Théorème central limite

Référence : Zuily-Queffelec : Analyse pour l'agrégation, p540+555+563

Lemme 1

Pour toute suite de complexes $(z_n)_n$ convergeant vers $z \in \mathbb{C}$ on a

$$\left(1+\frac{z_n}{n}\right)^n \underset{n}{\longrightarrow} \mathrm{e}^z.$$

Preuve du Lemme :

Par la formule du binôme de Newton, on a $\mathrm{e}^{z_n} - \left(1+\frac{z_n}{n}\right)^n = \sum_{k=0}^{+\infty} \frac{z_n^k}{k!} - \sum_{k=0}^{n} \binom{n}{k}\frac{z_n^k}{n^k} = \sum_{k=0}^{\infty} a_k^{(n)} z_n^k$ 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 positifs [1] . Donc :

$$\begin{aligned}\left|\mathrm{e}^{z_n} - \left(1+\frac{z_n}{n}\right)^n\right| = \left|\sum_{k=0}^{\infty} a_k^{(n)} z_n^k\right| \leqslant \sum_{k=0}^{\infty} a_k^{(n)} |z_n|^k &= \mathrm{e}^{|z_n|} - \left(1+\frac{|z_n|}{n}\right)^n = \mathrm{e}^{|z_n|} - \mathrm{e}^{\left(n\ln(1+\frac{|z_n|}{n})\right)} \\ &\underset{(\star)}{\leqslant} \mathrm{e}^{|z_n|} - \mathrm{e}^{n\left(\frac{|z_n|}{n} - \frac{|z_n|^2}{2n^2}\right)} = \mathrm{e}^{|z_n|}\left(1 - \mathrm{e}^{\left(-\frac{|z_n|^2}{2n}\right)}\right) \\ &\underset{(\star\star)}{\leqslant} \frac{|z_n|^2}{2n}\mathrm{e}^{|z_n|}\end{aligned}$$

$(\star)$ et $(\star\star)$ découlant à chaque fois d'une simple étude de fonction ou d'une formule de Taylor [2].
Cette inégalité nous permet alors de conclure car :

$$\left|\mathrm{e}^z - \left(1+\frac{z_n}{n}\right)\right| \leqslant |\mathrm{e}^z - \mathrm{e}^{z_n}| + \left|\mathrm{e}^{z_n} - \left(1+\frac{z_n}{n}\right)\right| \leqslant |\mathrm{e}^z - \mathrm{e}^{z_n}| + \frac{|z_n|^2}{2n}\mathrm{e}^{|z_n|}$$

avec le terme de droite tendant vers 0 puisque la suite (z_n) converge vers z (et que exp est continue) et z_n est alors bornée à partir d'un certain rang. ■

Lemme 2 (Fonction caractéristique de la loi normale)

Si $X \sim \mathcal{N}(0,1)$. Alors X a pour fonction caractéristique $\varphi_X(t) = \mathrm{e}^{-t^2/2}$.

Preuve (Nourdin p.82) :

1. $1 - \frac{n(n-1)\dots(n-k+1)}{n^k} = 1 - \underbrace{\overbrace{\frac{n}{n}}^{\leqslant 1}\overbrace{\frac{n-1}{n}}^{\leqslant 1}\cdots\overbrace{\frac{n-k+1}{n}}^{\leqslant 1}}_{\leqslant 1} \geqslant 0$
2. Pour $x > 0$ $\ln(1+x) \geqslant x - \frac{x^2}{2}$ et pour tout x, $1 - \mathrm{e}^{-x} \leqslant x$

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

Nous allons faire un peu plus et calculer, pour $z \in \mathbb{C}$, $G(z) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \underbrace{\mathrm{e}^{zx - x^2/2}}_{f(z,x)}\,\mathrm{d}x$

On veut appliquer le théorème d'holomorphie sous le signe intégral pour montrer que G est holomorphe sur $\mathbb{C}$. On va pour cela montrer que G est holomorphe sur $D(0,R)$, $R > 0$.

— $\forall z \in D(0,R)$, $x \to f(z,x)$ est intégrable sur $\mathbb{R}$.

— $\forall x \in \mathbb{R}$ $z \to f(z,x)$ est holomorphe dans $D(0,R)$.

— $\forall x \in \mathbb{R}, \forall z \in \mathcal{D}(0,R), |f(z,x)| = \left|\mathrm{e}^{zx}\mathrm{e}^{-\frac{x^2}{2}}\right| = \mathrm{e}^{x\mathrm{Re}z}\mathrm{e}^{-\frac{x^2}{2}} \leqslant \underbrace{\mathrm{e}^{Rx}\mathrm{e}^{-\frac{x^2}{2}}}_{\text{intégb et indép de } z}$

Donc G est holomorphe sur $D(0,R)$ et donc sur $\mathbb{C}$ (car holomorphe en tout point).
Calculons maintenant G sur $\mathbb{R}$. Soit $u \in \mathbb{R}$. On a :

$$G(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{ux - \frac{x^2}{2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{\frac{u^2}{2}}\mathrm{e}^{-\frac{(x-u)^2}{2}}\,\mathrm{d}x = \mathrm{e}^{\frac{u^2}{2}}$$

Ainsi, G et $z \to \mathrm{e}^{z^2/2}$ sont des fonctions holomorphes qui coïncident sur $\mathbb{R}$ (ouvert non vide de $\mathbb{C}$ donc sur $\mathbb{C}$ (ouvert connexe) par principe de prolongement analytique.
En particulier, si $t \in \mathbb{R}$, $\varphi_X(t) = G(\mathrm{i}t) = \mathrm{e}^{-t^2/2}$. ■

Théorème (Théorème central limite)

Soit $(X_n)_{n\geqslant 1}$ une suite de va iid dans L^2.
On note $S_n = \displaystyle\sum_{i=1}^{n} X_i$, $m = \mathbb{E}[X_1]$ et $\sigma^2 = \mathrm{Var}(X_1) > 0$. Alors on a :

$$\frac{S_n - nm}{\sqrt{n\sigma^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

<u>Preuve :</u>

Quitte à considérer les variables aléatoires $Y_n = \dfrac{X_n - m}{\sigma}$ on peut supposer $m = 0$ et $\sigma = 1$.
On cherche donc à montrer que $\dfrac{S_n}{\sqrt{n}}$ converge en loi vers une gaussienne centrée réduite. Pour cela, on va utiliser le théorème de Lévy : on va montrer

$$\varphi_{\frac{S_n}{\sqrt{n}}}(t) \underset{n}{\longrightarrow} \mathrm{e}^{-\frac{t^2}{2}} \quad \forall t \in \mathbb{R}$$

car si $N \sim \mathcal{N}(0,1)$, alors $\varphi_N(t) = \mathrm{e}^{-\frac{t^2}{2}}$.

$\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \varphi_{S_n}\left(\dfrac{t}{\sqrt{n}}\right) = \left(\varphi_{X_1}\left(\dfrac{t}{\sqrt{n}}\right)\right)^n$ car les va X_i sont iid.

De plus, comme X_1 est dans L^2, on en déduit [3] que φ est de classe C^2 et on connait ses dérivées en fonctions des moments de X_1 :
$\varphi'(0) = \mathrm{i}\mathbb{E}[X^1] = \mathrm{i}m = 0$
$\varphi''(0) = \mathrm{i}^2\mathbb{E}[X^2] = -(\mathrm{Var}[X] + \underbrace{\mathbb{E}[X]^2}_{0}) = -\sigma^2 = -1$

On peut donc écrire un développement de Taylor de φ à l'ordre 2 en l'origine :

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

On en déduit alors que $\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \dfrac{t^2}{2n} + \dfrac{\varepsilon_n}{n}\right)^n$. On voudrait maintenant passer à la limite quand n tend vers l'infini, mais la quantité à l'intérieur de la parenthèse n'est pas réelle (une fonction caractéristique est à

3. Prop 12 de la leçon 261 : Si X v.a. réelle admet un moment d'ordre n, alors φ est $\mathcal{C}^n$

valeurs complexes, c'est caché dans le ε_n ici). Nous allons donc utiliser le **Lemme 1** qui traite le cas complexe. On l'applique à l'expression trouvée pour $\varphi_{\frac{S_n}{\sqrt{n}}}$, avec $z_n = -\frac{t^2}{2} + \varepsilon_n$,on obtient :

$$\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + \varepsilon_n\right)^n \underset{n}{\longrightarrow} \mathrm{e}^{-\frac{t^2}{2}}$$

ce qui termine la démonstration du théorème.

■

Application : calcul d'intervalle de confiance asymptotique

Pour la leçon 249 ; Introduction : Les Bernouilli c'est la vie. Cadre-Vial p.1 ou Rivoirard-Stoltz p.1
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. Pour les application numériques dans la suite on prendra $n = 1000$.

Méthode nulle : Tchebychev (Cadre-Vial p.3)

Pour tout $\varepsilon > 0$

$$\mathbb{P}\left(\left|\bar{X}_n - p\right| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} \leqslant \frac{1}{4n\varepsilon^2}$$

Donc si on veut un intervalle de confiance à 95% il faut prendre $\frac{1}{4n\varepsilon^2} \leqslant 0.05$. Donc $\varepsilon = 0.08$.
Par suite,

$$\mathbb{P}\left(p \in [\bar{X}_n - 0.08, \bar{X}_n + 0.08]\right) \geqslant 0.95$$

Méthode bien : TCL ($\approx$ Rivoirard-Stoltz p.25)

D'après le TCL on a : $\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

Donc en notant $\bar{X}_n = \frac{S_n}{n}$ on a : $\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

De plus, la LGNf donne : $\sqrt{\bar{X}_n(1-\bar{X}_n)} \xrightarrow{\mathbb{P}} \sqrt{p(1-p)}$.

Donc par le théorème de Slutsky on a finalement : $\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

Ainsi, en notant q le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0,1)$ (ie $\mathbb{P}(N \leqslant q) = 1 - \frac{\alpha}{2}$ si $N \sim \mathcal{N}(0,1)$) on a :

$$\mathbb{P}\left(-q \leqslant \frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \leqslant q\right) \simeq \mathbb{P}(N \leqslant q) - \mathbb{P}(N \leqslant -q) = 2\mathbb{P}(N \leqslant q) - 1 = 1 - \alpha$$

pour n assez grand.

Donc pour n assez grand, p est dans l'intervalle $\left[\bar{X}_n - \frac{q}{\sqrt{n}}\sqrt{\bar{X}_n(1-\bar{X}_n)}; \bar{X}_n + \frac{q}{\sqrt{n}}\sqrt{\bar{X}_n(1-\bar{X}_n)}\right]$ avec une probabilité proche de $1 - \alpha$.
En pratique ; on prend souvent $\alpha = 0.05$ et dans ce cas $q = 1,96$.
Application numérique : on obtient dans notre cas

$$\frac{q}{\sqrt{1000}}\sqrt{\bar{X}_n(1-\bar{X}_n)} \approx \frac{1.96}{33}\frac{1}{\sqrt{4}} \approx 0.03$$

Donc 0.03 mieux que 0.08 !

Notes :

✓ **A l'oral**, on fait fonction caractéristique de la loi normale + intervalle de confiance. 8'49 TCL + fct caract ; 14'38 au total. Cependant, dans le leçon 249 on fait les 2 trucs d'intervalle de confiance + TCL seulement (parce que c'est la meilleure méthode !)
✓ Le TCL est un théorème clé en théorie des probabilités. Il souligne le rôle central des variables gaussiennes,

qui peuvent être vues comme le comportement global d'une multitude de petits phénomènes. Par exemple, les chocs de molécules d'eau sur une molécule de pollen ou les effets des conditions atmosphériques sur le plan de vol d'un avion peuvent être modélisés par des variables gaussiennes. En pratique, quand $n \geqslant 30$, le TCL fournit une bonne approximation de la situation (et on peut estimer son erreur grâce à l'inégalité de Berry-Esseen).
✓ Rappel : Théorème de continuité de Lévy.

$$\{\forall t \in \mathbb{R} : \varphi_n(t) \to \varphi(t)\} \quad \Leftrightarrow \quad \left\{X_n \xrightarrow{\mathcal{L}} X\right\}$$

✓ On peut également prendre dans le Ouvrard 2 mais le lemme n'est pas fait pour une suite z_n. De plus, il est fait en multidimensionnel ce qui complique les notations.