math.univ-lyon1.fr/homes-www/lerouvillois
lerouvillois@math.univ-lyon1.fr
poudevigne@math.univ-lyon1.fr

**Correction exercice 2 :**

1. Les $(X_i)_{i\in\mathbb{N}*}$ sont i.i.d dont les $(Y_i)_{i\in\mathbb{N}*}$ également (car $Y_i = f(X_i)$ avec $f(x) := \dfrac{x-\mu}{\sigma}$ qui est mesurable). De plus,

$$\mathbb{E}[Y_i] = \frac{1}{\sigma}\left(\mathbb{E}[X_i] - \mu\right) = 0,$$

et

$$\mathrm{Var}(Y_i) = \frac{1}{\sigma^2}\mathrm{Var}(X_i - \mu) = \frac{1}{\sigma^2}\mathrm{Var}(X_i) = 1.$$

2. On a pour tout $\xi \in \mathbb{R}$ :

$$\begin{aligned}
\phi_{Z_n}(\xi) &= \mathbb{E}\left[e^{i\xi Z_n}\right] \\
&= \mathbb{E}\left[e^{i\frac{\xi}{\sqrt{n}}\sum_{i=1}^n Y_i}\right] \\
&= \mathbb{E}\left[\prod_{i=1}^n e^{i\frac{\xi}{\sqrt{n}}\sum_{i=1}^n Y_i}\right] \\
&= \prod_{i=1}^n \mathbb{E}\left[e^{i\frac{\xi}{\sqrt{n}}Y_i}\right] \qquad \text{par indépendance} \\
&= \prod_{i=1}^n \phi_{Y_i}\left(\frac{\xi}{\sqrt{n}}\right) \\
&= \left(\phi_{Y_1}\left(\frac{\xi}{\sqrt{n}}\right)\right)^n \qquad \text{car les } Y_i \text{ ont même loi.}
\end{aligned}$$

3. D'après le cours : si une variable aléatoire réelle a un moment d'ordre $k$, alors sa fonction caractéristique $\phi_X$ est de classe $\mathcal{C}^k$. De plus, pour tout $l \le k$ et pour tout $\xi \in \mathbb{R}$,

$$\phi_X^{(l)}(\xi) = \mathbb{E}\left[(iX)^l e^{i\xi X}\right].$$

Remarque : cela revient à dériver directement dans l'espérance (le théorème de dérivation sous l'intégrale donne le résultat).
Ici, $Y_1$ est de carré intégrable donc $\phi_{Y_1}$ est de classe $\mathcal{C}^2$. De plus,

$$\phi_{Y_1}(0) = \mathbb{E}[1] = 1,$$

$$\phi'_{Y_1}(0) == \mathbb{E}\left[iY_1\right] = 0,$$

et

$$\phi''_{Y_1}(0) = \mathbb{E}\left[(iY_1)^2\right] = -1.$$

D'après la formule de Taylor-Young à l'ordre 2 au voisinage de zéro :

$$\phi_{Y_1}(x) = \phi_{Y_1}(0) + x\phi'_{Y_1}(0) + \frac{1}{2}x^2\phi''_{Y_1}(0) + \underset{x\to 0}{o}(x^2).$$

En appliquant cela à $x = \xi/\sqrt{n}$, on obtient le résultat :

$$\phi_{Y_1}\left(\frac{\xi}{\sqrt{n}}\right) = 1 - \frac{1}{2}\frac{\xi^2}{n} + \underset{n\to\infty}{o}\left(\frac{1}{n}\right).$$

4. On en déduit que pour tout $\xi \in \mathbb{R}$,

$$\begin{aligned}
\phi_{Z_n}(\xi) &= \left(1 - \frac{1}{2}\frac{\xi^2}{n} + \underset{n\to\infty}{o}\left(\frac{1}{n}\right)\right)^n \\
&= \exp\left(n\ln\left(1 - \frac{1}{2}\frac{\xi^2}{n} + \underset{n\to\infty}{o}\left(\frac{1}{n}\right)\right)\right) \\
&= \exp\left(n\left(-\frac{1}{2}\frac{\xi^2}{n} + \underset{n\to\infty}{o}\left(\frac{1}{n}\right)\right)\right) \qquad \text{car } \log(1+x) = x + \underset{x\to 0}{o}(x) \\
&= \exp\left(-\frac{\xi^2}{2} + \underset{n\to\infty}{o}(1)\right) \\
&\underset{n\to\infty}{\longrightarrow} \exp\left(-\frac{\xi^2}{2}\right).
\end{aligned}$$

D'après le théorème de Paul Lévy, on en conclut que :

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} = \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1)$$

<u>**Correction exercice 3**</u> **:**

1. On s'intéresse à la limite de

$$\begin{aligned}\mathbb{P}\left(S_n \in [0, an]\right) &= \mathbb{P}\left(\frac{S_n}{n} \leq a\right)\\ &= F_{S_n/n}(a).\end{aligned}$$

La convergence d'une fonction de répartition peut être obtenue par la convergence en loi. Or, les $Y_i$ sont i.i.d et intégrables donc d'après la Loi Forte des Grands Nombres,

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{p.s} \mathbb{E}[Y_1] = p.$$

Or la convergence en loi implique la convergence p.s donc

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{loi} p.$$

On en déduit que

$$F_{S_n/n}(x) \xrightarrow[n\to\infty]{} F_Z(x),$$

en tout point de continuité de $F_Z$ où $Z$ est la variable aléatoire constante égale à $p$. On a $F_Z(x) = \mathbf{1}_{x\geq p}$, qui est continue partout sauf en $p$. On en déduit que

— Si $a < p$, $F_{S_n/n}(a) \xrightarrow[n\to\infty]{} 0$.

— Si $a > p$, $F_{S_n/n}(a) \xrightarrow[n\to\infty]{} 1$.

Pour résoudre le cas $a = p$, on ne peut pas conclure directement. Nous allons utiliser le théorème central limite et l'égalité :

$$\begin{aligned}\mathbb{P}\left(S_n \leq pn\right) &= \mathbb{P}\left(S_n - pn \leq 0\right)\\ &= \mathbb{P}\left(\frac{S_n - pn}{\sqrt{p(1-p)n}} \leq 0\right)\\ &= F_{\frac{S_n - pn}{\sqrt{p(1-p)n}}}(0).\end{aligned}$$

Or, les $Y_i$ sont i.i.d et de carré intégrables donc d'après le TCL :

$$\frac{S_n - pn}{\sqrt{p(1-p)n}} \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1).$$

La fonction de répartition de la loi normale est continue est tout point (car c'est une loi à densité) donc on peut conclure que

$$F_{\frac{S_n - pn}{\sqrt{p(1-p)n}}}(0) \xrightarrow[n\to\infty]{} F_{\mathcal{N}(0,1)}(0) = 1/2.$$

La réponse à la question 1 est donc 0 si $a < p$, 1/2 si $a = p$ et 1 si $a > p$.

2. On utilise le TCL et l'égalité

$$\mathbb{P}\left(S_n \in [pn - 1.96\sqrt{p(1-p)n}, pn + 1.96\sqrt{p(1-p)n}]\right)$$
$$= \mathbb{P}(-1.96 \leq \frac{S_n - pn}{\sqrt{p(1-p)n}} \leq 1.96)$$
$$= F_{\frac{S_n - pn}{\sqrt{p(1-p)n}}}(1.96) - F_{\frac{S_n - pn}{\sqrt{p(1-p)n}}}(-1.96)$$
$$\xrightarrow[n\to\infty]{} F_{\mathcal{N}(0,1)}(1.96) - F_{\mathcal{N}(0,1)}(-1.96) \qquad \text{par convergence en loi due au TCL}$$
$$= 1 - \mathbb{P}\left(|Z| \geq 1.96\right),$$

où $Z \sim \mathcal{N}(0,1)$. En lisant dans la table de statistique de la loi normale, on trouve que $\mathbb{P}\left(|Z| \geq 1.96\right) \simeq 0.05$ et donc que :

$$\lim_{n\to\infty} \mathbb{P}\left(S_n \in I_n\right) \simeq 0.95.$$

**<u>Correction exercice 4</u> :** On note $X_i = 1$ si la $i$ème personne interrogée compte voter pour le maire et 0 sinon. Comme on interroge $n$ individus aux hasards et que l'on note $p$ la proportion de la population souhaitant voter pour le maire, on fait l'hypothèse que les $(X_i)_{0\leq i\leq n}$ sont indépendant et de loi $\mathcal{B}(p)$. La fréquence empirique $F_n$ est la proportion des réponses favorables à la réélection du maire :

$$F_n = \frac{\sum_{i=1}^n X_i}{n}.$$

1. D'après l'inégalité de Bienaymé-Tchebychev, on a que

$$\mathbb{P}\left(|F_n - p| \geq 0.01\right) \leq \frac{\mathrm{Var}(F_1)}{0.01^2} = \frac{10000p(1-p)}{n} \leq \frac{2500}{n},$$

   car $p(1-p) \leq 1/4$ pour tout $p \in [0,1]$.

   Il faut donc interroger $n = \dfrac{2500}{0.05} = 50000$ personnes pour estimer $p$ avec une précision de 1% et avec une précision de 95%.

2. Le TCL donne la convergence en loi suivante :

$$\sqrt{n}\frac{F_n - p}{\sqrt{p(1-p)}} \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1).$$

   On en déduit l'approximation

$$\begin{aligned}
\mathbb{P}\left(|F_n - p| \geq 0.01\right) &= \mathbb{P}\left(\sqrt{n}\frac{|F_n - p|}{\sqrt{p(1-p)}} \geq \sqrt{n}\frac{0.01}{\sqrt{p(1-p)}}\right) \\
&= \mathbb{P}\left(\sqrt{n}\frac{|F_n - p|}{\sqrt{p(1-p)}} \geq \sqrt{n}\frac{0.01}{\sqrt{1/4}}\right) \\
&\simeq \mathbb{P}\left(|Z| \geq 2\sqrt{n}0.01\right) \qquad \text{avec } Z \sim \mathcal{N}(0,1).
\end{aligned}$$

   Or, d'après la table de statistique, $\mathbb{P}(|Z| \geq 1.96) = 0.05$. On en déduit donc qu'il faut choisir $n$ de sorte que $2\sqrt{n}0.01 = 1.96$ soit $n = 9604$. Il faut donc interroger environ 9604 personnes pour estimer $p$ avec une précision de 1% et avec une précision de 95%. Le théorème central limite permet donc d'estimer avec plus de précision la proportion $p$ que l'inégalité de Bienaymé-Tchebychev.