ctre-exemple à la loi des grands nombres

variable symétrique : X tq X a même loi que $-X$

si $\mathbb{E}(X)$ existe, alors $\mathbb{E}(X) = \mathbb{E}(-X) = -\mathbb{E}(X) = 0$

Mais il est possible que X n'existe pas et dans ce cas, on peut avoir un contre-exemple à la loi des grands nombres

. X à valeurs dans $\mathbb{Z}$

$$\mathbb{P}(X = k) = \mathbb{P}(X = -k) = \frac{c}{k^\alpha} \qquad 1 < \alpha < 2$$

On peut montrer que si $X_1 \dots X_n$ sont indép, de même loi que X, alors $\boxed{\frac{X_1 + \dots + X_n}{n^{\alpha-1}} \xrightarrow{(l)} Y}$ où Y est une variable aléatoire non triviale. On dit que Y est stable

$\Rightarrow \frac{X_1 + \dots + X_n}{n}$ ne converge pas vers δ_0

(Feller. An Introduction to Proba theory)

Prop 10.3.2 : régularité des mesures boréliennes

$X_n \xrightarrow{(l)} X \iff \forall O$ ouvert, $\liminf \mathbb{P}(X_n \in O) \geq \mathbb{P}(X \in O)$

$\iff \forall F$ fermé, $\limsup \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F)$

$\iff \forall B$ borélien tq $\mathbb{P}(X \in \partial B) = 0$,

$\mathbb{P}(X_n \in B) \xrightarrow[n \to \infty]{} \mathbb{P}(X \in B)$

En dimension finie, la notion de convergence est assez simple

En dimension infinie, c'est plus compliqué

En analyse fonctionnelle, on définit souvent la convergence en prenant des fonctions tests

Pour la convergence en loi, on fait la même chose, on prend comme fonctions tests les fonctions continues bornées.

<u>Mais</u> les fonctions indicatrices ne sont pas continues (sauf dans le cas discret). Donc $\mathbb{P}(X_n \in A)$ ne converge pas toujours vers $\mathbb{P}(X \in A)$ même si $X_n \xrightarrow{(\ell)} X$

. Si X est à valeurs dans $\mathbb{R}^d$ et a une densité f, alors $\forall B$,

$\mathbb{P}(X \in \partial B) = \int_{\partial B} f(x)dx = 0 \to$ Si $X_n \xrightarrow{(\ell)} X$, alors $\forall B$

$$\mathbb{P}(X_n \in B) \underset{n \to \infty}{\longrightarrow} \mathbb{P}(X \in B)$$

. Si $X_n \xrightarrow{(\ell)} X$ et si on appelle μ_n loi de X_n

μ —— X

on dit que (μ_n) converge étroitement vers μ

[en anglais : "weak convergence"]

Th de Lévy :

<u>déf</u> : si X est à valeurs dans $\mathbb{R}^d$, sa fonction caractéristique est la fonction : $\mathbb{R}^d \xrightarrow{\Phi_X} \mathbb{C}$

$$\xi \mapsto \mathbb{E}[\exp(i\langle \xi, X\rangle)]$$

[transformée de Fourier]

<u>NB</u> Φ_X est bornée & continue $\Rightarrow$ Si $X_n \xrightarrow{(\ell)} X$ alors $\forall \xi \in \mathbb{R}^d$,

$$\Phi_{X_n}(\xi) \underset{n\to\infty}{\longrightarrow} \Phi_X(\xi)$$

$\mathbb{E}(g_\xi(X))$ où $g_\xi(y) = \exp(i\langle \xi, y\rangle)$

$\forall \xi \in \mathbb{R}^d$, g_ξ est une fonction continue bornée $\Rightarrow \mathbb{E}(g_\xi(X_n)) \to \mathbb{E}(g_\xi(X))$

$\mathbb{E}(g_\xi(X_n)) = \Phi_{X_n}(\xi)$, $\quad \mathbb{E}(g_\xi(X)) = \Phi_X(\xi)$

Réciproque :

<u>Th de Lévy</u> (X_n) v.a. à valeurs dans $\mathbb{R}^d$

$$X_n \xrightarrow{(\ell)} X \iff \forall \xi \in \mathbb{R}^d,\ \Phi_{X_n}(\xi) \to \Phi_X(\xi)$$

$$\Rightarrow$$

<u>Th</u> (TCL) X_n variables aléatoires iid (indépendantes, identiquement distribuées) telles que $\mathbb{E}(X_n) = m$

$$\operatorname{var}(X_n) = \sigma^2$$

$$\sqrt{n}\left(\frac{X_1 + X_2 \cdots + X_n}{n} - m\right) \xrightarrow[n\to\infty]{(\ell)} \mathcal{N}(0, \sigma^2)$$

Rappel $X \sim \mathcal{N}(\mu, \sigma^2)$ si X a une densité sur $\mathbb{R}$

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right)$$

En particulier $\forall a, b \in \mathbb{R}$ $\mathbb{P}\left(\sqrt{n}\left(\frac{X_1 + \dots + X_n}{n} - m\right) \in [a,b]\right) \to \mathbb{P}(X \in [a,b])$

$X \sim \mathcal{N}(0, \sigma^2)$

dém $Y_n = \frac{X_1 + X_2 \dots + X_n}{n} - m = \frac{(X_1 - m) + (X_2 - m) \dots + (X_n - m)}{n}$

$$\Phi_{Y_n}(\xi) = \mathbb{E}\left[\exp\left(i\xi\left[\frac{(X_1-m)}{n} + \left(\frac{X_2-m}{n}\right) \dots + \left(\frac{X_n-m}{n}\right)\right]\right)\right]$$

$$\overset{\text{indép}}{=} \mathbb{E}\left(\exp\left(i\xi\left(\frac{X_1-m}{n}\right)\right)\right) \cdot \mathbb{E}\left(\exp i\xi\left(\frac{X_2-m}{n}\right)\right) \dots$$

$$= \mathbb{E}\left(\exp i\xi\left(\frac{X_1-m}{n}\right)\right)^n$$

$$\Phi_{(X_1-m)}(\xi) = 1 + i\xi \underbrace{\mathbb{E}(X_1-m)}_{0} - \frac{1}{2}\xi^2 \operatorname{var}(X_1-m) + o(\xi^2)$$

$$= 1 - \frac{1}{2}\xi^2 \operatorname{var}(X_1) + o(\xi^2)$$

$$\log \Phi_{\sqrt{n}Y_n}(\xi) = n \log\left(\mathbb{E} \exp\left[i\frac{\xi}{\sqrt{n}}(X_1-m)\right]\right) = n \log\left(1 - \frac{1}{2n}\xi^2 \operatorname{var} X + o\left(\frac{\xi}{n}\right)\right)$$

$$= n\left(-\frac{1}{2}\frac{\xi^2}{n}\sigma^2 + o\left(\frac{\xi}{n}\right)\right) = -\frac{1}{2}\xi^2\sigma^2 + o(1)$$

$$\Phi_{\sqrt{n}Y_n}(\xi) \xrightarrow[n\to\infty]{} \exp\left(-\frac{1}{2}\sigma^2\xi^2\right) = \mathbb{E}(\exp i\xi X) = \Phi_X(\xi)$$

où $X \sim \mathcal{N}(0, \sigma^2)$

donc $\sqrt{n}\, Y_n \xrightarrow{(P)} X$

or $\sqrt{n}\, Y_n = \underbrace{\sqrt{n}} \left(\frac{X_1 + X_2 \cdots + X_n}{n} - m \right) \xrightarrow{(\ell)} X$

Remarque : s'il n'existe pas de variance, on peut trouver des contre-exemples [lois stables]

On peut relâcher l'hypothèse d'indépendance, $\mathrm{cov}(X_n, X_{n+k}) \leq \exp(-ck)$ $c > 0$

Informellement, quand on a beaucoup de petites sources d'aléa, et indépendance on a à la limite une loi gaussienne $\to$ les lois gaussiennes sont très utilisées pour modéliser de nombreux phénomènes en science

- $Y_n \xrightarrow{(P)} Y \Rightarrow Y_n \xrightarrow{(\ell)} Y$ aléa : hasard

$$\not\Leftarrow$$

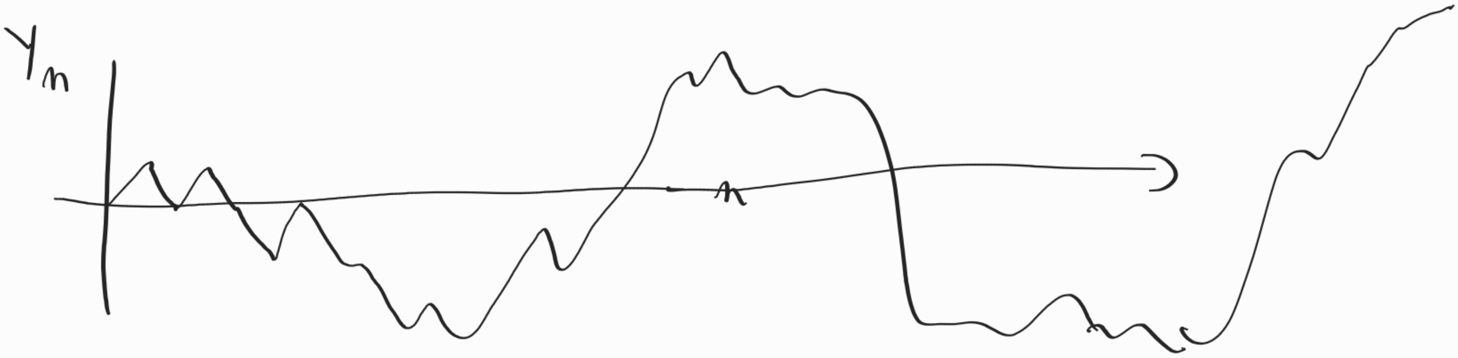
(X_n) indép de même loi, $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$

$$Y_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}} \qquad \begin{cases} m = 0 \\ \sigma^2 = \frac{1}{4} \end{cases}$$

TCL : $Y_n \xrightarrow{\ell} \mathcal{N}(0, \frac{1}{4})$

Mais il n'existe pas de variable aléatoire Y telle que $Y_n \xrightarrow{(P)} Y$

Loi du (0,1) : p.s. $\limsup \sqrt{n}\, Y_n = +\infty$ et $\liminf \sqrt{n}\, Y_n = -\infty$



par l'absurde, s'il existait Y telle que $Y_n \xrightarrow{P} Y$

on devrait avoir $\mathbb{P}(|Y_n - Y| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0$ pour tout $\varepsilon > 0$

(suite demain)

Convergence des mesures empiriques

X variable aléatoire dans $\mathbb{R}^d$

X : opinion d'un électeur tiré au hasard
vitesse d'un fluide dans un écoulement turbulent
température en un point

échantillonnage ("sampling" en anglais) : $X_1, X_2 \quad X_n$

variables aléatoires iid de même loi que X.

mesure empirique : mesure de probabilité sur $\mathbb{R}^d$ qui vaut

$$\frac{1}{n}\left(\delta_{X_1} + \delta_{X_2} \cdots + \delta_{X_n}\right) = \mu_n$$

<u>Th</u> avec probabilité 1, (μ_n) converge étroitement vers la loi de X

De manière équivalente soit Y_n définie par : U_n loi uniforme sur $\{1, 2 \ldots n\}$ et $Y_n = \delta_{X_{U_n}}$ (donc la loi de Y_n est μ_n). Alors $Y_n \xrightarrow{(\ell)} X$ presque sûrement

<u>dém</u> Soit H un sous-ensemble dénombrable dense de $\underbrace{\mathcal{C}_c(\mathbb{R}^1)}$ ↖ fonctions à support compact

soit $\varphi \in H$, on considère les v.a $\varphi(X_i)$: indép, de m loi, bornées → intégrables

Loi des grands nombres : <u>avec proba 1</u>, $\underbrace{\frac{\varphi(X_1) + \ldots + \varphi(X_n)}{n}} \to \mathbb{E}(\varphi(X_1))$

$$\int \varphi(x)\, d\mu_n(x) \xrightarrow[n \to \infty]{} \int \varphi(x)\, dP_X(x)$$

il existe une ensemble Ω_φ négligeable : $\mathbb{P}(\Omega_\varphi) = 0$ tel que sur Ω_φ^c, $\int \varphi(x)\, d\mu_n(x) \xrightarrow[n \to \infty]{} \int \varphi(x)\, dP_X(x)$

soit $\breve{\Omega} = \bigcup_{\varphi \in H} \Omega_\varphi$, H est dénombrable,

$\mathbb{P}(\breve{\Omega}) = 0$

sur $\tilde{\Omega}^c$, $\forall \varphi \in H$, $\int \varphi(x)\, d\mu_n(x) \xrightarrow[n\to\infty]{} \int \varphi(x)\, dP_X(x)$

→ convergence avec toutes les fonctions $\varphi \in H$, et $\overline{H} = \mathcal{C}_c(\mathbb{R}^d)$

d'après Prop 10.3.3, on a cv pour toutes les fonctions continues bornées. donc (μ_n) cv étroitement vers $\mathbb{P}_X$

Vecteurs gaussiens :

covariance : si X, Y sont des v.a. dans $\mathbb{R}$,

$cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

en particulier, si $X = Y$, $cov(X,X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$
$= var(X)$

Si $(X_1, \dots X_d)$ sont des v.a dans $\mathbb{R}$,

on dit que la matrice de covariance de $(X_1, \dots X_d)$ est la matrice $(M_{ij})_{\substack{1\le i\le d\\ 1\le j\le d}}$, $M_{ij} = cov(X_i, X_j)$

remarque : quand les X_i sont dans L^2, la matrice de cov existe car par Cauchy-Schwarz, $|\mathbb{E}(X_i X_j)| \leq \sqrt{\mathbb{E}(X_i^2)\mathbb{E}(X_j^2)}$

Prop si la matrice de cov existe, c'est une matrice symé-trique positive

Prop si M est une matrice carrée $d \times d$ symétrique positive il existe un vecteur $(X_1, \dots, X_d)$ aléatoire tel que

$$\boxed{\Phi_{(X_1 \dots X_d)}(\xi) = \mathbb{E}(\exp(i\langle \xi, X\rangle)) = \exp\left(-\frac{1}{2}\, {}^t\xi M \xi\right)}$$

Le vecteur $(X_1 \dots X_d)$ est appelé vecteur gaussien de matrice de covariance M. notation : $(X_1 \dots X_d) \sim \mathcal{N}(0, M)$

Th (TCL vectoriel) : soit (X_n) des variables aléatoires i.i.d à valeurs dans $\mathbb{R}^d$ de matrice de covariance M

alors $$\sqrt{n}\left(\frac{X_1 + \dots + X_n}{n} - \mathbb{E}(X_1)\right) \xrightarrow{(L)} \mathcal{N}(0, M)$$