Théorème central limite

- Bernis, *Analyse pour l'agrégation de mathématiques, 40 développements.* (207-215)

Lemme : Soit $n \in \mathbb{N}^*$. Soient $a, b \in \overline{D(0,1)}$. Alors, $|a^n - b^n| \leqslant n|a-b|$.

Démonstration. $|a^n - b^n| = \left|(a-b)\sum_{k=0}^{n-1} a^{n-1-k}b^k\right| \leqslant |a-b|\sum_{k=0}^{n-1} \underbrace{|a^{n-1-k}b^k|}_{\leqslant 1} = n|a-b|$ □

Théorème central limite : Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d. de moment d'ordre 2. Alors, $\dfrac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\mathrm{Var}(X_1)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

Démonstration. Quitte à centrer et à réduire en remplaçant X_n par $Y_n = \dfrac{X_n - \mathbb{E}[X_1]}{\sqrt{\mathrm{Var}(X_1)}}$.
On peut se ramener sans perte de généralité au cas où $\mathbb{E}[X_1] = 0$ et $\mathrm{Var}(X_1) = 1$.
D'après le théorème de Lévy, il suffit de montrer que $\forall t \in \mathbb{R}$, $\varphi_{\frac{S_n}{\sqrt{n}}}(t) \underset{n\to+\infty}{\longrightarrow} e^{-\frac{t^2}{2}}$.
Soit $n \in \mathbb{N}^*$. Soit $t \in \mathbb{R}$.

$$\begin{aligned}
\varphi_{\frac{S_n}{\sqrt{n}}}(t) &= \mathbb{E}\left[\exp\left(it\frac{S_n}{\sqrt{n}}\right)\right] = \mathbb{E}\left[\exp\left(i\frac{t}{\sqrt{n}}\sum_{k=1}^{n} X_k\right)\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{n} e^{it\frac{X_k}{\sqrt{n}}}\right] = \prod_{k=1}^{n}\mathbb{E}\left[e^{it\frac{X_k}{\sqrt{n}}}\right] \qquad \text{car les } X_k \text{ sont indépendants} \\
&= \left(\mathbb{E}\left[e^{it\frac{X_1}{\sqrt{n}}}\right]\right)^n = \varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n \qquad \text{car les } X_k \text{ ont même loi}
\end{aligned}$$

Comme X_1 admet un moment d'ordre 2 alors φ_{X_1} est de classe $\mathcal{C}^2$ avec

$$\varphi'_{X_1}(0) = \mathbb{E}[iX_1] = 0 \quad \text{et} \quad \varphi''_{X_1}(0) = \mathbb{E}[-X_1^2] = -\mathrm{Var}(X_1) = -1$$

Donc, φ_{X_1} admet un développement limité d'ordre 2 au voisinage de 0 de la forme

$$\varphi_{X_1}(h) = 1 - \frac{h^2}{2} + \underset{h\to 0}{o}(h^2) = 1 - \frac{h^2}{2} + h^2\varepsilon(h) \quad \text{où } \varepsilon : \mathbb{R} \to \mathbb{C} \text{ et } \lim_{h\to 0}\varepsilon(h) = 0$$

Choisissons $N \in \mathbb{N}^*$ tel que $\forall n \geqslant N$, $\left|1 - \dfrac{t^2}{2n}\right| \leqslant 1$. Comme $|\varphi_{X_1}(h)| = |\mathbb{E}[e^{ihX_1}]| \leqslant 1$ alors

$$\begin{aligned}
\left|\varphi_{\frac{S_n}{\sqrt{n}}}(t) - \left(1 - \frac{t^2}{2n}\right)^n\right| &= \left|\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| \\
&\leqslant n\left|\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) - \left(1 - \frac{t^2}{2n}\right)\right| \qquad \text{par le lemme} \\
&= t^2\left|\varepsilon\left(\frac{t}{\sqrt{n}}\right)\right| \underset{n\to+\infty}{\longrightarrow} 0
\end{aligned}$$

Or $\left(1 - \frac{t^2}{2n}\right)^n \underset{n\to+\infty}{\longrightarrow} e^{-\frac{t^2}{2}} = \varphi_X(t)$ où $X \sim \mathcal{N}(0,1)$.
Par l'inégalité triangulaire, on trouve que $\forall t \in \mathbb{R}$, $\varphi_{\frac{S_n}{\sqrt{n}}}(t) \underset{n\to+\infty}{\longrightarrow} \varphi_X(t)$.
Donc d'après le théorème de Lévy, on a $\dfrac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. □

Application : Intervalle de confiance asymptotique
Soit $(X_1, \dots, X_n)$ un échantillon de loi $\mathcal{B}(p)$ avec $p \in]0,1[$. Soit $\alpha \in]0,1[$.
Un intervalle de confiance asymptotique au niveau $1-\alpha$ est

$$I_{1-\alpha} = \left[\bar{X}_n - q_\alpha \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + q_\alpha \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \right]$$

Démonstration. Par la loi faible des grands nombres, $\bar{X}_n \xrightarrow{\mathbb{P}} p$ donc par continuité, $\sqrt{\bar{X}_n(1-\bar{X}_n)} \xrightarrow{\mathbb{P}} \sqrt{p(1-p)}$. De plus, par le TCL, $\dfrac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.
D'après le lemme de Slutsky (si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathbb{P}} c$ alors $X_nY_n \xrightarrow{\mathcal{L}} cX$)

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} = \frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1-p)}} \times \frac{\sqrt{p(1-p)}}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \sim Z$$

Par convergence en loi, $\forall t \in \mathbb{R},\ F_{Z_n}(t) \to F_Z(t)$ donc $\mathbb{P}(Z_n \in [-t,t]) \to \mathbb{P}(Z \in [-t,t])$.
Posons $q_\alpha = F_Z^{-1}\left(1 - \frac{\alpha}{2}\right)$ le quantile d'ordre $1 - \frac{\alpha}{2}$ de la $\mathcal{N}(0,1)$. Donc,

$$\mathbb{P}(Z_n \in [-q_\alpha, q_\alpha]) \underset{n \to +\infty}{\longrightarrow} \mathbb{P}(Z \in [-q_\alpha, q_\alpha]) = 1 - \alpha$$

Or, $\mathbb{P}(Z_n \in [-q_\alpha, q_\alpha]) = \mathbb{P}\left(\bar{X}_n - q_\alpha \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \leqslant p \leqslant \bar{X}_n + q_\alpha \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \right)$. $\square$