L'inégalité de Berry-Esséen par la méthode de Stein et une version quantitative du théorème central limite

Erik Thomas *

Résumé

Le but de cet article est de présenter une preuve de l'inégalité de Berry-Esséen en utilisant l'approche développée par Stein.

En particulier, nous résolvons l'équation de Stein pour les lois normales.

Mots clés : théorème central limite, inégalité de Berry-Esséen, équation de Stein

1 Introduction

Le théorème central limite est l'un des théorèmes les plus profonds des Mathématiques :

Théorème 1.1. *Théorème central limite.*

Si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires i.i.d. centrées (espérance nulle) et réduites (variance égale à 1), alors

$$\lim_{n\to+\infty} \frac{X_1+\cdots+X_n}{\sqrt{n}} \underset{=}{\text{loi}} \gamma_1 \tag{1}$$

où γ_1 désigne la loi normale centrée réduite. On notera dans toute la suite Φ la fonction de répartition de γ_1 : elle est donnée par

$$\forall x\in\mathbb{R},\quad \Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\mathrm{d}t.$$

Ainsi, si $(X_n)_{n\in\mathbb{N}^*}$ est une suite i.i.d. de variables aléatoires réelles centrées et réduites, la reformulation de (1) *est*

$$\forall x\in\mathbb{R},\quad \lim_{n\to+\infty}\mathbb{P}\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\leq x\right)=\Phi(x).$$

Ainsi, le théorème central limite permet d'approcher une somme de variables aléatoires i.i.d. par une loi normale. La preuve de ce résultat est très simple si l'on connnaît le résultat de Lévy qui assure que la convergence en loi est équivalente à la convergence simple des fonctions caractéristiques. Ainsi, il suffit de faire un développement limité de la fonction caractéristique de $\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k$. Nous renvoyons à [4] pour une preuve détaillée.

Le théorème central limite a cependant le problème de ne pas quantifier la vitesse de convergence, même en probabilité. L'inégalité de Berry-Esseén donne une réponse à ce problème moyennant une hypothèse supplémentaire sur la suite $X_1,\ldots,X_n$.

Théorème 1.2. *Inégalité de Berry-Esséen.*

Soit $n\in\mathbb{N}^*$ et soit $X_1,\ldots,X_n$ une suite de variables aléatoires i.i.d., centrées (d'espérance nulle) ayant toute un moment d'ordre 2 et telles que $\sum_{k=1}^{n}\mathbb{E}\left(X_k^2\right)=1$. Si l'on pose $W:=\sum_{k=1}^{n}X_k$, on a

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}(W\leq x)-\Phi(x)\right|\leq 7\sum_{k=1}^{n}\mathbb{E}\left|X_k\right|^3.$$

*erik.thomas@ens-rennes.fr

On notera que $\displaystyle\sum_{k=1}^{n}\mathbb{E}\left|X_k\right|^3 \in \mathbb{R}_+ \cup \{+\infty\}$.

Le théorème 1.2 permet de déduire facilement le corollaire suivant :

Corollaire 1.1. *Version quantifiée du théorème central limite.*

Soit $n \in \mathbb{N}^*$ *et soit* $X_1, \ldots, X_n$ *une suite de variables aléatoires i.i.d., centrées et réduites. On suppose qu'il existe* $\tau \geq 0$ *tel que pour tout* $i \in [\![1, n]\!]$, $\mathbb{E}\left|X_i\right|^3 \leq \tau$.

Alors, si $S_n := \dfrac{X_1 + \cdots + X_n}{\sqrt{n}}$, *on a :*

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(S_n \leq x\right) - \Phi\left(x\right)\right| \leq \frac{7\tau}{\sqrt{n}}.$$

Remarque 1. La constante obtenue ici n'est pas optimale. Dans [5], E. Gobet indique que la constante optimale est comprise entre 0,4097 et 0,4748. Par contre, la décroissance en $\sqrt{n}$ est optimale : il suffit de prendre pour les X_k des variables de Rademacher ($X_k(\Omega) = \{\pm 1\}$ et $\mathbb{P}(X_k = 1) = \mathbb{P}(X_k = -1) = 1/2$).

Nous prouvons le corollaire 1.1 à partir du théorème 1.2.

Démonstration. Soit $X_1, \ldots, X_n$ une suite i.i.d. de variables aléatoires centrées, réduites telles que $\mathbb{E}\left|X_k\right|^3 \leq \tau$ pour tout $k \in [\![1, n]\!]$.

On pose $\widetilde{X_k} := \dfrac{X_k}{\sqrt{n}}$. Pour tout $k \in [\![1, n]\!]$, on a $\mathbb{E}\left(\widetilde{X_k}\right) = 0$ et par indépendance, on a $\displaystyle\sum_{k=1}^{n}\mathbb{E}\left(\widetilde{X}_k^2\right) = 1$.

De plus,

$$\sum_{k=1}^{n}\mathbb{E}\left|\widetilde{X_k}\right|^3 = \frac{1}{n\sqrt{n}}\sum_{k=1}^{n}\mathbb{E}\left|X_k\right|^3 \leq \frac{\tau}{\sqrt{n}}.$$

L'inégalité de Berry-Esséen et la ligne précédente donnent alors

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(S_n \leq x\right) - \Phi\left(x\right)\right| \leq 7\sum_{k=1}^{n}\mathbb{E}\left|X_k\right|^3 \leq \frac{7\tau}{\sqrt{n}}.$$

□

2 Résolution de l'équation de Stein

Un point crucial de la preuve est la résolution de l'équation de Stein pour la loi normale. Le point important est que si X est une variable aléatoire réelle, alors

$$\mathbb{E}\left(f'\left(X\right)\right) = \mathbb{E}\left(Xf\left(X\right)\right)$$

pour toute fonction f de classe $\mathscr{C}^1$ telle que l'un des membres ait un sens (nombre fini ou infini), alors X suit la loi normale centrée réduite. On renvoie à [6].

On introduit alors l'opérateur S sur $\mathscr{C}^1(\mathbb{R})$ par

$$\forall f \in \mathscr{C}^1(\mathbb{R}), \quad S(f) := f' - xf.$$

On introduit aussi l'ensemble $\mathscr{C}^1_{\mathrm{cb}}(\mathbb{R}) := \left\{f : \mathbb{R} \longrightarrow \mathbb{R},\ f \text{ est continue et de classe } \mathscr{C}^1 \text{ par morceaux}\right\}$.

Lemme 2.1. *Résolution de l'équation de Stein.*

Soit $f \in \mathscr{C}^1_{\mathrm{cb}}(\mathbb{R})$ *avec* $f' \in L^1(\mathbb{R})$. *On pose* $N(f) := \displaystyle\int_{\mathbb{R}} f\,\mathrm{d}\gamma_1$.

Alors, il existe $h \in \mathscr{C}^1$ *telle que* $S(h) = f - N(f)$. *De plus,* h *est donnée par*

$$\forall x \in \mathbb{R}, \quad h(x) = \mathrm{e}^{x^2/2}\int_{-\infty}^{x}\left(f - N(f)\right)(t)\,\mathrm{e}^{-t^2/2}\mathrm{d}t.$$

De plus, cette solution h *vérifie*

$$\left\|h\right\|_\infty^{\mathbb{R}} \leq \sqrt{\frac{\pi}{8}}\left\|f'\right\|_{L^1(\mathbb{R})} \quad \text{et} \quad \left\|h'\right\|_\infty^{\mathbb{R}} \leq \left\|f'\right\|_{L^1(\mathbb{R})}.$$

Démonstration. Comme $f' \in L^1(\mathbb{R})$, on remarque que f est bornée, ainsi $N(f)$ est bien définie.

Il est facile de vérifier que

$$\forall x \in \mathbb{R}, \quad h'(x) = xh'(x) + f(x) - N(f).$$

On en déduit donc que $S(h) = f - N(f)$.

Il reste à vérifier les estimées annoncées.

- *Première inégalité.*
  On commence par écrire

$$f(t) = -\int_t^{+\infty} f'(y)\,\mathrm{d}y = -\int_{\mathbb{R}} \mathbb{1}_{]-\infty,y]}(t)\, f'(y)\,\mathrm{d}y$$

car $\mathbb{1}_{]-\infty,y]}(t) = \mathbb{1}_{[t,+\infty[}(y)$.
Une intégration par parties donne

$$-N(f) = -\int_{\mathbb{R}} f(y)\,\mathrm{d}\gamma_1(y) = \int_{\mathbb{R}} f'(y)\,\Phi(y)\,\mathrm{d}y.$$

On en déduit alors que

$$(f - N(f))(t) = -\int_{\mathbb{R}} \psi_y(x)\, f'(y)\,\mathrm{d}y$$

où l'on a posé $\psi_y(t) := \mathbb{1}_{]-\infty,y]}(x) - \Phi(y)$. Il s'ensuit que

$$\begin{aligned}
h(x) &= \mathrm{e}^{x^2/2} \int_{-\infty}^{x} (f - N(f))(t)\,\mathrm{d}t \\
&= \mathrm{e}^{x^2/2} \int_{-\infty}^{x} \left(-\int_{\mathbb{R}} \psi_y(t)\, f'(y)\,\mathrm{d}y \right) \mathrm{e}^{-t^2/2}\mathrm{d}t \\
&= -\mathrm{e}^{x^2/2} \int_{-\infty}^{x} \left(\int_{\mathbb{R}} \psi_y(t)\, f'(y)\, \mathrm{e}^{-t^2/2}\mathrm{d}y \right) \mathrm{d}t.
\end{aligned}$$

Comme $f' \in L^1(\mathbb{R})$, $y \longmapsto \mathrm{e}^{-y^2/2} \in L^1(\mathbb{R})$ et $(y,t) \in \mathbb{R}^2 \longmapsto \psi_y(t)$ est bornée sur $\mathbb{R}^2$, on a $(y,t) \in \mathbb{R}^2 \longmapsto \psi_y(t)\, f'(y)\,\mathrm{e}^{-t^2/2} \in L^1(\mathbb{R}^2)$. Par le théorème de Fubini, on a :

$$\begin{aligned}
h(x) &= -\mathrm{e}^{x^2/2} \int_{\mathbb{R}} \left(\int_{-\infty}^{x} \psi_y(t)\, f'(y)\, \mathrm{e}^{-t^2/2}\mathrm{d}t \right) \mathrm{d}y \\
&= -\mathrm{e}^{x^2/2} \int_{\mathbb{R}} f'(y) \left(\int_{-\infty}^{x} \left(\mathbb{1}_{]-\infty,y]}(t) - \Phi(y) \right) \mathrm{e}^{-t^2/2}\mathrm{d}t \right) \mathrm{d}y \\
&= -\mathrm{e}^{x^2/2} \int_{\mathbb{R}} f'(y) \left(\int_{-\infty}^{\min\{x,y\}} \mathrm{e}^{-t^2/2}\mathrm{d}t - \Phi(y) \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\mathrm{d}t \right) \mathrm{d}y \\
&= -\mathrm{e}^{x^2/2}\sqrt{2\pi} \int_{\mathbb{R}} f'(y) \left(\Phi(\min\{x,y\}) - \Phi(x)\,\Phi(y) \right) \mathrm{d}y \\
&= -\int_{\mathbb{R}} f'(y)\, k_y(x)\,\mathrm{d}y
\end{aligned}$$

où l'on a posé $k_y(x) := \sqrt{2\pi}\mathrm{e}^{x^2/2} \left(\Phi(\min\{x,y\}) - \Phi(x)\,\Phi(y) \right)$.
On remarque que comme Φ est croissante sur $\mathbb{R}$ et à valeurs dans $]0,1[$, on a $k_y \geq 0$. Toujours en utilisant la croissance de Φ, on a :

$$\Phi(\min\{x,y\}) - \Phi(x)\,\Phi(y) \leq \Phi(x) - \Phi(x)^2 = \frac{1}{4}\left(1 - 4\left(\Phi(x) - \frac{1}{2} \right)^2 \right).$$

Or,

$$\left(\Phi(x) - \frac{1}{2} \right)^2 = \left(\int_0^{|x|} \mathrm{e}^{-t^2/2} \frac{\mathrm{d}t}{\sqrt{2\pi}} \right)^2 \geq \frac{1}{2\pi} \times \frac{1}{4} \int_{u^2+v^2 \leq x^2} \mathrm{e}^{-\left(u^2+v^2\right)}\mathrm{d}u\mathrm{d}v.$$

Le changement de variable en coordonnées polaires donne

$$\int_{u^2+v^2\leq x^2} \mathrm{e}^{-\left(u^2+v^2\right)}\mathrm{d}u\mathrm{d}v = \int_0^{2\pi}\int_0^{|x|} r\mathrm{e}^{-r^2}\mathrm{d}r\mathrm{d}\theta = 2\pi\left(1-\mathrm{e}^{-x^2/2}\right).$$

On en déduit que

$$k_y(x) \leq \sqrt{2\pi}\mathrm{e}^{x^2/2}\times\frac{1}{4}\left(1-4\left(\frac{1}{2\pi}\times\frac{1}{4}\times 2\pi\left(1-\mathrm{e}^{-x^2/2}\right)\right)\right) = \sqrt{\frac{\pi}{8}}.$$

Finalement, on a bien $\|h\|_\infty^{\mathbb{R}} \leq \sqrt{\frac{\pi}{8}}\,\|f'\|_{L^1(\mathbb{R})}$.

- *Seconde inégalité.*
Comme $h'(x) = xh(x) - (f-N(f))(x)$, on a donc

$$\begin{aligned} h'(x) &= -\int_{\mathbb{R}} \left(xk_y(x)+\psi_y(x)\right) f'(y)\,\mathrm{d}y \\ &= -\int_{\mathbb{R}} \left(\sqrt{2\pi}x\mathrm{e}^{-x^2/2}\left(\Phi(\min\{x,y\}) - \Phi(x)\Phi(y)\right) + \mathbb{1}_{]-\infty,y]}(x) - \Phi(y)\right) f'(y)\,\mathrm{d}y \\ &= -\int_{\mathbb{R}} h_y(x) f'(y)\,\mathrm{d}y \end{aligned}$$

où l'on a posé

$$h_y(x) := \sqrt{2\pi}x\mathrm{e}^{-x^2/2}\left(\Phi(\min\{x,y\}) - \Phi(x)\Phi(y)\right) + \mathbb{1}_{]-\infty,y]}(x) - \Phi(y).$$

On distingue quatre cas :

  - ⋆ Si $x>0$. Déjà, une intégration par parties donne

$$\begin{aligned} 1-\Phi(x) &= \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} \mathrm{e}^{-t^2/2}\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} \frac{1}{t}\times t\mathrm{e}^{-t^2/2}\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}}\left(\left[-\frac{\mathrm{e}^{-x^2/2}}{2}\right]_x^{+\infty} - \int_x^{+\infty}\frac{\mathrm{e}^{-t^2/2}}{t^2}\mathrm{d}t\right) \\ &\leq \frac{\mathrm{e}^{-x^2/2}}{x\sqrt{2\pi}}. \end{aligned}$$

Pour estimer $h'(x)$, on distingue deux cas :

i) Si $y\geq x>0$. Dans ce cas, en utilisant la croissance de Φ, on a

$$\begin{aligned} h_y(x) &= \sqrt{2\pi}x\mathrm{e}^{-x^2/2}\left(1-\Phi(y)\right)\Phi(x) + 1 - \Phi(y) \\ &\leq \sqrt{2\pi}x\mathrm{e}^{-x^2/2}\left(1-\Phi(x)\right)\Phi(x) + 1 - \Phi(y). \end{aligned}$$

Or, $0\leq (1-\Phi(x)) \leq \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}x}$ et $1-\Phi(x)\leq 1-\Phi(y)\leq 1$, on en déduit donc que

$$0\leq h_y(x)\leq 1.$$

ii) Si $0\leq y\leq x$. Dans ce cas, on a

$$h_y(x) = \sqrt{2\pi}x\mathrm{e}^{x^2/2}\left(1-\Phi(x)\right)\Phi(y) - \Phi(y).$$

Comme $0\leq 1-\Phi(x)\leq \frac{x\sqrt{2\pi}}{\mathrm{e}^{-x^2/2}}$, on a $0\leq\sqrt{2\pi}x\mathrm{e}^{x^2/2}\left(1-\Phi(x)\right)\Phi(y)\leq\Phi(y)$, d'où

$$-1\leq h_y(x)\leq 0.$$

⋆ Si $x \leq 0$.
On traite ce cas de manière analogue au premier cas : on montre que $|h_y(x)| \leq 1$.

On en déduit que
$$\forall x \in \mathbb{R}, \quad \left|h'(x)\right| \leq \int_{\mathbb{R}} |h_y(x)| \left|f'(y)\right| \mathrm{d}y \leq \|f\|_{L^1(\mathbb{R})}.$$

□

3 Preuve de l'inégalité de Berry-Esséen

On note K_n la meilleure constante dans l'inégalité de Berry-Esséen, i.e. la plus petite constante telle que pour toute suite de variables aléatoires i.i.d. $X_1, \ldots, X_n$ centrées et telles que $\sum_{i=1}^{n} \mathbb{E}\left(X_i^2\right) = 1$
$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq x\right) - \Phi(x)\right| \leq K_n \sum_{i=1}^{n} \mathbb{E}|X_i|^3.$$

Remarquons que K_n est fini. En effet si l'on suppose $\sum_{i=1}^{n} \mathbb{E}|X_i|^3 < +\infty$ (sinon l'inégalité est claire), alors en utilisant l'inégalité de Hölder, on a
$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq x\right) - \Phi(x)\right| \leq 2 = 2\left(\sum_{i=1}^{n} \mathbb{E}\left(X_i^2\right)\right)^{3/2} \leq 2\sqrt{n} \sum_{i=1}^{n} \mathbb{E}|X_i|^3.$$
Ainsi, K_n est fini et $K_n \leq 2\sqrt{n}$.

Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d. centrées telles que $\sum_{i=1}^{n} \mathbb{E}\left(X_i^2\right)$. On pose :
$$\sigma_i^2 = \mathbb{E}\left(X_i^2\right), \quad W = \sum_{i=1}^{n} X_i \quad \text{et} \quad W_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} X_j = W - X_i$$
de sorte que l'on ait
$$\sum_{i=1}^{n} \sigma_i^2 = 1, \quad \mathbb{E}(W) = 0 \quad \text{et} \quad \operatorname{Var}(W) = 1.$$
On notera que W_i et X_i sont indépendantes. Ainsi, on estime
$$\sup_{x\in\mathbb{R}} |\mathbb{P}(W \leq x) - \Phi(x)|.$$
Si l'on pose $f := \mathbb{1}_{]-\infty,x]}$, on remarque que
$$\mathbb{P}(W \leq x) = \mathbb{E}(f(W)) \quad \text{et} \quad N(f) := \int_{\mathbb{R}} f \mathrm{d}\gamma_1 = \Phi(x), \tag{2}$$
donc
$$\mathbb{P}(W \leq x) - \Phi(x) = \mathbb{E}(f(W)) - N(f) = \mathbb{E}((f - N(f))(W)).$$
On aimerait utiliser le lemme 2.1 pour f, mais f n'est pas continue. On approche alors f par des fonctions continues. Soit $\varepsilon > 0$ fixé. Pour $x \in \mathbb{R}$, on introduit la fonction $f_{x,\varepsilon}$ par :
$$\forall x \in \mathbb{R}, \quad f_{x,\varepsilon}(y) := \begin{cases} 1 & \text{si } y \leq x - \varepsilon \\ -\frac{1}{2\varepsilon}(y - x - \varepsilon) & \text{si } y \in [x-\varepsilon, x+\varepsilon] \\ 0 & \text{si } y \geq x + \varepsilon \end{cases}.$$

On note que $f_{x,\varepsilon} \in \mathscr{C}^1_{\text{cb}}(\mathbb{R})$ avec $f'_{x,\varepsilon} = \mathbb{1}_{]x-\varepsilon,x+\varepsilon[}$ sur $\mathbb{R}\setminus\{x \pm \varepsilon\}$. En particulier, $\left\|f'_{x,\varepsilon}\right\|_{L^1(\mathbb{R})} = 1$. D'après le lemme 2.1, il existe $h_{x,\varepsilon} \in \mathscr{C}^1(\mathbb{R})$ telle que $S(h_{x,\varepsilon}) = f_{x,\varepsilon} - N(f_{x,\varepsilon})$. $h_{x,\varepsilon}$ vérifie

$$\|h_{x,\varepsilon}\|_\infty^{\mathbb{R}} \leq \sqrt{\frac{\pi}{8}} \quad \text{et} \quad \left\|h'_{x,\varepsilon}\right\|_\infty^{\mathbb{R}} \leq 1.$$

Comme $f \leq \mathbb{1}_{x+\varepsilon,\varepsilon}$, on en déduit

$$\mathbb{P}(W \leq x) \leq \mathbb{E}(f_{x+\varepsilon,\varepsilon}(W)) \tag{3}$$

De plus, comme Φ est lipschitzienne de rapport $\frac{1}{\sqrt{2\pi}}$ car $0 \leq \Phi'(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2} \leq \frac{1}{\sqrt{2\pi}}$, ainsi

$$\Phi(x) \geq \Phi(x+\varepsilon) - \frac{\varepsilon}{\sqrt{2\pi}}. \tag{4}$$

Les lignes (2), (3) et (4) donnent

$$\mathbb{P}(W \leq x) - \Phi(x) \leq \mathbb{E}(f_{x+\varepsilon,\varepsilon}(W)) - N(f) + \frac{\varepsilon}{\sqrt{2\pi}} \tag{5}$$

Un raisonnement analogue donne

$$\mathbb{P}(W \leq x) - \Phi(x) \geq \mathbb{E}(f_{x-\varepsilon,\varepsilon}(W)) - N(f) - \frac{\varepsilon}{\sqrt{2\pi}} \tag{6}$$

L'idée de la preuve du théorème de Berry-Esséen est alors d'établir une inégalité entre K_n et K_{n-1}. Le lemme suivant est fondamental en ce sens :

Lemme 3.1. *Lemme technique.*

Soit $f := f_{z_0,\varepsilon}$ pour un certain $z_0 \in \mathbb{R}$ et $\varepsilon > 0$. Si l'on pose $\beta := \sum_{i=1}^{n} \mathbb{E}|X_i|^3$, on a

$$|\mathbb{E}(f(W)) - N(f)| \leq \left(5 + \frac{3\beta K_{n-1}}{2\varepsilon}\right)\beta.$$

Démonstration. Soit φ donnée par le lemme 2.1 : on a $S(\varphi) = f - N(f)$. On a $\|\varphi\|_\infty^{\mathbb{R}} \leq \sqrt{\frac{\pi}{8}}$ et $\|\varphi'\|_\infty^{\mathbb{R}} \leq 1$.

Soit $\widetilde{\varphi}(x) := x\varphi(x)$. Il est clair que

$$\forall (x,y) \in \mathbb{R}^2, \quad \widetilde{\varphi}(x) - \widetilde{\varphi}(y) = \varphi(x)(x-y) + (\varphi(x) - \varphi(y))\,y.$$

L'inégalité des accroissements finis et les majorations de $\|\varphi\|_\infty^{\mathbb{R}}$ et $\|\varphi'\|_\infty^{\mathbb{R}}$ établies au lemme 2.1 donnent :

$$\forall (x,y) \in \mathbb{R}^2, \quad |\widetilde{\varphi}(x) - \widetilde{\varphi}(y)| \leq \left(\sqrt{\frac{\pi}{8}} + |y|\right)|x-y|.$$

On a donc :

$$\begin{aligned}
\mathbb{E}(f(W)) - N(f) &= \mathbb{E}\left(\varphi'(W) - \varphi(W)\,W\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left(\varphi'(W)\,\sigma_i^2 - \varphi(W)\,X_i\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left(\varphi'(W)\,\sigma_i^2 - \varphi(W_i)\,X_i - (\varphi(W) - \varphi(W_i))\,X_i\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left(\varphi'(W)\,\sigma_i^2 - (\varphi(W) - \varphi(W_i))\,X_i\right)
\end{aligned}$$

où la dernière égalité résulte du fait que $\mathbb{E}(\varphi(W_i)\,X_i) = \mathbb{E}(\varphi(W_i))\,\mathbb{E}(X_i) = 0$ par indépendance.

Soit θ suivant la loi uniforme sur $[0,1]$ indépendante de $X_1,\ldots,X_n$. Comme $\varphi\in\mathscr{C}^1_{\mathrm{cb}}(\mathbb{R})$, on a

$$\begin{aligned}\varphi(W)-\varphi(W_i) &= (W-W_i)\int_0^1\varphi'(W_i+t(W-W_i))\,\mathrm{d}t\\ &= X_i\int_0^1\varphi'(W_i+tX_i)\,\mathrm{d}t\\ &= \mathbb{E}\left(\varphi'(W_i+\theta X_i)\right).\end{aligned}$$

On en déduit que

$$\mathbb{E}(f(W))-N(f)=\sum_{i=1}^n\mathbb{E}\left(\varphi'(W)\,\sigma_i^2-\varphi'(W_i+\theta X_i)\,X_i^1\right).$$

Pour tout $g\in\mathscr{C}^1_{\mathrm{cb}}(\mathbb{R})$, on pose

$$R_i(g):=\mathbb{E}\left(g(W)\,X_i^2-g(W_i+\theta X_i)\,X_i^2\right).$$

Ainsi, comme $S(\varphi)=\varphi'-x\varphi=f-N(f)$, on a $\varphi'(x)=\widetilde{\varphi}(x)+f(x)-N(f)$, on a alors

$$\mathbb{E}(f(W))-N(f)=\sum_{i=1}^n\left(R_i(\widetilde{\varphi})+R_i(f)\right).\tag{7}$$

Si l'on pose, pour $g\in\mathscr{C}^1_{\mathrm{cb}}(\mathbb{R})$,

$$R_{i,1}(g):=\mathbb{E}\left((g(W)-g(W_i))\,\sigma_i^2\right)\quad\text{et}\quad R_{i,2}(g):=\mathbb{E}\left((g(W_i+\theta X_i)-g(W_i))\,X_i^2\right),\tag{8}$$

en utilisant encore l'indépendance de W_i et X_i, on a

$$R_i(g)=R_{i,1}(g)-R_{i,2}(g).$$

Nous allons donc majorer $R_i(\widetilde{\varphi})$ et $R_i(f)$.

- *Majoration de* $R_i(\widetilde{\varphi})$.
  L'inégalité (7) utilisée avec $(x,y):=(W,W_i)$ et $(x,y):=(W_i+\theta W_i,W_i)$ donne

  $$|\widetilde{\varphi}(W)-\widetilde{\varphi}(W_i)|\leq\left(\sqrt{\frac{\pi}{8}}+|W_i|\right)|W-W_i|=\left(\sqrt{\frac{\pi}{8}}+|W_i|\right)|X_i|\tag{9}$$

  et

  $$|\widetilde{\varphi}(W)-\widetilde{\varphi}(W_i+\theta X_i)|\leq\left(\sqrt{\frac{\pi}{8}}+|W_i|\right)|\theta X_i|.\tag{10}$$

  On obtient donc

  $$\begin{aligned}|R_i(\widetilde{\varphi})| &= \left|\mathbb{E}(R_{i,1}(\widetilde{\varphi}))-\mathbb{E}(R_{i,2}(\widetilde{\varphi}))\right|\\ &\leq \left|\mathbb{E}(R_{i,1}(\widetilde{\varphi}))\right|+\left|\mathbb{E}(R_{i,2}(\widetilde{\varphi}))\right|\\ &\leq \mathbb{E}\left(|\widetilde{\varphi}(W)-\widetilde{\varphi}(W_i)|\,\sigma_i^2\right)+\mathbb{E}\left(|\widetilde{\varphi}(W_i+\theta X_i)-\widetilde{\varphi}(W_i)|\,X_i^2\right)\end{aligned}$$

  En utilisant l'indépendance des variables aléatoires X_i, W_i et θ, le fait que $\mathbb{E}|\theta|=1/2$ et les inégalités (9) et (10), on a

  $$|R_i(\widetilde{\varphi})|\leq\left(\sqrt{\frac{\pi}{8}}+\mathbb{E}|W_i|\right)\sigma_i^2\,\mathbb{E}|W_i|+\frac{1}{2}\left(\sqrt{\frac{\pi}{8}}+\mathbb{E}|W_i|\right)\mathbb{E}|X_i|^3.$$

  L'inégalité de Hölder donne $\mathbb{E}|W_i|\leq\mathbb{E}\left(W_i^2\right)^{1/2}$. Or, $\mathbb{E}\left(W_i^2\right)=\sum_{\substack{j=1\\j\neq i}}^n\sigma_j^2\leq\sum_{j=1}^n\sigma_j^2=\mathbb{E}\left(W^2\right)=1$, il s'ensuit que $\mathbb{E}|W_i|\leq 1$.
  De plus, l'inégalité de Hölder assure encore que

  $$\sigma_i^2=\mathbb{E}\left(X_i^2\right)\leq(\mathbb{E}|X_i|)^2.$$

L'inégalité de Jensen utilisée avec la fonction convexe $x \in \mathbb{R}_+ \longmapsto x^3$ assure alors que $(\mathbb{E}\,|X_i|)^3 \leq \mathbb{E}\,|X_i|^3$, ce qui donne

$$\sigma_i^2\,\mathbb{E}\,|X_i| \leq \mathbb{E}\,|X_i|^3\,. \tag{11}$$

On récupère finalement

$$|R_i(\widetilde{\varphi})| \leq \frac{3}{2}\left(\sqrt{\frac{\pi}{8}}+1\right)\mathbb{E}\,|X_i|^3\,. \tag{12}$$

- *Majoration de $R_i(f)$.*
On commence par remarquer que comme $0 \leq f \leq 1$, on a

$$|R_i(f)| \leq \sigma_i^2. \tag{13}$$

Pour s'en convaincre, il suffit d'écrire

$$R_i(f) = \mathbb{E}\left(f(W)\,X_i^2 - f(W_i+\theta X_i)\,X_i^2\right) = \mathbb{E}\left(\left(f(W)-\frac{1}{2}\right)\sigma_i^2\right) - \mathbb{E}\left(\left(f(W_i+\theta X_i)-\frac{1}{2}\right)X_i^2\right)$$

et constater que $\left|f(W)-\frac{1}{2}\right| \leq \frac{1}{2}$ et $\left|f(W_i+\theta X_i)-\frac{1}{2}\right| \leq \frac{1}{2}$. Comme $f = f_{z_0}$, pour tout $x, y \in \mathbb{R}$, on a

$$f(x) - f(y) = -\frac{1}{2\varepsilon}\mathbb{1}_{[z_0-\varepsilon, z_0, \varepsilon]}\left(y+\theta_2(y-x)\right)(x-y) \tag{14}$$

où $\theta_2 \hookrightarrow \mathscr{U}([0,1])$ et est indépendante de $X_1, \ldots, X_n$ et θ.
On utilise (14) avec $(x,y) := (W = W_i + X_i, W_i)$ et $(x,y) := (W_i + \theta X_i, W_i)$ pour obtenir

$$f(W) - f(W_i) = -\frac{1}{2\varepsilon}\mathbb{1}_{[z_0-\varepsilon, z_0+\varepsilon]}(W_i+\theta_2 X_i)\,X_i \tag{15}$$

et

$$f(W_i+\theta X_i) - f(W_i) = -\frac{1}{2\varepsilon}\mathbb{1}_{[z_0-\varepsilon, z_0+\varepsilon]}(W_i+\theta_2\theta X_i)\,\theta X_i \tag{16}$$

En utilisant les définitions de $R_{i,1}(f)$ et $R_{i,2}(f)$ données à ligne (8), on a :

$$\left|R_{i,1}(f)\right| = \left|\mathbb{E}\left((f(W)-f(W_i))\,\sigma_i^2\right)\right| = \left|\mathbb{E}\left(-\frac{1}{2\varepsilon}\mathbb{1}_{[z_0-\varepsilon, z_0+\varepsilon]}(W_i+\theta_2 X_i)\,X_i\sigma_i^2\right)\right|$$

et

$$\left|R_{i,2}(f)\right| = \left|\mathbb{E}\left((f(W_i+\theta X_i)-g(W_i))\,X_i^2\right)\right| = \left|\mathbb{E}\left(-\frac{1}{2\varepsilon}\mathbb{1}_{[z_0-\varepsilon, z_0+\varepsilon]}(W_i+\theta_2\theta X_i)\,\theta X_i^3\right)\right|.$$

En calculant les espérances conditionnelles par rapport aux variables aléatoires indépendantes X_j $(j \neq i)$, X_i, θ et θ_2, on obtient :

$$\left|R_{i,1}(f)\right| \leq \frac{1}{2\varepsilon}\,\mathbb{E}_{X_i,\theta_2}\left(\mathbb{P}_{X_i,\theta_2}(z_0-\varepsilon \leq W_i+\theta_2 X_i \leq z_0+\varepsilon)\,\sigma_i^2\,|X_i|\right) \tag{17}$$

et

$$\left|R_{i,2}(f)\right| \leq \frac{1}{2\varepsilon}\,\mathbb{E}_{X_i,\theta,\theta_2}\left(\mathbb{P}_{X_i,\theta,\theta_2}(z_0-\varepsilon \leq W_i+\theta_2\theta X_i \leq z_0+\varepsilon)\,|X_i|^3\,|\theta|\right) \tag{18}$$

Par suite, il suffit d'estimer $\mathbb{P}(a \leq W_i \leq b)$ où a et b sont fixés. On pose $\alpha_i := \mathbb{E}\left(W_i^2\right) = 1-\sigma_i^2$. On a

$$\begin{aligned}
\mathbb{P}(a \leq W_i \leq b) &= \mathbb{P}\left(\frac{W_i}{\alpha_i} \leq \frac{b}{\alpha_i}\right) - \mathbb{P}\left(\frac{W_i}{\alpha_i} < \frac{a}{\alpha_i}\right) \\
&= \left(\mathbb{P}\left(\frac{W_i}{\alpha_i} \leq \frac{b}{\alpha_i}\right) - \Phi\left(\frac{b}{\alpha_i}\right)\right) - \left(\mathbb{P}\left(\frac{W_i}{\alpha_i} < \frac{a}{\alpha_i}\right) - \Phi\left(\frac{a}{\alpha_i}\right)\right) + \left(\Phi\left(\frac{b}{\alpha_i}\right) - \Phi\left(\frac{a}{\alpha_i}\right)\right).
\end{aligned}$$

En remarquant que Φ est lipschitzienne de rapport $\frac{1}{\sqrt{2\pi}}$ (car pour tout $x \in \mathbb{R}$, $|\Phi'(x)| = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2} \leq \frac{1}{\sqrt{2\pi}}$), on en déduit que

$$\mathbb{P}(a \leq W_i \leq b) \leq \frac{b-a}{\alpha_i\sqrt{2\pi}} + 2\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{W_i}{\alpha_i} \leq x\right) - \Phi(x)\right| \tag{19}$$

Si pour tout $j \in [\![1,n]\!] \setminus \{i\}$, on pose $\widetilde{X_j} = \frac{X_j}{\alpha_i}$, alors on a $W_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \widetilde{X}_j$ avec

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} \mathbb{E}\left(\widetilde{X_j}^2\right) = \frac{1}{\alpha_i^2} \sum_{\substack{j=1 \\ j \neq i}} \mathbb{E}\left(X_j^2\right) = 1,$$

l'inégalité de Berry-Esséen appliquée à l'ordre $n-1$ donne

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(\frac{W_i}{\alpha_i} \leq x\right) - \Phi(x) \right| \leq \frac{K_{n-1}}{\alpha_i^3} \sum_{\substack{j=1 \\ j \neq i}}^{n} \mathbb{E}\left|X_i\right|^3 \leq \frac{K_{n-1}}{\alpha_i^3} \sum_{j=1}^{n} \mathbb{E}\left|X_j\right|^3 .$$

On pose $\beta := \sum_{j=1}^{n} \mathbb{E}\left|X_j\right|^3$. La ligne précédente utilisée avec la ligne (19) donne

$$\mathbb{P}\left(a \leq W_i \leq b\right) \leq \frac{b-a}{\alpha_i \sqrt{2\pi}} + \frac{K_{n-1}\beta}{\alpha_i^3}. \tag{20}$$

La ligne (20) utilisée avec la ligne (17) donne

$$\begin{aligned} \left|R_{i,1}(f)\right| &\leq \frac{1}{2\varepsilon} \mathbb{E}_{X_i,\theta_2}\left(\frac{2\varepsilon}{\alpha_i\sqrt{2\pi}} + 2K_{n-1}\frac{\beta}{\alpha_i^3}\sigma_i^2\left|X_i\right|\right) \\ &\leq \frac{\sigma_i^2\,\mathbb{E}\left|X_i\right|}{\alpha_i\sqrt{2\pi}} + \frac{K_{n-1}\beta\sigma_i^2\,\mathbb{E}\left|X_i\right|}{\varepsilon\alpha_i^3} \\ &\leq \left(\frac{1}{\alpha_i\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon\alpha_i^3}\right)\mathbb{E}\left|X_i\right|^3 \end{aligned}$$

où l'on a utilisé la majoration $\sigma_i^2\,\mathbb{E}\left|X_i\right| \leq \mathbb{E}\left|X_i\right|^3$ déjà prouvée à la ligne (11).
De même, on a

$$\begin{aligned} \left|R_{i,2}(f)\right| &\leq \frac{1}{2\varepsilon} \mathbb{E}_{X_i,\theta,\theta_2}\left(\left(\frac{2\varepsilon}{\alpha_i\sqrt{2\pi}} + \frac{2K_{n-1}\beta}{\alpha_i^3}\right)\left|X_i\right|^3\left|\theta\right|\right) \\ &\leq \frac{1}{2\varepsilon}\left(\frac{2\varepsilon}{\alpha_i\sqrt{2\pi}} + \frac{2K_{n-1}\beta}{\alpha_i^3\varepsilon}\right)\mathbb{E}\left|X_i\right|^3 \times \frac{1}{2} \\ &\leq \left(\frac{1}{2\alpha_i\sqrt{2\pi}} + \frac{1}{2} \times \frac{K_{n-1}\beta}{\alpha_i^3\varepsilon}\right)\mathbb{E}\left|X_i\right|^3 . \end{aligned}$$

On en déduit que

$$\begin{aligned} \left|R_i(f)\right| &\leq \left|R_{i,1}(f)\right| + \left|R_{i,2}(f)\right| \\ &\leq \frac{3}{2}\left(\frac{1}{\alpha_i\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\alpha_i^3\varepsilon}\right)\mathbb{E}\left|X_i\right|^3 \\ &\leq \frac{3}{2}\left(\frac{\alpha_i^2}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\alpha_i^{-3}\,\mathbb{E}\left|X_i\right|^3 \\ &\leq \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\alpha_i^{-3}\,\mathbb{E}\left|X_i\right|^3 \end{aligned}$$

où, dans la dernière inégalité, on a utilisé le fait que $\alpha_i^2 = 1 - \sigma_i^2 \leq 1$. En utilisant cette précédente inégalité avec (13) pour obtenir

$$\left|R_i(f)\right| \leq \min\left\{\sigma_i^2, \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\alpha_i^{-3}\,\mathbb{E}\left|X_i\right|^3\right\} = \min\left\{\sigma_i^2, \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\left(1-\sigma_i^2\right)^{-3/2}\mathbb{E}\left|X_i\right|^3\right\} \tag{21}$$

On admet provisoirement l'inégalité suivante (prouvée à la fin de l'article pour ne pas alourdir)

$$\forall (a,\sigma) \in [0,1]^2, \ \forall x \geq 0, \quad \min\left\{a, x\left(1-\sigma^2\right)^{-3/2}\right\} \leq x + \sqrt{\frac{3}{2}}a\sigma. \tag{22}$$

L'inégalité (22) utilisée avec $a := \sigma_i^2$, $a := \sigma_i$ et $x := \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\mathbb{E}\,|X_i|^3$ dans (21) donne

$$|R_i(f)| \leq \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon}\right)\mathbb{E}\,|X_i|^3 + \sqrt{\frac{3}{2}}\sigma_i^3. \tag{23}$$

L'inégalité de Hölder donne

$$\sigma_i^2 = \mathbb{E}\left(X_i^2\right) = \mathbb{E}\left(1 \times X_i^2\right) \leq \mathbb{E}\left(1^3\right)^{1/3}\mathbb{E}\left(|X_i|^3\right)^{2/3}.$$

En particulier, $\sigma_i^3 \leq \mathbb{E}\,|X_i|^3$, ainsi l'inégalité (23) devient

$$|R_i(f)| \leq \frac{3}{2}\left(\frac{1}{\sqrt{2\pi}} + \frac{K_{n-1}\beta}{\varepsilon} + 1\right)\mathbb{E}\,|X_i|^3 \tag{24}$$

En utilisant les majorations (12) et (24) dans (7), on a

$$|\mathbb{E}(f(W)) - N(f)| \leq \left(C + \frac{3}{2} \times \frac{K_{n-1}\beta}{\varepsilon}\right)\sum_{i=1}^{n}\mathbb{E}\,|X_i|^3 \tag{25}$$

où $C := \frac{3}{2}\left(\sqrt{\frac{\pi}{8}} + 2 + \frac{1}{\sqrt{2\pi}}\right) \leq 5$.

□

Nous prouvons maintenant le théorème de Berry-Esséen.

Démonstration. En appliquant le lemme 3.1 aux fonctions $f_{x-\varepsilon,\varepsilon}$ et $f_{x+\varepsilon,\varepsilon}$, les inégalités (5) et (6) donnent

$$\sup_{x\in\mathbb{R}}|\mathbb{P}(W \leq x) - \Phi(x)| \leq \left(5 + \frac{3\beta K_{n-1}}{2\varepsilon}\right)\beta + \frac{\varepsilon}{\sqrt{2\pi}}.$$

En prenant $\varepsilon := \sqrt{2\pi K_{n-1}}\beta$ dans la ligne précédente, on obtient

$$\begin{aligned}
\sup_{x\in\mathbb{R}}|\mathbb{P}(W \leq x) - \Phi(x)| &\leq \left(5 + \frac{3\beta K_{n-1}}{2\sqrt{2\pi K_{n-1}}\beta}\right)\beta + \frac{\sqrt{2\pi K_{n-1}}\beta}{\sqrt{2\pi}} \\
&\leq \left(5 + \frac{3\sqrt{K_{n-1}}}{2\sqrt{2\pi}}\right)\beta + \sqrt{K_{n-1}}\beta \\
&\leq \left(5 + \frac{3}{2\sqrt{2\pi}}\sqrt{K_{n-1}}\right)\beta \\
&\leq \left(5 + 0{,}6\sqrt{K_{n-1}}\right)\beta
\end{aligned}$$

où l'on a utilisé la majoration $\frac{3}{2\sqrt{2\pi}} \leq 0{,}6$. On en déduit que $K_n \leq 5 + 0{,}6\sqrt{K_{n-1}}$. Une simple récurrence permet alors de conclure que $K_n \leq 7$ (car $7 \leq 5 + 0{,}6\sqrt{7}$).

□

Revenons maintenant sur la preuve de (22).

Démonstration. On remarque que si $a = 0$, l'inégalité à prouver est claire. On suppose $a > 0$. On remarque que, en factorisant par a, l'inégalité à prouver est équivalente à

$$\min\left\{1, \frac{x}{a}\left(1-\sigma^2\right)^{-3/2}\right\} \leq \frac{x}{a} + \sqrt{\frac{3}{2}}\sigma.$$

On pose $u := \dfrac{x}{a}$. Ainsi, cela revient à prouver

$$\min\left\{1, u\left(1-\sigma^2\right)^{-3/2}\right\} \leq u + \sqrt{\frac{3}{2}}\sigma. \tag{26}$$

Si $u \geq 1$ ou si $u = 0$, l'inégalité précédente est claire. On suppose donc $0 < u < 1$.

Soit $f : \sigma \in [0,1[\longmapsto u\left(1-\sigma^2\right)^{-3/2}$. Un simple calcul donne

$$f'(\sigma) = 3u\sigma\left(1-\sigma^2\right)^{-5/2} > 0 \quad \text{et} \quad f''(\sigma) = 3u\left(1-\sigma^2\right)^{-7/2}\left(4\sigma^2+1\right) > 0.$$

Ainsi, f est croissante et convexe sur $[0,1[$.

De plus, un simple calcul montre que

$$u\left(1-\sigma^{-3/2}\right)^2 \leq 1 \iff \sigma \leq u_0 := \sqrt{1-u^{2/3}}.$$

Par convexité de f sur $[0,1[$, $\mathscr{C}_f$ est en dessous de la corde passant par les points de coordonnées $(0,u)$ et $(u_0, f(u_0) = 1)$, ainsi

$$\forall \sigma \in [0, u_0], \quad u\left(1-\sigma^2\right)^{-3/2} \leq u + \frac{1-u}{u_0 - 0}\sigma = u + \frac{1-u}{\sqrt{1-u^{2/3}}}\sigma.$$

Or,

$$\frac{1-u}{\sqrt{1-u^{2/3}}} \leq \sqrt{\frac{3}{2}} \iff \frac{1-u^{2/3}}{1-u} \geq \frac{2}{3}(1-u).$$

On pose $g(t) := t^{2/3}$. D'après le théorème des accroissements finis, il existe $c \in]u,1[$ tel que

$$1-u^{2/3} = g(1) - g(u) = \frac{2}{3}c^{-1/3}(1-u) \iff \frac{1-u^{2/3}}{1-u} \geq \frac{2}{3}c^{-1/3}.$$

Comme $c \in]0,1[$, on a $c^{-1/3} \geq 1 \geq 1-u$, ainsi

$$\forall \sigma \in [0, u_0], \quad \min\left\{1, u\left(1-\sigma^2\right)^{-3/2}\right\} = u\left(1-\sigma^2\right)^{-3/2} \leq u + \sqrt{\frac{3}{2}}\sigma.$$

Pour terminer la preuve de (26), il suffit de remarquer, en utilisant la croissance de f et la définition de u_0, que :

$$\forall \sigma \in [u_0, 1[, \quad \min\left\{1, u\left(1-\sigma^2\right)^{-3/2}\right\} = 1 \quad \text{et} \quad u + \sqrt{\frac{3}{2}}\sigma \geq 1.$$

$\square$

Références

[1] A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independant variates.* Trans. Amer. Math. Soc., vol 49, no. 1, pp. 122-136, 1941.

[2] D. Cordero-Erausquin, `https://webusers.imj-prg.fr/~dario.cordero/Docs/M2/2009_2010/chap_Berry_Esseen.pdf`.

[3] C.-G. Esséen, *On the Liapunoff limit of error in the theory of probability.* Arkiv for Matematik Astronomi och Fysik, vol. A28, no. 9, pp. 1-19, 1942.

[4] W. Feller, *An Introduction to Probability Theory and Its Applications.* Statistical Science, vol. 1, no. 1, 1986.

[5] E. Gobet, *Méthodes de Monte-Carlo et processus stochastiques : du linéaire au non-linéaire.* Les Éditions de l'École Polytechnique, 2013.

[6] C. Stein, *A bound for the error in the normal approximation to the distribution of the sum of dependant random variables.* Sixth Berkeley Symposium, 1972.