

Théorème central limite

Thomas CHEN

Le but de cette feuille est de démontrer le théorème suivant.

Théorème 1. Soit $(X_n)_n$ suite de variables aléatoires indépendantes et identiquement distribuées admettant des moments d'ordre 2. On note $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \text{Var}(X_1)$. Alors

$$\frac{S_n - \mu n}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

On rappelle qu'une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire X lorsque pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée, on a

$$\mathbb{E}[f(X_n)] \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[f(X)].$$

1 Un peu de topologie

Soit $N \in \mathbb{N}^*$ et Ω un ouvert de \mathbb{R}^N . On note $\mathcal{C}_b(\omega)$ l'ensemble des fonctions continues bornées sur Ω . C'est un espace vectoriel que l'on munit de la norme infinie. Pour toute fonction $f : \Omega \rightarrow \mathbb{R}$, on appelle *support de f* l'ensemble $\text{supp}(f)$ défini par

$$\text{supp}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

Par exemple, le support de \sin est \mathbb{R} puisque l'adhérence de $\mathbb{R} \setminus \pi\mathbb{Z}$ est \mathbb{R} . Autre exemple, pour toute partie U de \mathbb{R} , le support de $\mathbf{1}_U$ est \bar{U} .

On définit $\mathcal{C}_c(\Omega)$ l'ensemble des fonctions de Ω dans \mathbb{R} à support compact (et comme un support est fermé par construction, il est équivalent à demander « fonctions à support inclus dans un compact »).

1. Montrer que $(\mathcal{C}_c(\Omega), \|\cdot\|_\infty)$ est un sous-espace vectoriel normé de $(\mathcal{C}_b(\Omega), \|\cdot\|_\infty)$.
2. Dans cette question, $N = 1$ et $\Omega =]-1, 1[$. On considère la fonction trapèze sur $[-1, 1]$ f définie par

$$f(x) = \begin{cases} 2(x+1) & \text{si } x \in]-1, -1/2[\\ 1 & \text{si } x \in [-1/2, 1/2] \\ 2(1-x) & \text{si } x \in [1/2, 1] \end{cases}$$

- (a) Quel est le support de f ?
- (b) Trouver une suite de fonctions $(f_n)_{n \in \mathbb{N}}$ dans $\mathcal{C}_c(\Omega)$ telle que pour tout $n \in \mathbb{N}$, f_n soit affine par morceaux sur Ω et $\text{supp}(f_n) = \left[-1 + \frac{1}{n}, 1 - \frac{1}{n}\right]$.
- (c) En déduire que $\mathcal{C}_c(\Omega)$ n'est pas fermé dans $\mathcal{C}_b(\Omega)$.

On définit alors $\mathcal{C}_0(\Omega)$ l'adhérence de $\mathcal{C}_c(\Omega)$ dans $\mathcal{C}_b(\Omega)$. Le but de la suite est de montrer que $f \in \mathcal{C}_0(\Omega)$ si, et seulement si, f est continue sur Ω et $\lim_{x \rightarrow \omega} f(x) = 0$ pour tout $\omega \in \partial\Omega$ (l'idée étant que $\mathcal{C}_0(\Omega)$ est l'ensemble des fonctions continues qui sont nulles sur le bord).

3. (Le lemme d'Urysohn) Soit K un compact de \mathbb{R}^N .
 - (a) Montrer qu'il existe V un ouvert borné de \mathbb{R}^N qui contient K .
 - (b) Montrer qu'il existe $\bar{x} \in K$ tel que $d(\bar{x}, \mathbb{R}^d \setminus V) = \inf_{x \in K} d(x, \mathbb{R}^d \setminus V) =: d$.
 - (c) Montrer que $d > 0$.

- (d) En déduire qu'il existe un ouvert borné U satisfaisant $K \subset U \subset \bar{U} \subset V$.
- (e) En déduire une fonction f qui vaut 1 sur K et qui vérifie $\text{supp}(f) \subset \bar{U}$. On écrira f à l'aide de $x \mapsto d(x, \mathbb{R}^d \setminus U)$ et $x \mapsto d(x, K)$.
- (f) En déduire le lemme d'Urysohn : si K est un compact de \mathbb{R}^N et V un ouvert borné de \mathbb{R}^N contenant K , alors il existe une fonction f qui vaut 1 sur K avec un support dans V .
4. Montrer que $f \in \mathcal{C}_0(\Omega)$ si, et seulement si, pour tout $\varepsilon > 0$, il existe un compact $K_\varepsilon \subset \Omega$ tel que $|f| < \varepsilon$ sur $\Omega \setminus K_\varepsilon$.

On définit $\mathcal{C}_c^\infty(\mathbb{R})$ comme étant $\mathcal{C}^\infty(\mathbb{R}) \cap \mathcal{C}_c(\mathbb{R})$. On énonce et admet deux résultats de densité :

- $(\mathcal{C}_c^\infty(\mathbb{R}), \|\cdot\|_1)^1$ est dense dans $(L^1(\mathbb{R}), \|\cdot\|_1)$.
- $(\mathcal{C}_c^\infty(\mathbb{R}), \|\cdot\|_\infty)$ est dense dans $\mathcal{C}_0(\mathbb{R})$.

Ces deux résultats sont assez pénibles à démontrer à la main. L'outil magique est la convolution et les suites régularisantes.

2 Vers le théorème faible de Lévy

5. Dans cette question, on souhaite montrer que $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire X lorsque pour toute fonction $f \in \mathcal{C}_0(\mathbb{R})$, on a

$$\mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[f(X)].$$

- (a) Traiter le sens direct.
- (b) On s'attelle maintenant au sens réciproque. Soit $f \in \mathcal{C}_b(\mathbb{R})$. Soit $\varepsilon > 0$. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire.
- i. Montrer qu'il existe $A > 0$ tel que $\mathbb{P}_X(\mathbb{R} \setminus [-A, A]) < \varepsilon$.
 - ii. Exhiber une fonction trapèze φ (c'est-à-dire qui est de la même forme que la question 2 sur un intervalle fermé borné et éventuellement nul en dehors de cet intervalle) de sorte que $\varphi_{[-A, A]} = 1$ et $\text{supp}(\varphi) \subset [-2A, 2A]$.
 - iii. Pour tout $n \in \mathbb{N}$, on note

$$A_n := \int_{\mathbb{R}} f(1 - \varphi) d\mathbb{P}_{X_n} ; B_n := \left(\int_{\mathbb{R}} f\varphi d\mathbb{P}_{X_n} - \int_{\mathbb{R}} f\varphi d\mathbb{P}_X \right) ; C_n := \int_{\mathbb{R}} f(1 - \varphi) d\mathbb{P}_X.$$

A. Montrer que $\lim_{n \rightarrow +\infty} B_n = 0$.

B. Montrer que $\forall n \in \mathbb{N}, |C_n| \leq \|f\|_\infty \varepsilon$.

C. Montrer que $(A_n)_{n \in \mathbb{N}}$ est majorée par une suite qui converge vers $\|f\|_\infty \varepsilon$.

iv. En déduire que $A_n + B_n + C_n \xrightarrow{n \rightarrow +\infty} 0$.

v. Conclure.

6. On souhaite démontrer le théorème faible de Lévy.

Théorème 2 (Lévy). Soit $X, X_1, \dots, X_n, \dots$ des variables aléatoires réelles. Alors

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \phi_{X_n} \xrightarrow[n \rightarrow +\infty]{C.S.} \phi_X.$$

- (a) Traiter l'implication directe.
- (b) On passe à l'implication réciproque.

1. On rappelle que $\|f\|_1 = \int |f| dx$ avec les notations adéquates

i. Soit φ une fonction $L^1(\mathbb{R})$. On note

$$f : x \in \mathbb{R} \mapsto \int_{\mathbb{R}} \varphi(t) e^{-itx} dt.$$

On dit que f est la transformée de Fourier de φ et on note $f = \widehat{\varphi}$. Montrer que $f \in \mathcal{C}_0(\mathbb{R})$. On pourra donc traiter le cas où φ est $\mathcal{C}_c^\infty(\mathbb{R})$ (penser à une intégration par parties) puis conclure par densité.

ii. Soit $f \in \mathcal{S}(\mathbb{R})$ la classe de Schwartz : $\mathcal{S}(\mathbb{R})$ désigne l'ensemble des fonctions f de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ où $\forall p, q \in \mathbb{N}^*, \forall x \in \mathbb{R}, |x|^p |f^{(q)}(x)| < +\infty$. Montrer que $\forall x \in \mathbb{R}, \widehat{\widehat{f}}(x) = f(-x)$.

iii. Que dire de $\widehat{\widehat{\widehat{f}}}$?

iv. En déduire que pour toute fonction $f \in \mathcal{S}(\mathbb{R})$, il existe $\varphi \in L^1(\mathbb{R})$ telle que $f = \widehat{\varphi}$.

v. Montrer que $\mathcal{C}_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$.

vi. Conclure par densité de $\mathcal{C}_c^\infty(\mathbb{R})$ dans $\mathcal{C}_0(\mathbb{R})$.

3 Preuve du théorème central limite

7. Un lemme complexe.

(a) Rappeler pourquoi pour tout $z \in \mathbb{C}$, la série $\sum_{n \geq 0} \frac{z^n}{n!}$ converge. On note la somme de la série e^z .

(b) Montrer que

$$\forall n \in \mathbb{N}^*, \forall z \in \mathbb{C}, \left| \left(1 + \frac{z}{n}\right)^n - e^z \right| \leq \sum_{k=n+1}^{+\infty} \frac{|z|^k}{k!} + \sum_{k=0}^n \left| \frac{1}{k!} - \binom{n}{k} \frac{1}{n^k} \right| |z|^k.$$

(c) Montrer que pour tout $n \in \mathbb{N}$, pour tout $k \in \llbracket 0, n \rrbracket$,

$$\frac{1}{k!} - \binom{n}{k} \frac{1}{n^k} \geq 0.$$

(d) En déduire que

$$\forall n \in \mathbb{N}^*, \forall z \in \mathbb{C}, \left| \left(1 + \frac{z}{n}\right)^n - e^z \right| \leq \left(1 + \frac{|z|}{n}\right)^n - e^{|z|}.$$

(e) Montrer que le membre de droite de l'inégalité précédente tend vers 0 quand $n \rightarrow +\infty$ puis conclure que

$$\left(1 + \frac{z}{n}\right)^n \xrightarrow{n \rightarrow +\infty} e^z.$$

8. On rappelle (ou on admet²) que si X admet un moment d'ordre 2, alors ϕ_X est de classe \mathcal{C}^2 sur \mathbb{R} (et la réciproque est vraie). On prend les notations du théorème et on note $X := X_1$.

(a) On suppose que l'on connaît le théorème central limite dans le cas où $\mu = 0$ et $\sigma = 1$. Conclure quant au théorème général. **Dans la suite, on se place donc dans le cas particulier où $\mu = 0$ et $\sigma = 1$.**

(b) Calculer la fonction caractéristique d'une loi normale centrée réduite.

(c) Calculer $\mathbb{E}[e^{itS_n/\sqrt{n}}]$ à l'aide de ϕ_X .

(d) Conclure avec le théorème faible de Lévy et du lemme précédent.

2. la preuve n'est pas compliquée