Théorème central limite

Référence : [Ouv09] p.314-315.

Théorème 0.1 (central limite) *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$, indépendantes, de même loi et admettant un moment d'ordre 2. Alors la suite de terme général Y_n définie pour tout $n \in \mathbb{N}^*$ par :*

$$Y_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (X_j - \mathbb{E}(X_j))$$

converge en loi vers la loi gaussienne $\mathcal{N}_{\mathbb{R}^d}(0, C_{X_1})$ où C_{X_1} est la matrice de covariance des X_j.

Démonstration $\rightsquigarrow$ Comme les variables aléatoires X_j sont indépendantes et de même loi, alors la fonction caractéristique de Y_n est donnée par, pour tout $t \in \mathbb{R}^d$:

$$\begin{aligned}
\phi_{Y_n}(t) &= \mathbb{E}\left(e^{i<t,Y_n>}\right) \\
&= \mathbb{E}\left(e^{i<t,\frac{1}{\sqrt{n}}\sum_{j=1}^n (X_j-\mathbb{E}(X_j))>}\right) \quad \text{par définition de } Y_n\,; \\
&= \prod_{j=1}^{n} \mathbb{E}\left(e^{i<\frac{t}{\sqrt{n}},(X_j-\mathbb{E}(X_j))>}\right) \quad \text{par indépendance des } X_j\,; \\
&= \prod_{j=1}^{n} \phi_{X_j-\mathbb{E}(X_j)}\left(\frac{t}{\sqrt{n}}\right) \quad \text{par définition d'une fonction caractéristique}\,; \\
&= \left(\phi_{<X_1-\mathbb{E}(X_1),t>}\left(\frac{1}{\sqrt{n}}\right)\right)^n \quad \text{car les } X_j \text{ sont de même loi.}
\end{aligned}$$

$\rightsquigarrow$ Comme la variable aléatoire $< X_1 - \mathbb{E}(X_1), t >$ est une variable aléatoire réelle centrée admettant un moment d'ordre 2, alors sa fonction caractéristique admet un développement limité à l'ordre 2 en zéro, ce qui nous donne :

$$\phi_{Y_n}(t) = \left(1 - \frac{1}{2n}\mathbb{E}\left(< X_1 - \mathbb{E}(X_1), t >^2\right) + o\left(\frac{1}{n}\right)\right)^n$$

$\rightsquigarrow$ De plus, d'après un lemme qui suit, on sait que la suite de terme général $\phi_{Y_n}(t)$ converge et que :

$$\lim_n \phi_{Y_n}(t) = \exp\left(-\frac{1}{2}\mathbb{E}\left(< X_1 - \mathbb{E}(X_1), t >^2\right)\right)$$

$\rightsquigarrow$ Or puisque $\mathbb{E}(< X_1 - \mathbb{E}(X_1), t >^2) =< C_{X_1}t, t >$, alors :

$$\lim_n \phi_{Y_n}(t) = \exp\left(-\frac{1}{2} < C_{X_1}t, t >\right)$$

$\rightsquigarrow$ Ainsi d'après le théorème de Levy, on sait que Y_n converge en loi vers $\mathcal{N}_{\mathbb{R}^d}(0, C_{X_1})$.

Trucs utilisés

Proposition 0.2 (sur les fonctions caractéristiques) *Soit X une variable aléatoire réelle et soit ϕ_X sa fonction caractéristique. On suppose que X admet un moment d'ordre $n \in \mathbb{N}^*$, alors ϕ_X est de classe C^n et pour tout k tel que $1 \leq k \leq n$, on a pour tout $t \in \mathbb{R}$:*

$$\phi_X^{(k)}(t) = i^k \int_\Omega X^k e^{itX} d\mathbb{P}$$

et en particulier $\phi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$.

Démonstration On a :

- pour tout $t \in \mathbb{R}$, la fonction $\omega \longmapsto e^{itX}$ est intégrable ;
- pour tout $\omega \in \Omega$, la fonction $t \longmapsto e^{itX}$ est de classe C^n et

$$\frac{d^k}{dt^k} e^{itX} = (iX)^k e^{itX}$$

- et :

$$\left| \frac{d^k}{dt^k} e^{itX} \right| \leq |X|^k$$

donc en appliquant le théorème de dérivabilité sous le signe intégral, on obtient le résultat voulu.

Lemme 0.1 *Pour tout $z \in \mathbb{C}$ et pour tout $n \in \mathbb{N}^*$, on a :*

$$\left| \exp(z) - \left(1 + \frac{z}{n}\right)^n \right| \leq \exp(|z|) - \left(1 + \frac{|z|}{n}\right)^n$$

Il en résulte que pour tout $z \in \mathbb{C}$, la suite de terme général $\left(1 + \frac{z}{n}\right)^n$ est convergente et que :

$$\lim_n \left(1 + \frac{z}{n}\right)^n = \exp(z)$$

Démonstration La formule du binôme et le développement de l'exponentielle donnent, pour tout $z \in \mathbb{C}$:

$$\exp(z) - \left(1 + \frac{z}{n}\right)^n = \sum_{j=0}^{+\infty} \frac{z^j}{j!} - \sum_{j=0}^{n} C_n^j \frac{z^j}{n^j}$$

Or (on obtient cette relation en développant son second membre) :

$$\frac{C_n^j}{n^j} = \frac{1}{j!} \prod_{k=0}^{j-1} \left(1 - \frac{k}{n}\right)$$

Donc en remplaçant cela dans l'égalité précédente :

$$\exp(z) - \left(1 + \frac{z}{n}\right)^n = \sum_{j=0}^{+\infty} \frac{z^j}{j!} - \sum_{j=0}^{n} \frac{z^j}{j!} \prod_{k=0}^{j-1} \left(1 - \frac{k}{n}\right) \tag{1}$$

$$= \sum_{j=n+1}^{+\infty} \frac{z^j}{j!} + \sum_{j=0}^{n} \frac{z^j}{j!} \left(1 - \prod_{k=0}^{j-1} \left(1 - \frac{k}{n}\right)\right) \tag{2}$$

Puisque $1 - \prod_{k=0}^{j-1} \left(1 - \frac{k}{n}\right) \geq 0$, il en résulte que :

$$\left| \exp(z) - \left(1 + \frac{z}{n}\right)^n \right| \leq \sum_{j=n+1}^{+\infty} \frac{|z|^j}{j!} + \sum_{j=0}^{n} \frac{|z|^j}{j!} \left(1 - \prod_{k=0}^{j-1} \left(1 - \frac{k}{n}\right)\right)$$

ce qui donne l'inégalité voulue en réutilisant la formule (2) pour $|z|$.
Enfin puisque :

$$ln \left(1 + \frac{|z|}{n}\right)^n = |z| + o(1)$$

on a :

$$\lim_n \left(1 + \frac{|z|}{n}\right)^n = \exp(|z|)$$

ce qui entraîne d'après l'inégalité de l'énoncé que :

$$\lim_n \left(1 + \frac{z}{n}\right)^n = \exp(z)$$

Définition 0.1 (Convergence étroite et convergence faible) *Une suite $(\mu_n)_{n\in\mathbb{N}}$ de mesures bornées converge vers la mesure μ étroitement (respectivement faiblement) si :*

$$\lim_n \int f d\mu_n = \int f d\mu$$

pour tout $f \in C_b^0(\mathbb{R}^d)$ (respectivement $C_0^0(\mathbb{R}^d)$ l'ensemble des fonctions continues tendant vers 0 à l'infini).

Définition 0.2 (Transformée de Fourier d'une mesure et fonction caractéristique) *On appelle transformée de Fourier de la mesure bornée μ, l'application $\widehat{\mu}$ de $\mathbb{R}^d$ dans $\mathbb{C}$ définie par, pour tout $t \in \mathbb{R}^d$:*

$$\widehat{\mu}(t) = \int_{\mathbb{R}^d} \exp(i < x, t >) d\mu(x)$$

On appelle fonction caractéristique de la variable aléatoire X la transformée de Fourier de sa loi $\mathbb{P}_X$, elle est notée ϕ_X.

Remarque : On déduit du lemme de transfert que pour tout $t \in \mathbb{R}^d$:

$$\phi_X(t) = \mathbb{E}(\exp(i < X, t >)$$

Théorème 0.3 (de Levy (version avec les mesures)) *Soit pour tout $n \in \mathbb{N}$, une mesure $\mu_n \in \mathcal{M}(b)$ qui est le sous-ensemble des mesures μ de masse inférieure ou égale à b (ie $\mu(\mathbb{R}^d) \leq b$). Alors :*

1. *Si la suite $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement vers μ, alors la suite $(\widehat{\mu_n})_{n\in\mathbb{N}}$ des transformées de Fourier de μ_n converge simplement vers $\widehat{\mu}$, transformée de Fourier de μ.*
2. *Inversement, si la suite $(\widehat{\mu_n})_{n\in\mathbb{N}}$ des transformées de Fourier des μ_n converge simplement vers une fonction ϕ continue en 0, alors il existe une unique mesure $\mu \in \mathcal{M}(b)$ telle que $\phi = \widehat{\mu}$; de plus, la suite $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement vers μ.*

Démonstration Observons tout d'abord que si la suite $(\mu_n)_{n\in\mathbb{N}}$ de mesures bornées sur $\mathbb{R}^d$ converge étroitement vers une mesure μ, on a :

$$\lim_n \int f d\mu_n = \int f d\mu$$

pour toute fonction $f \in C_b^0(\mathbb{R}^d)$ et à valeurs complexes : il suffit de remarquer que la convergence a lieu pour $Re(f)$ et $\Im(f)$.
$\rightsquigarrow$ (1) Pour tout $t \in \mathbb{R}^d$, la fonction $\exp i < ., t >$ est continue bornée, donc la suite de terme général $\widehat{\mu_n}(t)$ converge vers $\widehat{\mu}(t)$ (par définition de la transformation de Fourier).
$\rightsquigarrow$ (2) Montrons d'abord que la suite $(\mu_n)_{n\in\mathbb{N}}$ est faiblement convergente. Puisque $\mathcal{M}(b)$ est métrisable et compact pour la topologie faible, pour qu'il en soit ainsi, il faut et il suffit que cette suite admette au plus une valeur d'adhérence faible (car on sait que dans un espace métrique compact, toute suite possède au moins un point adhérent et qu'une suite qui n'admet qu'un seul point adhérent converge vers ce point).
Soit donc μ une valeur d'adhérence faible de la suite $(\mu_n)_{n\in\mathbb{N}}$ et soit $(\mu_{\psi(n)})_{n\in\mathbb{N}}$ une sous-suite qui converge faiblement vers μ.
On va montrer que $\mu_{\psi(n)}$ tend étroitement vers μ quand n tend vers l'infini, ce qui d'après le (1) assurera que la suite de terme général $\widehat{\mu_{\psi(n)}}$ converge simplement vers $\widehat{\mu}$. Puisque par hypothèse $\widehat{\mu_n}$ tend simplement vers ϕ quand n tend vers l'infini, il en est de même pour toute sous-suite et on aura donc $\phi = \widehat{\mu}$. L'unicité d'une valeur d'adhérence faible μ résultera alors de l'injectivité de la transformation de Fourier et on aura ainsi démontré la convergence faible de la suite $(\mu_n)_{n\in\mathbb{N}}$ vers μ.
Il faut donc montrer que la suite $(\mu_{\psi(n)})_{n\in\mathbb{N}}$ converge étroitement vers μ, et pour cela il suffit, puisqu'on a déjà la convergence faible, de montrer que :

$$\lim_n \mu_{\psi(n)}(\mathbb{R}^d) = \mu(\mathbb{R}^d)$$

Mais on sait par hypothèse que :

$$\lim_n \mu_{\psi(n)}(\mathbb{R}^d) = \lim_n \mu_n(\mathbb{R}^d) = \lim_n \widehat{\mu_n}(0) = \phi(0)$$

Puisque $\mu(\mathbb{R}^d) = \widehat{\mu}(0)$, il suffit donc de montrer que $\widehat{\mu}(0) = \phi(0)$.
Pour cela, on observe d'abord que pour $\epsilon > 0$, on a :

$$\lim_n \int_{[0,\epsilon]^d} \widehat{\mu_n}(t)dt = \int_{[0,\epsilon]^d} \phi(t)dt \tag{3}$$

En effet puisque $\widehat{\mu_n}$ tend simplement vers ϕ quand n tend vers l'infini et que les fonctions $\widehat{\mu_n}$ sont bornées en module par b, cela résulte du théorème de convergence dominée.
Utilisons maintenant le lemme qui suit :

Lemme 0.2 *Soit $\epsilon > 0$. Alors il existe une fonction $f_\epsilon \in C_0^0(\mathbb{R}^d)$ (les fonctions continues et tendant vers 0 à l'infini) telle que pour toute mesure bornée ν sur $\mathbb{R}^d$, on ait :*

$$\int_{[0,\epsilon]^d} \widehat{\nu}(t)dt = \int_{\mathbb{R}^d} f_\epsilon d\nu$$

Démonstration du lemme On a d'après le théorème de Fubini :

$$\int_{[0,\epsilon]^d} \widehat{\nu}(t)dt = \int_{[0,\epsilon]^d} \left(\int_{\mathbb{R}^d} e^{i<x,t>} d\nu(x) \right) dt = \int_{\mathbb{R}^d} \left(\int_{[0,\epsilon]^d} e^{i<x,t>} dt \right) d\nu(x)$$

Or toujours d'après le théorème de Fubini, on a :

$$\int_{[0,\epsilon]^d} e^{i<x,t>} dt = \prod_{j=1}^{d} \left(\int_0^\epsilon e^{ix_j t_j} dt_j \right)$$

On obtient donc le résultat en posant pour $u \in \mathbb{R}$:

$$g_\epsilon(u) = \int_0^\epsilon e^{iut} dt = \begin{cases} \dfrac{e^{i\epsilon u} - 1}{iu} & \text{si } u \neq 0\,; \\ \epsilon \quad \text{si } u = 0. \end{cases}$$

et pour $x \in \mathbb{R}^d$:

$$f_\epsilon(x) = \prod_{j=1}^{d} g_\epsilon(x_j)$$

On a alors que $f_\epsilon \in C_0^0(\mathbb{R}^d)$ (d'après le théorème de Riemann-Lebesgue).

Suite de la démonstration du théorème de Levy Puisque la suite de terme général $\mu_{\psi(n)}$ converge faiblement vers μ et que $f_\epsilon \in C_0^0(\mathbb{R}^d)$, on a :

$$\lim_n \int f_\epsilon d\mu_{\psi(n)} = \int f_\epsilon d\mu$$

D'où d'après le lemme précédent :

$$\lim_n \int_{[0,\epsilon]^d} \widehat{\mu_{\psi(n)}}(t)dt = \int_{[0,\epsilon]^d} \widehat{\mu}(t)dt$$

D'après l'égalité (3) appliqué à la sous-suite $(\mu_{\psi(n)})_{n\in\mathbb{N}}$, on a :

$$\frac{1}{\epsilon^d} \int_{[0,\epsilon]^d} \widehat{\mu}(t)dt = \frac{1}{\epsilon^d} \int_{[0,\epsilon]^d} \phi(t)dt$$

Grâce à la continuité de $\widehat{\mu}$ et de ϕ en zéro, on obtient en prenant la limite quand ϵ tend vers 0 dans les deux membres de l'égalité :

$$\widehat{\mu}(0) = \phi(0)$$

On a donc montré la convergence faible de la suite $(\mu_n)_{n\in\mathbb{N}}$.
Enfin il résulte de la dernière égalité et de la convergence simple de la suite $\widehat{\mu_n}$ vers ϕ que :

$$\lim_n \widehat{\mu_n}(0) = \widehat{\mu}(0)$$

ou autrement dit :

$$\lim_n \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$$

Ce qui achève la démonstration de la convergence étroite de la suite $(\mu_n)_{n\in\mathbb{N}}$ vers sa limite faible μ.

Théorème 0.4 (de Levy (version convergence en loi)) *Soit pour tout $n \in \mathbb{N}$, une variable aléatoire X_n définie sur un espace probabilisé $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ à valeurs dans $\mathbb{R}^d$, de fonction caractéristique ϕ_{X_n}. Alors :*

1. *Si la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X, où X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$, alors la suite $(\phi_{X_n})_{n\in\mathbb{N}}$ des fonctions caractéristiques converge simplement vers la fonction caractéristique ϕ_X de X ;*
2. *Inversement, si la suite $(\phi_{X_n})_{n\in\mathbb{N}}$ des fonctions caractéristiques converge simplement vers une fonction ϕ continue en 0, alors ϕ est la transformée de Fourier d'une probabilité μ sur $\mathbb{R}^d$, et la suite des variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi vers μ.*

De plus, il existe une variable aléatoire (non unique) X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$, telle que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.

Démonstration Ce n'est qu'une reformulation du théorème de Levy pour la convergence étroite des mesures bornées, une fois rappelé que ϕ_{X_n}est, par définition, la transformée de Fourier de la loi de X_n. Seul le dernier point de la réciproque nécessite un éclaircissement : d'après le théorème de Levy, la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers la probabilité μ telle que $\widehat{\mu} = \phi$ (μ est bien une probabilité, puisque $\lim_n \phi_{X_n}(0) = 1 = \phi(0) = \widehat{\mu}(0)$) ; on considère alors l'application identique X de $\mathbb{R}^d$ sur lui-même ; c'est une variable aléatoire définie sur l'espace probabilisé $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$ à valeurs dans $\mathbb{R}^d$ de loi μ et telle que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.

Références

[Ouv09] Jean-Yves Ouvrard. *Probabilités 2.* Cassini, 2009.