Probabilités L3S6, 2021

Université de Picardie Jules Verne

Probabilités (XI)

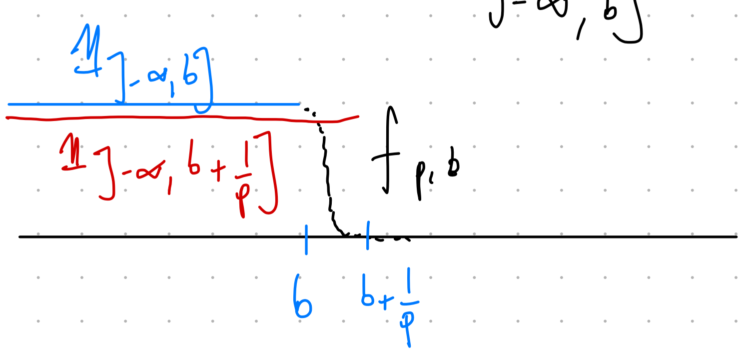
Prop. : Soient $(X_n)_n$ et X des variables aléatoires réelles de fonctions de répartition respectives F_n et F. Pour que $X_n \xrightarrow{\mathcal{L}} X$, il faut et il suffit que

$$F_n(x) \underset{n \to +\infty}{\longrightarrow} F(x) \text{ pour tout } x \text{ où } F \text{ est continue.}$$

dém. : a) Supposons $X_n \xrightarrow{\mathcal{L}} X$.

Soit a où F est continue, i.e., $F(a^-) = F(a)$.
Pour tout entier $p \geq 1$, et tout $b \in \mathbb{R}$, il existe une fonction $f_{p,b}$ continue bornée sur $\mathbb{R}$ t.q.

$$\mathbb{1}_{]-\infty, b]} \leq f_{p,b} \leq \mathbb{1}_{]-\infty, b+\frac{1}{p}]} \qquad (*)$$



Alors $\mathbb{E}(f_{p,b}(X_n))$ CV vers $\mathbb{E}(f_{p,b}(X))$ quand $n \to +\infty$.

On déduit de $(*)$ que $F_n(a) = \mathbb{P}(X_n \leq a) = \mathbb{E}(\mathbb{1}_{]-\infty,a]}(X_n)) \leq \mathbb{E}(f_{p,a}(X_n))$

et $\mathbb{E}(f_{p,a}(X)) \leq F(a+\frac{1}{p})$, donc $F_n(a) \leq F(a+\frac{1}{p})$ pour tout p, donc

ainsi $\limsup_n F_n(a) \leq F(a+\frac{1}{p})$ pour tout p.

En passant à la limite $p \to +\infty$, $F(a+\frac{1}{p}) \xrightarrow[p\to+\infty]{} F(a)$ car F est continue en a,

et alors $\limsup_n F_n(a) \leq F(a)$.

Nous avons également que $F_n(a) \geq \mathbb{E}(f_{p,a-\frac{1}{p}}(X_n))$

et $\mathbb{E}(f_{p,a-\frac{1}{p}}(X)) \geq F(a-\frac{1}{p})$ (car $F(a^-) = F(a)$)

donc $\liminf_n F_n(a) \geq F(a-\frac{1}{p})$ pour tout p, donc ainsi $\liminf_n F_n(a) \geq F(a)$.

Des deux inégalités en rouge on déduit que

$$\liminf_n F_n(a) = \limsup_n F_n(a) = \lim_n F_n(a) = F(a).$$

b) Supposons que $F_n(x) \longrightarrow F(x)$ pour tout $x \in T$ où T est une partie dense de $\mathbb{R}$ (rappel : F est continue sur $\mathbb{R} \setminus D$ où D est une partie dénombrable (au plus) donc on peut choisir $T \subset (\mathbb{R} \setminus D)$ dense)

Soit f une fonction continue bornée sur $\mathbb{R}$ et $\varepsilon > 0$.

Soit $a, b \in T$ avec $F(a) \leq \varepsilon$ et $F(b) \geq 1 - \varepsilon$. $(*)$

Il existe n_0 t.q. $n \geq n_0 \Rightarrow P\left(X_n \notin]a, b]\right) = 1 - F_n(b) + F_n(a)$

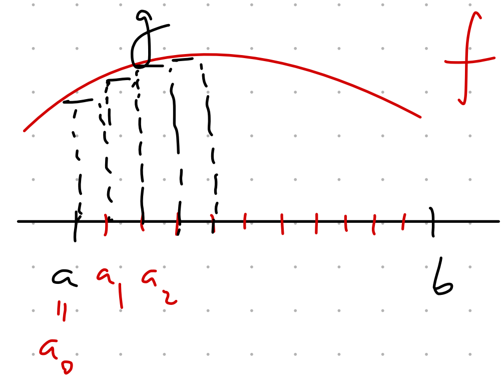
$\leq 3\varepsilon$ (d'après $(*)$)

La fonction f est continue donc uniformément continue sur $[a, b]$,

donc il existe un nombre fini de points $a_0 = a < a_1 < \dots < a_k = b$ dans T

t.q. $|f(x) - f(a_i)| \leq \varepsilon$ si $x \in [a_{i-1}, a_i]$.

Donc $$g(x) = \sum_{i=1}^{k} f(a_i)\, \mathbb{1}_{]a_{i-1}, a_i]}(x)$$



vérifie $|f-g| \le \varepsilon$ sur $]a,b]$.

Si $M = \sup_x |f(x)|$, on a alors

$$\left| \mathbb{E}(f(X_n)) - \mathbb{E}(g(X_n)) \right| \le M\, \mathbb{P}\left(X_n \notin]a,b]\right) + \varepsilon \qquad (\cdot)$$

et de même pour X.

Enfin, $$\mathbb{E}(g(X_n)) = \sum_{i=1}^{k} f(a_i) \Big(\underbrace{F_n(a_{i-1}) - F_n(a_i)}_{\mathbb{P}\left(X_n \in]a_{i-1}, a_i]\right)} \Big)$$ et de même pour X.

Comme $F_n(a_i) \xrightarrow[n \to +\infty]{} F(a_i)$ $\quad (a_i \in T\ \forall i)$

on déduit l'existence de n_1 t.q. $n \geq n_1 \Rightarrow \left| \mathbb{E}(g(X_n)) - \mathbb{E}(g(X)) \right| \le \varepsilon.$

$(\cdot\cdot)$

Par $(\cdot)$ et $(\cdot\cdot)$ on déduit que

$$n \geq \sup(n_0, n_1) \implies \left| \mathbb{E}(f(X_n)) - \mathbb{E}(f(X)) \right| \leq 3\varepsilon + 5M\varepsilon.$$

Mais $\varepsilon > 0$ est arbitraire donc $\mathbb{E}(f(X_n)) \xrightarrow[n\to+\infty]{} \mathbb{E}(f(X))$. ∎

Corollaire : Si $(X_n)_n$ suite de r.a.r. t.q. $X_n \xrightarrow{\mathcal{L}} X$ à densité,

alors pour tout $a < b$,

$$P(X_n \in]a,b]) \xrightarrow[n\to+\infty]{} P(X \in]a,b]).$$

dém. : X est à densité $\Rightarrow$ la fonction de répartition F de X est continue en tout point.

$\forall n$, $P(X_n \in]a,b]) = F_n(b) - F_n(a)$ (où F_n est la fct de répart. de X_n)

$$\xrightarrow[n\to+\infty]{} F(b) - F(a) = P(X \in]a,b]).$$ ∎

d'après le rés. précédent

(PL)

Prop. : Soient X_n et X des vecteurs aléatoires.

Pour que $X_n \xrightarrow{\mathcal{L}} X$ il faut et il suffit que $\mathbb{E}(f(X_n)) \longrightarrow \mathbb{E}(f(X))$

pour toute fonction f lipschitzienne bornée.

dém. : Il suffit de m.q. dans la preuve de la proposition précédente

on peut remplacer les fonctions continues $f_{p,L}$ par des fonctions lipschitziennes bornées. ∎

Théorème : (Lemme de Slutsky)

Soient $(X_n)_n$ et (Y_n) deux suites de vecteurs aléatoires.

Supposons que $\begin{cases} (X_n) \text{ CV en loi vers } X \\ X_n - Y_n \text{ CV en proba. vers } 0 \end{cases}$

Alors (Y_n) CV en loi vers X.

dém. : grâce à la proposition (PL) il suffit de m.q. $\mathbb{E}(f(Y_n)) \xrightarrow[n\to+\infty]{} \mathbb{E}(f(X))$ $\forall f$ bornée lipschitzienne.

Nous avons alors que $|f(x)| \leq k$, $\forall x$ pour des constantes $k, C > 0$.

et $|f(x) - f(y)| \leq C|x-y|$

Soit $\varepsilon > 0$. Nous écrivons

$$\left|\mathbb{E}(f(Y_n)) - \mathbb{E}(f(X_n))\right| \leq \mathbb{E}\left(\left|f(Y_n) - \mathbb{E}(f(X_n))\right|\right)$$

$$\leq \underbrace{C\varepsilon}_{\text{sur } \{|X_n - Y_n| \leq \varepsilon\}} + \underbrace{2k\, P(|X_n - Y_n| > \varepsilon)}_{\text{sur } \{|X_n - Y_n| > \varepsilon\}}$$

Quand $n \to +\infty$, $P(|X_n - Y_n| > \varepsilon) \longrightarrow 0$ (CV en proba.).

Comme $\varepsilon > 0$ est arbitrairement petit, on déduit que :

$$\left|\mathbb{E}(f(Y_n)) - \mathbb{E}(f(X_n))\right| \xrightarrow[n\to+\infty]{} 0.$$

■

Théorème : (Théorème de Lévy)

(X_n) des v.a. à valeurs dans $\mathbb{R}^d$.

a) Si $X_n \xrightarrow{\mathcal{L}} X$ alors $(\phi_{X_n})_n$ CV simplement vers ϕ_X.

b) Si $(\phi_{X_n})_n$ CV simplement vers une fonction (complexe) ϕ sur $\mathbb{R}^d$

et que ϕ est continue en 0, alors c'est la fonction caractéristique d'une v.a. X, et on a $X_n \xrightarrow{\mathcal{L}} X$.

[(b) est difficile]

Dém de (a) : $\phi_{X_n}(u) = \mathbb{E}(g_u(X_n))$ et $\phi_X(u) = \mathbb{E}(g_u(X))$

avec $g_u : x \longmapsto e^{i(u,x)}$ (continue bornée)

et on applique la définition de la CV en loi en séparant parties réelle et imaginaire.

Théorème de la limite centrale

Soit $(X_n)_n$ suite de v.a. indépendantes, de même loi, et dans $\mathcal{L}^2$.

Notons $m = E(X_n)$ et $\sigma^2 = V(X_n)$

l'espérance et la variance communes aux v.a. X_n, $n \geq 1$, et

$$S_n = X_1 + \dots + X_n \quad , \quad M_n = \frac{S_n}{n}.$$

On a vu que M_n CV vers m p.s. (et en moyenne).

On va s'intéresser à la vitesse de CV.

Pour cela, on va étudier la limite éventuelle de $n^\alpha \left(\frac{S_n}{n} - m\right)$ pour différents α.

En particulier, on espère que pour un α (et alors α est nécessairement unique),

$n^\alpha \left(\frac{S_n}{n} - m\right)$ CV vers une limite $\neq \infty, 0$.

La réponse est partiellement négative :

en fait $n^{\alpha}\left(\frac{S_n}{n} - m\right)$ ne CV pas au sens presque sûr (ni même en proba.) pour aucun α !

Par contre, pour la CV en loi on va voir que $\alpha = \frac{1}{2}$ convient (universel) :

Théorème : (Thm de la limite centrale)

Si les $(X_n)_n$ sont des v.a.r. indépendantes, de même loi, de carré intégrable, de moyenne m et variance $\sigma^2 > 0$, alors les variables

$$\frac{S_n - nm}{\sigma\sqrt{n}} \text{ CV en loi vers une v.a. de loi normale } \mathcal{N}(0,1).$$

(i.e. $\sqrt{n}\left(\frac{S_n}{n} - m\right)$ CV en loi vers une v.a. de loi $\mathcal{N}(0,\sigma^2)$)

Dém. : (preuve repose le thm de Lévy)

Soit ϕ la fonction caractéristique de $X_n - m$, et $Y_n = \dfrac{S_n - nm}{\sigma\sqrt{n}}$.

Comme les X_n sont indépendants, et d'après les résultats sur les fonctions carac.,

la fonction caractéristique de Y_n, notée ϕ_n, est

$$\phi_n : u \longmapsto \phi_n(u) = \phi\left(\frac{u}{\sigma\sqrt{n}}\right)^n.$$

Comme $\mathbb{E}(X_n - m) = 0$ et $\mathbb{E}\left((X_n - m)^2\right) = \sigma^2$, on a

$$(\bullet) \qquad \phi(u) = 1 - \frac{u^2\sigma^2}{2} + u^2 o(|u|) \quad \text{quand } u \to 0.$$

Comme $\phi(0) = 1$ et ϕ est continue en 0, pour u fixé et n assez grand, on a :

$$\left|\phi\left(\frac{u}{\sigma\sqrt{n}}\right) - 1\right| \leq \frac{1}{2}.$$

$$\left(\sim \left|\frac{u^2}{2n}\right|\right)$$

Il est possible de généraliser la notion de logarithme aux complexes $z \in \mathbb{C}$ t.q. $|z-1| \leq \frac{1}{2}$

et la fonction $\ln z$ définie sur le disque $\{ z \in \mathbb{C} : |z-1| \leq \frac{1}{2} \}$ a le même développement limité au voisinage de $z=1$ que le logarithme réel.

$\left(\ln(z) = z-1 + o(|z-1|) \right)$

On a alors : $\Phi_n(u) = \Phi\left(\frac{u}{\sigma\sqrt{n}}\right)^n = \exp\left(n \ln \underbrace{\Phi\left(\frac{u}{\sigma\sqrt{n}}\right)}_{1 - \frac{u^2}{2n} + \frac{1}{n}\varepsilon_n(u)}\right)$ d'après $(\bullet)$

(avec $\varepsilon_n(u) \to 0$ quand $n \to +\infty$)

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

$$\Rightarrow \quad \Phi_n(u) \xrightarrow[n \to +\infty]{} \exp\left(-\frac{u^2}{2}\right).$$

On reconnaît alors la fonction caractéristique de $\mathcal{N}(0,1)$.

Par le thm de Lévy on déduit que Y_n CV en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$. $\blacksquare$