

ESTIMATION DE DENSITÉS

POL LABARBARIE ¹, MEHDY HOUNKONNOU ¹, GAUTHIER THURIN ¹ et BENJAMIN
MENENDEZ ¹

¹Master 1 Modélisation Statistique et Stochastique, Université de Bordeaux

16 juin 2020

TABLE DES MATIÈRES

I	Introduction	3
II	Contexte	3
III	Définition de nos estimateurs	4
I	Terme général de l'estimateur de Revesz	4
II	Trois estimateurs particuliers	5
IV	Convergences, martingales et compagnie	6
I	Une décomposition, une martingale et son crochet	6
II	Convergence presque sûre	8
III	Normalité asymptotique	9
V	Visualisations	13
VI	Algorithme de Revesz moyennisé	16
I	Un estimateur convergent	16
II	Simulations	17
VII	Bilan	18
VIII	Bibliographie	19

I. INTRODUCTION

L'objectif de ce rapport est de travailler sur l'estimateur de Revesz pour l'estimation de densités dans \mathbb{R} deux fois dérivables. L'estimation par noyaux est une des méthodes les plus utilisées pour l'estimation non paramétrique de la densité. L'estimateur de Revesz est construit comme un algorithme stochastique à pas simple, mais peut être traduit en un terme général qui, selon les valeurs de ses coefficients, définit toute une classe d'estimateurs récursifs par noyaux. Parmi eux, on retrouve les estimateurs de Wolverton-Wagner-Yamato, de Wegman-Davies et de Deheuvels, pour lesquels on précisera tous nos résultats de convergence. Après avoir établi le contexte de notre travail, nous partirons de la définition par récurrence de l'estimateur de Revesz pour retrouver son terme général et faire le parallèle avec nos trois cas particuliers. Puis, nous montrerons les convergences ps et en loi de nos estimateurs. Pour finir, nous considérerons l'algorithme de Revesz moyennisé. Nous accompagnerons nos résultats de simulations et de codes Python.

II. CONTEXTE

Soit X une variable aléatoire réelle continue de densité de probabilité f inconnue. On suppose que f est deux fois dérivable à dérivée continue bornée. On se donne (X_1, \dots, X_n) un n -échantillon iid de même loi que X et K une fonction positive paire et bornée, appelée noyau, telle que

$$\int_{\mathbb{R}} K(x)dx = 1, \quad \int_{\mathbb{R}} xK(x)dx = 0, \quad \int_{\mathbb{R}} K(x)^2dx = \xi^2.$$

Pour nos simulations, nous utiliserons le noyau gaussien, défini comme :

$$K(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

La fenêtre h_n est une suite déterministe, strictement positive et décroissante vers zéro, telle que nh_n tend vers l'infini. On considère ici, pour simplifier les calculs, que $h_n = 1/n^\alpha$ où $\alpha \in [0, 1]$.

L'estimateur de Revesz est défini par analogie avec l'algorithme de Robbins-Monro comme suit :

Pour tout $x \in \mathbb{R}$,

$$\widehat{f}_n(x) = (1 - \gamma_n)\widehat{f}_{n-1}(x) + \gamma_n W_n(x) \quad \text{avec} \quad W_n(x) = \frac{1}{h_n} K\left(\frac{x - X_n}{h_n}\right) \quad (1)$$

où $\widehat{f}_0(x) = 0$ et où le pas (γ_n) est une suite déterministe, positive et décroissante vers zéro, dont la série est divergente. On définit aussi (a_n) une suite de réels appartenant à l'intervalle $]0, 1]$, telle que :

$$\gamma_n = \frac{a_n}{A_n} \quad \text{et} \quad A_n = \sum_{k=1}^n a_k.$$

De cette définition par récurrence de l'estimateur de Revesz, il est assez simple de déduire son terme général. Et, de son terme général, il sera immédiat d'établir des parallèles avec les estimateurs de Wolverton-Wagner-Yamato, de Wegman-Davies et de Deheuvels. C'est l'objectif de la section qui suit.

III. DÉFINITION DE NOS ESTIMATEURS

I. Terme général de l'estimateur de Revesz

On définit, pour tout $n \geq 1$, $B_n = \prod_{k=1}^n (1 - \gamma_k)$.

On montre d'abord, pour tout $x \in \mathbb{R}$, que

$$\widehat{f}_n(x) = B_n \sum_{k=1}^n \frac{\gamma_k}{B_k} W_k(x). \quad (2)$$

Soit la proposition $P_n : \forall n \geq 1, \widehat{f}_n(x) = B_n \sum_{k=1}^n \frac{\gamma_k}{B_k} W_k(x)$.

Initialisation : pour $n = 1$:

$$\widehat{f}_1(x) = (1 - \gamma_1)\widehat{f}_0(x) + \gamma_1 W_1(x) = \gamma_1 W_1(x) = B_1 \frac{\gamma_1}{B_1} W_1(x)$$

Hérédité : on suppose que P_n est vraie. Montrons P_{n+1} .

$$\begin{aligned} \widehat{f}_{n+1}(x) &= (1 - \gamma_{n+1})\widehat{f}_n(x) + \gamma_{n+1} W_{n+1}(x) \\ &= (1 - \gamma_{n+1})B_n \sum_{k=1}^n \frac{\gamma_k}{B_k} W_k(x) + \gamma_{n+1} W_{n+1}(x) \end{aligned}$$

Or $(1 - \gamma_{n+1})B_n = B_{n+1}$ et $\gamma_{n+1} W_{n+1}(x) = B_{n+1} \frac{\gamma_{n+1}}{B_{n+1}} W_{n+1}(x)$, d'où $\widehat{f}_{n+1}(x) = B_{n+1} \sum_{k=1}^{n+1} \frac{\gamma_k}{B_k} W_k(x)$.

Ainsi, P_n est vraie pour tout n .

A partir de l'équation (2), nous pouvons aboutir à une réécriture de $\widehat{f}_n(x)$.

En effet,

$$\widehat{f}_n(x) = B_n \sum_{k=1}^n \frac{\gamma_k}{B_k} W_k(x) = B_n \sum_{k=1}^n \frac{a_k}{B_k A_k} W_k(x) = \frac{1}{A_n} \sum_{k=1}^n \frac{B_n A_n}{B_k A_k} a_k W_k(x)$$

Mais, pour tout k ,

$$\begin{aligned} B_k &= \prod_{i=1}^k (1 - \gamma_i) = \prod_{i=1}^k \left(1 - \frac{a_i}{A_i}\right) = \prod_{i=1}^k \left(\frac{A_{i-1}}{A_i}\right) \\ &= \frac{A_0 A_1 A_2 \dots A_{k-1}}{A_1 A_2 A_3 \dots A_{k-1} A_k} = \frac{A_0}{A_k} \\ \text{ie } B_k A_k &= A_0 \end{aligned}$$

D'où, pour tout n , pour tout $k \leq n$,

$$\frac{B_n A_n}{B_k A_k} = \frac{A_0}{A_0} = 1$$

Nous obtenons ainsi le terme général de notre estimateur de Revesz :

$$\widehat{f}_n(x) = \frac{1}{A_n} \sum_{k=1}^n a_k W_k(x). \quad (3)$$

Avant de faire un parallèle entre (3) et les estimateurs de Parzen-Rosenblatt, Wolverton-Wagner-Yamato et Wegman-Davies, montrons que $\lim_{n \rightarrow +\infty} A_n = +\infty \iff \sum_{n=1}^{\infty} \gamma_n = +\infty$.

D'après la définition de γ_n , nous avons que :

$$\sum_{n=1}^{\infty} \gamma_n = \sum_{n=1}^{\infty} \frac{a_n}{A_n} = \sum_{n=1}^{\infty} \frac{A_n - A_{n-1}}{A_n}$$

Et il suffit de faire un dessin pour visualiser que, pour tout $n \geq 1$,

$$\frac{A_n - A_{n-1}}{A_n} \leq \int_{A_{n-1}}^{A_n} \frac{1}{x} dx,$$

Mais alors

$$\begin{aligned} \sum_{n=1}^{\infty} \gamma_n &\leq \sum_{n=1}^{\infty} \ln A_n - \ln A_{n-1} \\ \lim_{n \rightarrow +\infty} \ln A_n &\geq \sum_{n=1}^{\infty} \gamma_n \end{aligned}$$

On en déduit la première implication : $\sum_{n=1}^{\infty} \gamma_n = +\infty \Rightarrow \lim_{n \rightarrow +\infty} A_n = +\infty$.

Réciproquement, il suffit d'appliquer le lemme de Kronecker, en raisonnant par l'absurde :

- (A_n) est positive croissante divergente.
- (a_n) est une suite de réels tels que $\sum_{n=1}^{\infty} \frac{a_n}{A_n} < +\infty$.

Sous ces conditions, Kronecker nous assure que $\frac{1}{A_n} \sum_{k=1}^n a_k$ tend vers 0, ce qui est absurde.

II. Trois estimateurs particuliers

Connaissant le terme général de notre estimateur de Revesz, on peut maintenant introduire trois estimateurs de f en donnant, à une équivalence près sur A_n , les valeurs pour a_n qui permettent de les classer comme des cas particuliers de Revesz :

Pour tout $x \in \mathbb{R}$,

Estimateurs	Expression	a_n
Estimateur de Parzen-Rosenblatt / Deheuvels	$\frac{1}{nh_n} \sum_{k=1}^n K\left(\frac{x-X_k}{h_k}\right)$	h_n
Estimateur de Wolverton-Wagner-Yamato	$\frac{1}{n} \sum_{k=1}^n \frac{1}{h_k} K\left(\frac{x-X_k}{h_k}\right)$	1
Estimateur de Wegman-Davies	$\frac{1}{nh_n^{1/2}} \sum_{k=1}^n \frac{1}{\sqrt{h_k}} K\left(\frac{x-X_k}{h_k}\right)$	$\sqrt{h_n}$

La section suivante est destinée à l'étude des convergences ps et en loi de ces estimateurs.

IV. CONVERGENCES, MARTINGALES ET COMPAGNIE

I. Une décomposition, une martingale et son crochet

Dans cette partie, on introduit une martingale issue d'une décomposition de $\widehat{f}_n(x) - f(x)$ et un équivalent de son crochet, ce qui nous permettra ensuite de justifier la convergence ps de nos estimateurs et une normalité asymptotique.

À x fixé, étudions la différence $\widehat{f}_n(x) - f(x)$:

$$\begin{aligned}\widehat{f}_n(x) - f(x) &= \frac{1}{A_n} \left(\sum_{k=1}^n a_k W_k(x) - A_n f(x) \right) \\ &= \frac{1}{A_n} \left(\sum_{k=1}^n a_k W_k(x) - \sum_{k=1}^n a_k \mathbb{E}[W_k(x)] + \sum_{k=1}^n a_k \mathbb{E}[W_k(x)] - \sum_{k=1}^n a_k f(x) \right).\end{aligned}$$

On pose :

$$M_n(x) = \sum_{k=1}^n a_k (W_k(x) - \mathbb{E}[W_k(x)]) \text{ et } R_n(x) = \sum_{k=1}^n a_k (\mathbb{E}[W_k(x)] - f(x))$$

ce qui nous amène à la décomposition comme suit :

$$\widehat{f}_n(x) - f(x) = \frac{1}{A_n} (M_n(x) + R_n(x)).$$

Pour toute la suite, comme on travaille à x fixé, on simplifiera les notations et on se permettra d'omettre les x . Par exemple, on préférera M_n à $M_n(x)$.

On vérifie facilement que (M_n) est une martingale de carré intégrable. Soit la tribu $\mathcal{F}_n = \sigma(W_1, \dots, W_n)$.

• Carré intégrable : À n fixé,

$$\begin{aligned}\mathbb{E}(|M_{n+1}|^2) &\leq \sum_{k=1}^n a_k^2 \mathbb{E}[(W_k - \mathbb{E}(W_k))^2] \\ &\leq \sum_{k=1}^n a_k^2 \text{Var}(W_k), \quad \text{car } W_k - \mathbb{E}(W_k) \text{ est centrée} \\ &\leq \sum_{k=1}^n \frac{a_k^2}{h_k^2} \text{Var} \left(K \left(\frac{x - X_k}{h_k} \right) \right) \\ &\leq \sum_{k=1}^n \frac{a_k^2}{h_k^2} C, \quad \text{car } K(x) \leq C, \text{ pour tout } x, \text{ pour un certain } C \geq 0\end{aligned}$$

• Adaptabilité : À n fixé, M_n est \mathcal{F}_n -mesurable par composition d'une fonction mesurable avec des variables aléatoires \mathcal{F}_n -mesurables.

• Formule de martingale :

Notons que K est intégrable donc mesurable, et les (X_k) sont indépendants, donc les (W_k) sont indépendants et W_k est indépendante de \mathcal{F}_{k-1} . Sachant cela et le point précédent,

$$\mathbb{E}(M_{n+1} | \mathcal{F}_n) = \mathbb{E}(M_n | \mathcal{F}_n) + a_{n+1} \mathbb{E}[(W_{n+1} - \mathbb{E}(W_{n+1})) | \mathcal{F}_n] = M_n,$$

Finalement, (M_n) est bien une martingale de carré intégrable.

Regardons son crochet $\langle M \rangle_n$.

$$\begin{aligned}
 \langle M \rangle_n &= \sum_{k=1}^n \mathbb{E}[(M_k(x) - M_{k-1}(x))^2 | \mathcal{F}_{k-1}] \\
 &= \sum_{k=1}^n \mathbb{E}[a_k^2 [W_k(x) - \mathbb{E}[W_k(x)]]^2 | \mathcal{F}_{k-1}], \quad \text{on note que } W_k(x) \text{ est indépendante de } \mathcal{F}_{k-1} \\
 &= \sum_{k=1}^n a_k^2 \text{Var}(W_k(x)) \\
 &= \sum_{k=1}^n a_k^2 (\mathbb{E}[W_k(x)^2] - \mathbb{E}[W_k(x)]^2)
 \end{aligned}$$

Alors, en développant ces deux espérances pour tout k , avec le changement de variable $z = \frac{x-y}{h_k}$, nous obtenons :

$$\langle M \rangle_n = \sum_{k=1}^n \frac{a_k^2}{h_k} \left[\int_{\mathbb{R}} K(z)^2 f(x + zh_k) dz - h_k \left(\int_{\mathbb{R}} K(z) f(x + zh_k) dz \right)^2 \right]$$

Alors, en séparant en deux blocs :

D'une part :

$$\int_{\mathbb{R}} K(z)^2 f(x + zh_k) dz \xrightarrow[k \rightarrow +\infty]{} \xi^2 f(x), \quad \text{par continuité de } f, \text{ sachant que } h_k \xrightarrow[k \rightarrow +\infty]{} 0$$

et

$$h_k \int_{\mathbb{R}} K(z) f(x + zh_k) dz \xrightarrow[k \rightarrow +\infty]{} 0.$$

Et, d'autre part :

$$\sum_{k=1}^n \frac{a_k^2}{h_k} \text{ diverge car } \frac{a_k^2}{h_k} \xrightarrow[k \rightarrow +\infty]{} +\infty.$$

Ainsi, d'après le Lemme de Toeplitz, on a que

$$\frac{\langle M \rangle_n}{v_n} \xrightarrow[k \rightarrow +\infty]{} \xi^2 f(x), \quad \text{avec } v_n = \sum_{k=1}^n \frac{a_k^2}{h_k}$$

II. Convergence presque sûre

Dans cette section notre objectif est de montrer que $\lim_{n \rightarrow \infty} \hat{f}_n(x) = f(x)$ p.s pour nos trois estimateurs.

D'abord, montrons que, dans chacun de ces trois cas, $M_n(x) = o(A_n)$ ps. Pour ce faire, on utilise la deuxième LFGN pour les martingales, dont on rappelle l'énoncé. Pour (M_n) une martingale de carré intégrable dont le crochet est divergent :

$$\mathbf{2e\ LFGN} : \forall \gamma > 0 : \frac{M_n^2}{\langle M \rangle_n} = o\left(\log(\langle M \rangle_n)^{(1+\gamma)}\right)$$

Distinguons les cas.

• **Wolverton-Wagner-Yamato** : Pour $a_k = 1$, pour $\gamma > 0$: $M_n^2 = o(n^{\alpha+1}(\log n)^{1+\gamma})$
 car $\langle M \rangle_n \sim v_n$ en $+\infty$ ie $\langle M \rangle_n \sim n^{\alpha+1}$.
 Alors, pour $\gamma = \alpha$ en particulier :

$$M_n^2 = o(n^{\alpha+1}(\log n)^{\alpha+1})$$

Or

$$(\log n)^{\frac{\alpha+1}{2}} = o(n^{\frac{1-\alpha}{2}})$$

Donc

$$M_n = o(n^{\frac{\alpha+1}{2}} n^{\frac{1-\alpha}{2}}) = o(n) = o(A_n)$$

• **Deheuvels** : Pour $a_n = h_n$, $v_n = A_n$. Par la LFGN pour les martingales, $M_n = o(\langle M \rangle_n) = o(v_n) = o(A_n)$.

• **Wegman-Davies** : Pour $a_n = \sqrt{h_n}$, $v_n = n$, et

$$A_n = \sum_{k=1}^n h_k^{\frac{1}{2}} \sim \int_1^n \frac{1}{x^{\frac{\alpha}{2}}} dx$$

$$\text{D'où } A_n \sim \frac{n^{-\frac{\alpha}{2}+1}}{1-\frac{\alpha}{2}}$$

En utilisant encore la deuxième LFGN pour les martingales, on a, pour tout $\gamma > 0$:

$$M_n^2 = o(n(\log n)^{1+\gamma})$$

$$\text{Or } (\log n)^{\frac{1+\gamma}{2}} = o(n^{\frac{1-\alpha}{2}})$$

$$\text{Donc } M_n = o(n^{\frac{1}{2}} n^{\frac{1-\alpha}{2}}) = o(n^{1-\frac{\alpha}{2}}) = o(A_n) \text{ ps.}$$

Finalement, on vient de montrer que $\frac{M_n}{A_n} \xrightarrow[n \rightarrow +\infty]{} 0$ p.s pour nos trois estimateurs.

De plus, $R_n(x) = o(A_n)$. En effet, par Toeplitz puisque $\mathbb{E}(W_k) \xrightarrow[k \rightarrow +\infty]{} f(x)$,

$$\frac{R_n(x)}{A_n} = \frac{\sum_{k=1}^n a_k [\mathbb{E}(W_k(x)) - f(x)]}{\sum_{k=1}^n a_k} \xrightarrow[n \rightarrow +\infty]{} 0$$

On déduit que $\hat{f}_n(x) \xrightarrow{} f(x)$ p.s.

III. Normalité asymptotique

L'objectif de cette partie est d'aboutir à la normalité asymptotique suivante : pour tout $x \in \mathbb{R}$ tel que $f(x) > 0$,

$$\frac{A_n}{\sqrt{v_n}}(\widehat{f_n}(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x))$$

Pour ce faire, nous voulons établir la normalité asymptotique :

$$\frac{M_n(x)}{\sqrt{v_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x))$$

Nous allons utiliser le TCL pour les martingales. Fixons x tel que $f(x) > 0$.

Premièrement, nous avons montré précédemment que :

$$\frac{\langle M \rangle_n}{v_n} \xrightarrow{p.s.} \xi^2 f(x)$$

La convergence en probabilités tombe alors directement et le premier point du TCL est vérifié.

Maintenant, montrons la condition de Lindeberg. Ecrivons les accroissements de (M_n) . $\forall k \in \mathbb{N}$:

$$|M_k - M_{k-1}| = \frac{a_k}{h_k} \left| K\left(\frac{x - X_k}{h_k}\right) - \mathbb{E}K\left(\frac{x - X_k}{h_k}\right) \right|$$

Comme K est bornée, on peut choisir $C > 0$ telle que $\forall x, |K(x) - \mathbb{E}K(x)| \leq C$. On obtient que :

$$\forall k, |M_k - M_{k-1}| \leq \frac{2Ca_k}{h_k} \text{ p.s.}$$

Fixons $\epsilon > 0$.

Séparons maintenant les cas pour montrer que $\psi_n \xrightarrow{\mathbb{P}} 0$ où ψ_n dépend de ϵ :

$$\psi_n = \frac{1}{v_n} \sum_{k=1}^n \mathbb{E} \left[|M_k - M_{k-1}|^2 \mathbf{1}_{|M_k - M_{k-1}| \geq \epsilon \sqrt{v_n}} \mid \mathcal{F}_{k-1} \right]$$

Dans tous les cas, d'après ce qui précède :

$$0 \leq \psi_n \leq \frac{1}{v_n} \sum_{k=1}^n \frac{4C^2 a_k^2}{h_k^2} \mathbb{P}(|M_k - M_{k-1}| \geq \epsilon \sqrt{v_n}) \text{ p.s.}$$

De plus, comme :

$$\{|M_k - M_{k-1}| \geq \epsilon \sqrt{v_n}\} \subseteq \left\{ \frac{2Ca_k}{h_k} \geq \epsilon \sqrt{v_n} \right\}$$

On a en fait, dans tous les cas :

$$0 \leq \psi_n \leq \frac{1}{v_n} \sum_{k=1}^n \frac{4C^2 a_k^2}{h_k^2} \mathbb{E}(\mathbf{1}_{2C \geq \epsilon \sqrt{v_n} \frac{h_k}{a_k}}) \text{ p.s.}$$

D'où

$$0 \leq \psi_n \leq \frac{1}{v_n} \sum_{k=1}^n \frac{4C^2 a_k^2}{h_k^2} \mathbf{1}_{2C \geq \epsilon \sqrt{v_n} \frac{h_k}{a_k}} \text{ p.s.}$$

Alors, dans les 3 cas, il suffira, par Toeplitz, de montrer que : $4C^2 \mathbf{1}_{2C \geq \epsilon \sqrt{v_n} \frac{h_k}{a_k}} \xrightarrow[k \rightarrow \infty]{} 0$

•**Wolverton-Wagner-Yamato** : $a_k = 1$

L'inégalité devient :

$$0 \leq \psi_n \leq \frac{1}{v_n} \sum_{k=1}^n \frac{4C^2}{h_k^2} \mathbf{1}_{2C \geq \epsilon \sqrt{v_n} \frac{h_k}{a_k}}$$

D'une part, nous avons que $v_n \sim \frac{n^{\alpha+1}}{\alpha+1}$, $\sqrt{v_n} h_n \sim n^{\frac{1-\alpha}{2}}$, d'où $\sqrt{v_n} h_n \xrightarrow[n \rightarrow \infty]{} \infty$, et

$$4C^2 \mathbf{1}_{2C \geq \epsilon \sqrt{v_k} h_k} \xrightarrow[k \rightarrow \infty]{} 0$$

D'autre part : v_n diverge. Donc par le lemme de Toeplitz :

$$\frac{1}{v_n} \sum_{k=1}^n \frac{1}{h_k} 4C^2 \mathbf{1}_{2C \geq \epsilon \sqrt{v_k} h_k} \xrightarrow[n \rightarrow \infty]{} 0$$

Ainsi, par le théorème des gendarmes, $\psi_n \xrightarrow[n \rightarrow \infty]{} 0$ p.s. D'où $\psi_n \xrightarrow{\mathbb{P}} 0$.

•**Wagner-Davies** : $a_k = \sqrt{h_k}$

Ici :

$$\sqrt{v_k} \frac{h_k}{a_k} = \sqrt{v_k} \frac{1}{k^{\frac{\alpha}{2}}} \sim \frac{k^{\alpha+1}}{\alpha+1} \frac{1}{k^{\frac{\alpha}{2}}} = \frac{k^{\frac{\alpha}{2}+1}}{\alpha+1}$$

Donc $\sqrt{v_k} \sqrt{h_k} \xrightarrow[h \rightarrow \infty]{} \infty$. A partir d'un certain rang, on a donc $\mathbf{1}_{(2C \geq \epsilon \sqrt{v_k} \sqrt{h_k})} = 0$ donc

$$4C^2 \mathbf{1}_{2C \geq \epsilon \sqrt{v_k} \frac{h_k}{a_k}} \xrightarrow[h \rightarrow \infty]{} 0$$

D'où $\psi_n \xrightarrow{\mathbb{P}} 0$ pour $n \rightarrow \infty$

•**Deheuvels** : $a_k = h_k$

On a maintenant :

$$\sqrt{v_k} \frac{h_k}{a_k} = \sqrt{v_k} \xrightarrow[h \rightarrow \infty]{} \infty$$

donc

$$4C^2 \mathbf{1}_{2C \geq \epsilon \sqrt{v_k} \frac{h_k}{a_k}} \xrightarrow[h \rightarrow \infty]{} 0$$

D'où $\psi_n \xrightarrow{\mathbb{P}} 0$ pour $n \rightarrow \infty$

Finalement, dans les trois cas, la condition de Lindeberg est vérifiée. Le théorème central limite nous dit alors :

$$\frac{M_n(x)}{\sqrt{v_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x))$$

Pour conclure à la normalité asymptotique

$$\frac{A_n}{\sqrt{v_n}} (\widehat{f_n}(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x)),$$

Il ne reste qu'à montrer que $\frac{R_n(x)}{\sqrt{v_n}}$ converge vers 0. Majorons cette quantité.

$$\mathbb{E}[W_k(x) - f(x)] = \int_{\mathbb{R}} K(z) [f(x - zh_k) - f(x)] dz$$

En faisant un développement de Taylor, avec $t_k = x - zh_k\theta$ pour un certain $\theta \in]0, 1[$,

$$\mathbb{E}[W_k(x) - f(x)] = \underbrace{\int_{\mathbb{R}} -K(z)zh_k f(x)dz}_{=0 \text{ car } \int_{\mathbb{R}} zK(z)dz=0} + \frac{h_k^2}{2} \int_{\mathbb{R}} z^2 K(z) f''(x - zh_k \theta) dz$$

D'où :

$$\begin{aligned} |\mathbb{E}[W_k] - f(x)| &\leq \frac{h_k^2}{2} \int_{\mathbb{R}} z^2 K(z) |f''(x - zh_k \theta)| dz \\ &\leq \frac{h_k^2}{2} \left(\int_{\mathbb{R}} z^2 K(z) dz \right) \sup_{x \in \mathbb{R}} |f''(x)| \end{aligned}$$

Notons $\mathcal{T} = \int_{\mathbb{R}} z^2 K(z) dz$.

On a alors montré que, pour un certain $\mathcal{T} > 0$:

$$|\mathbb{E}[W_k - f(x)]| \leq \frac{h_k^2}{2} \mathcal{T} \sup_{x \in \mathbb{R}} |f''(x)|$$

Nous pouvons enfin montrer que $\frac{R_n}{\sqrt{v_n}} \xrightarrow[n \rightarrow \infty]{} 0$ dans les 3 cas. On a :

$$\frac{R_n}{\sqrt{v_n}} = \frac{1}{\sqrt{v_n}} \sum_{k=1}^n (\mathbb{E}[W_k(x) - f(x)])$$

Avec ce qui précède :

$$\begin{aligned} \left| \frac{R_n}{\sqrt{v_n}} \right| &\leq \frac{1}{\sqrt{v_n}} \sum_{k=1}^n |\mathbb{E}[W_k(x) - f(x)]| \\ \left| \frac{R_n}{\sqrt{v_n}} \right| &\leq \frac{1}{\sqrt{v_n}} \sum_{k=1}^n \frac{h_k^2}{2} \mathcal{T} \sup_{x \in \mathbb{R}} |f''(x)| \\ \left| \frac{R_n}{\sqrt{v_n}} \right| &\leq \frac{\mathcal{T} \sup_{x \in \mathbb{R}} |f''(x)|}{2\sqrt{v_n}} \sum_{k=1}^n h_k^2 \end{aligned}$$

Distinguons les cas.

Wolverton-Wagner-Yamato : $a_k = 1$

$$\begin{aligned} \sum_{k=1}^n h_k^2 &\sim \frac{n^{-2\alpha+1}}{-2\alpha+1} \\ \text{et } \sqrt{v_n} &\sim \sqrt{\frac{n^{\alpha+1}}{\alpha+1}} \end{aligned}$$

Donc, dès que $1 > \alpha > \frac{1}{5}$:

$$\sum_{k=1}^n h_k^2 = o(\sqrt{v_n})$$

et

$$\frac{R_n}{\sqrt{v_n}} \xrightarrow[n \rightarrow \infty]{} 0$$

Deheuvels : $a_k = h_k$

$$\sum_{k=1}^n h_k^2 \sim n^{-2\alpha+1}$$

$$\sqrt{v_n} \sqrt{\sum_{k=1}^n h_k} \sim \sqrt{n^{-\alpha+1}}$$

Alors :

$$\frac{1}{\sqrt{v_n}} \sum_{k=1}^n h_k^2 \sim n^{-\frac{3\alpha}{2} + \frac{1}{2}}$$

Donc $R_n = o(\sqrt{v_n})$, comme précédemment, dès que $\frac{1}{3} < \alpha < 1$.

Wegman-Davies : $a_k = \sqrt{h_k}$

On a :

$$\sqrt{v_n} = \sqrt{n}$$

$$\sum_{k=1}^n h_k^2 \sim \frac{n^{-2\alpha+1}}{-2\alpha+1}$$

D'où

$$\frac{1}{\sqrt{v_n}} \sum_{k=1}^n h_k^2 \sim n^{-2\alpha+1/2}.$$

Ce terme converge vers 0 dès que $\frac{1}{4} < \alpha < 1$.

Donc dans les trois cas, selon les valeurs de α , on a montré que pour tout $x \in \mathbb{R}$, (x tel que $f(x) > 0$)

$$\frac{R_n}{\sqrt{v_n}} \xrightarrow{n \rightarrow \infty} 0$$

$$\frac{M_n}{\sqrt{v_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x))$$

De ces deux points, on déduit, pour nos trois estimateurs :

$$\frac{A_n}{\sqrt{v_n}} (\widehat{f_n}(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x))$$

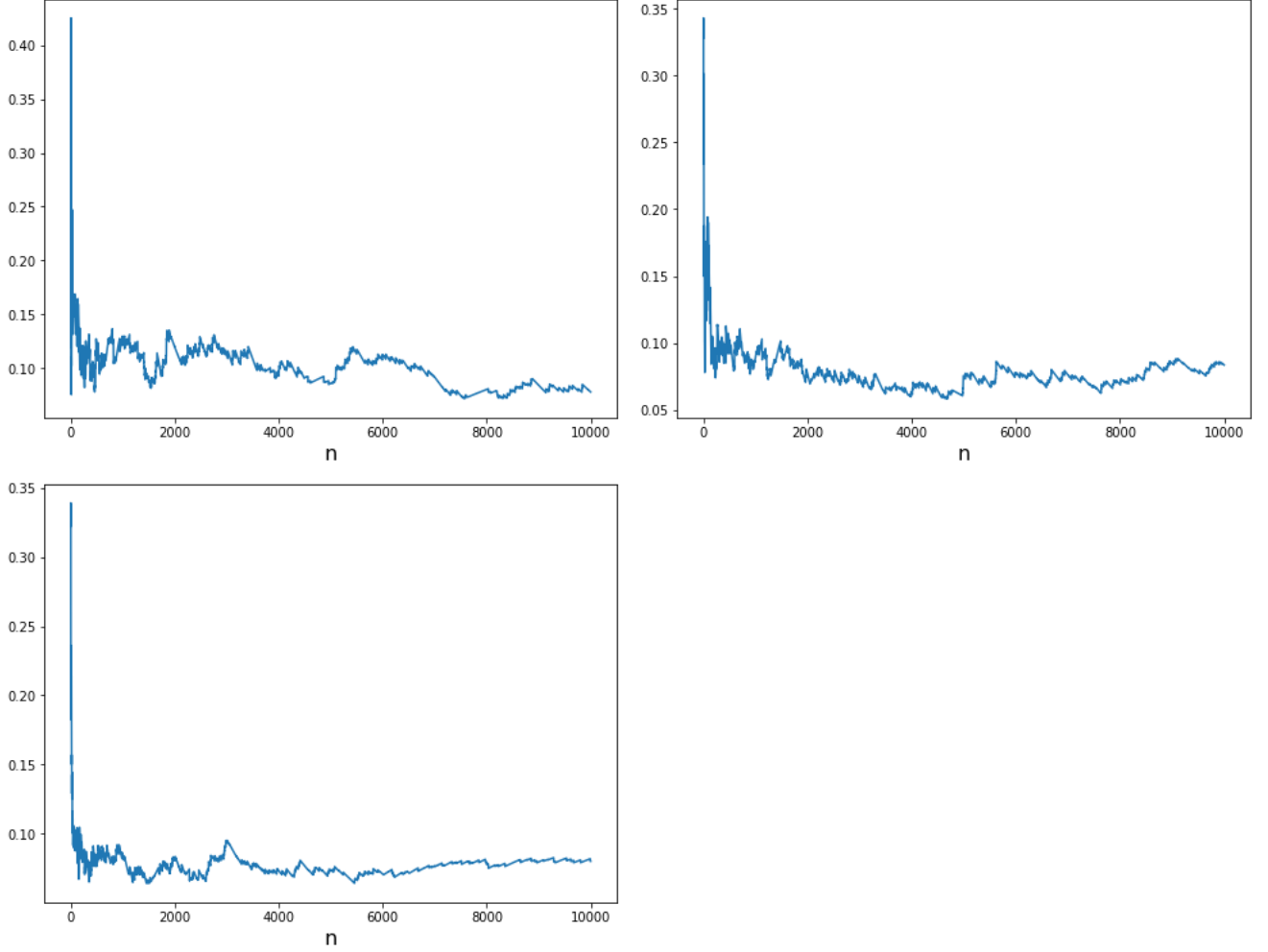
Le tableau suivant synthétise les conditions sur α , respectivement pour les estimateurs de Deheuvels, Wolverton-Wagner-Yamato et Wegman-Davies, nécessaires à cette normalité asymptotique.

Estimateur $\widehat{f_n}(x)$	Valeur de a_n	Condition sur α
Deheuvels : $\frac{1}{nh_n} \sum_{k=1}^n K\left(\frac{x-X_k}{h_k}\right)$	h_n	$1/3 < \alpha < 1$
WWY : $\frac{1}{n} \sum_{k=1}^n \frac{1}{h_k} K\left(\frac{x-X_k}{h_k}\right)$	1	$1/5 < \alpha < 1$
WD : $\frac{1}{nh_n^{1/2}} \sum_{k=1}^n \frac{1}{\sqrt{h_k}} K\left(\frac{x-X_k}{h_k}\right)$	$\sqrt{h_n}$	$1/4 < \alpha < 1$

V. VISUALISATIONS

Dans cette partie, nous donnons des illustrations des convergences démontrées plus haut. Pour Wolverton-Wagner-Yamato, nous prenons $\alpha = 3/5$, pour Deheuvels $\alpha = 3/4$ et pour Wegman-Davies $\alpha = 1/2$.

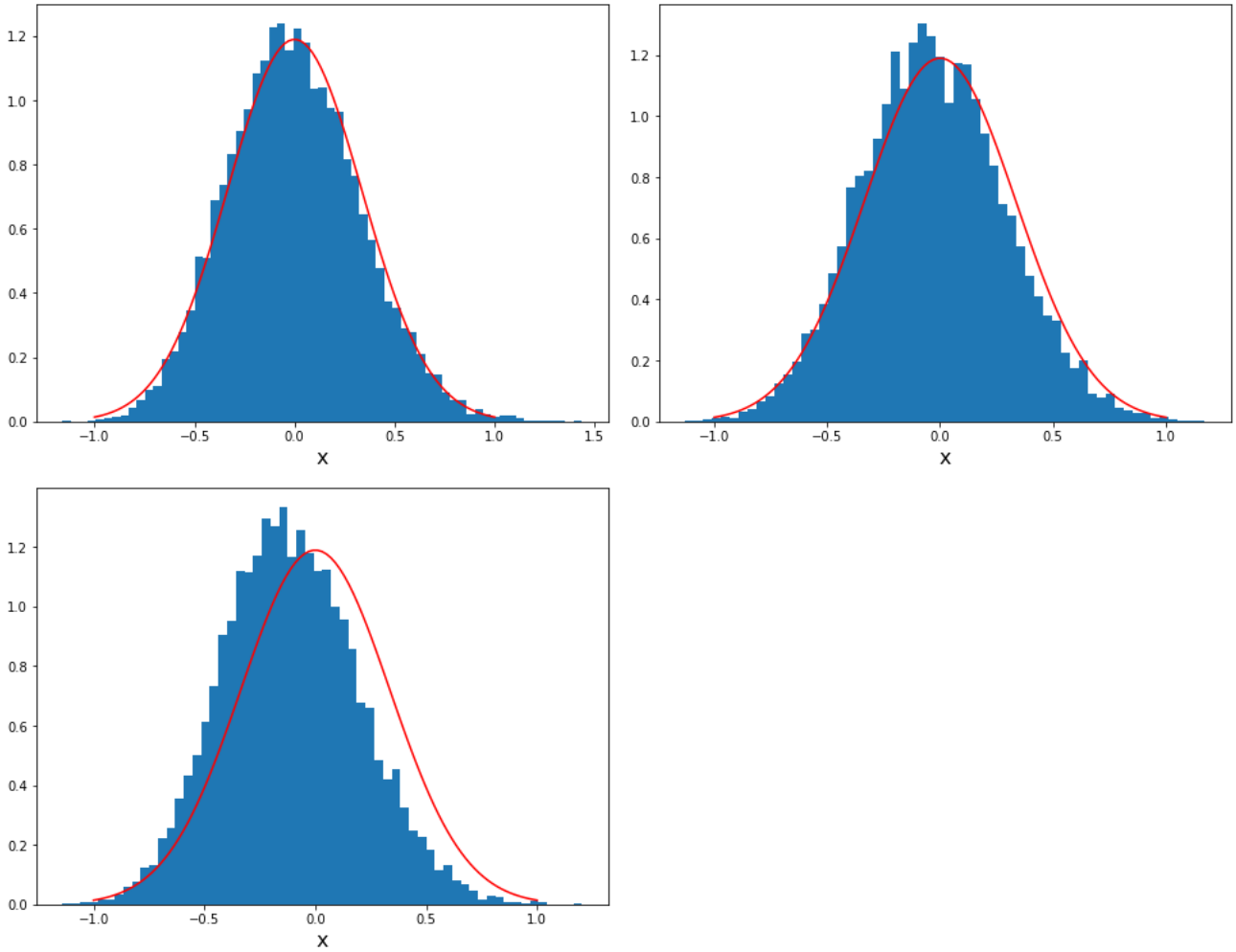
• D'abord, pour nos trois estimateurs, nous illustrons la convergence ps vers $f(x)$, à x fixé en traçant l'évolution du sup de la différence $\widehat{f}_n(x) - f(x)$. On obtient, respectivement pour Wolverton-Wagner-Yamato, Wegman-Davies et Deheuvels :



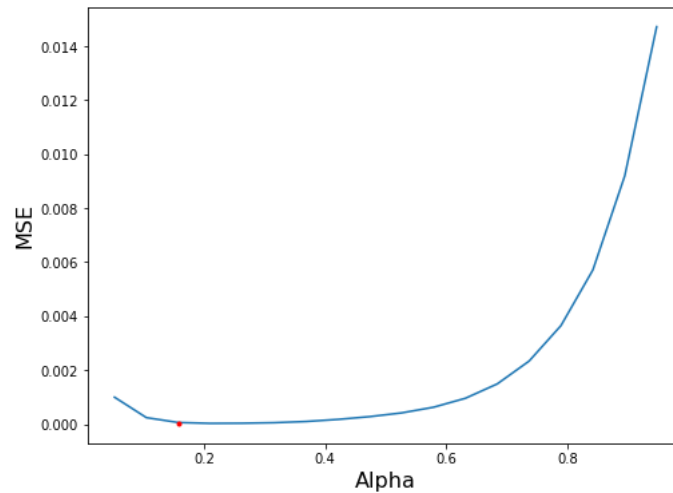
• Ensuite, nous illustrons la convergence en loi :

$$\frac{A_n}{\sqrt{v_n}}(\widehat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x)).$$

On obtient, respectivement pour Wolverton-Wagner-Yamato, Wegman-Davies et Deheuvels :



• On compare également les vitesses d'exécution de nos estimateurs, en comparant leurs temps de convergence vers $f(x)$, à x fixé, avec pour critère d'arrêt : $\widehat{f}_n(x) - f(x) < 10^{-5}$. Afin de mettre nos algorithmes sur un pied d'égalité avant cette course contre la montre, nous avons optimisé α dans chaque cas, au sens où ce α optimum minimise $\left(\widehat{f}_n(x) - f(x)\right)^2$. Le graphique suivant montre l'évolution du MSE en fonction des valeurs de α et la position du minimum (dans le cas Wolverton-Wagner-Yamato).



Pour obtenir plus de précision, nous faisons une moyenne, sur 100 exécutions du code, du temps de calcul. Nos résultats sont :

Estimateur $\widehat{f}_n(x)$	Temps de calcul moyen
Deheuvels : $\frac{1}{nh_n} \sum_{k=1}^n K\left(\frac{x-X_k}{h_n}\right)$	0.38756
WWY : $\frac{1}{n} \sum_{k=1}^n \frac{1}{h_k} K\left(\frac{x-X_k}{h_n}\right)$	0.16561
WD : $\frac{1}{nh_n^{1/2}} \sum_{k=1}^n \frac{1}{\sqrt{h_k}} K\left(\frac{x-X_k}{h_n}\right)$	2.55032

Wolverton-Wagner-Yamato est ainsi l'estimateur qui fournit l'algorithme le plus rapide.

VI. ALGORITHME DE REVESZ MOYENNISÉ

L'algorithme de Revesz moyennisé est défini, pour tout $x \in \mathbb{R}$, par :

$$\overline{f}_n(x) = \frac{1}{C_n} \sum_{k=1}^n c_k \widehat{f}_k(x) \quad \text{avec} \quad C_n = \sum_{k=1}^n c_k$$

où (c_n) est une suite déterministe et positive, dont la série est divergente. On considère, pour simplifier, que $c_n = 1/n^\gamma$ avec $\gamma \in [0, 1]$. Dans cette section, on s'attachera à montrer la convergence ps et en loi de cet estimateur, puis nous illustrerons ces résultats.

I. Un estimateur convergent

Fixons x un réel tel que $f(x) > 0$.

Rappelons que $\widehat{f}_k(x)$ converge ps vers $f(x)$. Alors la convergence ps de $\overline{f}_n(x)$ vers $f(x)$ est immédiate, en appliquant Toeplitz à :

$$\overline{f}_n(x) - f(x) = \frac{1}{C_n} \sum_{k=1}^n c_k (\widehat{f}_k(x) - f(x)).$$

Montrons maintenant la normalité asymptotique suivante, où (V_n) reste à expliciter :

$$\frac{C_n}{\sqrt{V_n}} (\overline{f}_n(x) - f(x)) \Rightarrow \mathcal{N}(0, \xi^2 f(x)).$$

Soit F_n la filtration associée aux $(W_n(x))$. On a d'abord la décomposition :

$$\overline{f}_n(x) - f(x) = \frac{1}{C_n} (L_n(x) + P_n(x))$$

où $P_n(x) := \sum_{k=1}^n c_k (\mathbb{E}(\widehat{f}_k(x)) - f(x))$ et $L_n(x) := \sum_{k=1}^n c_k (\widehat{f}_k(x) - \mathbb{E}(\widehat{f}_k(x)))$ est une martingale de carré intégrable. En effet :

1. Pour tout n , $\widehat{f}_n(x)$ étant $(F_n(x))$ -adapté, $L_n(x)$ l'est aussi.
2. Pour tout n , $L_n(x)$ est centrée et de carré intégrable : d'une part, par récurrence descendante, $\text{Var}(L_n(x)) = \sum_{k=1}^n c_k^2 \text{Var}(\widehat{f}_k(x))$, et, d'autre part, à k fixé et K étant bornée, $W_k(x)$ est bornée dans L^2 et a fortiori $\text{Var}(\widehat{f}_n(x)) = \frac{1}{A_n^2} \sum_{k=1}^n a_k^2 \text{Var}(W_k(x))$ est bornée.
3. Du terme général de $\widehat{f}_n(x)$ et puisque les $(W_k(x))$ sont indépendants, on déduit que $\widehat{f}_n(x)$ est indépendant de $F_{n-1}(x)$ pour tout n . D'où on peut déduire que $\mathbb{E}(L_{n+1}(x) | F_n(x)) = L_n(x)$.

Dans toute la suite, comme il n'y a pas d'ambiguïté dans nos notations (on travaille à x fixé), on se permettra d'omettre les x dans l'écriture de nos processus. Par exemple, on préférera écrire L_n au lieu de $L_n(x)$.

Utilisons encore le TCL pour les martingales, appliqué à (L_n) . Il est rapide de vérifier que le crochet associé à (L_n) vérifie :

$$\langle L \rangle_n = \sum_{k=1}^n c_k^2 \text{Var}(\widehat{f}_k(x)). \quad \text{Mais } \text{Var}(\widehat{f}_k(x)) = \frac{1}{A_k^2} \langle M \rangle_k. \quad \text{Donc } \langle L \rangle_n = \sum_{k=1}^n \frac{c_k^2}{A_k^2} v_k \left(\frac{\langle M \rangle_k}{v_k} \right).$$

Alors, par Toeplitz et d'après la convergence ps de $\frac{\langle M \rangle_n}{v_n}$, en notant $V_n := \sum_{k=1}^n \frac{c_k^2}{A_k^2} v_k$,

$$\lim_{n \rightarrow +\infty} \frac{\langle L \rangle_n}{V_n} = \xi^2 f(x) \text{ ps}$$

(V_n) étant une suite croissante de réels strictement positifs, le premier point du TCL pour les martingales est vérifié si (V_n) diverge. Pour justifier ceci, on peut distinguer selon que $a_n = h_n$, $a_n = \sqrt{h_n}$, ou $a_n = 1$, et obtenir

que dans tous ces cas : $V_n \sim n^{-2\gamma+\alpha}$ en $+\infty$. Alors, pour peu que $\alpha > 2\gamma$, le premier point du TCL pour les martingales est vérifié.

Reste à vérifier la condition de Lindeberg. Fixons $\varepsilon > 0$. On note $\psi_n := \frac{1}{V_n} \sum_{k=1}^n \mathbb{E} \left[(L_k - L_{k-1})^2 \mathbf{1}_{|L_k - L_{k-1}| \geq \varepsilon \sqrt{V_n}} | \mathcal{F}_{k-1} \right]$. L'indépendance des (W_k) justifie qu'on puisse remplacer les espérances conditionnelles de cette somme par des espérances. Encadrons alors ψ_n :

K étant bornée, on peut choisir $C > 0$ tel que $(L_k - L_{k-1})^2 \leq C \frac{c_k^2}{A_k^2} \left(\sum_{i=1}^k \frac{a_i}{h_i} \right)^2$.

D'où on peut déduire que :

$$0 \leq \psi_n \leq \frac{1}{V_n} \sum_{i=1}^k \frac{c_k^2}{A_k^2} v_k \left[\frac{C}{v_k} \left(\sum_{i=1}^k \frac{a_i}{h_i} \right)^2 \mathbb{P} \left(|L_k - L_{k-1}| \geq \varepsilon \sqrt{V_n} \right) \right] ps$$

Mais alors, d'après Toeplitz, il suffit de montrer que $\frac{C}{v_k} \left(\sum_{i=1}^k \frac{a_i}{h_i} \right)^2 \mathbb{P} \left(|L_k - L_{k-1}| \geq \varepsilon \sqrt{V_n} \right)$ tend vers 0.

Pour ce faire, remarquons que l'on peut écrire $L_k - L_{k-1} = c_k \left(\frac{M_k}{A_k} \right)$, avec $M_k = o(A_k)$ et $c_k = 1/k^\gamma$ pour obtenir que $\lim_{k \rightarrow +\infty} L_k - L_{k-1} = 0$. Alors, ayant supposé $\alpha > 2\gamma$, comme (V_n) est divergente, $\mathbb{P} \left(|L_k - L_{k-1}| \geq \varepsilon \sqrt{V_k} \right)$ est nulle à partir d'un certain rang.

Ainsi, la condition de Lindeberg est vérifiée et, au sens de la convergence en loi :

$$\frac{L_n}{\sqrt{V_n}} \Rightarrow \mathcal{N} \left(0, \xi^2 f(x) \right).$$

Du reste,

$$\frac{P_n}{\sqrt{V_n}} = \frac{C_n}{\sqrt{V_n}} \left[\frac{1}{C_n} \sum_{k=1}^n c_k \mathbb{E} \left(\hat{f}_k(x) - f(x) \right) \right].$$

Mais alors, par Toeplitz et convergence ps de $\hat{f}_k(x)$, et puisque $C_n \sim n^{-\gamma}$ et $\sqrt{V_n} \sim n^{-\gamma+\alpha/2}$, on a également, par produit, que :

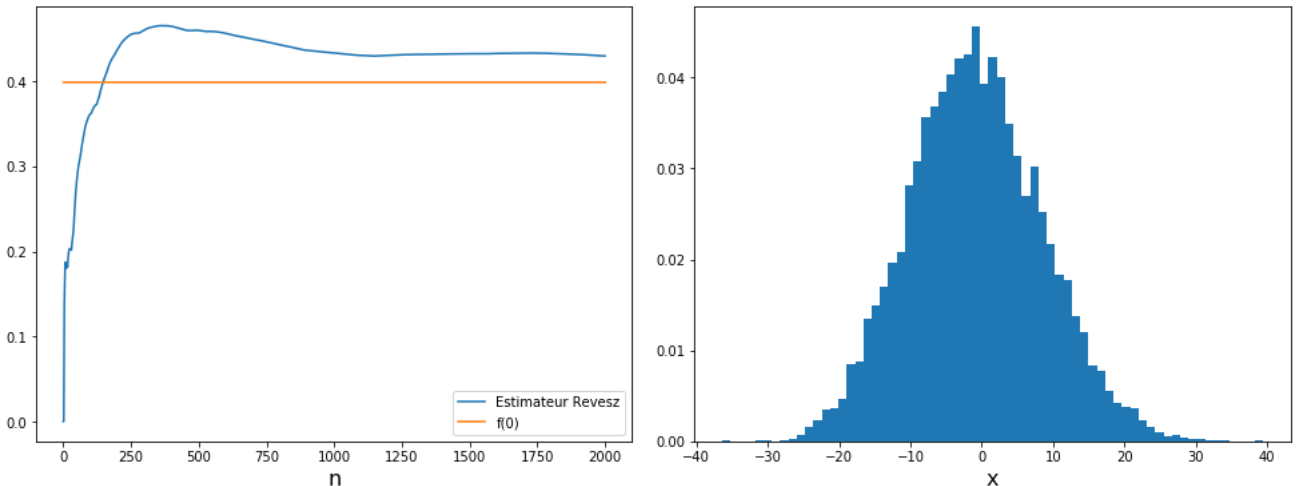
$$\lim_{n \rightarrow +\infty} \frac{P_n}{\sqrt{V_n}} = 0 \text{ ps.}$$

Finalement, on bien a montré que, au sens de la convergence en loi :

$$\frac{C_n}{\sqrt{V_n}} (\bar{f}_n(x) - f(x)) \Rightarrow \mathcal{N} \left(0, \xi^2 f(x) \right).$$

II. Simulations

Nous illustrons la convergence ps et la normalité asymptotique de la section précédente, en moyennisant l'estimateur de WWY, pour $\alpha = 3/5$ et $\gamma = 3/10$.



VII. BILAN

Récapitulons les convergences montrées et les conditions particulières à ces convergences. Nous avons montré que nos estimateurs de Deheuvels, Wolverton-Wagner-Yamato et Wegman-Davies, ainsi que l'estimateur de Revesz moyennisé sont convergents ps. Nous avons aussi montré des normalités asymptotiques pour ces estimateurs. Le tableau suivant regroupe les conditions sur α , respectivement pour les estimateurs de Deheuvels, Wolverton-Wagner-Yamato et Wegman-Davies, qui permettent d'assurer la convergence en loi :

$$\frac{A_n}{\sqrt{v_n}}(\widehat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x)).$$

Estimateur $\widehat{f}_n(x)$	Valeur de a_n	Condition sur α
Deheuvels : $\frac{1}{nh_n} \sum_{k=1}^n K\left(\frac{x-X_k}{h_k}\right)$	h_n	$1/3 < \alpha < 1$
WWY : $\frac{1}{n} \sum_{k=1}^n \frac{1}{h_k} K\left(\frac{x-X_k}{h_k}\right)$	1	$1/5 < \alpha < 1$
WD : $\frac{1}{nh_n^{1/2}} \sum_{k=1}^n \frac{1}{\sqrt{h_k}} K\left(\frac{x-X_k}{h_k}\right)$	$\sqrt{h_n}$	$1/4 < \alpha < 1$

En outre, pour peu que $\gamma < \alpha/2$, nous avons démontré la convergence en loi :

$$\frac{C_n}{\sqrt{V_n}}(\overline{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \xi^2 f(x)).$$

Nos simulations numériques corroborent ces résultats. Du reste, les vitesses de calcul constatées pour nos algorithmes récursifs permettent de classer Wolverton-Wagner-Yamato comme celui dont la vitesse d'exécution est la plus rapide.

VIII. BIBLIOGRAPHIE

1. B. Bercu D.Chafai. Modélisation stochastique et simulation. Version 2.5066282 (2007-07-09).
2. Yousri Slaoui. Application des methodes d'approximations stochastiques a l'estimation de la densite et de la regression. Mathematiques [math]. Universite de Versailles-Saint Quentin en Yvelines, 2006. Francais. tel-00131964