Théorème Central Limite [1]

Leçons : 218, 249 [2], 261, 262, 263, *223, 245, 260*

[ZQ], section XIII.II.3.d)
[Nou], proposition 1.19.10

Théorème

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de vaiid L^2, $S_n = \sum_{i=1}^{n} X_i$, $m = \mathbb{E}[X_1]$ et $\sigma^2 = \mathrm{Var}\ (X_1) > 0$.
Alors on a : $\dfrac{S_n - nm}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$.

Démonstration :

Quitte à travailler avec $\dfrac{X_i - m}{\sigma}$ en lieu et place de X_i, on peut supposer que $m = 0$ et $\sigma^2 = 1$.

On va donc montrer que $\dfrac{S_n}{\sqrt{n}}$ converge en loi vers une gaussienne centrée réduite.

On va utiliser le théorème de Lévy et montrer que : $\forall t \in \mathbb{R}, \varphi_{\frac{S_n}{\sqrt{n}}}(t) \xrightarrow[n\to\infty]{} \mathrm{e}^{-\frac{t^2}{2}}$.

En effet, on dispose du lemme suivant.

Lemme 1

Si $X \sim \mathcal{N}(0,1)$,
Alors $\forall t \in \mathbb{R}, \varphi_X(t) = \mathrm{e}^{-\frac{t^2}{2}}$.

Démonstration du lemme 1 :

Par le lemme de transfert, $\varphi_X(t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-\frac{x^2}{2}}\, \mathrm{d}x$.

En fait, on va calculer $G(z) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \underbrace{\mathrm{e}^{zx}\mathrm{e}^{-\frac{x^2}{2}}}_{f(z,x)}\, \mathrm{d}x$.

G est bien définie sur $\mathbb{C}$, on va appliquer le théorème d'holomorphie sous le signe intégrale pour montrer que G est holomorphe.
Plus précisément, on va l'appliquer sur tout disque $\mathcal{D}(0,R)$ où $R > 0$:
- $\forall z \in \mathcal{D}(0,R), x \mapsto f(z,x)$ est intégrable sur $\mathbb{R}$;
- $\forall x \in \mathbb{R}, z \mapsto f(z,x)$ est holomorphe sur $\mathcal{D}(0,R)$;

---

1. On dispose de l'application suivante.
Soit une pièce qu'on lance n fois ; on symbolise les résultats des lancers par une suite $(X_i)_{i\in\mathbb{N}^*}$ de vaiid de loi $b(p)$. On cherche à estimer le paramètre p.
D'après le TCL, on a : $\dfrac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to\infty]{b(p)-\mathcal{L}} \mathcal{N}(0,1)$.
En notant $\overline{X_n} = \dfrac{S_n}{n}$, cela revient à dire $\sqrt{\dfrac{n}{p(1-p)}}\left(\overline{X_n} - p\right) \xrightarrow[n\to\infty]{b(p)-\mathcal{L}} \mathcal{N}(0,1)$.
Mais, par la loi faible des grands nombres : $\sqrt{\overline{X_n}\left(1-\overline{X_n}\right)} \xrightarrow[n\to\infty]{b(p)-\mathbb{P}} \sqrt{p(1-p)}$.
Donc, en utilisant le lemme de Slutsky : $\sqrt{\dfrac{n}{\overline{X_n}\left(1-\overline{X_n}\right)}}\left(\overline{X_n} - p\right) \xrightarrow[n\to\infty]{b(p)-\mathcal{L}} \mathcal{N}(0,1)$.
Soit $N \sim \mathcal{N}(0,1)$, q le quantile d'ordre $1 - \dfrac{\alpha}{2}$ de $\mathcal{N}(0,1)$, on a :
$$\mathbb{P}_p\left(-q \leqslant \sqrt{\frac{n}{\overline{X_n}\left(1-\overline{X_n}\right)}}\left(\overline{X_n} - p\right) \leqslant q\right) \simeq \mathbb{P}(N \leqslant q) - \mathbb{P}(N \leqslant -q) = 2\mathbb{P}(N \leqslant q) - 1 = 1 - \alpha.$$
Donc, pour n assez grand, on dispose de l'intervalle de confiance asymptotique $\left[\overline{X_n} - \dfrac{q}{\sqrt{n}}\sqrt{\overline{X_n}\left(1-\overline{X_n}\right)}, \overline{X_n} + \dfrac{q}{\sqrt{n}}\sqrt{\overline{X_n}\left(1-\overline{X_n}\right)}\right]$ qui contient p avec une probabilité proche de $1 - \alpha$.

2. Pour la leçon 249, on peut ajouter une comparaison de calculs d'intervalles de confiance asymptotiques, entre l'inégalité de Tchebychev et le théorème central limite ; pour cela, on peut prendre exemple sur le document de Laura GAY.

– $\forall x \in \mathbb{R}, \forall z \in \mathcal{D}(0,R), |f(z,x)| = \left|e^{zx}e^{-\frac{x^2}{2}}\right| = e^{x\mathrm{Re}z}e^{-\frac{x^2}{2}} \leqslant e^{Rx}e^{-\frac{x^2}{2}}$, majorant qui a eu la bonne idée d'être à la fois intégrable et indépendant de z.

Donc $\forall R > 0$, G est holomorphe sur $\mathcal{D}(0,R)$ et donc sur $\mathbb{C}$.

Soit $u \in \mathbb{R}$, $G(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ux-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\frac{u^2}{2}} e^{-\frac{(x-u)^2}{2}}\, dx = e^{\frac{u^2}{2}}$.

G et $z \mapsto e^{\frac{z^2}{2}}$ sont des fonctions holomorphes qui coïncident sur $\mathbb{R}$, donc, par prolongement analytique, elles sont égales sur $\mathbb{C}$.

En particulier, $\forall t \in \mathbb{R}, \varphi_X(t) = G(\mathrm{i}t) = e^{-\frac{t^2}{2}}$. ■

Comme $X_1 \in \mathrm{L}^2$, φ_{X_1} est de classe $\mathcal{C}^2$; de plus :

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

D'autre part, $\forall t \in \mathbb{R}, \varphi_{\frac{S_n}{\sqrt{n}}}(t) = \varphi_{S_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$, car les variables X_i sont iid.

φ_{X_1} étant de classe $\mathcal{C}^2$, on a le développement limité :

$$\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = \varphi_{X_1}(0) + \frac{t}{\sqrt{n}}\varphi'_{X_1}(0) + \frac{t^2}{2n}\varphi''_{X_1}(0) + \frac{\varepsilon_n}{n} = 1 - \frac{t^2}{2n} + \frac{\varepsilon_n}{n}, \text{ où } \varepsilon_n \underset{n\to\infty}{\longrightarrow} 0.$$

On en déduit alors que $\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + \frac{\varepsilon_n}{n}\right)^n$.

Lemme 2

Si $(z_n)_{n\in\mathbb{N}^*}$ est une suite de nombres complexes qui tend vers $z \in \mathbb{C}$, Alors $\lim\limits_{n\to\infty}\left(1 + \frac{z_n}{n}\right)^n = e^z$.

Démonstration du lemme 2 :

On a : $e^{z_n} - \left(1 + \frac{z_n}{n}\right)^n = \sum\limits_{k=0}^{\infty} a_{k,n} z_n^k$ avec $a_{k,n} = \begin{cases} \frac{1}{k!}\left(1 - \frac{n(n-1)\dots(n-k+1)}{n^k}\right) & \text{si } k \leqslant n \\ \frac{1}{k!} & \text{sinon} \end{cases}$.

Les $a_{k,n}$ sont donc tous positifs, on en déduit :

$$\begin{aligned}
\left|e^{z_n} - \left(1 + \frac{z_n}{n}\right)^n\right| &\leqslant \sum_{k=0}^{\infty} a_{k,n} |z_n|^k \\
&= e^{|z_n|} - \left(1 + \frac{|z_n|}{n}\right)^n \\
&= e^{|z_n|} - \exp\left(n \ln\left(1 + \frac{|z_n|}{n}\right)\right) \\
&\leqslant e^{|z_n|} - \exp\left(n\left(\frac{|z_n|}{n} - \frac{|z_n|^2}{2n^2}\right)\right) \quad \text{car } \ln(1+x) \geqslant x - \frac{x^2}{2} \text{ (pour } x > 0) \\
&\leqslant e^{|z_n|}\left(1 - \exp\left(-\frac{|z_n|^2}{2n}\right)\right) \\
&\leqslant e^{|z_n|}\frac{|z_n|^2}{2n} \quad \text{car } 1 - e^x \leqslant x \text{ (pour } x \in \mathbb{R})
\end{aligned}$$

Donc $\left|e^z - \left(1 + \frac{z_n}{n}\right)^n\right| \leqslant |e^z - e^{z_n}| + e^{|z_n|}\frac{|z_n|^2}{2n} \underset{n\to\infty}{\longrightarrow} 0$. ■

On applique le lemme 2 avec $z_n = -\frac{t^2}{2} + \varepsilon_n$, d'où $\varphi_{\frac{S_n}{\sqrt{n}}}(t) \underset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}$. ■

Références

[ZQ] H. QUEFFÉLEC et C. ZUILY – *Analyse pour l'agrégation*, 4e éd., Dunod, 2013.
[Nou] I. NOURDIN – *Agrégation de mathématique - Épreuve orale*, 2e éd., Dunod, 2006.