

THÉORÈME CENTRAL LIMITE

- 234, 235, 241, 262, 209, 239, 250, 261 -

Nous allons démontrer ici le célèbre théorème central limite dont l'utilité n'est pas plus à vanter. En chemin, nous aurons à montrer divers résultats sur la convergence en loi, en particulier le très important théorème de Lévy qui permet d'étudier la convergence en loi à l'aide de la fonction caractéristique.

Pour une suite de variables aléatoires (X_n) convergeant en loi vers une variable aléatoire X de loi $\mathcal{L}$, on notera :

$$X_n \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} X \quad \text{ou} \quad X_n \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} \mathcal{L} \quad (1)$$

Lorsque X est une variable aléatoire à valeurs dans $\mathbb{R}$, on notera φ_X sa fonction caractéristique et $\mathbb{P}_X$ sa loi.

On notera C_b^0 l'ensemble des fonctions continues bornées de $\mathbb{R}$ dans $\mathbb{C}$, C_0^0 celui des fonctions continues tendant vers 0 à l'infini et C_c^∞ celui des fonctions C^∞ à support compact.

Lorsque f est L^1 , on notera :

$$\mathcal{F}[f] : \xi \mapsto \int_{\mathbb{R}} f(x) e^{-i\xi x} dx \quad (2)$$

sa transformée de Fourier.

Théorème de Lévy [1], p.536

Le résultat suivant est dû à Paul Lévy. Il montre que la convergence de $\mathbb{E}[f(X_n)]$ vers $\mathbb{E}[f(X)]$ pour tout $f \in C_b^0$ est obtenue simplement par la convergence de la famille à un paramètre $(x \mapsto e^{-itx})_{t \in \mathbb{R}}$, ce qui est loin d'être trivial !

Théorème 1 (Lévy). Soit (X_n) une suite de variables aléatoires réelles et X une autre variable aléatoire réelle. Les assertions suivantes sont équivalentes :

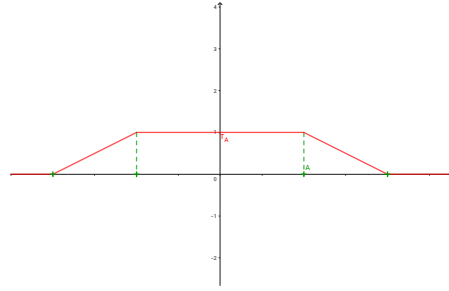
1. $X_n \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} X$
2. $\forall t \in \mathbb{R}, \varphi_{X_n}(t) \xrightarrow[n \rightarrow +\infty]{} \varphi_X(t)$

La démonstration du théorème de Lévy passera par le lemme suivant :

Lemme 1 ([1], II.18, p. 533). *Les assertions suivantes sont équivalentes :*

1. $\forall f \in C_b^0, \mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[f(X)]$ (i.e. $X_n \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} X$)
2. $\forall f \in C_0^0, \mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[f(X)]$

Démonstration. Il est évident que (1) $\implies$ (2) puisque $C_0^0 \subset C_b^0$. Soit donc $f \in C_b^0$. On va approcher f par des fonctions de C_0^0 à l'aide de fonctions trapèze. Etant donné un réel $A > 0$, on pose T_A la fonction trapézoïdale valant 1 sur $[-A, A]$ et 0 sur $\mathbb{R} \setminus [-2A, 2A]$.



Fixons $\varepsilon > 0$. Il existe alors $A \in \mathbb{R}$ tel que :

$$\mathbb{P}(|X| \geq A) = \int_{\mathbb{R}} \mathbf{1}_{|x| \geq A} d\mathbb{P}_X(x) \leq \varepsilon \quad (3)$$

On a alors :

$$|\mathbb{E}[f(X) - \mathbb{E}[f(X_n)]]| \leq \underbrace{|\mathbb{E}[f(X) - T_A f(X)]|}_{\alpha} + \underbrace{|\mathbb{E}[T_A f(X) - T_A f(X_n)]|}_{\beta_n} + \underbrace{|\mathbb{E}[T_A f(X_n) - f(X_n)]|}_{\gamma_n} \quad (4)$$

Contrôlons séparément chacun de ces trois termes.

•

$$\alpha = \left| \int_{\mathbb{R}} (1 - T_A(x)) f(x) d\mathbb{P}_X(x) \right| \quad (5)$$

$$\leq \|f\|_{\infty} \int_{\mathbb{R}} \mathbf{1}_{|x| \geq A} d\mathbb{P}_X(x) \quad (6)$$

$$\leq \|f\|_{\infty} \varepsilon \quad (7)$$

- $T_A f \in C_0^0$, donc par hypothèse :

$$\limsup_{n \rightarrow +\infty} \beta_n = 0 \quad (8)$$

•

$$\gamma_n = \left| \int_{\mathbb{R}} (T_A(x) - 1) f(x) d\mathbb{P}_{X_n}(x) \right| \quad (9)$$

$$\leq \|f\|_{\infty} \int_{\mathbb{R}} 1 - T_A(x) d\mathbb{P}_{X_n}(x) \quad (1 - T_A \geq 0) \quad (10)$$

$$= \|f\|_{\infty} \left(1 - \int_{\mathbb{R}} T_A(x) d\mathbb{P}_{X_n}(x) \right) \quad (11)$$

$$\xrightarrow{n \rightarrow +\infty} \|f\|_{\infty} \left(1 - \int_{\mathbb{R}} T_A(x) d\mathbb{P}_X(x) \right) \quad \text{car } T_A \in C_0^0 \quad (12)$$

$$\leq \|f\|_{\infty} \varepsilon \quad (13)$$

d'où on déduit :

$$\limsup_{n \rightarrow +\infty} \gamma_n \leq \|f\|_{\infty} \varepsilon \quad (14)$$

Finalement, en recombinaison ces inégalités, on obtient :

$$0 \leq \liminf_{n \rightarrow +\infty} |\mathbb{E}[f(X)] - \mathbb{E}[f(X_n)]| \leq \limsup_{n \rightarrow +\infty} |\mathbb{E}[f(X)] - \mathbb{E}[f(X_n)]| \leq 2\|f\|_{\infty} \varepsilon \quad (15)$$

et donc, en laissant tendre ε vers 0, on obtient bien $\mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[f(X)]$. $\square$

On va maintenant démontrer le théorème de Lévy. Comme il est évident que (1) $\implies$ (2), on se contente du sens réciproque. Commençons par montrer la convergence lorsque $f \in \mathcal{F}(L^1(\mathbb{R}))$, c'est-à-dire lorsque f est la transformée de Fourier d'une fonction $\psi \in L^1$, c'est-à-dire :

$$f(x) = \int_{\mathbb{R}} \psi(t) e^{ixt} dt \quad (16)$$

On a alors :

$$\mathbb{E}[f(X_n)] = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(t) \exp(itx) dt d\mathbb{P}_{X_n}(x) \quad (17)$$

$$= \int_{\mathbb{R}} \psi(t) \int_{\mathbb{R}} \exp(itx) d\mathbb{P}_{X_n}(x) dt \quad (\text{théorème de Fubini}) \quad (18)$$

$$= \int_{\mathbb{R}} \psi(t) \varphi_{X_n}(t) dt \quad (\text{par définition de l'espérance}) \quad (19)$$

$$\xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}} \psi(t) \varphi_X(t) dt \quad (\text{par convergence dominée}) \quad (20)$$

$$= \mathbb{E} \left[\int_{\mathbb{R}} \psi(t) \exp(itX) dt \right] \quad (\text{en appliquant encore Fubini}) \quad (21)$$

$$= \mathbb{E}[f(X)] \quad (22)$$

Or, la transformée de Fourier est bijective sur l'espace de Schwartz⁽ⁱ⁾. Donc toutes les fonctions Schwartz sont dans $\mathcal{F}(L^1)$, et en particulier, pour toute fonction $g \in C_c^{\infty}$, on a $\mathbb{E}[g(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[g(X)]$. Mais C_c^{∞} est dense dans C_0^0 pour la topologie de la norme ∞ ⁽ⁱⁱ⁾.

(i). Attention, ça n'est pas du tout évident à démontrer, et je pense qu'il faut au moins avoir une bonne idée de la preuve. Ça découle de la stabilité de l'espace de Schwartz par transformation de Fourier et de la formule d'inversion dans L^1 . Alternativement, on peut le démontrer à partir de la formule sommatoire de Poisson.

(ii). Voir annexe.

Donc si f est une fonction de C_0^0 , pour tout $\varepsilon > 0$, il existe $g \in C_c^\infty$ telle que $\|f - g\|_\infty \leq \varepsilon$. On a alors :

$$|\mathbb{E}[f(X)] - \mathbb{E}[f(X_n)]| \leq |\mathbb{E}[f(X) - g(X)]| + |\mathbb{E}[g(X) - g(X_n)]| + |\mathbb{E}[g(X_n) - f(X_n)]| \quad (23)$$

$$\leq 2\|f - g\|_\infty + |\mathbb{E}[g(X) - g(X_n)]| \quad (24)$$

et donc

$$0 \leq \limsup_{n \rightarrow +\infty} |\mathbb{E}[f(X)] - \mathbb{E}[f(X_n)]| \leq 2\varepsilon \quad (25)$$

ce quelque soit ε . Il ne reste qu'à appliquer le lemme pour montrer le théorème de Lévy !

Remarque : Pour appuyer le recasage dans la 261, on peut observer ce corollaire immédiat :

Théorème 2. Soient X, Y deux variables aléatoires réelles. Alors les assertions suivantes sont équivalentes :

1. $\mathbb{P}_X = \mathbb{P}_Y$
2. $\varphi_X = \varphi_Y$

Démonstration. Il est clair que (i) $\implies$ (ii). Supposons donc que $\varphi_X = \varphi_Y$. On peut alors considérer la suite de variables aléatoires (X_n) constante égale à X . Alors φ_{X_n} converge simplement vers φ_Y et donc $X_n \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} Y$. Par unicité de la limite pour la topologie de la convergence en loi, on a alors $\mathbb{P}_X = \mathbb{P}_Y$. $\square$

Théorème central limite

On va maintenant démontrer le théorème central limite :

Théorème 3 (Central limite). Soit (X_n) une suite de variables aléatoires réelles i.i.d. admettant une variance, notée $\sigma^2 \in \mathbb{R}_+$. Alors :

$$\frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^n (X_i - \mathbb{E}[X_i]) \xrightarrow[n \rightarrow +\infty]{(\mathcal{L})} \mathcal{N}(0, 1) \quad (26)$$

où $\mathcal{N}(0, 1)$ désigne la loi normale standard.

Pour cela, on aura besoin du petit lemme :

Lemme 2. Soit $z \in \mathbb{C}$. Alors :

$$\left(1 + \frac{z}{n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow[n \rightarrow +\infty]{} e^z \quad (27)$$

Démonstration. Pour n suffisamment grand, $|\frac{z}{n} + o(\frac{1}{n})| \leq \frac{1}{2}$. Désignons par $\log$ la détermination principale du logarithme holomorphe ⁽ⁱⁱⁱ⁾. A partir d'un certain rang, on a donc :

$$\left(1 + \frac{z}{n} + o\left(\frac{1}{n}\right)\right)^n = \exp\left(n \log\left(1 + \frac{z}{n} + o\left(\frac{1}{n}\right)\right)\right) \quad (28)$$

$$= \exp(z + o(1)) \quad (29)$$

$$= e^z(1 + o(1)) \quad (30)$$

ce qui prouve le lemme. $\square$

Tous les ingrédients sont maintenant réunis pour prouver le théorème central limite. Notons φ la fonction caractéristique commune des (X_i) et S_n le terme de gauche dans (26). Sans perte de généralité, on suppose les (X_i) centrées et réduites.

$$\forall t \in \mathbb{R}, \varphi_{S_n}(t) = \varphi_{\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i}(t) \quad (31)$$

$$= \prod_{i=1}^n \varphi\left(\frac{t}{\sqrt{n}}\right) \quad (32)$$

$$= \left(1 + i\mathbb{E}[X_1] \frac{t}{\sqrt{n}} - \mathbb{E}[(X_1)^2] \frac{t^2}{2n} + o\left(\frac{t}{n}\right)\right)^n \quad (33)$$

$$= \left(1 - \frac{t^2}{2n} + o\left(\frac{t}{n}\right)\right)^n \quad (34)$$

$$\xrightarrow[n \rightarrow +\infty]{} e^{-\frac{t^2}{2}} \quad (35)$$

en appliquant le lemme ^(iv). Ainsi, la fonction caractéristique de S_n converge simplement vers celle de la loi normale centrée réduite. Par le théorème de Lévy, on en déduit le théorème central limite.

Annexe

On va montrer ici un théorème utilisé plus tôt :

Théorème 4. C_c^∞ est dense dans C_0^0 pour la topologie de la norme ∞ .

Je n'ai pas trouvé (en toute honnêteté, je n'ai pas cherché) de référence qui fasse la démonstration. Dans [1], on ne trouve qu'un flegmatique "Stone-Weierstraß".

Démonstration. Soit $f \in C_0^0$, soit $\varepsilon > 0$. Soit $A > 0$ tel que :

$$\forall x \in \mathbb{R}, |x| \geq A \implies |f(x)| \leq \varepsilon \quad (36)$$

(iii). Il est possible de faire une preuve plus élémentaire de ce résultat à l'aide d'une simple inégalité. Je trouve personnellement que la preuve avec le log holomorphe est plus élégante et efficace, et puisque celui-ci est au programme, on aurait tort de se priver. Toutefois, il faut bien avoir conscience que ce résultat, très simple à prouver dans le cas réel, nécessite des arguments plus raffinés pour fonctionner dans le cas complexe.

(iv). Ici, la variable t est réelle, et donc on pourrait être tenté de penser qu'on pourrait se contenter du lemme dans le cas réel, et donc évitant ainsi le log holomorphe. Attention ici, car c'est le o qui est un complexe ! On a donc vraiment besoin du lemme dans le cas complexe, le cas réel ne suffit pas.

(qui existe car f tend vers 0 en $\pm\infty$). On fixe alors χ une fonction plateau telle que :

1. $\chi \in C_c^\infty$
2. $0 \leq \chi \leq 1$
3. $\forall x \in \mathbb{R}, (|x| \leq A \implies \chi(x) = 1) \wedge (|x| \geq A+1 \implies \chi(x) = 0)$

□

On a alors :

$$\|f - \chi f\|_\infty \leq \varepsilon \quad (37)$$

Par ailleurs, le théorème d'approximation de Stone-Weierstraß indique qu'il existe P une fonction polynomiale telle que :

$$\|(f - P)\mathbb{1}_{[-A-1, A+1]}\|_\infty \leq \varepsilon \quad (38)$$

ce qui donne :

$$\|\chi f - \chi P\|_\infty \leq \varepsilon \quad (39)$$

Finalement, on a bien :

$$\|f - \chi P\|_\infty \leq 2\varepsilon \quad (40)$$

avec $\chi P \in C_c^\infty$.

Références

- [1] Hervé QUEFFÉLEC et Claude ZUILY. *Analyse pour l'agrégation*. DUNOD.