Travail d'Étude et de Recherche

Méthode de Stein et inégalité de Berry-Esseen

Emeline DEVAUD

Encadrée par Julien CHEVALLIER

Année 2021

Table des matières

Durant tout le mémoire, les variables aléatoires utilisées sont considérées sur un même espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ et, pour de telles variables aléatoires, l'espérance et la variance seront respectivement notées $\mathbb{E}$ et $\mathbb{V}$.

Introduction

Dans la première moitié du XXᵉ siècle, la théorie des probabilités a largement été dominée par la formulation et la démonstration des théorèmes limites classiques (lois des grands nombres, théorème central limite, ...) pour des sommes de variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*). Le Théorème Central Limite [1] en particulier est l'un des résultats les plus utilisés en probabilités et en statistiques. On rappelle son énoncé :

Théorème (Théorème Central Limite (TCL))**.**
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d., admettant un moment d'ordre 2, avec $\mathbb{E}[X_1] = \mu$ et $\mathbb{V}[X_1] = \sigma^2 > 0$. Alors

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1), \qquad \text{où } S_n := \sum_{i=1}^{n} X_i.$$

Cependant, l'approche classique du TCL reposait fortement sur les méthodes de Fourier et l'étude des fonctions caractéristiques. Ce n'est seulement dans les années 1940 que Andrew Campbell Berry et Carl-Gustav Esseen ont donné une approximation plus explicite de la loi normale centrée réduite, grâce à des inégalités permettant de quantifier la vitesse de convergence vers une telle loi.
Le but de ce mémoire sera alors de démontrer le résultat suivant :

Théorème (Inégalités de Berry-Esseen)**.**
Soient $\xi_1, \ldots, \xi_n$ des variables aléatoires indépendantes, admettant un moment d'ordre 3, avec $\mathbb{E}[\xi_i] = 0$ pour tout $1 \leq i \leq n$ et $\sum_{i=1}^{n} \mathbb{E}[\xi_i^2] = 1$. On note $W := \sum_{i=1}^{n} \xi_i$, et on pose $\gamma := \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3]$. Alors on a :

- ***Inégalité de Berry-Esseen uniforme*** *: il existe une constante universelle C_0 telle que*

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq C_0\gamma.$$

- ***Inégalité de Berry-Esseen non-uniforme*** *: il existe une constante universelle C_1 telle que*

$$\forall z \in \mathbb{R}, \quad |\mathbb{P}(W \leq z) - \Phi(z)| \leq C_1 \frac{1}{1+|z|^3}\gamma.$$

où Φ désigne la fonction de répartition associée à la loi normale centrée réduite.

Souvent, pour montrer qu'une variable aléatoire converge en loi vers une autre, on étudie leur fonction caractéristique. C'est alors que Charles Stein a développé une autre technique, présentée en 1972, aujourd'hui appelée **méthode de Stein** [2]. Cette dernière est une méthode directe dont le point de départ

1. Théorème dû à Pierre-Simon de Laplace, vers 1809.
2. La méthode de Stein est devenue particulièrement importante, non seulement pour le développement de la théorie des probabilités, mais également dans de nombreux autres domaines (les mathématiques financières par exemple, ou encore la théorie des épidémies).

est une caractérisation de la loi normale centrée réduite à l'aide de l'espérance et d'une fonction absolument continue. Cette méthode semble alors magique pour donner une borne pour l'approximation de la loi normale centrée réduite [3]. Nous nous appuierons majoritairement sur les ouvrages [1] et [2] dont l'un des auteurs, Louis H.Y. Chen, était étudiant en thèse sous la direction de Charles Stein.
Après quelques rappels de résultats utiles du cours de L3 notamment, nous verrons quelles sont les idées principales de la méthode de Stein, puis quelques exemples d'application. Enfin, nous utiliserons cette dernière pour démontrer les inégalités de Berry-Esseen (uniforme et non-uniforme) dans le cas de variables aléatoires indépendantes.

Figure 1 – Charles Stein (1920-2016), Carl-Gustav Esseen (1918-2001)

3. Originellement développée pour approcher la loi normale centrée réduite, la méthode de Stein s'est très rapidement étendue : approximation de la loi de Poisson par exemple.

1 Préliminaires

Dans cette partie, nous allons rappeler des résultats qui nous seront grandement utiles lors de nombreuses démonstrations.

1.1 Quelques rappels d'analyse

Définition 1.1. *Soit $I = [a, b]$ un intervalle compact de $\mathbb{R}$. Une fonction $f : I \to \mathbb{R}$ est dite absolument continue sur I si pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour tout $n \geq 1$ et toute suite finie $([a_k, b_k])_{1\leq k\leq n}$ de sous-intervalles de I d'intérieurs disjoints, on a :*

$$\sum_{k=1}^{n} |b_k - a_k| \leq \eta \implies \sum_{k=1}^{n} |f(b_k) - f(a_k)| \leq \varepsilon.$$

De plus, $f : \mathbb{R} \to \mathbb{R}$ est dite absolument continue si elle est absolument continue sur tout intervalle compact de $\mathbb{R}$.

Propriété 1.1. *Soit $I = [a, b]$ un intervalle compact de $\mathbb{R}$. Une fonction $f : I \to \mathbb{R}$ est absolument continue sur I **si et seulement si** f est dérivable en presque tout point de I, de dérivée f' (définie presque partout), telle que $f' \in L^1(\lambda)$ et*

$$\forall x \in I, \quad f(x) - f(a) = \int_a^x f' d\lambda,$$

où λ correspond à la mesure de Lebesgue.

Démonstration. On se réfèrera à [3, théorème 8.21, p 161] ∎

Définition 1.2. *Soit $K > 0$. Une fonction $h : \mathbb{R} \to \mathbb{R}$ est dite K-lipschitzienne si*

$$\forall x, y \in \mathbb{R}, \quad |f(x) - f(y)| \leq K|x - y|.$$

Dans la suite, pour toute fonction $f : \mathbb{R} \mapsto \mathbb{R}$, $\|f\|_\infty$ désigne le **supremum essentiel** de f, *i.e.*

$$\|f\|_\infty = \inf\{M \in \mathbb{R}, |f| \leq M \quad p.p.\}.$$

Propriété 1.2. *Une fonction $h : \mathbb{R} \to \mathbb{R}$ est K-lipschitzienne **si et seulement si** h est absolument continue et $\|h'\|_\infty < K$.*

Démonstration.
$\boxed{\Rightarrow}$ On suppose que h est K-lipschitzienne. Soient $I = [a, b]$ un intervalle compact de $\mathbb{R}$, $\varepsilon > 0$ et une suite finie $([a_k, b_k])_{1\leq k\leq n}$ de sous-intervalles de I d'intérieurs disjoints.

Posons $\eta = \dfrac{\varepsilon}{K}$, tel que $\displaystyle\sum_{k=1}^{n} |b_k - a_k| \leq \eta$. h étant K-lipschitzienne, on a :
$\displaystyle\sum_{k=1}^{n} |h(b_k) - h(a_k)| \leq \sum_{k=1}^{n} K|b_k - a_k| \leq K\eta \leq \varepsilon$. Ainsi, h est absolument continue sur $[a, b]$. Cet intervalle étant arbitraire, on en déduit que h est absolument continue.
En outre, h étant K-lipschitzienne, on a $\forall x, y \in \mathbb{R}, \left|\dfrac{h(x) - h(y)}{x - y}\right| \leq K$.

On a donc, en tout $y \in \mathbb{R}$ où $\displaystyle\lim_{x\to y} \left|\frac{h(x) - h(y)}{x - y}\right|$ existe, $|h'(y)| \leq K$. Cette dernière quantité est alors définie presque partout, et comme $\|.\|_\infty$ désigne le supremum essentiel, on déduit ainsi $\|h'\|_\infty < K$.

$\boxed{\Leftarrow}$ On suppose que h est absolument continue et $\|h'\|_\infty < K$. Soit $x, y \in \mathbb{R}$. D'après la propriété 1.1, on a presque partout,

$$|h(x) - h(y)| \leq \int_y^x |h'| d\lambda \leq |x-y| \|h'\|_\infty = |x-y|K.$$

Donc h est $K-$lipschitzienne.

■

1.2 Quelques rappels de probabilités

Définition 1.3. *La loi normale centrée réduite, notée $\mathcal{N}(0,1)$, est la loi qui a pour densité*

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \forall x \in \mathbb{R}.$$

Remarque. On notera souvent Z une variable aléatoire qui suit la loi $\mathcal{N}(0,1)$.

Propriété 1.3. *Soit $Z \sim \mathcal{N}(0,1)$. La fonction de répartition de Z, notée Φ, est telle que*

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

En outre, Φ est lipschitzienne de rapport $\dfrac{1}{\sqrt{2\pi}}$.

Propriété 1.4 (Inégalité de Hölder pour l'espérance)**.** *Soient p, q des réels strictement positifs tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soit X, Y deux variables aléatoires. Alors*

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}}.$$

Corollaire 1.1. *Soient p, q, r des réels strictement positifs tels que $p, r < q$. Soit X une variable aléatoire. Alors $\mathbb{E}[|X|^p] \leq \mathbb{E}[|X|^q]^{\frac{p}{q}}$. En particulier, $\mathbb{E}[|X|^p]\mathbb{E}[|X|^r] \leq \mathbb{E}[|X|^{p+r}]$.*

Démonstration. On a $\frac{p}{q} + \frac{q-p}{q} = 1$ donc d'après propriété 1.4, on a

$$\mathbb{E}[|X|^p] = \mathbb{E}[(|X|^p \times 1)^1] \leq \mathbb{E}[(|X|^p)^{\frac{q}{p}}]^{\frac{p}{q}} \mathbb{E}[1^{\frac{q}{q-p}}]^{\frac{q-p}{q}} = \mathbb{E}[|X|^q]^{\frac{p}{q}}.$$

De même, $\mathbb{E}[|X|^r] \leq \mathbb{E}[|X|^q]^{\frac{r}{q}}$. Donc, pour $q = p + r$, on a $\mathbb{E}[|X|^p]\mathbb{E}[|X|^r] \leq \mathbb{E}[|X|^{p+r}]$. ■

Propriété 1.5 ("Interversion espérance / intégrale")**.** *Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. Soit $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ une fonction telle que $\mathbb{E}\left[\int_{\mathbb{R}} |f(t, X)| dt\right] < +\infty$. Alors :*

$$\mathbb{E}\left[\int_{\mathbb{R}} f(t, X) dt\right] = \int_{\mathbb{R}} \mathbb{E}[f(t, X)] dt.$$

Démonstration. On a $\mathbb{E}\left[\int_{\mathbb{R}} f(t, X) dt\right] = \int_{\Omega} \left(\int_{\mathbb{R}} f(t, X(w)) dt\right) d\mathbb{P}(w)$.
Or, f étant intégrable, on a d'après le théorème de Fubini,

$$\begin{aligned}\int_{\Omega} \left(\int_{\mathbb{R}} f(t, X(w)) dt\right) d\mathbb{P}(w) &= \int_{\mathbb{R}} \left(\int_{\Omega} f(t, X(w)) d\mathbb{P}(w)\right) dt \\ &= \int_{\mathbb{R}} \mathbb{E}[f(t, X)] dt.\end{aligned}$$

D'où,

$$\mathbb{E}\left[\int_{\mathbb{R}} f(t, X) dt\right] = \int_{\mathbb{R}} \mathbb{E}[f(t, X)] dt.$$

■

Propriété 1.6 (Inégalité de Markov)**.** *Soit X une variable aléatoire réelle supposée positive ou nulle presque sûrement. Alors*

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Corollaire 1.2 (Corollaire de l'inégalité de Markov)**.** *Soit f une fonction croissante positive ou nulle sur un ensemble I. Soit Y une variable aléatoire réelle telle que l'image est contenue dans I. Alors*

$$\forall b > 0 \mid f(b) > 0, \quad \mathbb{P}(Y \geq b) \leq \frac{E[f(Y)]}{f(b)}.$$

Démonstration. On pose $X = f(Y)$ et $a = f(b)$, alors X est bien une variable aléatoire positive ou nulle et $a > 0$. D'après l'inégalité de Markov, on a $\mathbb{P}(f(Y) \geq f(b)) \leq \dfrac{\mathbb{E}[f(Y)]}{f(b)}$.
Or, f étant croissante, on a $\{Y \geq b\} \subset \{f(Y) \geq f(b)\}$ et donc

$$\forall b > 0, \mathbb{P}(Y \geq b) \leq \mathbb{P}(f(Y) \geq f(b)) \leq \frac{\mathbb{E}[f(Y)]}{f(b)}.$$

■

Nous pouvons à présent nous plonger au cœur du sujet, commençons par parler de la bien nommée méthode de Stein.

2 Fondamentaux de la méthode de Stein

Dans tout ce qui suit, nous noterons Z pour désigner une variable aléatoire qui suit la loi $\mathcal{N}(0,1)$.
Ici, nous allons donner des détails sur la méthode de Stein. Tout d'abord, nous évoquerons la caractérisation de la loi normale centrée réduite, puis nous verrons quelques propriétés des solutions de l'équation de Stein. Enfin, nous pourrons construire l'identité de Stein.

2.1 Caractérisation de la loi normale centrée réduite

Le point de départ de la méthode de Stein repose sur une caractérisation de la loi normale centrée réduite. Avant d'énoncer la véritable caractérisation, nous avons besoin de quelques résultats préliminaires.

2.1.1 Premiers résultats

Lemme 2.1. *Soit $z \in \mathbb{R}$. L'unique solution bornée $f := f_z$ de l'équation*

$$f'(w) - wf(w) = \mathbb{1}_{\{w \leq z\}} - \Phi(z) \tag{2.1}$$

est donnée par :

$$f_z(w) = \begin{cases} \sqrt{2\pi} e^{\frac{w^2}{2}} \Phi(w)[1 - \Phi(z)] & \text{si } w \leq z \\ \sqrt{2\pi} e^{\frac{w^2}{2}} \Phi(z)[1 - \Phi(w)] & \text{si } w > z. \end{cases} \tag{2.2}$$

Démonstration. On va utiliser la méthode de variation de la constante. Tout d'abord, la solution homogène associée à (2.1) est $w \mapsto \lambda e^{\frac{w^2}{2}}, \lambda \in \mathbb{R}$.

Variation de la constante : Considérons une fonction $g : \mathbb{R} \to \mathbb{R}$ telle que $g(w) = \lambda(w)e^{\frac{w^2}{2}}$.
On obtient alors : $\lambda'(w)e^{\frac{w^2}{2}} = \mathbb{1}_{w\leq z} - \Phi(z)$, d'où $\lambda'(w) = e^{-\frac{w^2}{2}}\left(\mathbb{1}_{w\leq z} - \Phi(z)\right)$.
Cette dernière quantité étant intégrable, on déduit :

$$\lambda(w) = \begin{cases} \displaystyle\int_{-\infty}^{w} e^{-\frac{x^2}{2}}\left(\mathbb{1}_{\{x\leq z\}} - \Phi(z)\right)dx \\ \displaystyle -\int_{w}^{+\infty} e^{-\frac{x^2}{2}}\left(\mathbb{1}_{\{x\leq z\}} - \Phi(z)\right)dx \end{cases}$$

où l'on a pris les constantes d'intégration égales à 0.
Une solution particulière de (2.1) est donc $g(w) = \lambda(w)e^{\frac{w^2}{2}}$, c'est-à-dire :

$$\begin{aligned} g(w) &= \begin{cases} e^{\frac{w^2}{2}}\left(\displaystyle\int_{-\infty}^{z} e^{-\frac{x^2}{2}}dx - \int_{-\infty}^{w} e^{-\frac{x^2}{2}}\Phi(z)dx\right) & \text{si } w > z \\ e^{\frac{w^2}{2}}\left(-\displaystyle\int_{w}^{z} e^{-\frac{x^2}{2}}dx + \int_{w}^{+\infty} e^{-\frac{x^2}{2}}\Phi(z)dx\right) & \text{si } w \leq z \end{cases} \\ &= \begin{cases} e^{\frac{w^2}{2}}\left(\sqrt{2\pi}\Phi(z) - \Phi(z)\displaystyle\int_{-\infty}^{w} e^{-\frac{x^2}{2}}dx\right) & \text{si } w > z \\ e^{\frac{w^2}{2}}\left(-\sqrt{2\pi}(\Phi(z) - \Phi(w)) + \sqrt{2\pi}\Phi(z)(1 - \Phi(w))\right) & \text{si } w \leq z \end{cases} \\ &= \begin{cases} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(z)[1 - \Phi(w)] & \text{si } w > z \\ \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1 - \Phi(z)] & \text{si } w \leq z. \end{cases} \end{aligned}$$

La solution totale est donc $f_z(w) = g(w) + de^{\frac{w^2}{2}}, d \in \mathbb{R}$. Et comme, d'après la propriété qui suit (2.1), la solution est bornée, il faut prendre $d = 0$.
Dans ce cas, $f_z(w) = g(w)$, ce qui est le résultat voulu.

■

Propriété 2.1 (Propriétés de f_z)**.** *Soit $z \in \mathbb{R}$ et soit f_z la fonction donnée dans le lemme 2.1. Alors, pour tous $u, v, w \in \mathbb{R}$, on a :*

$$|wf_z(w)| \leq 1 \quad et \quad |wf_z(w) - uf_z(u)| \leq 1, \tag{2.3}$$

$$|f'_z(w)| \leq 1 \quad et \quad |f'_z(w) - f'_z(u)| \leq 1, \tag{2.4}$$

$$0 \leq f_z(w) \leq \min\left(\frac{\sqrt{2\pi}}{4}, \frac{1}{|z|}\right), \tag{2.5}$$

$$|(w+u)f_z(w+u) - (w+v)f_z(w+v)| \leq \left(|w| + \frac{\sqrt{2\pi}}{4}\right)(|u| + |v|)\,. \tag{2.6}$$

Remarque. Nous aurons souvent recours aux résultats précédents dans les sections suivantes.

Démonstration. La preuve de cette propriété est assez longue et calculatoire ; comme les arguments utilisés ne relèvent pas de la méthode de Stein, le lecteur se réfèrera à [1, appendice chapitre 2] s'il le souhaite. ■

2.1.2 Caractérisation

Nous pouvons maintenant évoquer le point clé de la méthode de Stein, qui sera à la genèse de nombreuses démonstrations.

Propriété 2.2 (Caractérisation de la loi normale centrée réduite)**.**
Soit W une variable aléatoire réelle.

(i) Si $W \sim \mathcal{N}(0,1)$, alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ absolument continue telle que $\mathbb{E}[f'(W)] < \infty$, on a : $\mathbb{E}[f'(W)] = \mathbb{E}[Wf(W)]$.

(ii) Réciproquement, si pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ bornée, continue, de classe $\mathcal{C}^1$ par morceaux telle que $\mathbb{E}[f'(W)] < \infty$, on a $\mathbb{E}[f'(W)] = \mathbb{E}[Wf(W)]$ alors : $W \sim \mathcal{N}(0,1)$.

Démonstration. (i) Supposons que $W \sim \mathcal{N}(0,1)$. Soit $f : \mathbb{R} \to \mathbb{R}$ absolument continue telle que $\mathbb{E}[f'(W)] < \infty$. Alors, par le théorème de transfert :

$$\begin{aligned}
\mathbb{E}[f'(W)] =& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(w) e^{-\frac{w^2}{2}} dw \\
=& \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} f'(w) e^{-\frac{w^2}{2}} dw + \int_{0}^{+\infty} f'(w) e^{-\frac{w^2}{2}} dw \right) \\
=& \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} f'(w) \left(\int_{-\infty}^{w} -xe^{-\frac{x^2}{2}} dx \right) dw + \int_{0}^{+\infty} f'(w) \left(\int_{w}^{+\infty} xe^{-\frac{x^2}{2}} dx \right) dw \right\}.
\end{aligned}$$

Et puisque $\mathbb{E}[f'(W)] < \infty$, $(w,x) \mapsto f'(w)xe^{-\frac{x^2}{2}}$ est intégrable. Le théorème de Fubini nous donne ainsi :

$$\begin{aligned}
\mathbb{E}[f'(W)] =& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \left(\int_{x}^{0} f'(w) dw \right) \left(-xe^{-\frac{x^2}{2}} \right) dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} \left(\int_{0}^{x} f'(w) dw \right) xe^{-\frac{x^2}{2}} dx \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(f(x) - f(0) \right) xe^{-\frac{x^2}{2}} dx \\
=& \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} f(x) xe^{-\frac{x^2}{2}} dx - f(0) \left[-e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} \right) \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) xe^{-\frac{x^2}{2}} dx \\
=& \mathbb{E}[Wf(W)].
\end{aligned}$$

(ii) Supposons que pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ bornée, continue, de classe $\mathcal{C}^1$ par morceaux telle que $\mathbb{E}[f'(W)] < \infty$, on a $\mathbb{E}[f'(W)] = \mathbb{E}[Wf(W)]$.
La fonction f_z donnée dans le lemme 2.1 vérifie ces hypothèses, donc on a en particulier :

$$\begin{aligned}
0 =& \mathbb{E}[f_z'(W) - Wf_z(W)] \\
=& \mathbb{E}[\mathbb{1}_{\{W \leq z\}} - \Phi(z)] \quad \text{car } f_z \text{ est solution de (2.1)} \\
=& \mathbb{P}(W \leq z) - \Phi(z).
\end{aligned}$$

Comme la fonction de répartition caractérise la loi, on en déduit que $W \sim \mathcal{N}(0,1)$.

■

2.2 Équation de Stein

Afin d'utiliser son observation sur la caractérisation de la loi normale centrée réduite, Stein s'est intéressé à une équation en particulier, qui conduira à ladite équation de Stein. On rappelle que $Z \sim \mathcal{N}(0,1)$.

2.2.1 Équation

Propriété 2.3. *Soit h une fonction mesurable donnée telle que $\mathbb{E}[h(Z)] < \infty$. On note $N(h) := \mathbb{E}[h(Z)]$. L'unique solution bornée $f := f_h$ de l'équation*

$$f'(w) - wf(w) = h(w) - N(h)$$

est donnée par :

$$f_h(w) = \begin{cases} e^{\frac{w^2}{2}} \displaystyle\int_{-\infty}^{w} (h(x) - N(h))\, e^{-\frac{x^2}{2}}\, dx \\ -\, e^{\frac{w^2}{2}} \displaystyle\int_{w}^{+\infty} (h(x) - N(h))\, e^{-\frac{x^2}{2}}\, dx. \end{cases} \tag{2.7}$$

Démonstration. On utilise la même démarche que celle de la preuve du lemme 2.1.
Le caractère borné de f_h sera vu ultérieurement dans la propriété 2.4 ■

Définition 2.1 (Équation de Stein). *Soit h une fonction mesurable donnée telle que $\mathbb{E}[h(Z)] < \infty$. On note toujours $N(h) := E[h(Z)]$.*
On appelle équation de Stein, l'équation donnée par :

$$f'(w) - wf(w) = h(w) - N(h). \tag{2.8}$$

Remarque. L'équation (2.8) est un cas particulier de l'équation (2.1) pour $h(w) = \mathbb{1}_{\{w \leq z\}}$.

2.2.2 Propriétés de la solution de l'équation de Stein

Propriété 2.4. *Soit h une fonction mesurable donnée telle que $\mathbb{E}[h(Z)] < \infty$, et soit f_h la solution de l'équation de Stein, donnée dans la propriété 2.3.*
Si h est absolument continue, alors :

$$\|f_h\|_\infty \leq 2\|h'\|_\infty, \tag{2.9}$$

$$\|f_h'\|_\infty \leq \sqrt{\frac{2}{\pi}}\|h'\|_\infty \quad \textit{et} \quad \|f_h''\|_\infty \leq 2\|h'\|_\infty, \tag{2.10}$$

$$\|f_h'\|_\infty \leq \min\left(2\|h(.) - N(h)\|_\infty, 4\|h'\|_\infty\right). \tag{2.11}$$

Remarque. Nous aurons souvent recours aux résultats précédents dans les sections suivantes.

Démonstration. De même que pour la propriété 2.1, cette preuve étant assez calculatoire, le lecteur est invité à se référer à [1, appendice chapitre 2] pour plus de détails. ■

2.3 Construction de l'identité de Stein

Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction absolument continue telle que $N(h) < \infty$. Partant de l'équation de Stein $f'(w) - wf(w) = h(w) - N(h)$, on veut montrer qu'une variable aléatoire réelle W telle que $E[W] = 0$ et $\mathbb{V}[W] = 1$ peut être approchée par une variable aléatoire de loi $\mathcal{N}(0, 1)$.
Étant donnée une telle variable W, d'après les majorations des propriétés 2.1 et 2.4, l'espérance dans l'équation de Stein est bien définie et l'on peut alors écrire

$$\mathbb{E}[f'(W) - Wf(W)] = \mathbb{E}[h(W)] - \mathbb{E}[h(Z)].$$

Dès lors, afin de montrer que la quantité $\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]$ tend vers 0 (pour avoir la convergence en loi), on va naturellement s'intéresser à $\mathbb{E}[f'(W) - Wf(W)]$.
Dans ce but et en nous intéressant à une certaine somme de variables aléatoires indépendantes, nous allons alors construire l'identité de Stein.

2.3.1 Somme de variables aléatoires indépendantes

Soient $\xi_1, \dots \xi_n$ des variables aléatoires indépendantes admettant un moment d'ordre 3 telles que :

$$\forall \quad 1 \leq i \leq n, \quad \mathbb{E}[\xi_i] = 0 \quad \text{et} \quad \sum_{i=1}^{n} \mathbb{E}[\xi_i^2] = 1.$$

Posons $W := \displaystyle\sum_{i=1}^{n} \xi_i$, et soit $W^i := W - \xi_i$ (alors $E[W] = 0$ et $\mathbb{V}[W] = 1$).

Remarque. Pour tout $1 \leq i \leq n$, $\quad \xi_i$ et W^i sont indépendantes.

On considère également une fonction K_i donnée par

$$\forall t \in \mathbb{R}, K_i(t) = \mathbb{E}[\xi_i(\mathbb{1}_{[0,\xi_i]}(t) - \mathbb{1}_{[\xi_i,0[}(t))].$$

Propriété 2.5 (Propriétés de la fonction K_i)**.** *La fonction K_i définie précédemment vérifie :*

(i) $\forall t \in \mathbb{R}, K_i(t) \geq 0,$

(ii) $\displaystyle\int_{-\infty}^{+\infty} K_i(t)dt = \mathbb{E}[\xi_i^2],$

(iii) $\displaystyle\int_{-\infty}^{+\infty} |t|K_i(t)dt = \frac{1}{2}\mathbb{E}[|\xi_i|^3].$

Démonstration.
(i) Soit $t \in \mathbb{R}$. Par définition, $K_i(t) = \mathbb{E}[\xi_i \mathbb{1}_{[0,\xi_i]}(t)] - \mathbb{E}[\xi_i \mathbb{1}_{[\xi_i,0[}(t)]$. On a donc 2 cas :

- ✓ Si $t \geq 0$, $K_i(t) = \mathbb{E}[\xi_i \mathbb{1}_{[0,\xi_i]}(t)] \geq 0$.
- ✓ Si $t < 0$, $K_i(t) = -E[\xi_i \mathbb{1}_{[\xi_i,0[}(t)] \geq 0$ puisque $\xi_i \mathbb{1}_{[\xi_i,0[}(t) < 0$.

D'où le résultat.

(ii)

$$\begin{aligned}
\int_{-\infty}^{+\infty} K_i(t)dt &= \int_{-\infty}^{+\infty} \mathbb{E}[\xi_i(\mathbb{1}_{[0,\xi_i]}(t) - \mathbb{1}_{[\xi_i,0[}(t))]dt \\
&= \int_{-\infty}^{0} -\mathbb{E}[\xi_i \mathbb{1}_{[\xi_i,0[}(t)]dt + \int_{0}^{+\infty} \mathbb{E}[\xi_i \mathbb{1}_{[0,\xi_i]}(t)]dt \\
&= -\,\mathbb{E}\left[\int_{\xi_i}^{0} \xi_i dt \mathbb{1}_{]-\infty,0]}(\xi_i)\right] + \mathbb{E}\left[\int_{0}^{\xi_i} \xi_i dt \mathbb{1}_{[0,+\infty[}(\xi_i)\right] \\
&= \mathbb{E}[\xi_i^2],
\end{aligned}$$

où l'on a utilisé la propriété "Interversion espérance / intégrale" dans l'avant-dernière égalité, car $(t,\xi_i) \mapsto \xi_i \mathbb{1}_{[\xi_i,0[}(t)$ (respectivement $(t,\xi_i) \mapsto \xi_i \mathbb{1}_{[0,\xi_i]}(t)$) est intégrable.

(iii)

$$\begin{aligned}
\int_{-\infty}^{+\infty} |t|K_i(t)dt &= \int_{-\infty}^{0} -tK_i(t)dt + \int_{0}^{+\infty} tK_i(t)dt \\
&= \int_{-\infty}^{0} -t\mathbb{E}[-\mathbb{1}_{[\xi_i,0[}(t))]dt + \int_{0}^{+\infty} tE[\xi_i(\mathbb{1}_{[0,\xi_i]}(t)]dt \\
&= \mathbb{E}\left[\int_{-\infty}^{0} t\xi_i \mathbb{1}_{[\xi_i,0[}(t)dt\right] + \mathbb{E}\left[\int_{0}^{+\infty} t\xi_i \mathbb{1}_{[0,\xi_i]}(t)dt\right] \\
&= \mathbb{E}\left[\int_{\xi_i}^{0} t\xi_i dt \mathbb{1}_{]-\infty,0]}(\xi_i)\right] + \mathbb{E}\left[\int_{0}^{\xi_i} t\xi_i dt 1_{[0,+\infty[}(\xi_i)\right] \\
&= \frac{1}{2}\mathbb{E}[|\xi_i|^3],
\end{aligned}$$

où l'on a encore utilisé la propriété 1.5 puisque $(t,\xi_i) \mapsto t\xi_i \mathbb{1}_{[0,\xi_i]}(t)$ (respectivement $(t,\xi_i) \mapsto \mathbb{1}_{[\xi_i,0[}(t))$) est intégrable.

■

À présent, comme indiqué dans le préambule, nous allons nous intéresser à la quantité $\mathbb{E}[f'(W) - Wf(W)]$. À ces fins, nous allons construire l'identité de Stein.
Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction absolument continue telle que $N(h) < \infty$.
Soit $f := f_h$ la solution de l'équation de Stein (2.8). D'après les propriétés 2.1 et 2.4, les quantités $\mathbb{E}[f'(W)]$ et $\mathbb{E}[Wf(W)]$ sont bien définies. On exprime alors chaque terme.

Etape 1 : Comme $\int_{-\infty}^{+\infty} K_i(t)dt = \mathbb{E}[\xi_i^2]$, alors $\sum_{i=1}^n \int_{-\infty}^{+\infty} K_i(t)dt = \sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$.

On a donc :

$$\begin{aligned}\mathbb{E}[f'(W)] = \mathbb{E}\left[\sum_{i=1}^n \mathbb{E}[\xi_i^2]f'(W)\right] &= \mathbb{E}\left[\sum_{i=1}^n \int_{-\infty}^{+\infty} K_i(t)f'(W)dt\right] \\ &= \sum_{i=1}^n \int_{-\infty}^{+\infty} K_i(t)\mathbb{E}[f'(W)]dt, \end{aligned} \tag{2.12}$$

où l'on a encore utilisé la propriété 1.5 puisque comme $\mathbb{E}[f'(W)] < \infty$, $(t, W) \mapsto K_i(t)f'(W)$ est intégrable.

Etape 2 : D'autre part, nous avons :

$$\begin{aligned}\mathbb{E}[Wf(W)] = \mathbb{E}\left[\sum_{i=1}^n \xi_i f(W)\right] &= \sum_{i=1}^n \mathbb{E}[\xi_i f(W)] \\ &= \sum_{i=1}^n \mathbb{E}[\xi_i(f(W) - f(W^i)] \quad \text{car } \mathbb{E}[\xi_i] = 0 \text{ et } \xi_i \text{ et } W^i \text{ sont indépendantes pour tout } 1 \leq i \leq n \\ &= \sum_{i=1}^n \mathbb{E}\left[\xi_i \int_0^{\xi_i} f'(W^i + t)dt\right] \\ &= \sum_{i=1}^n \mathbb{E}\left[\int_{-\infty}^{+\infty} f'(W^i + t)\xi_i(\mathbb{1}_{[0,\xi_i]}(t) - \mathbb{1}_{[\xi_i,0[}(t))dt\right] \\ &= \sum_{i=1}^n \int_{-\infty}^{+\infty} \mathbb{E}[f'(W^i + t)]K_i(t)dt, \end{aligned} \tag{2.13}$$

où on a encore utilisé la propriété 1.5 et l'indépendance de ξ_i et W^i pour tout $1 \leq i \leq n$.
Ainsi, en faisant la différence des expressions (2.12) et (2.13), on obtient :

$$\mathbb{E}[f'(W) - Wf(W)] = \sum_{i=1}^n \int_{-\infty}^{+\infty} K_i(t)\mathbb{E}[f'(W) - f'(W^i + t)]dt. \tag{2.14}$$

2.3.2 Identité de Stein

On se place exactement dans les mêmes conditions que dans la section 2.3.1. Le résultat obtenu en (2.14) nous conduit alors à la définition suivante.

Définition 2.2 (Identité de Stein)**.** *On appelle identité de Stein, l'expression :*

$$\mathbb{E}[f'(W) - Wf(W)] = \sum_{i=1}^n \int_{-\infty}^{+\infty} K_i(t)\mathbb{E}[f'(W) - f'(W^i + t)]dt, \tag{2.15}$$

où $f := f_h$ est la solution de l'équation de Stein (2.8).

Cette dernière identité a donc été obtenue par ce qu'on appelle communément la **méthode de Stein pour la loi normale centrée réduite**, à savoir : une caractérisation de la loi normale centrée réduite, puis l'introduction d'une certaine somme de variables aléatoires indépendantes combinée à la solution de l'équation de Stein.

Remarque. L'identité de Stein (ainsi que tous les autres résultats que l'on a vus dans cette partie) va être largement utilisée dans de nombreuses démonstrations.

3 Utilisation de la méthode de Stein

Il est maintenant temps de mettre en application la méthode de Stein. Pour ce faire, nous considèrerons une somme de variables aléatoires indépendantes, afin de contrôler la quantité $\mathbb{E}[h(W)] - N(h)$ pour une fonction h d'une certaine classe. L'idée principale sera d'utiliser l'équation de Stein (2.8) ainsi que l'identité de Stein (2.15). Par ailleurs, nous verrons également que la méthode de Stein permet de démontrer le Théorème Central Limite de Lindeberg, et aussi de donner une borne inférieure pour l'approximation de la loi $\mathcal{N}(0,1)$. On se place ici dans le cas de variables indépendantes, mais les idées introduites permettent (même si nous ne le verrons pas ici) de traiter des situations plus générales.
Dans toute cette section, on se place dans les conditions suivantes (déjà utilisées en partie 2.3.1).

Cadre de travail 1

Soient $\xi_1, \dots \xi_n$ des variables aléatoires indépendantes, admettant un moment d'ordre 3, telles que :

$$\forall \quad 1 \leq i \leq n, \quad \mathbb{E}[\xi_i] = 0, \quad \text{et} \quad \sum_{i=1}^{n} \mathbb{E}[\xi_i^2] = 1.$$

On pose $\gamma := \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3]$. En outre, on note $W := \sum_{i=1}^{n} \xi_i$, $W^i := W - \xi_i$, et on considère

$$\forall t \in \mathbb{R}, \quad K_i(t) = \mathbb{E}[\xi_i(\mathbb{1}_{[0,\xi_i]}(t) - \mathbb{1}_{[\xi_i,0[}(t))].$$

Par ailleurs, $Z \sim \mathcal{N}(0,1)$ et $N(h) := \mathbb{E}[h(Z)]$ pour une fonction lisse h convenable.

Remarque. On rappelle que des résultats intéressants sur la fonction K_i ont été vus dans la propriété 2.5.
De plus, ξ_i et W^i sont indépendantes pour tout $1 \leq i \leq n$ (ce qui sera utilisé dans de très nombreux calculs).

3.1 Approximation de $\mathcal{N}(0,1)$ pour des fonctions lipschitziennes

On se place dans le Cadre de travail 1 précédent. Le but de cette partie est d'estimer la quantité $\mathbb{E}[h(W)] - N(h)$ pour une fonction h lipschitzienne telle que $\|h'\|_\infty < \infty$. Dans ces conditions, on va alors obtenir diverses majorations.

Théorème 3.1.
Soit h une fonction lipschitzienne. Alors :

$$|\mathbb{E}[h(W)] - N(h)| \leq 3\|h'\|_\infty \gamma. \tag{3.1}$$

Démonstration. On considère f_h la solution de l'équation de Stein (2.8). h étant lipschitzienne, on a par la propriété 1.2, que h est absolument continue. Alors, par le résultat (2.10), on a $\|f_h''\|_\infty \leq 2\|h'\|_\infty$. Par l'identité de Stein (2.15), il vient :

$$\begin{aligned}|\mathbb{E}[h(W)] - N(h)| = |\mathbb{E}[f_h'(W) - Wf_h(W)]| \leq& \sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mathbb{E}\left[|f_h'(W) - f_h'(W^i + t)|\right] K_i(t)dt \\ \leq& 2\|h'\|_\infty \sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mathbb{E}[|t| + |\xi_i|]K_i(t)dt,\end{aligned}$$

car d'après l'inégalité des accroissements finis : $|f_h'(W) - f_h'(W^i + t)| \leq \|f_h''\|_\infty |t - \xi_i|$.
En utilisant la propriété 2.5, on obtient alors :

$$\begin{aligned}|\mathbb{E}[h(W)] - N(h)| \leq& 2\|h'\|_\infty \left(\sum_{i=1}^{n} \frac{1}{2}\mathbb{E}[|\xi_i|^3] + \mathbb{E}[|\xi_i|]\mathbb{E}[\xi_i^2]\right) \\ \leq& 3\|h'\|_\infty \sum_{i=1}^{n} E[|\xi_i|^3] \qquad (3.2) \\ =& 3\|h'\|_\infty \gamma,\end{aligned}$$

où l'on a utilisé le corollaire 1.1 pour (3.2) avec $p = 1, r = 2$. D'où le résultat. ■

Par ailleurs, le théorème qui suit montre que l'on peut obtenir une majoration de $|\mathbb{E}[h(W)] - N(h)|$ en omettant l'hypothèse d'existence du moment d'ordre 3 introduite dans le Cadre de travail 1.

Théorème 3.2.
Pour toute fonction h lipschitzienne, on a :

$$|\mathbb{E}[h(W)] - N(h)| \leq \|h'\|_\infty (4\beta_2 + 3\beta_3), \qquad (3.3)$$

où $\beta_2 = \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>1\}}]$ *et* $\beta_3 = \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3 \mathbb{1}_{\{|\xi_i|\leq 1\}}]$.

Démonstration. On va utiliser l'identité de Stein (2.15). Pour cela, on considère f_h la solution de l'équation de Stein. h étant lipschitizienne, on a par la propriété 1.2, que h est absolument continue et $\|h'\|_\infty < \infty$.

Etape 1 : Par le résultat (2.11), on a :

$$\begin{aligned}|f_h'(W) - f_h'(W^i + t)| \leq& \min(2\|f_h'\|_\infty, \|f_h''\|_\infty |t - \xi_i|) \\ \leq& 2\|h'\|_\infty \min(1, |t - \xi_i|) \\ \leq& 2\|h'\|_\infty (\min(|t|, 1) + \min(|\xi_i|, 1)).\end{aligned}$$

Par l'identité de Stein (2.15), on a en outre :

$$\begin{aligned}|\mathbb{E}[h(W)] - N(h)| =& |\mathbb{E}[f_h'(W) - Wf_h'(W)]| \\ \leq& \sum_{i=1}^{n} \int_{-\infty} \mathbb{E}\left[|f_h'(W) - f_h'(W^i + t)|\right] K_i(t)dt \\ \leq& 2\|h'\|_\infty \sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mathbb{E}[\min(|t|, 1) + \min(|\xi_i|, 1)]K_i(t)dt. \qquad (3.4)\end{aligned}$$

Etape 2 : On cherche alors à exprimer la borne précédente en fonction de β_2 et β_3. À ces fins, remarquons que :

$$\int_{-\infty}^{+\infty} \min(|t|,1)\left(\mathbb{1}_{[0,x]}(t) - \mathbb{1}_{[-x,0]}(t)\right) dt = \begin{cases} \frac{1}{2}|x|t^2 + |x|(|x|-1) & \text{si } |x| > 1 \\ \frac{1}{2}|x|^3 & \text{si } |x| \leq 1. \end{cases} \tag{3.5}$$

On peut alors écrire :

$$\begin{aligned}\int_{-\infty}^{+\infty} \mathbb{E}[\min(|t|,1) + \min(|\xi_i|,1)]K_i(t)dt =& \mathbb{E}[|\xi_i|(|\xi_i|-1)\mathbb{1}_{\{|\xi_i|>1\}}] + \frac{1}{2}\mathbb{E}[|\xi_i|\min(|\xi_i|,1)^2] \\ &+ \mathbb{E}[\xi_i^2\mathbb{E}[\min(|\xi_i|,1)]] \\ =& \mathbb{E}[\xi_i^2\mathbb{1}_{\{|\xi_i|>1\}}] - \frac{1}{2}\mathbb{E}[|\xi_i|\mathbb{1}_{\{|\xi_i|>1\}}] \\ &+ \frac{1}{2}\mathbb{E}[|\xi_i|^3\mathbb{1}_{\{|\xi_i|\leq 1\}}] + \mathbb{E}[\xi_i^2\mathbb{E}[\min(|\xi_i|,1)]] \\ \leq& \mathbb{E}[\xi_i^2\mathbb{1}_{\{|\xi_i|>1\}}] + \frac{1}{2}\mathbb{E}[|\xi_i|^3\mathbb{1}_{\{|\xi_i|\leq 1\}}] \\ &+ \mathbb{E}[\xi_i^2\mathbb{E}[\min(|\xi_i|,1)]].\end{aligned}$$

En sommant pour i allant de 1 à n et en utilisant (3.4), on a :

$$|\mathbb{E}[h(W) - N(h)]| \leq 2\|h'\|_\infty \left\{\beta_2 + \frac{1}{2}\beta_3 + \sum_{i=1}^{n} \mathbb{E}[\xi_i^2\mathbb{E}[\min(|\xi_i|,1)]]\right\}. \tag{3.6}$$

Etape 3 : Or, $x \mapsto x^2$ et $x \mapsto \min(x,1)$ sont croissantes pour $x \geq 0$. De plus, si g et h sont des fonctions croissantes, alors pour toute variable aléatoire X, $\mathbb{E}[g(X)]E[h(X)] \leq \mathbb{E}[g(X)h(X)]$ (cf. Appendice section 3, lemme 6.1). Donc pour tout $\xi_i \in \mathbb{R}$, comme $|\xi_i| \geq 0$, on a :

$$\mathbb{E}[\xi_i^2\mathbb{E}[\min(|\xi_i|,1)]] = \mathbb{E}[\xi_i^2]\mathbb{E}[\min(|\xi_i|,1)] \leq \mathbb{E}[\xi_i^2\min(|\xi_i|,1)] = \mathbb{E}[|\xi_i|^3\mathbb{1}_{\{|\xi_i|\leq 1\}}] + \mathbb{E}[\xi_i^2\mathbb{1}_{\{|\xi_i|>1\}}].$$

En sommant pour i de 1 à n et en injectant cette inégalité dans (3.6), on obtient :

$$|\mathbb{E}[h(W)] - N(h)| \leq 2\|h'\|_\infty \left(\beta_2 + \frac{1}{2}\beta_3 + \beta_2 + \beta_3)\right) = \|h'\|_\infty(4\beta_2 + 3\beta_3),$$

d'où le résultat.

■

Que ce soit avec le théorème 3.1 ou le théorème 3.2, on ne peut pas déduire une inégalité de Berry-Esseen pour de telles fonctions lipschitziennes. Néanmoins, on peut arriver à une borne supérieure pour $\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$, comme le montre le théorème suivant.

Théorème 3.3 (R.V. Erickson, 1974)**.**
On suppose qu'il existe δ tel que pour toute fonction h lipschitzienne, on ait $|\mathbb{E}[h(W)] - N(h)| \leq \delta\|h'\|_\infty$. Alors :

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 2\sqrt{\delta}. \tag{3.7}$$

Remarque. Bien que la preuve de ce théorème n'utilise pas la méthode de Stein, il nous sera utile par la suite. Sa démonstration est alors présentée dans l'Appendice section 3.

Comme nous venons de le constater, la méthode de Stein s'avère alors très utile et efficace pour donner des bornes d'approximation de la loi normale centrée réduite en utilisant l'équation de Stein et l'identité de Stein, avec des fonctions lipschitziennes. Dans la suite, nous retrouverons fréquemment les raisonnements développés précédemment.

3.2 Théorème Central Limite de Lindeberg

Etudions à présent une autre conséquence de la méthode de Stein. En effet, nous allons voir que cette dernière (où plutôt les résultats de la section 3.1 qui en découlent) permet de démontrer un résultat énoncé par Jarl Waldemar Lindeberg [4] en 1920. Pour cela, on se place toujours dans le Cadre de travail 1 introduit en début de partie.

Définition 3.1 (Condition de Lindeberg)**.** *On appelle Condition de Lindeberg, la condition suivante :*

$$\forall \varepsilon > 0, \lim_{n\to+\infty} \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>\varepsilon\}}] = 0. \tag{3.8}$$

Théorème 3.4 (Théorème Central Limite de Lindeberg)**.**
Si la Condition de Lindeberg (3.8) est vérifiée, alors :

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \xrightarrow[n\to+\infty]{} 0. \tag{3.9}$$

Remarque. Nous verrons ultérieurement (cf. Théorème 3.5) que, sous certaines hypothèses, la Condition de Lindeberg peut aussi être nécessaire.

Démonstration. L'idée est d'utiliser les résultats vus dans la section 3.1. Pour cela, on reprend, comme dans le Théorème 3.2, $\beta_2 = \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 1_{\{|\xi_i|>1\}}]$ et $\beta_3 = \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3 \mathbb{1}_{\{|\xi_i|\leq 1\}}]$.
Soit maintenant $0 < \varepsilon < 1$, alors :

$$\begin{aligned}
\beta_2 + \beta_3 &= \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>1\}}] + \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3 \mathbb{1}_{\{|\xi_i|\leq 1\}}] \\
&\leq \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>1\}}] + \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{\varepsilon<|\xi_i|\leq 1\}}] + \sum_{i=1}^{n} \varepsilon \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|\leq \varepsilon\}}] \\
&\leq \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>\varepsilon\}}] + \varepsilon \quad \text{comme } \sum_{i=1}^{n} \mathbb{E}[\xi_i^2] = 1 \\
&\xrightarrow[n\to+\infty]{} 0 \quad \text{car la Condition de Lindeberg (3.8) est vérifée et car } \varepsilon \text{ est arbitraire.}
\end{aligned}$$

Ainsi, par les Théorèmes 3.3 et 3.2 et en prenant $\delta = (4\beta_2 + 3\beta_3)$,
on a :

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 4\sqrt{\beta_2 + \frac{3}{4}\beta_3} \leq 4\sqrt{\beta_2 + \beta_3} \xrightarrow[n\to+\infty]{} 0,$$

d'où le résultat. ∎

3.3 Borne inférieure

Avant de s'intéresser à l'inégalité de Berry-Esseen (qui donnera une borne supérieure pour $\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$), nous allons voir que la méthode de Stein permet de donner une borne inférieure pour cette quantité. Le théorème suivant est dû à Barbour et Hall (1984), et c'est grâce à la méthode de Stein qu'ils ont pu l'établir. On utilise toujours le Cadre de travail 1.

4. mathématicien finlandais, 1876-1932.

Théorème 3.5 (Barbour et Hall).
Soient $\xi_1, \dots, \xi_n$ *des variables aléatoires indépendantes telles que pour tout* $1 \leq i \leq n$, $\mathbb{E}[\xi_i] = 0$, $\mathbb{E}[\xi_i^2] = \sigma_i^2 < \infty$ *et* $\sum_{i=1}^{n} \sigma_i^2 = 1$. *Alors il existe une constante universelle* C *telle que, pour tout* $\varepsilon > 0$ *:*

$$\left(1 - e^{-\frac{\varepsilon^2}{4}}\right) \sum_{i=1}^{n} \mathbb{E}\left[\xi_i^2 \mathbb{1}_{\{|\xi_i| > \varepsilon\}}\right] \leq C \left(\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| + \sum_{i=1}^{n} \sigma_i^4 \right),$$

où $\Phi(z) = \mathbb{P}(Z \leq z)$.

Remarque. En 1935, Feller et Lévy ont prouvé que si $\max_{1 \leq i \leq n} \sigma_i^2 \xrightarrow[n \to \infty]{} 0$ (appelée condition de Feller-Lévy) alors la Condition de Lindeberg est nécessaire. Grâce au Théorème 3.5, Barbour et Hall ont pu donner une autre preuve de cette nécessité. En effet, on a par la condition de Feller-Lévy :

$$\sum_{i=1}^{n} \sigma_i^4 \leq \left(\sum_{i=1}^{n} \sigma_i^2 \right) \max_{1 \leq i \leq n} \sigma_i^2 = \max_{1 \leq i \leq n} \sigma_i^2 \xrightarrow[n \to \infty]{} 0.$$

Donc si $\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \xrightarrow[n \to +\infty]{} 0$, alors par le théorème de Barbour et Hall,
$\sum_{i=1}^{n} \mathbb{E}[\xi_i^2] \mathbb{1}_{\{|\xi_i| \geq \varepsilon\}} \xrightarrow[n \to \infty]{} 0$ pour tout $\varepsilon > 0$. Et ainsi, la Condition de Lindeberg est nécessaire.

Démonstration. On rappelle que $Z \sim \mathcal{N}(0, 1)$. Soit h une fonction absolument continue telle que $N(h) = E[h(Z)] < \infty$, que l'on devra choisir plus tard.
Une fois n'est pas coutume, on considère $f := f_h$ la solution de l'équation de Stein (2.8) et d'après le propriété 2.4, on peut écrire que $\mathbb{E}[f'(W) - Wf(W)] = \mathbb{E}[h(W)] - N(h)$. La preuve nécessite plusieurs étapes, qui ne seront pas toutes détaillées ici.

Etape 1 : Si l'on suppose que les variables considérées sont à densité, on a d'une part :

$$\begin{aligned}
|\mathbb{E}[h(W)] - N(h)| &= \left| \int_{-\infty}^{+\infty} h(x)(f_W(x) - f_Z(x))dx \right| \quad \text{où } f_W \text{ et } f_Z \text{ sont les densités respectives de } W \text{ et } Z \\
&\leq \lim_{M \to +\infty} \left| [h(x)(\mathbb{P}(W \leq x) - \Phi(x))]_{-M}^{M} \right| \\
&+ \lim_{M \to +\infty} \left| \int_{-M}^{M} h'(x)(\mathbb{P}(W \leq x) - \Phi(x))dx \right| \quad \text{par intégration par parties} \\
&= \left| \int_{-\infty}^{+\infty} h'(x)(\mathbb{P}(W \leq x) - \Phi(x))dx \right| \quad \text{car } h \text{ est bornée} \\
&\leq \Delta \int_{-\infty}^{+\infty} |h'(x)|dx, \quad \text{où } \Delta := \sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|. \qquad (3.10)
\end{aligned}$$

Remarque. Bien que le calcul précédent ait été réalisé pour des variables à densité, on peut toutefois le généraliser en utilisant l'intégrale de Stieltjes [5] .

5. Si Φ est la fonction de répartition d'une variable aléatoire réelle X, alors $\int_{-\infty}^{+\infty} f(x)d\Phi(x) = \mathbb{E}[f(X)]$.

Etape 2 : Comme $\mathbb{E}[\xi_i] = 0$ et ξ_i et W^i sont indépendantes pour tout $1 \leq i \leq n$, on a d'une part :

$$\begin{aligned} \mathbb{E}[Wf(W)] &= \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 f'(W^i)] + \sum_{i=1}^{n} \mathbb{E}[\xi_i(f(W^i + \xi_i) - f(W^i) - \xi_i f'(W^i))] \\ &= \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[f'(W^i)] + \sum_{i=1}^{n} \mathbb{E}[\xi_i(f(W^i + \xi_i) - f(W^i) - \xi_i f'(W^i))], \end{aligned} \tag{3.11}$$

où l'on a utilisé l'indépendance des ξ_i et W^i pour $1 \leq i \leq n$. D'autre part, on a :

$$\mathbb{E}[f'(W)] = \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[f'(W^i)] + \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[f'(W) - f'(W^i)]. \tag{3.12}$$

Donc,

$$E[f'(W) - Wf(W)] = \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[f'(W) - f'(W^i)] - \sum_{i=1}^{n} \mathbb{E}[\xi_i(f(W^i + \xi_i) - f(W^i) - \xi_i f'(W^i))].$$

Etape 3 : À présent, le but est de borner $|\mathbb{E}[f'(W) - Wf(W)]| = |\mathbb{E}[h(W)] - N(h)|$, afin de donner une borne inférieure pour $\Delta := \sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$. La clé va être de considérer une fonction g, définie par

$$g(w, y) = -y^{-1}(f(w + y) - f(w) - yf'(w)).$$

Dès lors, on a :

$$\begin{aligned} \left| \mathbb{E}[f'(W) - Wf(W)] - \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 g(W^i, \xi_i)] \right| &= \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[f'(W) - f'(W^i)] \\ &\leq \frac{1}{2} \|f'''\|_\infty \sum_{i=1}^{n} \sigma_i^4, \end{aligned} \tag{3.13}$$

où l'on a utilisé l'inégalité de Taylor-Lagrange pour écrire :

$$|f'(W) - f'(W^i) - \xi_i f''(W^i)| \leq \frac{\xi_i^2}{2} \|f'''\|_\infty.$$

Ainsi, puisque $\mathbb{E}[\xi_i] = 0$ et ξ_i et W^i sont indépendants pour tout $1 \leq i \leq n$, cela justifie (3.13).

Etape 4 : Alors, en combinant (3.10) et (3.13), puisque $\mathbb{E}[f'(W) - Wf(W)] = \mathbb{E}[h(W)] - N(h)$, on voit que l'on pourrait donner une borne inférieure pour $\Delta := \sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$, à condition que l'on puisse trouver une fonction h telle que $\displaystyle\int_{-\infty}^{+\infty} |h'(w)| dw < \infty$, et aussi à condition que $\|f'''\|_\infty < \infty$ (où $f := f_h$).

On doit donc trouver une fonction f qui vérifie les conditions précédentes. En pratique, on peut assez facilement trouver une telle fonction et ensuite considérer $h(w) = f'(w) - wf(w)$, de sorte que $N(h)$ tend vers 0 à l'infini, $\displaystyle\int_{-\infty}^{+\infty} |h'(w)| dw < \infty$ et $\|f'''\|_\infty < \infty$.

Toutefois, on remarque que si f est linéaire, alors g est nulle et si f est paire, alors $\mathbb{E}[g(Z, y)]$ est antisymétrique en y. On veut donc utiliser une fonction f qui ne soit ni linéaire, ni paire et qui vérifie les conditions précédentes.

En choisissant $f(w) = we^{-\frac{w^2}{2}}$, il est possible d'arriver au résultat attendu.

➜ Les arguments et calculs qui suivent consistent à borner différentes quantités, en utilisant des résultats d'analyse, dans le but d'arriver à la majoration voulue. Ceci étant calculatoire et n'étant pas directement lié à la méthode de Stein, nous ne donnerons pas plus de détails ici. Toutefois, si le lecteur le souhaite, il peut se référer à l'Appendice section 3.

■

Ces quelques applications de la méthode de Stein nous ont permis de constater que la démarche est souvent la même et que la genèse de toute preuve est la très utile Identité de Stein. Dans la section suivante, nous nous intéresserons alors au cœur de notre problème : la fameuse inégalité de Berry-Esseen qui, elle, donnera une borne supérieure pour $\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$.

4 Inégalité de Berry-Esseen uniforme pour des variables aléatoires indépendantes

Les idées développées précédemment vont en partie être réutilisées ici. Par ailleurs, on utilise toujours les mêmes conditions que celles introduites dans la section 3 ; on les rappelle toutefois pour des raisons pratiques.

Cadre de travail 1

Soient $\xi_1, \dots \xi_n$ des variables aléatoires indépendantes telles que :

$$\forall \quad 1 \leq i \leq n, \quad \mathbb{E}[\xi_i] = 0, \quad \text{et} \quad \sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1.$$

On pose $\gamma := \sum_{i=1}^n \mathbb{E}[|\xi_i|^3]$. En outre, on note $W := \sum_{i=1}^n \xi_i$, $W^i := W - \xi_i$, et on considère

$$\forall t \in \mathbb{R}, \quad K_i(t) = \mathbb{E}[\xi_i(\mathbb{1}_{[0,\xi_i]}(t) - \mathbb{1}_{[\xi_i,0[}(t))].$$

Par ailleurs, $Z \sim \mathcal{N}(0,1)$ et $N(h) = \mathbb{E}[h(Z)]$ pour une fonction h lisse convenable.

Remarque. On a toujours ξ_i et W^i indépendantes pour tout $1 \leq i \leq n$.

Le but de cette partie est de démontrer l'inégalité de Berry-Esseen dans le cas uniforme. On rappelle son énoncé :

Théorème (Inégalité de Berry-Esseen uniforme)**.**
Dans les conditions précédentes, il existe une constante universelle C telle que

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq C\gamma, \tag{4.1}$$

où Φ désigne la fonction de répartition associée à la loi normale centrée réduite.

Pour ce faire, nous étudierons alors deux cas, selon que les variables aléatoires sont bornées ou non.

4.1 Cas de variables aléatoires bornées

Dans la section précédente, on constate (dans le théorème 3.3 par exemple) que les bornes sont en $O(\delta)$ pour les fonctions lipschitziennes, alors qu'elles sont seulement en $O(\sqrt{\delta})$ pour la norme $\sup|.|$. Le but de cette partie est donc d'obtenir une borne pour la norme $\sup|.|$, d'un ordre comparable aux bornes obtenues pour les fonctions lipschitziennes.
On va ici obtenir une majoration de $\sup_{z\in\mathbb{R}}|\mathbb{P}(W \le z) - \Phi(z)|$ lorsque les variables aléatoires considérées sont bornées ; on aura alors une pseudo-inégalité de Berry-Esseen pour des variables aléatoires bornées.

Théorème 4.1 (Inégalité de Berry-Esseen uniforme, cas borné)**.**
Dans les conditions précédentes, on a :

(i)

$$\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i + t \le z)K_i(t)dt - \Phi(z)\right| \le 2,44\gamma. \tag{4.2}$$

(ii) Si de plus, $|\xi_i| \le \delta_0$ *pour* $1 \le i \le n$ *alors,*

$$\sup_{z\in\mathbb{R}}|\mathbb{P}(W \le z) - \Phi(z)| \le 3,3\delta_0. \tag{4.3}$$

Remarque. L'inégalité *(ii)* n'est pas optimale puisque la borne ne dépend pas de γ, comme c'est normalement le cas dans l'inégalité de Berry-Esseen classique, mais elle permet toutefois d'approcher la loi normale centrée réduite.

Démonstration. Encore une fois, l'idée va être d'utiliser la méthode de Stein, on va particulièrement s'inspirer du raisonnement utilisé pour la Construction de l'identité de Stein.
(i) Soit $z \in \mathbb{R}$, soit $f := f_z$ la solution de l'équation (2.1). Par le résultat (2.3) de la propriété 2.1, on sait que $\mathbb{E}[Wf(W)]$ est bien définie.

Etape 1 : Grâce au résultat obtenu en (2.12), on a de façon identique :

$$\begin{aligned}\mathbb{E}[Wf(W)] &= \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}[f'(W^i + t)]K_i(t)dt\\ &= \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}[(W^i + t)f(W^i + t) + \mathbb{1}_{\{W^i+t\le z\}} - \Phi(z)]K_i(t)dt, \quad \text{car } f \text{ est solution de (2.1)}.\end{aligned}$$

En outre, comme $\sum_{i=1}^{n}\int_{-\infty}^{+\infty}K_i(t)dt = \sum_{i=1}^{n}\mathbb{E}[\xi_i^2] = 1$ (cf. propriété 2.5) et $\mathbb{E}[\mathbb{1}_{\{W^i+t\le z\}} - \Phi(z)] = \mathbb{P}(\{W^i + t\} \le z) - \Phi(z)$, on déduit de l'égalité précédente :

$$\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(\{W^i + t\} \le z)K_i(t)dt - \Phi(z) = \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}[Wf(W) - (W^i + t)f(W^i + t)]K_i(t)dt. \tag{4.4}$$

Etape 2 : On va alors borner la quantité précédente.

$$\begin{aligned}
&\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}[Wf(W)-(W^i+t)f(W^i+t)]K_i(t)dt\right| \\
\leq&\sum_{i=1}^{n}\int_{-\infty}^{+\infty}|\mathbb{E}[Wf(W)-(W^i+t)f(W^i+t)]|K_i(t)dt \quad \text{car } K_i(t)\geq 0 \quad \forall t\in\mathbb{R} \\
\leq&\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}[|(W^i+\xi_i)f(W^i+\xi_i)-(W^i+t)f(W^i+t)|]K_i(t)dt \\
\leq&\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}\left[\left(|W^i|+\frac{\sqrt{2\pi}}{4}\right)(|\xi_i|+|t|)\right]K_i(t)dt \quad \text{d'après l'inégalité (2.6).}
\end{aligned}$$

Ainsi, en utilisant l'égalité (4.4), et comme $\mathbb{E}[(W^i)^2]\leq 1$ pour tout $1\leq i\leq n$, on a :

$$\begin{aligned}
\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}((W^i+t)\leq z))K_i(t)dt-\Phi(z)\right| \leq&\left(1+\frac{\sqrt{2\pi}}{4}\right)\sum_{i=1}^{n}\mathbb{E}[|\xi_i|]\mathbb{E}[\xi_i^2]+\frac{1}{2}\mathbb{E}[|\xi_i|^3] \\
\leq&\frac{3}{2}\left(1+\frac{\sqrt{2\pi}}{4}\right)\sum_{i=1}^{n}\mathbb{E}[|\xi_i|^3] \\
\leq&2,44\gamma,
\end{aligned}$$

où on a encore utilisé les Propriétés de la fonction K_i pour la première inégalité et le corollaire 1.1 pour la seconde (même argument que pour (3.2)). D'où le résultat.

(ii) On remarque que si l'on pouvait directement remplacer $\mathbb{P}(W^i+t\leq z)$ par $\mathbb{P}(W\leq z)$ dans *(i)*, alors l'inégalité de Berry-Esseen serait démontrée (car $\sum_{i=1}^{n}\int_{-\infty}^{+\infty}K_i(t)dt=1$).
Cependant, comme $|\xi_i|\leq\delta_0$, alors $K_i(t)=0,\forall|t|>\delta_0$. On va donc considérer le cas où $|t|,|\xi_i|<\delta_0$. On a alors :

$$\mathbb{P}(W^i+t\leq z)=\mathbb{P}(W-\xi_i+t\leq z)\begin{cases}\geq\mathbb{P}(W\leq z-2\delta_0)\\ \leq\mathbb{P}(W\leq z+2\delta_0).\end{cases}$$

Etape 1 : On remplace z par $z+2\delta_0$ dans l'expression obtenue en *(i)*.
Alors, comme Φ est $\frac{1}{\sqrt{2\pi}}$-lipschitzienne, on a :

$$\begin{aligned}
2,44\gamma\geq&\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}((W^i+t)\leq z+2\delta_0))K_i(t)dt-\Phi(z+2\delta_0) \\
\geq&\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W\leq z)K_i(t)dt-\Phi(z+2\delta_0) \\
\geq&\mathbb{P}(W\leq z)-\Phi(z)-\frac{2\delta_0}{\sqrt{2\pi}}, \quad \text{car } \sum_{i=1}^{n}\int_{-\infty}^{+\infty}K_i(t)dt=1.
\end{aligned}$$

De plus, comme $|\xi_i|\leq\delta_0$, alors $\gamma=\sum_{i=1}^{n}\mathbb{E}[|\xi_i|^3]\leq\delta_0\sum_{i=1}^{n}\mathbb{E}[\xi_i^2]=\delta_0$.
D'où, en utilisant l'inégalité précédente, on obtient,

$$\mathbb{P}(W\leq z)-\Phi(z)\leq 2,44\gamma-\frac{2\delta_0}{\sqrt{2\pi}}\leq 2,44\delta_0-\frac{2\delta_0}{\sqrt{2\pi}}\leq 3,3\delta_0.$$

Etape 2 : Pour obtenir la borne inférieure, on utilise exactement le même raisonnement mais en remplaçant cette fois z par $z - 2\delta_0$. On obtient alors $\mathbb{P}(W \leq z) - \Phi(z) \geq -3,3\delta_0$.

Ainsi, en combinant les deux résultats précédents et par passage au sup, on obtient aisément

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 3,3\delta_0,$$

d'où le résultat. ∎

4.2 Cas général

On a vu dans la partie précédente comment obtenir une sorte d'inégalité de Berry-Esseen lorsque que les variables en jeu sont bornées. On va ici oublier toute restriction de bornitude et l'on verra comment démontrer l'inégalité générale dans ce cas. On utilise une fois encore le même Cadre de travail 1.
On remarque que le résultat (i) du Théorème 4.1 reste vrai si les variables aléatoires considérées ne sont pas bornées. Pour traiter le cas général, nous allons alors nous appuyer sur ce dernier résultat. L'idée sera d'utiliser, une fois encore, l'identité de Stein (2.15) et la caractérisation de la loi $\mathcal{N}(0,1)$ (propriété 2.2). Avant toute chose, nous avons besoin d'introduire une inégalité de concentration [6].

4.2.1 Inégalité de concentration

Propriété 4.1 (Inégalité de concentration uniforme)**.** *Pour tous $a < b$ réels et tout $1 \leq i \leq n$, on a :*

$$\mathbb{P}(a \leq W^i \leq b) \leq \sqrt{2}(b-a) + (1+\sqrt{2})\gamma. \tag{4.5}$$

Démonstration. On va s'inspirer de la fonction K_i introduite lors de la Construction de l'identité de Stein pour une somme de variables aléatoires indépendantes.
Soit $\delta = \frac{\gamma}{2}$. Posons

$$f(w) = \begin{cases} -\dfrac{1}{2}(b-a) - \delta & \text{si } w < a - \delta \\ w - \dfrac{1}{2}(b+a) & \text{si } a - \delta \leq w \leq b + \delta \\ \dfrac{1}{2}(b-a) + \delta & \text{si } w > b + \delta. \end{cases}$$

Alors, $f' = \mathbb{1}_{[a-\delta, b+\delta]} \geq 0$ et $\|f\|_\infty = \frac{1}{2}(b-a) + \delta$. On considère ensuite, à l'image de la fonction K_i, pour tout $t \in \mathbb{R}$,

$$\begin{cases} \overline{M_j(t)} = \xi_j \left(\mathbb{1}_{[-\xi_j,0]}(t) - \mathbb{1}_{[0,-\xi_j]}(t)\right) \geq 0 & \text{(mêmes arguments que pour le \textit{(i)} de la propriété 2.5)} \\ \overline{M(t)} = \displaystyle\sum_{j=1}^{n} \overline{M_j(t)} \geq 0 & \\ M(t) = \mathbb{E}[\overline{M(t)}]. & \end{cases}$$

Par ailleurs, on rappelle que : ✓ ξ_j et $W^i - \xi_j$ sont indépendants pour $1 \leq i, j \leq n \quad (j \neq i)$,
✓ ξ_i et W^i sont indépendants pour $1 \leq i \leq n$.

On va chercher à encadrer la quantité $\mathbb{E}[W^i f(W^i)] - \mathbb{E}[\xi_i f(W^i - \xi_i)]$ où ξ_i et W^i ont déjà été définies dans le Cadre de travail 1.

6. Les inégalités de concentration fournissent des bornes sur la probabilité qu'une variable aléatoire dévie d'une certaine valeur.

Etape 1 : Par définition de W^i, par indépendance de ξ_j et $W^i - \xi_j$ et comme $\mathbb{E}[\xi_j] = 0$ pour $1 \leq j \leq n$, il vient :

$$\begin{aligned}
\mathbb{E}[W^i f(W^i)] - \mathbb{E}[\xi_i f(W^i - \xi_i)] &= \sum_{j=1}^{n} \mathbb{E}[\xi_j (f(W^i) - f(W^i - \xi_j))] \\
&= \sum_{j=1}^{n} \mathbb{E}\left[\xi_j \int_{-\xi_j}^{0} f'(W^i + t) dt\right] \\
&= \sum_{j=1}^{n} \mathbb{E}\left[\int_{-\infty}^{+\infty} f'(W^i + t) \overline{M_j(t)} dt\right] \\
&= \mathbb{E}\left[\int_{-\infty}^{+\infty} f'(W^i + t) \overline{M(t)} dt\right], \quad \text{par définition de } \overline{M(t)}.
\end{aligned}$$

Comme $f' \geq 0$ et $\overline{M(t)} \geq 0 \quad \forall t \in \mathbb{R}$, on a de surcroît :

$$\begin{aligned}
\mathbb{E}\left[\int_{-\infty}^{+\infty} f'(W^i + t) \overline{M(t)} dt\right] &\geq \mathbb{E}\left[\int_{|t|<\delta} f'(W^i + t) \overline{M(t)} dt\right] \\
&\geq \mathbb{E}\left[\mathbb{1}_{[a,b]}(W^i) \int_{|t|<\delta} \overline{M(t)} dt\right] \\
&= \mathbb{E}\left[\mathbb{1}_{[a,b]}(W^i) \sum_{j=1}^{n} |\xi_j| \min(\delta, |\xi_j|)\right] \\
&\geq H_{1,1} - H_{1,2}, \qquad (4.6)
\end{aligned}$$

avec $H_{1,1} := \mathbb{P}(a \leq W^i \leq b) \sum_{j=1}^{n} \mathbb{E}\left[|\xi_j| \min(\delta, |\xi_j|)\right]$,

$$H_{1,2} := \mathbb{E}\left[\left|\sum_{j=1}^{n} |\xi_j| \min(\delta, |\xi_j|) - \mathbb{E}[|\xi_j| \min(\delta, |\xi_j|)]\right|\right].$$

Etape 2 : Minorons tout d'abord $H_{1,1}$.
On utilise pour cela le fait que pour $x, y > 0$, $\min(x, y) \geq x - \frac{x^2}{4y}$(cf. Appendice section 4, lemme 6.3). Cela implique alors que

$$\sum_{j=1}^{n} \mathbb{E}[|\xi_j| \min(\delta, |\xi_j|)] \geq \sum_{j=1}^{n} \left(\mathbb{E}[\xi_j^2] - \frac{\mathbb{E}[|\xi_j|^3]}{4\delta}\right) = \frac{1}{2} \quad (\text{car } \delta = \frac{\gamma}{2}), \qquad (4.7)$$

et à fortiori,

$$H_{1,1} \geq \frac{1}{2} \mathbb{P}(a \leq W^i \leq b). \qquad (4.8)$$

Etape 3 : Il nous faut maintenant majorer $H_{1,2}$. Comme $H_{1,2}$ est l'espérance d'une variable aléatoire centrée, par l'inégalité de Cauchy-Schwarz, on a en outre,

$$H_{1,2} \leq \mathbb{V}\left[\sum_{j=1}^{n} |\xi_j| \min(\delta, |\xi_j|)\right]^{\frac{1}{2}} \leq \left(\sum_{j=1}^{n} \mathbb{E}[\xi_j^2 \min(\delta, |\xi_j|)^2]\right)^{\frac{1}{2}} \leq \delta \left(\sum_{j=1}^{n} \mathbb{E}[\xi_j^2]\right)^{\frac{1}{2}} = \delta. \qquad (4.9)$$

Ainsi, puisque $\mathbb{E}[W^i f(W^i)] - \mathbb{E}[\xi_j f(W^i - \xi_i)] = \mathbb{E}\left[\int_{-\infty}^{+\infty} f'(W^i + t)\overline{M(t)}dt\right]$, avec (4.6), (4.8) et (4.9), on déduit :

$$\mathbb{E}[W^i f(W^i)] - \mathbb{E}[\xi_j f(W^i - \xi_i)] \geq \frac{1}{2}\mathbb{P}(a \leq W^i \leq b) - \delta. \tag{4.10}$$

Etape 4 : Pour pouvoir majorer $\mathbb{P}(a \leq W^i \leq b)$, on va donner une borne supérieure pour l'espérance précédente. Pour cela, remarquons que, comme $\|f\|_\infty = \frac{1}{2}(b-a) + \delta$, on a :

$$\begin{aligned}
\mathbb{E}[W^i f(W^i)] - \mathbb{E}[\xi_j f(W^i - \xi_i)] \leq & (\frac{1}{2}(b-a) + \delta)(\mathbb{E}[|W^i|] + \mathbb{E}[|\xi_i|]) \\
\leq & \frac{1}{\sqrt{2}}(\mathbb{E}[|W^i|]^2 + \mathbb{E}[|\xi_i|]^2)^{\frac{1}{2}}(b-a+2\delta) \\
\leq & \frac{1}{\sqrt{2}}(\mathbb{E}[|W^i|^2] + \mathbb{E}[|\xi_i|^2])^{\frac{1}{2}}(b-a+2\delta) \quad \text{par l'inégalité de Jensen, avec } x \mapsto x^2 \text{ convexe} \\
= & \frac{1}{\sqrt{2}}(\mathbb{E}[|W^i + \xi_i|^2])^{\frac{1}{2}}(b-a+2\delta) \quad \text{car } \xi_i \text{ et } W^i \text{ sont indépendantes et } \mathbb{E}[\xi_i] = 0 \text{ pour tout } 1 \leq i \leq n \\
= & \frac{1}{\sqrt{2}}(\mathbb{E}[|W|^2])^{\frac{1}{2}}(b-a+2\delta) \\
= & \frac{1}{\sqrt{2}}(\mathbb{V}[W])^{\frac{1}{2}}(b-a+2\delta) \quad \text{car } \mathbb{E}[W] = 0 \\
= & \frac{1}{\sqrt{2}}(b-a+2\delta).
\end{aligned} \tag{4.11}$$

Etape 5 : Il ne reste plus qu'à combiner les résultats précédents pour obtenir l'encadrement souhaité. Dès lors, avec (4.10) et (4.11), on obtient :

$$\frac{1}{2}\mathbb{P}(a \leq W^i \leq b) - \delta \leq \frac{1}{\sqrt{2}}(b-a+2\delta),$$

et à fortiori

$$\begin{aligned}
\mathbb{P}(a \leq W^i \leq b) \leq & \sqrt{2}(b-a) + 2(1+\sqrt{2})\delta \\
= & \sqrt{2}(b-a) + 2(1+\sqrt{2})\gamma \quad \text{car } \delta = \frac{\gamma}{2}.
\end{aligned}$$

D'où l'Inégalité de concentration uniforme.

■

4.2.2 Résultat final

On peut à présent démontrer l'inégalité uniforme générale de Berry-Esseen, pour des variables aléatoires indépendantes.

Théorème 4.2 (Inégalité de Berry-Esseen uniforme)**.**
Soient $\xi_1, \ldots \xi_n$ des variables aléatoires indépendantes, admettant un moment d'ordre 3, telles que :

$$\forall 1 \leq i \leq n, \quad \mathbb{E}[\xi_i] = 0 \quad et \quad \sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1.$$

Posons $W := \sum_{i=1}^n \xi_i, \quad W^i := W - \xi_i, \quad$ *et* $\quad \gamma := \sum_{i=1}^n \mathbb{E}[|\xi_i|^3]$. *Alors*

$$\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 7\gamma. \tag{4.12}$$

Remarque. La constante a été améliorée au fil des années. Initialement, Berry et Esseen avaient prouvé que $C \leq 7$. Cette valeur a très rapidement chuté et Ilya Tyurin a montré en 2010 que $C \leq 0,4785$.

Démonstration. On utilise toujours les notations du Cadre de travail 1. On va user de nombreux résultats précédents, notamment l'Inégalité de concentration uniforme et le Théorème 4.1. Afin de s'inspirer de la démonstration de ce dernier, on va s'intéresser à la quantité $\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt-\mathbb{P}(W\leq z)\right|$.
Comme $\mathbb{P}(W \leq z) = \mathbb{P}(W^i \leq z - \xi_i)$, on a pour $t \in \mathbb{R}$,

$$(\ast)\quad |\mathbb{P}(W^i+t\leq z)-\mathbb{P}(W\leq z)| \leq \mathbb{P}\left(z-\max(t,\xi_i)\leq W^i\leq z-\min(t,\xi_i)\right).$$

Etape 1 : Comme $\sum_{i=1}^{n}\int_{-\infty}^{+\infty}K_i(t)dt = 1$, et en utilisant $(\ast)$, on a d'une part :

$$\begin{aligned}
&\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt-\mathbb{P}(W\leq z)\right|\\
&\leq \sum_{i=1}^{n}\int_{-\infty}^{+\infty}|\mathbb{P}(W^i+t\leq z)-\mathbb{P}(W\leq z)|K_i(t)dt\\
&\leq \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}\left(z-\max(t,\xi_i)\leq W^i\leq z-\min(t,\xi_i)\right)K_i(t)dt\\
&= \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}\left[\mathbb{P}\left(z-\max(t,\xi_i)\leq W^i\leq z-\min(t,\xi_i)\mid \xi_i\right)\right]K_i(t)dt
\end{aligned}\tag{4.13}$$

En effet, pour $1 \leq i \leq n$ fixé, en posant $A := \{z - \max(t, \xi_i) \leq W^i \leq z - \min(t, \xi_i)\}$, on a :

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[\mathbb{1}_A \mid \xi_i]] = \mathbb{E}[\mathbb{P}(A \mid \xi_i)],\quad \text{et la dernière écriture de } \mathbb{P}(A) \text{ justifie (4.13).}$$

Or, en considérant $f(s,w) := \mathbb{1}_{\{z-\max(t,s)\leq w\leq z-\min(t,s)\}}$, puisque ξ_i et W^i sont indépendantes, on a d'après un résultat vu en TD de processus stochastiques [4] :

$$\begin{aligned}
\mathbb{E}[f(\xi_i,W^i)\mid \xi_i=s]=\mathbb{E}[f(s,W^i)] &=\mathbb{E}[\mathbb{1}_{\{z-\max(t,s)\leq W^i\leq z-\min(t,s)\}}]\\
&=\mathbb{P}\left(z-\max(t,s)\leq W^i\leq z-\min(t,s)\right).
\end{aligned}\tag{4.14}$$

Ainsi, avec la définition de A, f et (4.14), on a :

$$\mathbb{E}[\mathbb{1}_A \mid \xi_i] = \mathbb{P}\left(z-\max(t,\xi_i)\leq W^i\leq z-\min(t,\xi_i)\right).$$

Dès lors, en utilisant, (4.13) et la remarque précédente, il vient par l'Inégalité de concentration uniforme :

$$\begin{aligned}
&\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt-\mathbb{P}(W\leq z)\right|\\
&\leq \sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{E}\left[\sqrt{2}(|t|+|\xi_i|)+(1+\sqrt{2})\gamma\right]K_i(t)dt\\
&= (1+\sqrt{2})\gamma+\sqrt{2}\sum_{i=1}^{n}\left(\frac{1}{2}\mathbb{E}[|\xi_i|^3]+\mathbb{E}[|\xi_i|]\mathbb{E}[\xi_i^2]\right)\quad \text{d'après les Propriétés de la fonction } K_i\\
&\leq \left(1+\frac{5}{2}\right)\gamma\quad \text{d'après le corollaire 1.1.}
\end{aligned}\tag{4.15}$$

Etape 2 : D'autre part, d'après le Théorème 4.1, on sait aussi que

$$\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt-\Phi(z)\right|\leq 2,44\gamma.$$

Ainsi, avec (4.15), on a par inégalité triangulaire :

$$\begin{aligned}|\mathbb{P}(W\leq z)-\Phi(z)|\leq&\left|\mathbb{P}(W\leq z)-\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt\right|\\&+\left|\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\mathbb{P}(W^i+t\leq z)K_i(t)dt-\Phi(z)\right|\\\leq&\left(1+\frac{5}{2}\right)\gamma+2,44\gamma\\\leq&7\gamma,\end{aligned}$$

d'où le résultat par passage au sup.

■

L'inégalité de Berry-Esseen uniforme est donc établie, et par les résultats qui ont été mis en œuvre pour la démontrer, on constate la grande utilité de la méthode de Stein, notamment l'identité de Stein.

5 Inégalité de Berry-Esseen non uniforme pour des variables aléatoires indépendantes

Il convient à présent de s'intéresser à l'inégalité de Berry-Esseen non uniforme. Comme précédemment, on va distinguer deux cas, selon que les variables aléatoires considérées sont bornées ou non. De même que pour le cas uniforme, nous travaillons avec des sommes de variables aléatoires indépendantes et nous allons manipuler des formules et des idées similaires, avec quelques variables supplémentaires. On reconnaîtra alors dans cette section des calculs similaires à certains déjà menés auparavant. Dans toute la suite, on considère les conditions suivantes.

Cadre de travail 2

Soient $\xi_1,\dots\xi_n$ des variables aléatoires indépendantes telles que :

$$\forall\quad 1\leq i\leq n,\quad \mathbb{E}[\xi_i]=0,\quad \text{et}\quad \sum_{i=1}^{n}\mathbb{E}[\xi_i^2]=1.$$

On pose $\gamma:=\sum_{i=1}^{n}\mathbb{E}[|\xi_i|^3]$ et l'on note $W:=\sum_{i=1}^{n}\xi_i, W^i:=W-\xi_i$.

En outre, on considère $\overline{\xi_i}:=\xi_i\mathbb{1}_{\{|\xi_i|\leq 1\}}$, $\overline{W}:=\sum_{i=1}^{n}\overline{\xi_i}$ et $\overline{W}^i:=\overline{W}-\overline{\xi_i}$.

On a toujours $Z\sim\mathcal{N}(0,1)$.

5.1 Cas de variables aléatoires bornées

On va ici s'intéresser aux variables $\overline{W}$ qui, par leur définition, s'avèrent être bornées. Pour utiliser les mêmes idées que dans la démonstration du cas uniforme, nous avons besoin d'une inégalité de concentration, non uniforme cette fois-ci.

5.1.1 Inégalité de concentration non uniforme

Propriété 5.1 (Inégalité de concentration non uniforme). *Dans les conditions précédentes, pour tous $a < b$ réels et tout $1 \le i \le n$, on a :*

$$\mathbb{P}(a \le \overline{W}^i \le b) \le e^{-\frac{a}{2}}(5(b-a) + 7\gamma). \tag{5.1}$$

Avant de démontrer cette propriété, nous avons besoin du résultat suivant.

Lemme 5.1. *Soient $\alpha > 0, B \in \mathbb{R}$. On considère $\eta_1, \dots, \eta_n$ des variables aléatoires indépendantes telles que pour tout $1 \le i \le n$, $\mathbb{E}[\eta_i] \le 0$, $\eta_i \le \alpha$ et $\sum_{i=1}^{n} \mathbb{E}[\eta_i^2] \le B^2$. On pose $S := \sum_{i=1}^{n} \eta_i$. Alors :*

$$\mathbb{E}\left[e^{tS}\right] \le \exp\left(\frac{e^{t\alpha} - 1 - t\alpha}{\alpha^2} B^2\right) \qquad \text{pour } t > 0 \quad \text{(Inégalité de Bennett)}, \tag{5.2}$$

$$\mathbb{P}(S \ge x) \le \exp\left(-\frac{B^2}{\alpha^2}\left[\left(1 + \frac{\alpha x}{B^2}\right)\ln\left((1 + \frac{\alpha x}{B^2}\right) - \frac{\alpha x}{B^2}\right]\right), \tag{5.3}$$

$$\mathbb{P}(S \ge x) \le \exp\left(-\frac{x^2}{2(B^2 + \alpha x)}\right) \qquad \text{pour } x > 0. \tag{5.4}$$

Démonstration. On ne démontre que **l'inégalité de Bennett**, puisque les deux autres en découlent. Leur démonstration est néanmoins présentée dans l'Appendice section 5, lemme 6.5.
Remarquons que la fonction $f : s \mapsto \dfrac{e^s - 1 - s}{s^2}$ est croissante sur $\mathbb{R}$, car prolongeable par continuité en 0 (cf. Appendice section 5, lemme 6.4).
Donc pour $s \le \alpha$ et $t > 0$, il vient

$$\frac{(e^{ts} - 1 - ts)}{(ts)^2} \le \frac{(e^{t\alpha} - 1 - t\alpha)}{(t\alpha)^2}, \quad \text{et ainsi} \quad e^{ts} \le 1 + ts + (ts)^2 \frac{(e^{t\alpha} - 1 - t\alpha)}{(t\alpha)^2}.$$

Puis on a :

$$\begin{aligned}
\mathbb{E}[e^{tS}] &= \prod_{i=1}^{n} \mathbb{E}[e^{t\eta_i}] \quad \text{puisque que les } \eta_i \text{ sont indépendantes} \\
&\le \prod_{i=1}^{n} \mathbb{E}\left[1 + t\eta_i + (t\eta_i)^2 \frac{(e^{t\alpha} - 1 - t\alpha)}{(t\alpha)^2}\right] \\
&\le \prod_{i=1}^{n} (1 + t\mathbb{E}[\eta_i] + \frac{1}{\alpha^2}(e^{t\alpha} - 1 - t\alpha)\mathbb{E}[\eta_i^2]) \\
&\le \prod_{i=1}^{n} (1 + \frac{1}{\alpha^2}(e^{t\alpha} - 1 - t\alpha)\mathbb{E}[\eta_i^2]) \quad \text{car } \mathbb{E}[\eta_i] \le \alpha \text{ et } t\alpha > 0 \\
&= \exp\left(\ln\left(\prod_{i=1}^{n} (1 + \frac{1}{\alpha^2}(e^{t\alpha} - 1 - t\alpha)\mathbb{E}[\eta_i^2])\right)\right)
\end{aligned}$$

$$\begin{aligned}\mathbb{E}[e^{tS}] \leq &\exp\left(\sum_{i=1}^{n}\ln(1+\frac{1}{\alpha^2}(e^{t\alpha}-1-t\alpha)\mathbb{E}[\eta_i^2])\right)\\ \leq &\exp\left(\sum_{i=1}^{n}(\frac{1}{\alpha^2}(e^{t\alpha}-1-t\alpha)\mathbb{E}[\eta_i^2])\right) \quad \text{par concavité de ln}\\ \leq &\exp\left(\frac{e^{t\alpha}-1-t\alpha}{\alpha^2}B^2\right), \quad \text{puisque } \textstyle\sum_{i=1}^{n}\mathbb{E}[\eta_i^2] \leq B^2.\end{aligned}$$

D'où l'inégalité de Bennett (5.2). ∎

Nous pouvons à présent démontrer l'Inégalité de concentration non uniforme.

Démonstration. On va fortement s'inspirer de la preuve de l'Inégalité de concentration uniforme. On se place dans le Cadre de travail 2 introduit en début de section.

Etape 1 : D'après le lemme 5.1, on voit qu'en prenant $\alpha = 1$ et $B^2 = 1$, on a :

$$\begin{aligned}\mathbb{P}(a \leq \overline{W}^i \leq b) \leq &e^{-\frac{a}{2}}\mathbb{E}\left[e^{\frac{\overline{W}^i}{2}}\right]\\ \leq &e^{-\frac{a}{2}}\exp(e^{\frac{1}{2}}-\frac{3}{2}) \quad \text{d'après l'inégalité de Bennett (5.2) avec } t = 1/2\\ \leq &1,19e^{-\frac{a}{2}}.\end{aligned}$$

On voit donc que si $7\gamma \geq 1,19$, le résultat est démontré. Supposons à présent que $\gamma \leq \dfrac{1,19}{7} = 0,17$. Soit $\delta = \frac{\gamma}{2}$ donc $\delta \leq 0,085$. Posons

$$f(w) = \begin{cases} 0 & \text{si } w < a-\delta\\ e^{\frac{w}{2}}(w-a+\delta) & \text{si } a-\delta \leq w \leq b+\delta\\ e^{\frac{w}{2}}(b-a+2\delta) & \text{si } w > b+\delta.\end{cases} \tag{5.5}$$

Alors $f'(w) \geq 0$ et $f'(w) \geq e^{\frac{w}{2}}$ pour $a-\delta \leq w \leq b+\delta$.
Posons ensuite, pour tout $t \in \mathbb{R}$, à l'image de la preuve de l'inégalité de concentration uniforme,

$$\begin{cases}\overline{M_i(t)} = \xi_i\left(\mathbb{1}_{[-\overline{\xi_i},0]}(t) - \mathbb{1}_{[0,-\overline{\xi_i}]}(t)\right) \geq 0 \quad \text{(mêmes arguments que pour le (i) de la propriété 2.5)}\\ \overline{M^i(t)} = \sum_{j\neq i}\overline{M_j(t)} \geq 0.\end{cases}$$

La démarche qui va suivre sera quelque peu similaire à celle de l' Inégalité de concentration uniforme.

Etape 2 : Par définition de $\overline{W}^i$, on obtient :

$$\begin{aligned}\mathbb{E}[W^i f(\overline{W}^i)] = &\sum_{j\neq i}\mathbb{E}\left[\xi_i(f(\overline{W}^i) - f(\overline{W}^i - \overline{\xi_i})\right]\\ = &\sum_{j\neq i}\mathbb{E}\left[\int_{-\infty}^{+\infty} f'(\overline{W}^i + t)\overline{M_j(t)}dt\right]\\ = &\mathbb{E}\left[\int_{-\infty}^{+\infty} f'(\overline{W}^i + t)\overline{M^i(t)}dt\right].\end{aligned}$$

Comme $f' \geq 0$ et $\overline{M^i(t)} \geq 0$, alors,

$$\begin{aligned}
\mathbb{E}\left[\int_{-\infty}^{+\infty} f'(\overline{W}^i + t)\overline{M^i(t)}dt\right] \geq & \mathbb{E}\left[\mathbb{1}_{[a,b]}(\overline{W}^i) \int_{|t|\leq\delta} f'(\overline{W}^i + t)\overline{M^i(t)}dt\right] \\
\geq & \mathbb{E}\left[\exp(\frac{\overline{W}^i - \delta}{2})\mathbb{1}_{[a,b]}(\overline{W}^i) \int_{|t|\leq\delta} \overline{M^i(t)}dt\right] \\
\geq & \mathbb{E}\left[\exp(\frac{\overline{W}^i - \delta}{2})\mathbb{1}_{[a,b]}(\overline{W}^i) \sum_{j\neq i} |\xi_j| \min(\delta, |\overline{\xi_j}|)\right] \\
\geq & e^{-\frac{\delta}{2}}(H_{2,1} - H_{2,2}), \qquad (5.6)
\end{aligned}$$

avec $H_{2,1} := \mathbb{E}\left[e^{\frac{\overline{W}^i}{2}}\mathbb{1}_{\{a\leq\overline{W}^i\leq b\}}\right]\sum_{j\neq i}\mathbb{E}[|\xi_j|\min(\delta, |\overline{\xi_j}|)]$,

$H_{2,2} := \mathbb{E}\left[e^{\frac{\overline{W}^i}{2}}\left|\sum_{j\neq i}|\xi_j|\min(\delta, |\overline{\xi_j}|) - \mathbb{E}[|\xi_j|\min(\delta, |\overline{\xi_j}|)]\right|\right]$.

Etape 3 : On minore d'abord $H_{2,1}$. On se rappelle que $\delta \leq 0,085$ et $\gamma \leq 0,17$, on va procéder de la même manière que pour l'Inégalité de concentration uniforme. On a :

$$\begin{aligned}
\sum_{j\neq i}\mathbb{E}[|\xi_j|\min(\delta, |\overline{\xi_j}|)] = & \sum_{j\neq i}\mathbb{E}[|\xi_j|(\min(\delta, |\xi_j|) - \delta\mathbb{1}_{\{\xi_j>1\}})] \quad \text{par définition des } \overline{\xi_i} \\
\geq & -\delta\mathbb{E}[\xi_i] + \sum_{j=1}^{n}\mathbb{E}[|\xi_j|\min(\delta, |\xi_j|)] - \delta\gamma \\
\geq & -\delta\gamma^{\frac{1}{3}} + \frac{1}{2} - \delta\gamma \qquad \text{de même que pour (4.7), en utilisant le lemme 6.3} \\
\geq & -0,085 \times (0,17)^{\frac{1}{3}} + \frac{1}{2} - 0,085 \times 0,17 \\
\geq & 0,43.
\end{aligned}$$

Dès lors, avec l'**étape 1** et la définition de $H_{2,1}$:

$$H_{2,1} \geq 0,43e^{\frac{a}{2}}\mathbb{P}(a \leq \overline{W}^i \leq b). \qquad (5.7)$$

Etape 4 : Il nous faut aussi majorer $H_{2,2}$. On a par le lemme 5.1,

$$\mathbb{E}\left[e^{\overline{W}^i}\right] \leq \exp(e-2).$$

Comme pour l'Inégalité de concentration uniforme, on a par l'inégalité de Cauchy-Schwarz,

$$\begin{aligned}
H_{2,2} \leq \mathbb{E}\left[e^{\overline{W}^i}\right]^{\frac{1}{2}}\left(\mathbb{V}\left[\sum_{j\neq i}|\xi_j|\min(\delta, |\overline{\xi_j}|)\right]\right)^{\frac{1}{2}} \leq & \exp(\frac{e}{2} - 1)\left(\sum_{j\neq i}\mathbb{E}[\xi_j^2\min(\delta, |\overline{\xi_j}|)^2]\right)^{\frac{1}{2}} \\
\leq & \exp(\frac{e}{2} - 1)\delta \\
\leq & 1,44\delta. \qquad (5.8)
\end{aligned}$$

En combinant (5.6), (5.7), (5.8), on obtient alors :

$$\mathbb{E}[W^i f(\overline{W}^i)] \geq e^{-\frac{\delta}{2}}(0,43e^{\frac{a}{2}}\mathbb{P}(a \leq \overline{W}^i \leq b) - 1,44\delta). \qquad (5.9)$$

Etape 5 : Afin de majorer $\mathbb{P}(a \leq \overline{W}^i \leq b)$, on détermine une borne inférieure pour l'espérance précédente. Pour cela, on utilise le fait que par la définition de f (5.5), on a :

$$\begin{aligned}
\mathbb{E}[W^i f(\overline{W}^i)] \leq& (b-a+2\delta)\mathbb{E}\left[|W^i| e^{\frac{\overline{W}^i}{2}}\right] \\
\leq& (b-a+2\delta)\mathbb{E}[|W^i|^2]^{\frac{1}{2}}\mathbb{E}\left[e^{\overline{W}^i}\right]^{\frac{1}{2}} \\
\leq& (b-a+2\delta)\exp(e-2) \\
\leq& 2,06(b-a+2\delta). \qquad (5.10)
\end{aligned}$$

On peut maintenant en déduire le résultat voulu. Ainsi, en utilisant (5.9) et (5.10), on a :

$$\begin{aligned}
\mathbb{P}(a \leq \overline{W}^i \leq b) \leq& \frac{e^{\frac{-a}{2}}}{0,43}(e^{\frac{\delta}{2}} \times 2,06(b-a+2\delta) + 1,44\gamma) \\
\leq& \frac{e^{\frac{-a}{2}}}{0,43}(e^{0,0425} \times 2,06(b-a+2\delta) + 1,44\gamma) \\
\leq& e^{\frac{-a}{2}}(5(b-a) + 13,4\delta) \\
\leq& e^{\frac{-a}{2}}(5(b-a) + 7\gamma), \quad \text{car } \delta = \tfrac{\gamma}{2}.
\end{aligned}$$

D'où l' Inégalité de concentration non uniforme.

■

5.1.2 Résultat final

On peut maintenant établir une inégalité de Berry-Esseen non uniforme dans le cas borné (par définition des $\overline{W}$).

Théorème 5.1 (Inégalité de Berry-Esseen non uniforme, cas borné)**.**
Avec les notations du Cadre de travail 2, il existe une constante universelle C telle que

$$\forall z \in \mathbb{R}, \quad |\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \leq \ C\gamma e^{-\frac{z}{2}}.$$

Démonstration. On va s'inspirer des démonstrations faites auparavant. Plus particulièrement, on va utiliser le même raisonnement que pour la Construction de l'identité de Stein. Sans perte de généralités, on peut supposer $z \geq 0$ (quitte à considérer $-\overline{W}$).

Etape 1 : Soit $f := f_z$ la solution de l'équation (2.1). Posons, à l'image de la fonction K_i,

$$\forall t \in \mathbb{R}, \overline{K_i(t)} = \mathbb{E}[\overline{\xi_i}(\mathbb{1}_{[0,\overline{\xi_i}]}(t) - \mathbb{1}_{[\overline{\xi_i},0[}(t))].$$

En s'inspirant du calcul fait pour la Construction de l'identité de Stein, on a :

$$\begin{aligned}
\mathbb{E}[\overline{W}f(\overline{W})] =& \sum_{i=1}^n \mathbb{E}[\overline{\xi_i} f(\overline{W}^i + \overline{\xi_i})] \\
=& \sum_{i=1}^n \mathbb{E}[\overline{\xi_i}(f(\overline{W}) - f(\overline{W}^i))] + \sum_{i=1}^n \mathbb{E}[\overline{\xi_i}]\mathbb{E}[f(\overline{W}^i)] \\
=& \sum_{i=1}^n \mathbb{E}\left[\overline{\xi_i}\int_{-\infty}^1 f'(\overline{W}^i + t)\overline{\xi_i}(\mathbb{1}_{[0,\overline{\xi_i}]}(t) - \mathbb{1}_{[\overline{\xi_i},0[}(t))dt\right] \\
&+ \sum_{i=1}^n \mathbb{E}[\overline{\xi_i}]\mathbb{E}[f(\overline{W}^i)] \quad \text{car par définition de } \overline{\xi_i}, \text{ on a } \overline{\xi_i} \leq 1 \\
=& \sum_{i=1}^n \int_{-\infty}^1 f'(\overline{W}^i + t)\overline{K_i(t)}dt + \sum_{i=1}^n \mathbb{E}[\overline{\xi_i}]\mathbb{E}[f(\overline{W}^i)]. \qquad (5.11)
\end{aligned}$$

Comme $\sum_{i=1}^{n}\int_{-\infty}^{1}\overline{K_i(t)}dt=\sum_{i=1}^{n}\mathbb{E}[\overline{\xi_i}^2]=1-\sum_{i=1}^{n}\mathbb{E}[\xi_i^2\mathbb{1}_{\{\xi_i>1\}}]$, on obtient, puisque f est solution de (2.1) :

$$\begin{aligned}\mathbb{P}(\overline{W}\leq z)-\Phi(z)=&\mathbb{E}[f'(\overline{W})]-\mathbb{E}[\overline{W}f(\overline{W})]\\ =&\sum_{i=1}^{n}\mathbb{E}[\xi_i^2\mathbb{1}_{\{\xi_i>1\}}]\mathbb{E}[f'(\overline{W})]\\ &+\sum_{i=1}^{n}\int_{-\infty}^{1}\mathbb{E}[f'(\overline{W}^i+\overline{\xi}_i)-f'(\overline{W}^i+t)]\overline{K_i(t)}dt\\ &+\sum_{i=1}^{n}\mathbb{E}[\xi_i\mathbb{1}_{\{\xi_i>1\}}]\mathbb{E}[f(\overline{W}^i)]\\ :=&R_1+R_2+R_3.\end{aligned}\tag{5.12}$$

Etape 2 : **Dans ce qui suit, C désigne une constante arbitraire, que l'on notera toujours pareil mais dont la valeur peut différer.**
On va donner des bornes pour R_1 et R_3. Comme $f:=f_z$ est solution de (2.1), on déduit de (2.2) que

$$f'(w)=\begin{cases}(\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)+1)[1-\Phi(z)] & \text{si } w\leq z\\ (\sqrt{2\pi}we^{\frac{w^2}{2}}(1-\Phi(w))-1)\Phi(z) & \text{si } w>z.\end{cases}\tag{5.13}$$

✓ Par les résultats (2.4),et (5.3), on a :

$$\begin{aligned}\mathbb{E}[|f'(\overline{W})|]=&\mathbb{E}[|f'(\overline{W})|\mathbb{1}_{\{\overline{W}\leq\frac{z}{2}\}}]+\mathbb{E}[|f'(\overline{W})|\mathbb{1}_{\{\overline{W}>\frac{z}{2}\}}]\\ \leq&\left(1+\sqrt{2\pi}\frac{z}{2}e^{\frac{z^2}{8}}\Phi(w)\right)(1-\Phi(z))+\mathbb{P}(\overline{W}\geq\frac{z}{2})\quad\text{car } |f'(\overline{W})|\leq 1 \text{ par (2.4) et par (5.13)}\\ \leq&\left(1+\sqrt{2\pi}\frac{z}{2}e^{\frac{z^2}{8}}\Phi(w)\right)(1-\Phi(z))+e^{-\frac{z}{2}}\mathbb{E}[e^{\overline{W}}]\quad\text{par le résultat (5.3)}\\ \leq&Ce^{-\frac{z}{2}}.\end{aligned}$$

On en déduit donc que

$$|R_1|\leq C\gamma e^{-\frac{z}{2}}.\tag{5.14}$$

✓ De la même façon, on écrit que $\mathbb{E}[f(\overline{W}^i)]=\mathbb{E}[|f(\overline{W}^i)|\mathbb{1}_{\{\overline{W}^i\leq\frac{z}{2}\}}]+\mathbb{E}[|f(\overline{W}^i)|\mathbb{1}_{\{\overline{W}^i>\frac{z}{2}\}}]$. Comme f est solution de (2.1), on utilise l'expression (2.2), et l'on déduit $\mathbb{E}[f(\overline{W}^i)]\leq Ce^{-\frac{z}{2}}$ et à fortiori

$$|R_3|\leq C\gamma e^{-\frac{z}{2}}.\tag{5.15}$$

Etape 3 : Reste à majorer R_2. Pour cela, écrivons $R_2:=R_{2,1}+R_{2,2}$ avec

$$R_{2,1}:=\sum_{i=1}^{n}\int_{-\infty}^{1}\mathbb{E}[\mathbb{1}_{\{\overline{W}^i+\overline{\xi}_i\leq z\}}-\mathbb{1}_{\{\overline{W}^i+t\leq z\}}]\overline{K_i(t)}dt,$$

$$R_{2,2}:=\sum_{i=1}^{n}\int_{-\infty}^{1}\mathbb{E}[(\overline{W}^i+\overline{\xi}_i)f(\overline{W}^i+\overline{\xi}_i)-(\overline{W}^i+t)f(\overline{W}^i+t)]\overline{K_i(t)}dt.$$

✓ Par l'Inégalité de concentration non uniforme, on a :

$$
\begin{aligned}
R_{2,1} \leq & \sum_{i=1}^{n} \int_{-\infty}^{1} \mathbb{E}[\mathbb{1}_{\{\overline{\xi_i} \leq t\}} \mathbb{P}(z - t \leq \overline{W}^i \leq z - \overline{\xi_i} \mid \overline{\xi_i})] \overline{K_i(t)} dt \\
\leq & C \sum_{i=1}^{n} \int_{-\infty}^{1} e^{-\frac{z-t}{2}} \mathbb{E}[|\overline{\xi_i}| + |t| + \gamma] \overline{K_i(t)} dt \\
\leq & C\gamma e^{-\frac{z}{2}}. \qquad (5.16)
\end{aligned}
$$

✓ Pour majorer $R_{2,2}$, on utilise le fait que, pour $s < t \leq 1$, on a (cf. [2, p47-48]) :

$$
\mathbb{E}[(\overline{W}^i + t) f(\overline{W}^i + t)] - \mathbb{E}[(\overline{W}^i + s) f(\overline{W}^i + s)] \leq C e^{-\frac{z}{2}} (|s| + |t|). \qquad (5.17)
$$

Il vient alors :

$$
\begin{aligned}
R_{2,2} \leq & \sum_{i=1}^{n} \int_{-\infty}^{1} \mathbb{E}\left[\mathbb{1}_{\{\overline{\xi_i} \geq t\}} \left(\mathbb{E}[(\overline{W}^i + \overline{\xi_i}) f(\overline{W}^i + \overline{\xi_i}) \mid \overline{\xi_i}] - \mathbb{E}[(W^i + t) f(W^i + t)]\right)\right] \overline{K_i(t)} dt \\
\leq & C e^{-\frac{z}{2}} \sum_{i=1}^{n} \int_{-\infty}^{1} \mathbb{E}[|\overline{\xi_i}| + |t|] \overline{K_i(t)} dt \\
\leq & C\gamma e^{-\frac{z}{2}}. \qquad (5.18)
\end{aligned}
$$

✓ Puisque $R_2 := R_{2,1} + R_{2,2}$, on a par (5.16) et (5.18), $R_2 \leq C\gamma e^{-\frac{z}{2}}$.
En procédant de façon similaire, on obtient $R_2 \geq -C\gamma e^{-\frac{z}{2}}$.
On obtient ainsi :

$$
|R_2| \leq C\gamma e^{-\frac{z}{2}}. \qquad (5.19)
$$

Etape 4 : En combinant (5.14), (5.19) et (5.15), et quitte à considérer $-\overline{W}$ on obtient

$$
|\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \leq C\gamma e^{-\frac{z}{2}},
$$

ce qui achève la démonstration.

■

5.2 Cas général

Afin de démontrer l'inégalité non uniforme de Berry-Esseen dans le cadre général, nous allons utiliser les résultats vus dans le cas borné. Avant de démontrer le théorème principal, nous avons besoin du lemme suivant.

Lemme 5.2 (Inégalité des moments). [7] *Soit $2 < p \leq 3$ et soit $\eta_1, \dots, \eta_n$ des variables aléatoires indépendantes telle que pour $1 \leq i \leq n, \mathbb{E}[\eta_i] = 0$ et $\mathbb{E}[|\eta_i|^p] < \infty$.*
On pose $B^2 := \sum_{i=1}^{n} \mathbb{E}[\eta_i^2]$ et $S := \sum_{i=1}^{n} \eta_i$. Alors :

$$
\mathbb{E}[|S|^p] \leq (p-1) B^p + \sum_{i=1}^{n} \mathbb{E}[|\eta_i|^p]. \qquad (5.20)
$$

Démonstration. Soit $S^i = S - \eta_i$ pour tout $1 \leq i \leq n$, $S^i \neq 0$.

Remarque. Pour tout $1 \leq i \leq n$, η_i et S^i sont indépendantes.

7. Les inégalités de moments de ce type ont été prouvées en premier par Rosenthal en 1970.

Etape 1 : On a en premier lieu :

$$
\begin{aligned}
\mathbb{E}[|S|^p] &= \sum_{i=1}^n \mathbb{E}[\eta_i S|S|^{p-2}] \quad \text{car } S := \textstyle\sum_{i=1}^n \eta_i \\
&= \sum_{i=1}^n \mathbb{E}[\eta_i(S|S|^{p-2} - S^i|S|^{p-2})] + \sum_{i=1}^n \mathbb{E}[\eta_i(S^i|S|^{p-2} - S^i|S^i|^{p-2})] \quad \text{car } \forall 1 \le i \le n,\ \eta_i \text{ et } S^i \text{ sont indépendantes et } \mathbb{E}[\eta_i]=0 \\
&\le \sum_{i=1}^n \mathbb{E}[\eta_i^2|S|^{p-2}] + \sum_{i=1}^n \mathbb{E}[|\eta_i||S^i|\{(|S^i|+|\eta_i|)^{p-2} - |S^i|^{p-2}\}] \\
&\le \sum_{i=1}^n \mathbb{E}[\eta_i^2(|\eta_i|^{p-2} + |S^i|^{p-2})] + \sum_{i=1}^n \mathbb{E}\left[|\eta_i||S^i|^{p-1}\{\left(1+\left(\frac{|\eta_i|}{|S^i|}\right)\right)^{p-2} - 1\}\right].
\end{aligned} \tag{5.21}
$$

Or, pour $x \ge 0$, on a $(1+x)^{p-2} - 1 \le (p-2)x$ (voir l'Appendice section 5, lemme 6.7). En utilisant cette observation et le résultat (5.21), il vient :

$$
\begin{aligned}
\mathbb{E}[|S|^p] &\le \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + \sum_{i=1}^n \mathbb{E}[\eta_i^2|S^i|^{p-2}] + \sum_{i=1}^n \mathbb{E}\left[|\eta_i||S^i|^{p-1}(p-2)\frac{|\eta_i|}{|S^i|}\right] \\
&= \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + (p-1)\sum_{i=1}^n \mathbb{E}[\eta_i^2]\mathbb{E}[|S^i|^{p-2}] \quad \text{car } \eta_i \text{ et } S^i \text{ sont indépendantes} \\
&\le \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + (p-1)\sum_{i=1}^n \mathbb{E}[\eta_i^2]\mathbb{E}\left[|S^i|^2\right]^{\frac{p-2}{2}},
\end{aligned} \tag{5.22}
$$

où on a utilisé le corollaire 1.1 avec $p = p-2$ et $q = 2$.

Etape 2 : Or, pour $1 \le i \le n$ fixé,

$$
\begin{aligned}
\mathbb{E}[|S^i|^2] = \mathbb{E}\left[S^2 - 2S\eta_i + \eta_i^2\right] &= \mathbb{E}[S^2 - 2\sum_{i\neq j}\eta_i\eta_j - 2\eta_i^2 + \eta_i^2] \\
&= \mathbb{E}[S^2] - 2\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbb{E}[\eta_i]\mathbb{E}[\eta_j] - \mathbb{E}[\eta_i^2] \quad \text{par indépendance des } \eta_i \\
&= \mathbb{E}\left[\left(\sum_{i=1}^n \eta_i\right)^2\right] - \mathbb{E}[\eta_i^2] \quad \text{car les } \eta_i \text{ sont centrées} \\
&= \mathbb{E}\left[\sum_{i=1}^n \eta_i^2\right] + 2\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbb{E}[\eta_i]\mathbb{E}[\eta_j] - \mathbb{E}[\eta_i^2] \\
&= \sum_{i=1}^n \mathbb{E}[\eta_i^2] - \mathbb{E}[\eta_i^2] \\
&= B^2 - \mathbb{E}[\eta_i^2] \le B^2.
\end{aligned}
$$

Ainsi, avec (5.22), on a :

$$\begin{aligned}\mathbb{E}[|S|^p] &\le \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + (p-1)\sum_{i=1}^n \mathbb{E}[\eta_i^2](B^2)^{\frac{p-2}{2}} \\ &\le \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + (p-1)B^{p-2}\sum_{i=1}^n \mathbb{E}[\eta_i^2] \\ &= \sum_{i=1}^n \mathbb{E}[|\eta_i|^p] + (p-1)B^p.\end{aligned}$$

■

À présent, on peut enfin démontrer l'inégalité non uniforme de Berry-Esseen, pour des variables aléatoires indépendantes. Nous considérons toujours les conditions du Cadre de travail 2.

Théorème 5.2 (Inégalité de Berry-Esseen non uniforme)**.**
Soient $\xi_1, \dots \xi_n$ des variables aléatoires indépendantes, admettant un moment d'ordre 3, telles que :

$$\forall 1 \le i \le n, \quad \mathbb{E}[\xi_i] = 0 \quad \text{et} \quad \sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1.$$

Posons $W := \sum_{i=1}^n \xi_i$, $\quad W^i := W - \xi_i$, $\quad$ et $\quad \gamma := \sum_{i=1}^n \mathbb{E}[|\xi_i|^3]$. Alors il existe une constante universelle C telle que :

$$\forall z \in \mathbb{R}, \quad |\mathbb{P}(W \le z) - \Phi(z)| \le C\frac{1}{1+|z|^3}\gamma. \tag{5.23}$$

Démonstration. Cette preuve repose en grande partie sur l'Inégalité de Berry-Esseen non uniforme, cas borné. Sans perte de généralités, nous pouvons supposer que $z \ge 0$ (quitte à considérer -W).

Etape 1 : Puisque $\{W \ge z\} \subset \{|W| \ge z\}$, par le Corollaire de l'inégalité de Markov appliqué à $f : [0, +\infty[\to]0, +\infty[, x \mapsto 1 + x^3$ (croissante et positive sur $[0, +\infty[$), on a

$$\mathbb{P}(W \ge z) \le \mathbb{P}(|W| \ge z) \le \frac{1 + \mathbb{E}[|W|^3]}{1+z^3} \le \frac{1+2+\gamma}{1+z^3}, \tag{5.24}$$

où l'on a utilisé l'Inégalité des moments (avec $p = 3$) pour la deuxième inégalité.
Si $\gamma \ge 1$, le théorème est alors démontré, d'après ce qui précède. On suppose alors $\gamma < 1$, et on considère pour tout $1 \le i \le n$, les variables $\overline{\xi_i}, \overline{W}, \overline{W}^i$ comme introduites dans le Cadre de travail,2.

Remarquons d'abord que $\{W \ge z\} = \{W \ge z, \max_{1\le i\le n} \xi_i > 1\} \cup \{W \ge z, \max_{1\le i\le n} \xi_i \le 1\}$
$\subset \{W \ge z, \max_{1\le i\le n} \xi_i > 1\} \cup \{\overline{W} \ge z\}$.

On a donc :

$$\mathbb{P}(W \ge z) \le \mathbb{P}(\overline{W} > z) + \mathbb{P}(W > z, \max_{1\le i\le n} \xi_i > 1). \tag{5.25}$$

Et comme par définition de $\overline{W}$, on a $W \ge \overline{W}$, on a aussi :

$$\mathbb{P}(\overline{W} > z) \le \mathbb{P}(W \ge z). \tag{5.26}$$

Etape 2 : En outre : $\mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1) \leq \sum_{i=1}^{n} \mathbb{P}(W \geq z, \xi_i > 1)$.

On peut alors distinguer 2 cas :

✓ Si $\xi_i > \frac{z}{2}$, alors $(W > z, \xi_i > 1, \xi_i > \frac{z}{2}) \implies \xi_i > \max(1, \frac{z}{2})$.

✓ Si $\xi_i < \frac{z}{2}$, alors $\{W \geq z, \xi_i > 1, \xi < \frac{z}{2}\} \subset \{W^i > \frac{z}{2}\}$ car $W^i = W - \xi_i$.

Ainsi,

$$\begin{aligned}
\mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1) &\leq \sum_{i=1}^{n} \mathbb{P}(W \geq z, \xi_i > 1) \\
&= \sum_{i=1}^{n} \mathbb{P}(\xi_i > \max(1, \frac{z}{2})) + \sum_{i=1}^{n} \mathbb{P}(W^i > \frac{z}{2})\mathbb{P}(\xi_i > 1), \quad \text{par indépendance de } \xi_i \text{ et de } W^i \text{ pour tout } 1 \leq i \leq n.
\end{aligned}$$

Etape 3 : Puisque $\mathbb{P}(\xi_i > \max(1, \frac{z}{2})) \leq \mathbb{P}(|\xi_i| > \max(1, \frac{z}{2}))$, en utilisant le Corollaire de l'inégalité de Markov dans la somme de gauche appliqué à $f : [0, +\infty[\to [0, +\infty[, x \mapsto x^3$ (croissante et positive sur $[0, +\infty[$), on obtient :

$$\mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1) \leq \frac{\gamma}{(\max(1, \frac{z}{2}))^3} + \sum_{i=1}^{n} \mathbb{P}(W^i > \frac{z}{2})\mathbb{P}(\xi_i > 1).$$

De plus, en reprenant les arguments de (5.24) avec W^i et $\frac{z}{2}$, on a que :

$$\begin{aligned}
\mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1) &\leq \sum_{i=1}^{n} \frac{(1 + \mathbb{E}[|W^i|^3])}{1 + \left(\frac{z}{2}\right)^3} \mathbb{E}[|\xi_i|^3] \\
&\leq \frac{\gamma}{1 + \left(\frac{z}{2}\right)^3} + \sum_{i=1}^{n} \frac{3 + \gamma}{1 + \left(\frac{z}{2}\right)^3} \mathbb{E}[|\xi_i|^3] \\
&= \frac{\gamma}{1 + \left(\frac{z}{2}\right)^3} + \frac{3\gamma + \gamma^2}{1 + \left(\frac{z}{2}\right)^3} \\
&\leq \frac{40\gamma}{1 + z^3}, \quad \text{car } \gamma{<}1 \text{ donc } \gamma^2 < \gamma.
\end{aligned}$$

Par (5.25) et (5.26), on a alors :

$$\mathbb{P}(\overline{W} > z) \leq \mathbb{P}(W > z) \leq \mathbb{P}(\overline{W} > z) + \mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1),$$

donc $|\mathbb{P}(W \leq z) - \mathbb{P}(\overline{W} \leq z)| \leq \mathbb{P}(W > z, \max_{1\leq i\leq n} \xi_i > 1) \leq \frac{40\gamma}{1 + z^3}$.

Enfin, par inégalité triangulaire, on a :

$$|\mathbb{P}(W \leq z) - \Phi(z)| \leq |\mathbb{P}(W \leq z) - \mathbb{P}(\overline{W} \leq z)| + |\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \leq \frac{40\gamma}{1 + z^3} + |\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \quad (\ast).$$

Etape 4 : Pour démontrer le théorème, il suffit alors de prouver que

$$|\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \leq \frac{C\gamma}{1 + z^3}, \tag{5.27}$$

où C est une constante arbitraire. Or, d'après l'Inégalité de Berry-Esseen non uniforme, cas borné, on a :

$$|\mathbb{P}(\overline{W} \leq z) - \Phi(z)| \leq C\gamma e^{-\frac{z}{2}} \leq C\frac{1}{1 + |z|^3}\gamma.$$

Avec $(\ast)$ et quitte à considérer $-W$, , on déduit alors l'Inégalité de Berry-Esseen non uniforme.

■

L'inégalité de Berry-Esseen non uniforme est donc établie et, une fois de plus, l'identité de Stein est à la genèse de sa démonstration.

6 Ouverture

Dans ce dossier, on a étudié la méthode de Stein pour la loi normale centrée réduite, et l'on a pu constater sa grande utilité et notamment le fait qu'elle permet de démontrer les inégalités de Berry-Esseen. Cependant, ce ne sont pas les seules applications de cette méthode "magique". En 1975, Louis H.Y Chen, un doctorant de Charles Stein, adapta la méthode de Stein pour obtenir des résultats sur l'approximation de la loi de Poisson. Cette méthode est aujourd'hui appelée **méthode de Chen-Stein**. Le point clé de cette méthode est, comme pour la loi normale centrée réduite, une caractérisation de la loi de Poisson en termes d'espérance.

Propriété 6.1 (Caractérisation de la loi de Poisson)**.**

$$X \sim \mathcal{P}(\lambda) \iff \mathbb{E}[Xf(X)] = \lambda \mathbb{E}[f(X+1)]$$

pour toute fonction f pour laquelle les espérances précédentes sont bien définies.

Et ainsi, on peut définir une équation de Chen-Stein, comme pour la loi normale centrée réduite.

Définition 6.1 (Equation de Chen-Stein)**.** *Soit h une fonction $\mathcal{P}(\lambda)$-intégrable. L'équation de Chen-Stein est donnée par :*

$$\lambda f(k+1) - kf(k) = h(k) - \sum_{k=0}^{+\infty} h(k) \frac{\lambda^k}{k!} e^{-\lambda}.$$

En suivant la méthode de Stein, Louis H.Y. Chen a alors utlisé la solution de l'équation de Chen-Stein dans divers calculs pour obtenir une borne sur l'erreur d'approximation de la loi de Poisson.

Annexe

Ici sont présentées les preuves de certains résultats utilisés lors de différentes démonstrations

Appendice section 3

Les conditions de travail sont toujours celles du Cadre de travail 1.

Lemme 6.1. *Soit X une variable aléatoire, si g et h sont des fonctions croissantes, alors :*

$$\mathbb{E}[g(X)]\mathbb{E}[h(X)] \leq \mathbb{E}[g(X)h(X)].$$

Démonstration. On considère deux variables aléatoires X et X' définies sur le même espace probabilisé Ω, indépendantes et de même loi.
Par croissance de g et h, on a toujours $(g(X) - g(X'))(h(X) - h(X')) \geq 0$ sur Ω. Alors, en utilisant le fait que X et X' sont indépendantes et de même loi, il vient :

$$\begin{aligned}
0 \leq &\mathbb{E}[(g(X) - g(X'))(h(X) - h(X'))] \\
= &\mathbb{E}[g(X)h(X)] + \mathbb{E}[g(X')h(X')] - \mathbb{E}[g(X')h(X)] - \mathbb{E}[g(X)h(X')] \\
= &2\mathbb{E}[g(X)h(X)] - \mathbb{E}[g(X')]\mathbb{E}[h(X)] - \mathbb{E}[g(X)]\mathbb{E}[h(X')] \\
= &2\mathbb{E}[g(X)h(X)] - 2\mathbb{E}[g(X)]\mathbb{E}[h(X)].
\end{aligned}$$

D'où, $\mathbb{E}[g(X)]\mathbb{E}[h(X)] \leq \mathbb{E}[g(X)h(X)]$. ■

On démontre maintenant le Théorème 3.3.

Théorème (R.V. Erickson).
On suppose qu'il existe δ tel que pour toute fonction h lipschitzienne, on ait $|\mathbb{E}[h(W)] - N(h)| \leq \delta \|h'\|_\infty$. Alors :

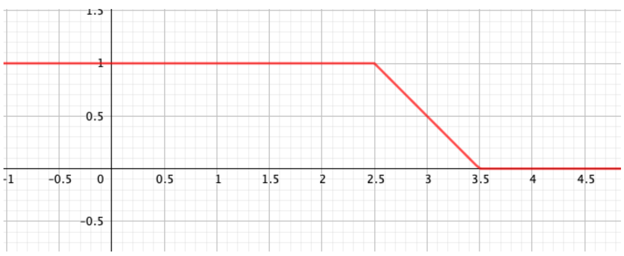
$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 2\sqrt{\delta}. \tag{6.1}$$

Démonstration. On suppose qu'un tel δ existe et que $\delta \leq \frac{1}{4}$ (sinon le résultat est trivial).

Etape 1 : Soit $\alpha \in \mathbb{R}^+$, que l'on déterminera plus tard. Pour $z \in \mathbb{R}$ fixé, on définit

$$h_\alpha(w) = \mathbb{1}_{\{w\leq z\}} + \left(1 + \frac{z-w}{\alpha}\right)\mathbb{1}_{\{z\leq w\leq z+\alpha\}}.$$

Alors h est lipschitzienne et l'on voit aisément que $\|h'\|_\infty = \frac{1}{\alpha}$. Par exemple, pour $z = 2,5$ et $\alpha = 1$, la courbe représentative de h_α est de la forme suivante.



On en déduit donc que $\forall w \in \mathbb{R}, \mathbb{1}_{\{w\leq z\}} \leq h_\alpha(w) \leq \mathbb{1}_{\{w\leq z+\alpha\}}$. Et ainsi :

$$\mathbb{P}(W \leq z) \leq \mathbb{E}[h_\alpha(W)] \leq \mathbb{P}(W \leq z + \alpha).$$

De même, puisque $N(h_\alpha) = \mathbb{E}[h_\alpha(Z)]$, on obtient : $\Phi(z) \leq N(h_\alpha) \leq \Phi(z) + \mathbb{P}(z \leq W \leq z + \alpha)$.

Etape 2 : Dès lors,

$$\begin{aligned}
\mathbb{P}(W \leq z) - \Phi(z) \leq & \mathbb{E}[h_\alpha(W)] - \Phi(z) \\
= & \mathbb{E}[h_\alpha(W)] - N(h_\alpha) + N(h_\alpha) - \Phi(z) \\
\leq & \frac{\delta}{\alpha} + N(h_\alpha) - \Phi(z) \quad \text{par hypothèse du théorème et } \|h'\|_\infty = \tfrac{1}{\alpha} \\
\leq & \frac{\delta}{\alpha} + \mathbb{P}(z \leq Z \leq z+\alpha) \\
= & \frac{\delta}{\alpha} + \Phi(z+\alpha) - \Phi(z) \\
\mathbb{P}(W \leq z) - \Phi(z) \leq & \frac{\delta}{\alpha} + \frac{\alpha}{\sqrt{2\pi}}, \quad \text{car } \Phi \text{ est } \tfrac{1}{\sqrt{2\pi}}\text{-lipschitzienne.}
\end{aligned}$$

En prenant $\alpha = \sqrt{\delta}(2\pi)^{1/4}$, on a $\dfrac{\delta}{\alpha} + \dfrac{\alpha}{\sqrt{2\pi}} = 2(2\pi)^{-1/4}\sqrt{\delta} \leq 2\sqrt{\delta}$.
Et ainsi,

$$\mathbb{P}(W \leq z) - \Phi(z) \leq 2\sqrt{\delta}. \tag{6.2}$$

Etape 3 : Afin d'obtenir la borne inférieure, on utilise la même démarche que précédemment. Pour $z \in \mathbb{R}$ fixé, on définit :

$$h_\alpha(w) = -\mathbb{1}_{\{w \leq z\}} - \left(1 + \frac{z-w}{\alpha}\right)\mathbb{1}_{\{z \leq w \leq z+\alpha\}}.$$

On remarque que $\mathbb{E}[h_\alpha(W)] \leq \mathbb{P}(W \leq z)$. Alors, en utilisant excatement le même raisonnement que précédemment, on trouve :

$$\mathbb{P}(W \leq z) - \Phi(z) \geq -2\sqrt{\delta}. \tag{6.3}$$

Ainsi, en combinant les résultats (6.2) et (6.3), et par passage au sup, on déduit :

$$\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| \leq 2\sqrt{\delta},$$

d'où le résultat.

■

On rappelle maintenant l'énoncé du Théorème de Barbour et Hall (3.5), dont nous n'avons donné qu'une partie de la démonstration.

Théorème (Barbour et Hall)**.**
Soient $\xi_1, \dots, \xi_n$ des variables aléatoires indépendantes telles que pour tout $1 \leq i \leq n$, $\mathbb{E}[\xi_i] = 0$, $\mathbb{E}[\xi_i^2] = \sigma_i^2 < \infty$ et $\sum_{i=1}^n \sigma_i^2 = 1$. Alors il existe une constante universelle C telle que, pour tout $\varepsilon > 0$:

$$\left(1 - e^{-\frac{\varepsilon^2}{4}}\right)\sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \mathbb{1}_{\{|\xi_i| > \varepsilon\}}\right] \leq C\left(\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| + \sum_{i=1}^n \sigma_i^4\right),$$

où $\Phi(z) = \mathbb{P}(Z \leq z)$.

Démonstration. On reprend la démonstration à partir de (3.13) et de ***l'étape 4***. On choisit alors $f(w) = we^{-\frac{w^2}{2}}$ (puisque $\int_{-\infty}^{+\infty} |h'(w)|dw < \infty$ et $\|f'''\|_\infty < \infty$, voir calculs ci-après), et ainsi, en prenant $h(w) = f'(w) - wf(w)$, on a :

$$c_1 := \int_{-\infty}^{+\infty} |h'(w)|dw \leq 18; \quad c_2 := \int_{-\infty}^{+\infty} |f''(w)|dw \leq 10; \quad c_3 := \sup_{w \in \mathbb{R}} |f'''(w)| = 3$$

En effet :

✓ $h'(w) = e^{-\frac{w^2}{2}}(2w^3 - 5w)$ et $f''(w) = e^{-\frac{w^2}{2}}(w^3 - 3w)$.
Pour calculer c_1 et c_2, on utilise l'inégalité triangulaire, la parité de $x \mapsto |x|$ et le fait que $\forall k \geq 0, I_{2k+1} := \displaystyle\int_0^{+\infty} w^{2k+1} e^{-\frac{w^2}{2}} dw = 2^k k!$ (cf.lemme 6.2).

✓ $f'''(w) = e^{-\frac{w^2}{2}}(-w^4 + 6w^2 - 3)$ donc $f^{(4)}(w) = e^{-\frac{w^2}{2}}(w^5 - 10w^3 + 15w)$.

$f^{(4)}$ s'annule en $0, w_1 = -\sqrt{5+\sqrt{10}}$, $\quad w_2 = -\sqrt{5-\sqrt{10}}$, $\quad w_3 = \sqrt{5-\sqrt{10}}$, $\quad$ et $w_4 = \sqrt{5-\sqrt{10}}$.
On a donc le tableau de variations suivant :

w	$-\infty$		w_1		w_2		0		w_3		w_4		$+\infty$
$f^{(4)}(w)$		$-$	0	$+$	0	$-$	0	$+$	0	$-$	0	$+$	
$f'''(w)$	0	$\searrow$	$f(w_1)$	$\nearrow$	$f(w_2)$	$\searrow$	-3	$\nearrow$	$f(w_3)$	$\searrow$	$f(w_4)$	$\nearrow$	0

où $f(w_1) \simeq -0,35$, $\quad f(w_2) \simeq 1,85$, $\quad f(w_3) \simeq 1,85$, $\quad f(w_4) \simeq -0,35$.
On en déduit donc que $c_3 := \sup_{w \in \mathbb{R}} |f'''(w)| = 3$.

Etape 5 : Ceci étant établi, on va utiliser la formule (3.13) pour pouvoir minorer $\Delta := \sup_{z\in\mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$. Intuitivement, on peut penser que si W et Z ont des lois proches alors,

$$R_1 := \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 g(W^i, \xi_i)] \quad \text{et} \quad R := \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 g(Z, \xi_i)]$$

seront aussi assez proches, où $Z \sim \mathcal{N}(0,1)$ est indépendante des ξ_i pour tout $1 \leq i \leq n$. Pour pouvoir utiliser cette intuition, nous allons estimer R et R_1 afin de borner correctement (3.13).
Alors, avec un tel choix de f comme précédemment et en utilisant $g(w,y) = -y^{-1}(f(w+y) - f(w) - yf'(w))$, on a en premier lieu :

$$\begin{aligned}
\mathbb{E}[g(Z,y)] &= -\frac{1}{y\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left((w+y)e^{-\frac{(w+y)^2}{2}} - we^{-\frac{w^2}{2}} - y(1-w^2)e^{-\frac{w^2}{2}} \right) e^{-\frac{w^2}{2}} dw \\
&= -\frac{1}{y} \left(\frac{y}{2\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}(w^2 + (w+y)^2) \right) dw - \frac{y}{2\sqrt{2}} \right) \\
&= -\frac{1}{y} \left(\frac{y}{2\sqrt{2\pi}} \sqrt{\pi} e^{-\frac{y^2}{4}} - \frac{y}{2\sqrt{2}} \right) \\
&= \frac{1}{2\sqrt{2}}(1 - e^{-\frac{y^2}{4}}) > 0, \quad \text{car } e^{-\frac{y^2}{4}} < 1.
\end{aligned}$$

Et si $|y| \geq \varepsilon$ pour $\varepsilon > 0$, alors $\mathbb{E}[g(Z,y)] \geq \frac{1}{2\sqrt{2}}(1 - e^{-\frac{\varepsilon^2}{4}})$. Donc, en reprenant l'expression de R, il vient :

$$R \geq \frac{1}{2\sqrt{2}}(1 - e^{-\frac{\varepsilon^2}{4}}) \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i| > \varepsilon\}}]. \tag{6.4}$$

Etape 6 : Il nous faut alors maintenant minorer R_1. Pour cela, on pose $R_2 = \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 g(W^*, \xi_i)]$, où W^* a la même loi que W, mais est indépendante des ξ_i pour $1 \leq i \leq n$.

✓ Ainsi, en reprenant l'expression de R_1 et de g, on a :

$$\begin{aligned}R_1 =& -\sum_{i=1}^n \mathbb{E}\left[\xi_i \int_0^1 (f'(W^i + t\xi_i) - \xi_i f'(W^i))dt\right] \\ =& R_2 + \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \int_0^1 (f'(W^* + t\xi_i) - f'(W^i + t\xi_i))dt\right] - \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \int_0^1 (f'(W^*) - f'(W^i))dt\right].\end{aligned}$$

✓ Pour $t \in [0,1]$ quelconque, comme $\mathbb{E}[\xi_i] = 0$ pour tout $1 \leq i \leq n$ et ξ_i et W^i sont indépendantes, on a de plus :

$$\begin{aligned}|\mathbb{E}[f'(W^* + t\xi_i) - f'(W^i + t\xi_i)]| =& |\mathbb{E}[f'(W^i + \xi_i + t\xi_i) - f'(W^i + t\xi_i)]| \\ =& |\mathbb{E}[f'(W^i + \xi_i + t\xi_i) - f'(W^i + t\xi_i) - \xi_i f''(W^i + t\xi_i)] \\ \leq& \frac{1}{2} c_3 \sigma_i^2, \end{aligned} \tag{6.5}$$

où on a utilisé l'inégalité de Taylor-Lagrange pour écrire :
$\exists Y \in]W^i + t\xi_i, W^i + \xi_i + t\xi_i[, |f'(W^i + \xi_i + t\xi_i) - f'(W^i + t\xi_i) - \xi_i f''(W^i + t\xi_i)| \leq \frac{\xi_i^2}{2} \|f'''\|_\infty$
et en déduire la majoration précédente.
On obtient donc, en utilisant (6.5) en $t \neq 0$ et en $t = 0$,

$$R_1 \geq R_2 - c_3 \sum_{i=1}^n \sigma_i^4. \tag{6.6}$$

✓ On va maintenant minorer R_2. En raisonnant de même que précédemment, on obtient :

$$R_2 = R + \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \int_0^1 (f'(Z + t\xi_i) - f'(W^* + t\xi_i))dt\right] - \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \int_0^1 (f'(Z) - f'(W^i))dt\right].$$

Et $\forall t \in [0,1]$, on a par intégration par parties, puisque f' est bornée à l'infini :

$$\begin{aligned}|\mathbb{E}[f'(W^* + t\xi_i)] - \mathbb{E}[f'(Z + t\xi_i]| =& \left|\int_{-\infty}^{+\infty} f''(w)(\mathbb{P}(W^* \leq w - t\xi_i) - \Phi(w - t\xi_i))dw\right| \\ \leq& c_2 \Delta.\end{aligned} \tag{6.7}$$

Ainsi, en utilisant (6.7) en $t \neq 0$ et en $t = 0$, on obtient :

$$R_2 \geq R - 2c_2\Delta. \tag{6.8}$$

Etape 7 : Il ne nous reste plus qu'à utiliser les résultats précédents pour obtenir le résultat voulu.

✓ En combinant (3.10), (3.13), (6.6) et (6.8), on obtient :

$$c_1\Delta \geq R_1 - \frac{1}{2}c_3 \sum_{i=1}^n \sigma_i^4 \geq R - \frac{3}{2}c_3 \sum_{i=1}^n \sigma_i^4 - 2c_2\Delta.$$

✓ Enfin, au regard de (6.4), on a :

$$\Delta(c_1 + 2c_2) + \frac{3}{2}c_3 \sum_{i=1}^n \sigma_i^4 \geq \frac{1}{2\sqrt{2}} \left(1 - e^{-\frac{\varepsilon^2}{4}}\right) \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \mathbb{1}_{\{|\xi_i| > \varepsilon\}}\right],$$

et comme $\Delta = \sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)|$, $c_1 \leq 18; c_2 \leq 10, c_3 = 3$, on obtient :

$$\left(1 - e^{-\frac{\varepsilon^2}{4}}\right) \sum_{i=1}^n \mathbb{E}\left[\xi_i^2 \mathbb{1}_{\{|\xi_i| > \varepsilon\}}\right] \leq C \left(\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leq z) - \Phi(z)| + \sum_{i=1}^n \sigma_i^4\right),$$

d'où le théorème.

■

Lemme 6.2. $\forall k \geq 0, I_{2k+1} := \int_0^{+\infty} w^{2k+1} e^{-\frac{w^2}{2}} = 2^k k!.$

Démonstration. Soit $k \geq 0$, soit $M > 0$:

$$\begin{aligned} I_{2k+1} := \int_0^{+\infty} w^{2k+1} e^{-\frac{w^2}{2}} dw = \int_0^{+\infty} w^{2k} w e^{-\frac{w^2}{2}} dw \\ = \lim_{M \to +\infty} \left[-w^{2k} e^{-\frac{w^2}{2}} \right]_0^M + 2k \int_0^M w^{2k-1} e^{-\frac{w^2}{2}} dw \quad & \text{par intégration par parties} \\ = 2k \int_0^{+\infty} w^{2k-1} e^{-\frac{w^2}{2}} dw = 2k \times I_{2k-1}. \end{aligned}$$

Par récurrence sur k, on a donc $I_{2k+1} = 2 \times 4 \times 6 \cdots \times (2k)$
$= 2^k \times k!.$ ■

Appendice section 4

Lemme 6.3. $\forall x, y > 0, \min(x, y) \geq x - \frac{x^2}{4y}.$

Démonstration. ✓ si $\min(x, y) = x$ alors le résultat est évident.

✓ si $\min(x, y) = y$, alors on fixe $y > 0$, on considère $x \geq y$ et on pose $f_y(x) = x - \frac{x^2}{4y}$.
Alors $f_y'(x) = 1 - \dfrac{x}{2y} > 0$ si $x \leq 2y$.
< 0 si $x \geq y$
f_y atteint donc son maximum en $x = 2y$ et $f_y(2y) = y$, on a donc $f_y(x) \leq y$, ce qui est le résultat voulu.

■

Appendice section 5

Lemme 6.4. $f : s \mapsto \dfrac{e^s - 1 - s}{s^2}$ *est croissante sur* $\mathbb{R}$.

Démonstration. D'après la formule de Taylor avec reste intégral, on a $\forall s \in \mathbb{R}$,

$$e^s - s - 1 = \int_0^1 s^2 e^{st} (1 - t) dt.$$

Donc $f(s) = \displaystyle\int_0^1 e^{st} (1 - t) dt.$

On pose $g(s, t) = e^{st}(1 - t)$. Alors :

✓ $\forall s \in \mathbb{R}, t \mapsto g(s, t)$ est continue et intégrable sur $[0, 1]$.

✓ $\forall (s, t) \in R \times [0, 1], \dfrac{\partial g}{\partial s}(s, t) = t e^{st}(1 - t).$

- ✓ Soit $]a, b[\in \mathbb{R}, a < b < \infty$, $\forall s \in]a, b[, \forall t \in [0, 1]$, on a $|te^{st}(1-t)| \leq e^b$, et $t \mapsto e^b$ est intégrable sur $[0, 1]$.
- ✓ Par le théorème de dérivation sous le signe intégral, on déduit que f est dérivable sur $]a, b[$.
- ✓ Ceci étant vrai pour tout intervalle $]a, b[$ de $\mathbb{R}$, on déduit que f est dérivable sur $\mathbb{R}$ et que $\forall s \in \mathbb{R}, f'(s) = \int_0^1 te^{st}(1-t)dt$.

Or, $\forall s \in \mathbb{R}, \quad t \mapsto te^{st}(1-t) \geq 0$ pour $t \in [0, 1]$, donc par positivité de l'intégrale, $f'(s) \geq 0, \forall s \in \mathbb{R}$. f est donc prolongée par continuité en 0 et on en déduit donc que f est croissante sur $\mathbb{R}$. ■

On démontre à présent les 2 résultats restants de Inégalité des moments.

Lemme 6.5. *Soient $\alpha > 0, B \in \mathbb{R}$. On considère $\eta_1, \ldots, \eta_n$ des variables aléatoires indépendantes telles que pour tout $1 \leq i \leq n$, $\mathbb{E}[\eta_i] \leq 0$, $\eta_i \leq \alpha$ et $\sum_{i=1}^n \mathbb{E}[\eta_i^2] \leq B^2$. On pose $S = \sum_{i=1}^{n} \eta_i$. Alors :*

$$\mathbb{E}\left[e^{tS}\right] \leq \exp\left(\frac{e^{t\alpha} - 1 - t\alpha}{\alpha^2} B^2\right) \qquad \textit{pour } t > 0 \quad \textit{(Inégalité de Bennett)}, \tag{6.9}$$

$$\mathbb{P}(S \geq x) \leq \exp\left(-\frac{B^2}{\alpha^2}\left[\left(1 + \frac{\alpha x}{B^2}\right)\ln\left((1 + \frac{\alpha x}{B^2}\right) - \frac{\alpha x}{B^2}\right]\right), \tag{6.10}$$

$$\mathbb{P}(S \geq x) \leq \exp\left(-\frac{x^2}{2(B^2 + \alpha x)}\right) \qquad \textit{pour } x > 0. \tag{6.11}$$

Démonstration. L'inégalité de Bennett a déjà été démontrée.

(6.10) Considérons maintenant $t = \frac{1}{\alpha} \ln\left(1 + \frac{\alpha x}{B^2}\right)$. Alors par (5.2) :

$$\mathbb{P}(S \geq x) \leq e^{-tx}\mathbb{E}[e^{tS}] \leq \exp(-tx + \frac{1}{\alpha^2}(e^{t\alpha} - 1 - t\alpha)B^2) = \exp\left(-\frac{B^2}{\alpha^2}\left[\left(1 + \frac{\alpha x}{B^2}\right)\ln\left(1 + \frac{\alpha x}{B^2}\right) - \frac{\alpha x}{B^2}\right]\right),$$

d'où le résultat (5.3).

(6.11) Il suffit de voir (cf. lemme 6.6) que, pour tout $s > 0$, $(1+s)\ln(1+s) - s \geq \frac{s^2}{2(1+s)}$.
Ainsi, en appliquant cette inégalité dans (5.3) pour $s > 0$, on obtient bien

$$\mathbb{P}(S \geq x) \leq \exp\left(-\frac{x^2}{2(B^2 + \alpha x)}\right).$$

■

Lemme 6.6. $\forall s > 0$,

$$(1+s)\ln(1+s) - s \geq \frac{s^2}{2(1+s)}.$$

Démonstration. Pour tout $s > 0$, notons $g(s) = (1+s)\ln(1+s) - s - \frac{s^2}{2(1+s)}$.
On a $g'(s) = \ln(1+s) - \frac{s}{1+s} - \frac{s^2}{2(1+s)^2}$ et $g''(s) = \frac{s^2}{(s+1)^3} > 0, \quad \forall s > 0$.

- ✓ g' est donc croissante et comme $g'(0) = 0$, alors $g'(s) \geq 0$ sur $]0, +\infty[$.

✓ g est donc croissante et comme $g(0) = 0$, $g(s) \geq 0$ sur $]0, +\infty[$.

On en déduit donc $(1+s)\ln(1+s) - s \geq \dfrac{s^2}{2(1+s)}$, $\forall s > 0$.

■

Lemme 6.7. *Soit* $2 < p \leq 3$, $\forall x \geq 0$, *on a :* $(1+x)^{p-2} - 1 \leq (p-2)x$.

Démonstration. Posons $f(x) = (p-2)x - (1+x)^{p-2} - 1$.
Alors $f'(x) = p - 2 - (p-2)(1+x)^{p-3} = (p-2)(1-(1+x)^{p-3})$.

Et comme $p > 2, p \leq 3, x \geq 0$, alors $(p-2) > 0, (1+x)^{p-3} \leq 1$ et donc $f'(x) \geq 0$.
Ainsi f est croissante et comme $f(0) = 0$, on a $f(x) \geq 0 \quad$ pour tout $\quad x \geq 0$,
i.e. $(1+x)^{p-2} - 1 \leq (p-2)x$. ■

Références

[1] Louis H.Y. Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein's Method.* Springer, 2010.

[2] A.D. Barbour and Louis H.Y. Chen. *An Introduction to Stein's Method*, volume 4. Singapore University Press and World Scientific Publishing, 2005.

[3] Walter Rudin. *Analyse réelle et complexe.* Masson, 1978.

[4] M1MG UFR IM2AG. Processus stochastiques, TD3- Espérance conditionnelle. Exercice 6, 2020-21.