}_2$ *p.s., pour une constante* $\mathcal{M}_2$ *convenable. Soit* $\{X_n : n \geq 1\}$ *une suite de variables aléatoires de fonction de répartition* F_X *dérivable au point* x *et de dérivée* $f_X(x) > 0$. *Les suites de variables* $\{Y_n^* : n \geq 1\}$ *et* $\{X_n : n \geq 1\}$ *sont supposées indépendantes. Supposons que la suite* $\{k(n) : n \geq 1\}$ *vérifie les conditions suivantes*

-i- *(i)* $0 < k(n) \leq n$, *(ii)* $k(n) \uparrow \infty$, *et (iii)* $n^{-1}k(n) \downarrow 0$ *lorsque* $n \to \infty$ *;*

-ii- $k(n)/\log_2 n \to \infty$ *lorsque* $n \to \infty$.

alors, pour tout réel t fixé, la suite de fonctions de $s \in [0,1]$ définie par

$$\Bigg\{ \Psi_{n,c}(s) = \frac{1}{\sqrt{2\log_2 n}} \Theta_{n,c} \Bigg(\left(\frac{k(n)}{n} f_X(x) \right)^{-1} \left(F_X \left(x + \frac{k(n)}{n} c \right) - F_X(x), \right. \\ \left. F_X \left(x + \frac{k(n)}{n}(c+s) \right) - F_X(x) \right), F_X(x), \frac{k(n)}{n} f_X(x) \Bigg) : n \geq 1 \Bigg\} \tag{7.67}$$

est presque sûrement relativement compacte dans l'ensemble des fonctions bornées sur l'intervalle $[0,1]$, muni de la topologie de la convergence uniforme. L'ensemble des points limites correspondant contient les fonctions de la forme

$$\psi(s) = \int_0^s \phi(u)\, du \ \textit{avec} \ \int_0^1 \phi(u)\, du \leq 1.$$

Rappelons que le processus empirique composé local au point t indexé par une certaine classe $\mathbb{C}$ se définit par

$$\begin{aligned} \Theta_{n,c}(C) &= \Theta_{n,c}\left(C, t, \frac{k(n)}{n} \right) \\ &= \left(\frac{k(n)}{n} \right)^{-1/2} \alpha_{n,c}\left(t + \left(\frac{k(n)}{n} \right) C \right), \end{aligned}$$

où $C \in \mathbb{C}$. Pour de plus amples détails sur le processus empirique composé local au point t indexé par une certaine classe $\mathbb{C}$, voir le chapitre 6.

Démonstration. Pour la démonstration de ce théorème, nous renvoyons au chapitre 6, théorème 6.5.2. □

En appliquant ce théorème, nous montrons que l'ensemble des points limites du membre de droite de (7.66) est presque sûrement égal à

$$\left\{ \pm\sqrt{\mathbb{E}[(Y^*)^2] f_X(x)} \int_{-M}^{M} \left\{ \int_{-M}^{u} \phi(s)\, ds \right\} d\left(-\widetilde{K}(u) \right) : \int_{-M}^{M} \phi^2(u)\, du \leq 1 \right\}.$$

En faisant maintenant une intégration par parties, nous montrons que l'ensemble des points limites du membre de droite de (7.66) est presque sûrement égal à

$$\left\{ \pm\sqrt{\mathbb{E}[(Y^*)^2] f_X(x)} \int_{-M}^{M} \phi(u) \widetilde{K}(u)\, du : \int_{-M}^{M} \phi^2(u)\, du \leq 1 \right\}.$$

Enfin en appliquant l'inégalité de Cauchy Schwarz, nous montrons que l'ensemble des points limites du membre de droite de (7.66) est presque sûrement égal à

$$\left\{ \pm\sqrt{\mathbb{E}[(Y^*)^2] f_X(x) \int_{\mathbb{R}} \widetilde{K}^2(u)\, du} : \int_{-M}^{M} \phi^2(u)\, du \leq 1 \right\}.$$

Quatrième étape. Maintenant étudions la quantité $\mathcal{J}$. D'après le lemme 7.7.1, remarquons que, pour tout entier $i \geq 1$,

$$Y_i - Y_i^* = G^{-1}\left(G^*(Y_i^*); X_i\right) - Y_i^*.$$

On pose alors

$$\Phi(Y_i^*) = G^{-1}\left(G^*(Y_i^*); X_i\right) - Y_i^*$$

Ainsi nous sommes ramenés à étudier

$$\limsup_{n\to\infty} \pm \frac{1}{\sqrt{2nh_n \log_2 n}} \left(\sum_{i=1}^n \Phi(Y_i^*) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[\sum_{i=1}^n \Phi(Y_i^*) K\left(\frac{x - X_i}{h_n}\right)\right]\right).$$

En procédant de la même manière que dans la troisième étape, nous obtenons que l'ensemble des points limites de

$$\pm \frac{1}{\sqrt{2nh_n \log_2 n}} \left(\sum_{i=1}^n \Phi(Y_i^*) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[\sum_{i=1}^n \Phi(Y_i^*) K\left(\frac{x - X_i}{h_n}\right)\right]\right)$$

est presque sûrement égal à

$$\left\{ \pm\sqrt{\mathbb{E}[(\Phi(Y^*))^2] f_X(x) \int_{\mathbb{R}} \widetilde{K}^2(u)\, du} : \int_{-M}^{M} \phi^2(u)\, du \leq 1 \right\}.$$

Cinquième étape. Maintenant, il faut noter que lorsque nous faisons tendre $n \to \infty$, la variable aléatoire Y_n^* se rapproche de la variable aléatoire Y_n. Par conséquent $\mathbb{E}[(\Phi(Y^*))^2]$ tend vers 0 et $\mathbb{E}[(Y^*)^2]$ tend vers $\mathbb{E}[Y^2|X = x]$. Ce qui achève la démonstration du théorème 7.7.1.

Montrons maintenant le théorème 7.7.2.

Première étape. Pour tout $x \in \mathbb{R}$, décomposons la quantité $r_n(x) - r(x)$. Nous avons alors

$$\begin{aligned} r_n(x) - r(x) = {} & f_X^{-1}(x)\left[(nh_n)^{-1}\widehat{m}_n(x) - r(x) f_n(x)\right] \\ & + f_X^{-1}(x)\left[r_n(x) - r(x)\right] \cdot \left[f_X(x) - f_n(x)\right], \end{aligned} \tag{7.68}$$

où la fonction $m_n(x)$ est définie en (7.65) et $f_n(x)$ est l'estimateur à noyau de la densité $f_X(x)$ introduit par Parzen [109] et Rosenblatt[115] et défini, pour tout entier $n \geq 1$ et pour tout $x \in \mathbb{R}$ par,

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right). \tag{7.69}$$

Deuxième étape. À cette étape de la démonstration, nous avons besoin de rappeler un théorème de Deheuvels et Mason [36] déjà cité dans le chapitre 6.

Théorème 7.7.4. *Supposons que la fonction de répartition F_X soit dérivable au point $x \in \mathbb{R}$ et admette une dérivée $f_X(x) > 0$. Supposons de plus que la suite $\{h_n : n \geq 1\}$ vérifie les conditions suivantes*

$$\begin{aligned}
&(S.1) \quad (i) \quad h_n \searrow 0; \quad (ii) \quad nh_n \nearrow \infty \quad \textit{lorsque} \quad n \to \infty. \\
&(S.2) \quad \frac{nh_n}{\log_2 n} \to \infty \quad \textit{lorsque} \quad n \to \infty.
\end{aligned}$$

Soit un noyau $K(\cdot)$ qui vérifie les propriétés (K.1–2–3). Alors nous avons

$$\limsup_{n\to\infty} \pm \sqrt{\frac{nh_n}{2\log_2 n}} \Big(f_n(x) - \mathbb{E}[f_n(x)] \Big) = \sqrt{f_X(x) \int_{\mathbb{R}} K^2(u)\, du} \quad p.s. \tag{7.70}$$

Démonstration. La démonstration de ce théorème se trouve dans le chapitre 6, p.164, théorème 6.5.3. Elle s'appuie essentiellement sur les arguments de l'article de Deheuvels et Mason [36]. □

D'après l'article de Noda [107], nous savons que

$$r_n(x) - r(x) = o(1) \quad \text{p.s.},$$

sous les hypothèses du thèorème 7.7.2. Ce dernier résultat et (7.70) impliquent alors que le second terme du membre de droite de la décomposition de la quantité $r_n(x) - r(x)$ ci-dessus est de l'ordre de $o\left(\sqrt{\frac{nh_n}{2\log_2 n}}\right)$ p.s.

Troisième étape. Le premier membre de droite de la décomposition de la quantité $r_n(x) - r(x)$ ci-dessus peut s'écrire sous la forme suivante

$$\begin{aligned}
&\frac{(nh_n)^{-1}(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)])}{f_X(x)} + \frac{(nh_n)^{-1}\mathbb{E}[\widehat{m}_n(x)] - r(x) f_X(x)}{f_X(x)} - \frac{r(x)\,(f_n(x) - \mathbb{E}[f_n(x)])}{f_X(x)} \\
&\qquad + \frac{r(x)\,(f_X(x) - \mathbb{E}[f_n(x)])}{f_X(x)},
\end{aligned}$$

où la fonction $m_n(x)$ est définie en (7.65) et l'estimateur de densité $f_n(x)$ en (7.69). En utilisant les mêmes arguments de la démonstration du théorème 7.7.1, il s'ensuit que les biais $\big((nh_n)^{-1}\mathbb{E}[\widehat{m}_n(x)] - r(x) f_X(x)\big)$ et $\big(\mathbb{E}[f_n(x)] - f_X(x)\big)$ tendent vers 0 lorsque $n \to \infty$. Pour les démonstrations de ce type de résultats, nous renvoyons au paragraphe 7.2 de ce chapitre, ou à l'article de Parzen [109].

Quatrième étape. Il reste donc à montrer que la quantité

$$(nh_n)^{-1}\,(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)]) - r(x)\,(f_n(x) - \mathbb{E}[f_n(x)])$$

suit la loi du logarithme itéré, i.e.

$$\begin{aligned}\limsup_{n\to\infty} \pm \sqrt{\frac{nh_n}{2\log_2 n}} &\left\{(nh_n)^{-1}\left(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)]\right) - r(x)\left(f_n(x) - \mathbb{E}[f_n(x)]\right)\right\} \\ &= \sqrt{V(x) f_X(x) \int_{\mathbb{R}} K^2(u)\,du} \quad \text{p.s.}\end{aligned}$$

Nous allons utiliser la même méthode que celle utilisée dans la démonstration du théorème 7.7.1, à savoir faire apparaître la suite $\{Y_i^* : i \geq 1\}$ de variables aléatoires indépendantes et identiquement distribuées, qui plus est, cette suite $\{Y_i^* : i \geq 1\}$ est indépendante de la suite de variables $\{X_i : i \geq 1\}$. Ensuite on applique des résultats que nous avons déjà établis préalablement dans les autres chapitres de cette thèse.
Remarquons que la quantité $\{(nh_n)^{-1}\left(\widehat{m}_n(x) - \mathbb{E}[\widehat{m}_n(x)]\right) - r(x)\left(f_n(x) - \mathbb{E}[f_n(x)]\right)\}$ peut s'écrire comme

$$\begin{aligned}(nh_n)^{-1} &\sum_{i=1}^{n}\left(Y_i K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[Y K\left(\frac{x - X}{h_n}\right)\right]\right) \\ &- (nh_n)^{-1} r(x) \sum_{i=1}^{n}\left(K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[K\left(\frac{x - X}{h_n}\right)\right]\right)\end{aligned}$$

Nous sommes donc ramenés à étudier la

$$\begin{aligned}\limsup_{n\to\infty} \pm \frac{1}{\sqrt{2nh_n\log_2 n}} &\left\{\sum_{i=1}^{n}\left(Y_i K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[Y K\left(\frac{x - X}{h_n}\right)\right]\right)\right. \\ &\left. - r(x)\sum_{i=1}^{n}\left(K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[K\left(\frac{x - X}{h_n}\right)\right]\right)\right\}.\end{aligned}$$

Après réarrangement des termes, en fait nous sommes ramenés à étudier la

$$\limsup_{n\to\infty} \pm \frac{1}{\sqrt{2nh_n\log_2 n}} \left\{\sum_{i=1}^{n}\left((Y_i - r(x)) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[(Y - r(x)) K\left(\frac{x - X}{h_n}\right)\right]\right)\right\}.$$

Étudions donc la quantité

$$d_n(x) = \sum_{i=1}^{n}\left((Y_i - r(x)) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[(Y - r(x)) K\left(\frac{x - X}{h_n}\right)\right]\right).$$

En introduisant la suite de variables $\{Y_i^* : i \geq 1\}$, nous avons

$$\begin{aligned}d_n(x) &= \sum_{i=1}^{n}\left((Y_i - Y_i^*) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[(Y - Y^*) K\left(\frac{x - X}{h_n}\right)\right]\right) \\ &\quad + \sum_{i=1}^{n}\left((Y_i^* - r(x)) K\left(\frac{x - X_i}{h_n}\right) - \mathbb{E}\left[(Y^* - r(x)) K\left(\frac{x - X}{h_n}\right)\right]\right) \\ &= \mathcal{D}_1 + \mathcal{D}_2.\end{aligned}$$

Étudions d'abord la quantité $\mathcal{D}_1$. En appliquant la même technique que celle utilisée dans la démonstration du théorème 7.7.1, nous montrons que l'ensemble des points limites de

$$\pm\frac{1}{\sqrt{2nh_n\log_2 n}}\sum_{i=1}^{n}\left((Y_i^*-r(x))K\left(\frac{x-X_i}{h_n}\right)-\mathbb{E}\left[(Y^*-r(x))K\left(\frac{x-X}{h_n}\right)\right]\right)$$

est presque sûrement égal à (après une intégration par parties et une application de l'inégalité de Cauchy Schwarz)

$$\left\{\pm\sqrt{f_X(x)\mathbb{E}[(Y^*-r(x))^2]\int_{\mathbb{R}}\tilde{K}^2(u)\,du]}:\int_{-M}^{M}\phi^2(u)\,du\leq 1\right\}.$$

Étudions maintenant la quantité $\mathcal{D}_2$. Avec la même technique, nous montrons que l'ensemble des points limites de

$$\pm\frac{1}{\sqrt{2nh_n\log_2 n}}\sum_{i=1}^{n}\left((Y_i-Y_i^*)K\left(\frac{x-X_i}{h_n}\right)-\mathbb{E}\left[(Y-Y^*)K\left(\frac{x-X}{h_n}\right)\right]\right)$$

est presque sûrement égal à (après une intégration par parties et une application de l'inégalité de Cauchy Schwarz)

$$\left\{\pm\sqrt{f_X(x)\mathbb{E}[(Y-Y^*)^2]\int_{\mathbb{R}}\tilde{K}^2(u)\,du]}:\int_{-M}^{M}\phi^2(u)\,du\leq 1\right\}.$$

Enfin, remarquons que lorsque $n\to\infty$, alors la variable Y_n^* se rapproche de la variable Y_n et par conséquent $\mathbb{E}[(Y-Y^*)^2]$ tend vers 0 et $\mathbb{E}[(Y^*-r(x))^2]$ tend vers $\mathbb{E}[(Y-r(x))^2|X=x]$, c'est-à-dire tend vers $\mathrm{Var}[Y|X=x]$. Ce qui achève la démonstration du théorème 7.7.2.

Bibliographie

[1] K. S. Alexander. The central limit theorem for empirical processes on Vapnik-Červonenkis classes. *Ann. Probab.*, 15 :178–203, 1987.

[2] G. Bennett. Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.*, 57 :33–45, 1962.

[3] P. Bickel and E. Lehman. Unbiased estimation in convex families. *Ann. Math. Statist.*, 40 :1523–1525, 1969.

[4] P. Bickel and M. Rosenblatt. On some global measures of the deviations of density function estimates. *Ann. Statist.*, 1 :1071–1091, 1973.

[5] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[6] J. Bleuez and D. Bosq. Etude d'une classe d'estimateurs non paramétriques de la densité. *Ann. Inst. H. Poincaré*, 14 :479–498, 1978.

[7] F. Bonvalot and N. Castelle. Strong approximation of uniform empirical process by Kiefer process. *Prépublications 91-41, Université de Paris-Sud Mathématiques, Bât 425, Orsay, France.*, 1991.

[8] F. Bonvalot and N. Castelle. Strong approximation of bivariate uniform empirical processes. *Ann. Inst. H. Poincaré*, 34 :425–480, 1999.

[9] D. Bosq. Contribution à la théorie de l'estimation fonctionnelle. *Public. Inst. Statist. Univ. Paris*, 19 fasc.2 et 3, 1970.

[10] D. Bosq. *Linear processes in function spaces.* Springer, New York, 2000.

[11] D. Bosq and J.-P. Lecoutre. *Théorie de l'estimation fonctionnelle.* Economica, Paris, 1987.

[12] J. Bretagnolle and P. Massart. Hungarian constructions from the nonasymptotic viewpoint. *Ann. Probab.*, 17 :239–256, 1989.

[13] Y. S. Chow and H. Teicher. *Probability Theory.* Springer-Verlag, New York, 1978.

[14] K.-L. Chung. An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.*, 67 :36–50, 1949.

[15] C. Cocozza-Thivent. *Processus stochastiques et fiabilité des systèmes.* Springer-Verlag, Berlin, 1997.

[16] G. Collomb. Estimation non-paramétrique de la régression par la métode du noyau. *Thèse, Université Paul Sabatier, Toulouse*, 1976.

[17] G. Collomb. Estimation non-paramétrique de la régression par la méthode du noyau : propriété de convergence asymptotiquement normale indépendante. *Annales Scientifiques de l'Université de Clermont*, pages 24–46, 1977.

[18] G. Collomb. Quelques propriétés de la méthode du noyau pour l'estimation non-paramétrique de la régression en un point fixé. *C. R. Acad. Sci. Paris A*, 285 :289–292, 1977.

[19] G. Collomb. Conditions nécessaires et suffisantes de convergence uniforme d'un estimateur de la régression, estimation des dérivées de la régression. *C. R. Acad. Sci. Paris A*, 288 :161–163, 1978.

[20] G. Collomb. Estimation non-paramétrique de la régression : régressogramme et méthode du noyau. *Publications du Laboratoire de Statistique et Probabilités de l'Université de Toulouse*, 07-78 :1–59, 1978.

[21] G. Collomb. Estimation non-paramétrique de la régression : Revue bibliographique. *International Statistical Review*, 49 :75–93, 1981.

[22] M. Csörgő, P. Deheuvels, and L. Horváth. An approximation of stopped sums with applications in queueing theory. *Adv. Appl. Prob.*, 19 :674–690, 1987.

[23] S. Csörgő. Limit behaviour of the empirical characteristic function. *Ann. Probab.*, 9 :130–144, 1981.

[24] M. Csörgő and L. Horváth. *Weighted approximations in probability and statistics.* Wiley, Chichester, 1993.

[25] M. Csörgő, L. Horváth, and P. Kokoszka. Approximation for bootstrapped empirical processes. *Proc. Amer. Math. Soc.*, 128 :2457–2464, 1999.

[26] M. Csörgő, L. Horváth, and J. Steinebach. Invariance principles for renewal processes. *Ann. Probab.*, 15 :1441–1461, 1987.

[27] M. Csörgő and P. Révész. *Strong Approximations in Probability and Statistics.* Academic Press, New York, 1981.

[28] A. De Acosta. Invariance principles in probability for triangular arrays of b-valued random vectors and some applications. *Ann. Probab.*, 10 :346–373, 1982.

[29] P. Deheuvels. Strong laws for local quantile processes. *Ann. Probab.*, 25 :2007–2054, 1997.

[30] P. Deheuvels and M. A. Lifshits. Strassen-type functional laws for strong topologies. *Probab. Th. Rel. Fields*, 97 :151–167, 1993.

[31] P. Deheuvels and M. A. Lifshits. Necessary and sufficient conditions for Strassen law of the iterated logarithm in nonuniform topologies. *Ann. Probab.*, 22 :1838–1856, 1994.

[32] P. Deheuvels and D. M. Mason. The asymptotic behavior of sums of exponential extreme values. *Bulletin des Sciences Math.*, 112 :211–233, 1988.

[33] P. Deheuvels and D. M. Mason. Nonstandard functional laws of the iterated logarithm for tail empirical and quantile processes. *Ann. Probab.*, 18 :1693–1722, 1990.

[34] P. Deheuvels and D. M. Mason. A tail empirical process approach to some non-standard laws of the iterated logarithm. *J. Theor. Probab.*, 4 :53–85, 1991.

[35] P. Deheuvels and D. M. Mason. Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.*, 20 :1248–1287, 1992.

[36] P. Deheuvels and D. M. Mason. Functional laws of the iterated logarithm for local empirical processeses indexed by sets. *Ann. Probab.*, 22 :1619–1661, 1994.

[37] P. Deheuvels and D. M. Mason. Nonstandard local empirical processeses indexed by sets. *J. Statist. Planning Inference*, 45 :91–112, 1995.

[38] P. Deheuvels and J. Steinebach. Exact Convergence Rates in Strong Approximation Laws for Large Increments of Partial Sums. *Probab. Th. Rel. Fields*, 76 :369–393, 1987.

[39] L. P. Devroye. The uniform convergence of the Nadaraya-Watson regression function estimate. *Canad. J. Statist.*, 6 :179–191, 1979.

[40] L. P. Devroye. On the almost everywhere convergence of nonparametric regression function estimates. *Ann. Statist.*, 9 :1310–1319, 1981.

[41] L. P. Devroye and T. Wagner. Distribution-free consistency results in nonparametric discrimination and regression function estimation. *Ann. Statist.*, 8 :231–239, 1979.

[42] L. P. Devroye and T. Wagner. On the L^1 convergence of kernel regression function estimators with applications in discrimination. *Z. Wahrsch. verw. Gebiete*, 51 :15–25, 1980.

[43] J. Diebolt. Testing functions defining a nonlinear autoregressive time series. *Stochastic Proc. Appl.*, 36 :85–106, 1990.

[44] J. Diebolt. A nonparametric test for the regression function. *Asymptotic theory. J. Statist. Plan. Inference*, 44 :1–17, 1995.

[45] J. Diebolt and N. Laïb. Un principe d'invariance faible pour l'étude d'un test nonparamétrique relatif à la fonction de régression. *C. R. Acad. Sci. Paris*, 312 :887–891, 1991.

[46] M. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6, 1951.

[47] R. M. Dudley. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis*, 1 :290–330, 1967.

[48] R. M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.*, 6 :899–929, 1978.

[49] U. Einmahl. *A refinement of the KMT-inequality for partial sum strong approximation.* Tech. Rep. Ser. Lab. Res. Stat. Probab. No 88, 1986.

[50] U. Einmahl. Extensions of results of Komlós, Major and Tusnády to the multivariate case. *J. Multivariate Anal.*, 28 :20–68, 1989.

[51] U. Einmahl and D. M. Mason. An empirical process approach to the uniform consistency of kernel-type function estimators. *J.Theor.Probab.*, 13 :1–37, 2000.

[52] R. C. Elandt-Johnson and N. L. Johnson. *Survival models and data analysis*. Wiley, New York, 1980.

[53] P. Erdős and A. Rényi. On a new law of large numbers. *J. Analyse Math.*, 23 :103–111, 1970.

[54] W. Feller. *An introduction to probability theory and its applications. Vol.1, second edition*. Wiley, New York, 1950.

[55] H. Finkelstein. The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.*, 42 :607–615, 1971.

[56] E. Fix and J. L. Hodges. *Discriminatory analysis. Nonparametric discrimination : consistency properties*. Technical Report 4, Project no.21-49-004, 1951.

[57] P. Gaenssler. *Empirical Processes*. IMS, Hayward, CA, 1982.

[58] E. Giné and J. Zinn. Some limit theorems for empirical processes. *Ann. Probab.*, 12 :929–989, 1984.

[59] E. Giné and J. Zinn. Lectures on the central limit theorem for empirical processes. In *Probability in Banach Spaces. Lecture Notes in Math*, volume 1221, pages 50–113. Springer, Berlin, 1986.

[60] W. Greblicki and A. Krzyzak. Asymptotic properties of kernel estimates of a regression function. *J. Statist. Planning Inference*, 4 :81–90, 1980.

[61] J. Hájek, Z. Sidák, and P. Sen. *Theory of Rank Tests*. Academic Press, San Diego, California, 1998.

[62] P. Hall. Laws of the iterated logarithm for nonparametric density estimators. *Z. Wahrsch. Verw. Gebiete*, 54 :47–61, 1981.

[63] P. Hall. *The Bootstrap and Edgeworth Expansion*. Springer, Berlin, 1992.

[64] W. Härdle. A law of the iterated logarithm for nonparametric regression function estimators. *Ann. Statist.*, 12 :624–635, 1984.

[65] W. Härdle, P. Janssen, and R. Serfling. Strong uniform consistency rates of estimators of conditional functionals. *Ann. Statist.*, 16 :1428–1449, 1988.

[66] W. Härdle and S. Luckhaus. Unifom consistency of a class of regression function estimators. *Ann.Statist.*, 12 :612–623, 1984.

[67] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58 :13–30, 1963.

[68] J. Hoffman-Jorgensen. Sums of independent Banach space valued random variables. *Studia Math.*, 52 :159–186, 1974.

[69] T. Höglund. A unified formulation of the central limit theorem for small and large deviations from the mean. *Z. Wahrsch. verw. Gebiete*, 49 :105–117, 1979.

[70] L. Horváth. Approximations for hybrids of empirical and partial sums processes. *J. Stat. Planning Inf.*, 88 :1–18, 2000.

[71] L. Horváth, P. Kokoszka, and J. Steinebach. Approximations for weighted bootstrap processes with an application. *Statist. Probab. Lett.*, 48 :59–70, 2000.

[72] L. Horváth and J. Steinebach. On the best approximation for bootstrapped empirical processes. *Statist. Probab. Lett.*, 41 :117–122, 1999.

[73] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin, 1987.

[74] N. C. Jain and M. B. Marcus. Integrability of infinite sums of independent vector-valued random variables. *Trans. Amer. Math. Soc.*, 212 :1–36, 1975.

[75] J.Kuelbs and R. M. Dudley. Log log laws for empirical measures. *Ann. Probab.*, 8 :405–418, 1980.

[76] N. J. Johnson and S. Kotz. *Discrete distributions.* Houghton-Mifflin, Boston, 1969.

[77] G. Johnston. Smooth nonparametric regression analysis. *Inst. of Stat. Mimeo Series No.1253, University of North Carolina*, 1979.

[78] G. Johnston. Probabilities of maximal deviation of nonparametric regression function estimation. *J. Multivariate Anal.*, 12 :402–414, 1982.

[79] J. Kahane. *Some random series of functions.* D. C. Heath, Lexington, Mass, 1968.

[80] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes.* Academic Press, San Diego, California, 1997.

[81] D. Kh. Fuk and S. V. Nagaev. Probability inequalities for sums of independent random variables. *Theory Probab. Appl.*, 16 :643–660, 1971.

[82] J. Kiefer. Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. In *Proc. Sixth Berkeley Symp. Math. Statist. Probab.*, volume 1, pages 227–244. Univ. California Press, Berkeley, 1972.

[83] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent r.v.'s and the sample d.f. I. *Z. Wahrsch. verw. Gebiete*, 32 :111–131, 1975.

[84] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent r.v.'s and the sample d.f. II. *Z. Wahrsch. verw. Gebiete*, 34 :33–58, 1976.

[85] V. Konakov. Asymptotic properties of some functions of nonparametric estimate of a density function. *J. Multivariate Anal.*, 5 :454–468, 1973.

[86] V. Konakov and V. Piterbarg. On the convergence rate of maximal deviation distribution for kernel regression estimates. *J. Multivariate Anal.*, 15 :279–294, 1984.

[87] T. L. Lai. Reproducing kernel Hilbert space and the law of the iterated logarithm for gaussian processes. *Z. Wahrschein. Verw. Gebiete*, 29 :7–19, 1974.

[88] M. Ledoux and M. Talagrand. Some applications of isoperimetric methods to strong limit theorems for sums of independant random variables. *Ann. Probab.*, 18 :754–789, 1990.

[89] P. Lévy. *Théorie de l'addition des variables aléatoires.* Gauthier-Villars, Paris, 1937.

[90] D. O. Loftsgaarden and C. P. Quesenberry. A nonparametric estimate of a multivariate density function. *Ann. Math. Statist.*, 36 :1049–1051, 1965.

[91] E. Lukacs. *Characteristics Functions.* Charles Griffin and Company Ltd, 1960.

[92] Y. Mack and B. Silverman. Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. verw. Gebiete*, 61 :405–415, 1982.

[93] P. Major. The approximation of partial sums of independent r.v.'s. *Z. Wahrsch. Verw. Gebiete*, 35 :213–220, 1976.

[94] D. M. Mason. A strong invariance theorem for the tail empirical process. *Ann. Inst. H. Poincaré Probab. Statist.*, 24 :491–506, 1988.

[95] D. M. Mason, G. R. Shorack, and J. A. Wellner. Strong limit theorems for the oscillation moduli of the uniform empirical process. *Z. Wahrsch. verw. Gebiete*, 65 :83–97, 1983.

[96] D. M. Mason and W. R. van Zwet. A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.*, 15 :871–884, 1987.

[97] P. Massart. Strong approximation for multivariate empirical and related processes, via KMT constructions. *Ann. Probab.*, 17 :266–291, 1989.

[98] M. Maumy. Le comportement des oscillations du processus empirique composé. *C. R. Acad. Sci. Paris, Ser.I*, 333 :1101–1104, 2001.

[99] M. Maumy. Sur les oscillations du processus de poisson composé. *C. R. Acad. Sci. Paris, Ser.I*, 334 :705–708, 2002.

[100] L. Menneteau. Functional laws of the iterated logarithm for Kiefer processes. *Math. Meth. in Stat*, 8 :1–21, 1999.

[101] A. A. Mogul'skii. On the law of the iterated logarithm in Chung's form for functional spaces. *Theory Probab. Appl.*, 24 :405–413, 1979.

[102] E. A. Nadaraya. On estimating regression. *Theor. Prob. Appl.*, 9 :141–142, 1964.

[103] E. A. Nadaraya. On nonparametric estimation of density function and regression. *Theor. Prob. Appl.*, 10 :186–190, 1965.

[104] E. A. Nadaraya. Remarks on nonparametric estimates for density functions and regression curves. *Theor. Prob. Appl.*, 15 :134–137, 1970.

[105] E. A. Nadaraya. *Nonparametric estimation of probability densities and regression curves.* Kluwer Academic Publishers, The Netherlands, 1989.

[106] Y. Nikitin. *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press, 1995.

[107] K. Noda. Estimation of a regression function by the Parzen kernel-type density estimators. *Ann. Inst. Math. Statist.*, 28 :221–234, 1976.

[108] S. Orey and W. E. Pruitt. Sample functions of the N-parameter Wiener process. *Ann. Probab.*, 1 :138–163, 1973.

[109] E. Parzen. On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33 :1065–1076, 1962.

[110] K. Pearson and A. Lee. On the laws of inheritance in a man. *Biometrika*, 2 :357–462, 1903.

[111] V. Petrov. *Limit theorems of probability theory.* Clarendon Press, 1995.

[112] P. Révész. On strong approximation of the multidimensional empirical process. *Ann. Probab.*, 4 :729–743, 1976.

[113] F. Riesz and B. Sz-Nagy. *Functional Analysis.* Frederick Unger, 1955.

[114] T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels. *Stochastic Processes for Insurance and Finance.* Wiley, 1999.

[115] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, 27 :832–837, 1956.

[116] M. Rosenblatt. *Multivariate Analysis II*, chapter Conditional probability density and regression estimators, pages 25–31. Academic Press, New York, 1969.

[117] L. Rüschendorf. Applications of empirical processes. *Communications lors des Journées sur les Propriétés Asymptotiques en Statistique Non-Paramétrique, Rouen*, juin 1979.

[118] F. H. Ruymgaart and J. A. Wellner. Some properties of weighted multivariate empirical processes. *Statistics and Decisions*, 2 :199–223, 1984.

[119] M. Schilder. Some asymptotic formulas for wiener integrals. *Trans. Amer. Math.*, 134 :193–216, 1966.

[120] E. F. Schuster. Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *Ann. Math. Statist.*, 43 :84–88, 1972.

[121] E. F. Schuster and S. Yakowitz. Contributions to the theory of nonparametric regression, with application to system identification. *Ann. Statist.*, 7 :1310–1319, 1979.

[122] D. W. Scott. *Multivariate Density Estimation : Theory, Practice and Visualization.* Wiley, New York, 1992.

[123] R. J. Serfling. *Approximation Theorems of Mathematical Statistics.* Wiley, New York, 1980.

[124] J. Shao and D. Tu. *The Jacknife and Bootstrap.* Springer-Verlag, New York, 1996.

[125] G. R. Shorack and J. A. Wellner. *Empirical Processes with Applications to Statistics.* Wiley, New York, 1986.

[126] A. V. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1 :261–290, 1956.

[127] N. Smirnoff. Sur les écarts de la courbe de distribution empirique. *Mat. Sbornik.*, 6 :3–26, 1939.

[128] C. Spiegelman and J. Sacks. Consistent window estimation in nonparametric regression. *Ann. Statist.*, 8 :240–246, 1980.

[129] C. Stone. Consistent nonparametric regression. *Ann. Statist.*, 5 :595–645, 1977.

[130] V. Strassen. An invariance principle for the Law of the Iterated Logarithm. *Z. Wahrsch. und verw. Gebiete*, 3 :221–226, 1964.

[131] W. Stute. The oscillation behaviour of empirical processes. *Ann. Probab.*, 10 :86–107, 1982.

[132] M. Talagrand. Donsker classes of sets. *Probab. Theory Related Fields*, 78 :169–191, 1988.

[133] M. S. Taqqu and C. Czado. A survey of functional laws of the iterated logarithm for self-similar processes. *Com. Statist. Stochastic Models.*, 1 :77–115, 1985.

[134] A. D. Ventsel. Rough limit theorems on large deviations for Markov processses. *Theory Probab. Appl.*, 21 :227–242, 499–512, 1976.

[135] G. Watson. Smooth regression analysis. *Sankhyã A*, 26 :359–372, 1964.