262 — Modes de convergence d'une suite de variables aléatoires. Exemples et applications.

I- Définitions et caractérisation des modes de convergence

1) Définitions

Définition - Soit (X_n) une suite de variables aléatoires. On dit que (X_n) converge presque sûrement vers X si $\mathbb{P}(\{\omega\in\Omega \,/\, X_n(\omega) \xrightarrow[n\to+\infty]{} X(\omega)\}) = 1$. On notera $\boxed{X_n \xrightarrow{p.s} X}$. encadrer mots

Exemple 1 Soit $(\varepsilon_k) \sim \mathcal{B}(\frac{1}{2})$ et indépendantes alors $X_n = \sum_{k=0}^{n} \frac{\varepsilon_k}{2^k}$ converge presque sûrement vers $X \sim \mathcal{U}([0;1])$ + un autre ex

Définition - (X_n) converge en probabilité vers X si $\forall\varepsilon>0$, $\lim \mathbb{P}(|X_n - X|>\varepsilon) = 0$. On note $\boxed{X_n \xrightarrow{\mathbb{P}} X}$

Exemple 2 $X_n = \mathcal{B}(\frac{1}{n})$ alors $X_n \xrightarrow{\mathbb{P}} 0$ mais $X_n \not\xrightarrow{p.s} 0$ si indépendantes

Définition - Si on note F_n les fonctions de répartition d'une suite de va (X_n) et F celle de X. Alors on dira que X_n converge en loi vers X si $\forall t\in\mathbb{R}$ / F continue en t alors $F_n(t) \xrightarrow[n\to+\infty]{} F(t)$

Exemple 3 En considérant (X_n) indépendantes de loi $\mathcal{B}(\,)$. Alors $X_n \Rightarrow X \sim \mathcal{B}(p)$ cependant $\mathbb{P}(|X_n - X|>\frac{1}{2}) = \frac{1}{2}$ donc $X_n \not\xrightarrow{\mathbb{P}} X$

Définition - (X_n) converge vers X en norme L^p si $\mathbb{E}[|X_n - X|^p] \xrightarrow[n\to+\infty]{} 0$

Exemple 4 $X_n(\omega) = \cos(n\omega^2)$ définie sur $]0;1[=\Omega$ converge vers 0 en norme L^p pour tout $p\in[1;+\infty[$ cependant pour presque chaque ω $X_n(\omega) \not\rightarrow 0$.

Contre-exemple 5 $X_n = \frac{1}{n}\mathbb{1}_{[0;n]}$ converge presque sûrement vers 0 mais pas en norme L^1.

2) Critères de convergences

Borel-Cantelli presque sûr - (X_n) suite de v.a. réelles.

1) Si $\forall\varepsilon>0$ $\sum_{n=0}^{\infty} \mathbb{P}(|X_n - X|\geq\varepsilon) < +\infty$ alors $X_n \xrightarrow{p.s} X$

2) Si les X_n sont mutuellement indépendants alors $X_n \xrightarrow{p.s} 0$ est équivalent à $\forall\varepsilon>0$ on a $\sum_{n=0}^{+\infty} \mathbb{P}(|X_n|\geq\varepsilon) < +\infty$

Remarque - Cela explique l' exemple 2. => mettre ça ici et mettre ça en annexe

Critère de Cauchy presque-sûr - (X_n) converge presque sûrement ssi $\forall\varepsilon>0$, $\mathbb{P}(\bigcup_{m\geq1}\bigcap_{n\geq1} |X_n - X_m|<\varepsilon) = 1$. (Cu) / φ

Propriété 1 $d_p(X,Y) = \mathbb{E}[\min\{|X-Y|, 1\}]$ est une distance qui métrise la convergence en proba. Ainsi $X_n \xrightarrow{\mathbb{P}} X$ ssi $d_p(X_n, X) \longrightarrow 0$

Remarque 1) D'autres distances telles que $(X,Y) \longmapsto \mathbb{E}[\frac{|X-Y|}{1+|X-Y|}]$ ou $(X,Y) \longmapsto \inf\{\varepsilon>0 \,/\, \mathbb{P}(|X-Y|>\varepsilon)\leq\varepsilon\}$ ~~métrisent~~ aussi la convergence proba.

2) Muni de cette distance l'espace des v.a. est complet donc on a aussi le critère de Cauchy qui s'applique.

Propriété 2 $X_n \Rightarrow X$ ssi $\forall g:\mathbb{R}\longrightarrow\mathbb{R}$ continue et bornée on a $\lim \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$.

Théorème de Paul-Lévy - On considère (φ_n) les fonctions caractéristiques d'une suite de v.a. (X_n) réelles. On suppose que $\forall t\in\mathbb{R}$, $\varphi_n(t)\longrightarrow\varphi(t)$ ~~telle que~~ φ continue en 0. Alors φ est la fonction caractéristique d'une v.a. X et $X_n \Rightarrow X$.

II - Construction du diagramme des relations entre les modes de convergence

1) Implications directes

Propriété 3 La convergence presque-sûre implique la convergence en probabilité.

Propriété 4 La convergence en probabilité implique la cvg en loi.

Remarque - On en déduit que la convergence presque sûre implique celle en loi cependant cela se montre assez facilement directement alors que la propriété 4, elle, est plus difficile.

- On a défini une famille de convergence L^p il est donc intéressant de savoir comment elles interagissent entre elles. Par Holder on a :

Propriété 5 Si $p > q$ alors $X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{L^q} X$

Par l'inégalité de Markov on a aussi :

Propriété 6 La convergence L^p implique la convergence en probabilité

On a ainsi construit un diagramme d'implications des modes de convergence. Il est cependant utile de savoir si des réciproques existent

DEV 1

2) Réciproques partielles du diagramme

Propriété 7 - On peut préciser la propriété 3 avec ce critère : (X_n) suite de variables aléatoires. Alors $X_n \xrightarrow{P} X$ ssi $\forall (n')$ suite croissante d'entiers, $\exists (n'')$ sous-suite de (n') telle que $X_{n''} \xrightarrow{p.s} X$

Théorème de Lévy - Soit (X_n) suite de v.a. réelles indépendantes. Alors la série $\sum X_n$ converge en probabilité ssi elle converge presque-sûrement.

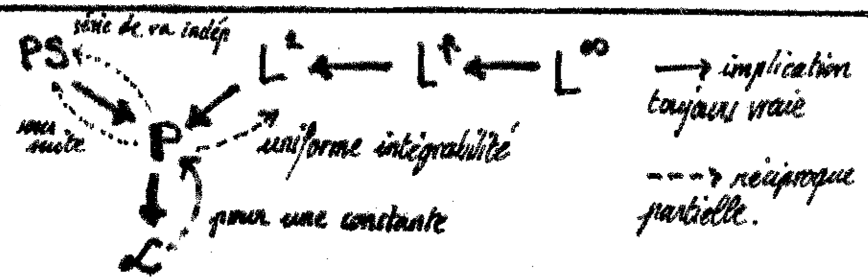
Définition - Une suite de v.a. intégrable (X_n) est dite uniformément intégrable si $\lim\limits_{c \to +\infty} \sup\limits_{n \geq 0} \mathbb{E}[|X_n| \mathbb{1}_{\{|X_n| \geq c\}}] = 0$

Théorème de Vitali - Soit $p \geq 1$, (X_n) suite de $L^p(\Omega, \mathcal{F}, \mathbb{P})$ alors (X_n) converge en L^p ssi (X_n) uniformément intégrable et converge en probabilité.

Contre-exemple - $X_n : [0;+\infty[\longrightarrow \mathbb{R}$, $x \longmapsto \frac{\mathbb{1}_{[n;+\infty[}(x)}{(nx)^2}$ intégrables, converge en probabilité mais ne converge pas en norme L^1

Propriété - Si $X_n \Longrightarrow c$ constante alors $X_n \xrightarrow{P} c$.

Remarque - On peut alors établir le diagramme d'implications toujours vraies et des réciproques partielles (voir p.)



III - Applications des modes de convergences

1) Théorèmes limites

Loi faible des grands nombres - Soit (X_n) une suite de v.a définies sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, admettant un moment d'ordre 2 et deux à deux non corrélées. On suppose la convergence des suites $\frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[X_j] \longrightarrow m$ et $\frac{1}{n^2}\sum_{j=1}^{n} \sigma_{X_j}^2 \xrightarrow[n \to +\infty]{} 0$

Alors la suite $\overline{X_n} = \frac{1}{n}\sum_{j=1}^{n} X_j$ converge en proba vers m.

Théorème de Kitchine - (X_n) suite de v.a deux à deux indépendantes et de même loi μ et admettant une moyenne. Alors la suite des variables aléatoire $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ converge en proba vers $\mathbb{E}[X_1]$, la moyenne commune.

DEV

Exemple d'application - On peut déterminer une quantité inconnue dont on sait qu'elle sera l'espérance d'une certaine variable aléatoire vérifiant les hypothèses du théorème en considérant de gros échantillons : pourcentage de présence d'un allèle dans une population, prévision d'une élection (sondage), équilibrage d'une pièce...

LGN forte - Soit (X_n) suite de v.a. iid. Il y a équivalence entre :

1) $X_1 \in L^1$

2) $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s} c \in \mathbb{R}$, $n \to +\infty$

Et dans ce cas on a $c = \mathbb{E}[X_1]$

Ainsi la loi des grands nombres fournit un résultat asymptotique consernant $\frac{1}{n}S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Cette convergence est précisée par le théorème central limite (TCL) qui indique que la vitesse de convergence est en $\sqrt{n}$ et que l'erreur commise suit, à la variance près, une loi centrée réduite gaussienne.

(appel)

Théorème central limite - (X_n) une suite de v.a réelles iid L^2 d'espérance m et de variance $\sigma^2 > 0$. Alors on a

$$\frac{S_n - nm}{\sqrt{n\sigma^2}} = \frac{\sqrt{n}}{\sigma^2}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - m\right) \Longrightarrow N(0;1)$$

(appel)

Théorème de Slutsky : Soit $X_n \xrightarrow{\mathbb{P}} X$ et $Y_n \Longrightarrow c$ constante

Alors $(X_n, Y_n) \xrightarrow{\mathbb{P}} (X, c)$

(appel)

2) Applications

Méthode delta - Soit $\phi : \mathbb{R} \to \mathbb{R}$ dérivable avec ϕ' continue en $a \in \mathbb{R}$, $(c_n) \in (\mathbb{R}_*^+)^{\mathbb{N}}$ / $c_n \to +\infty$ et (X_n) suite de v.a / $c_n(X_n - a) \Rightarrow X$. Alors $c_n(\phi(X_n) - \phi(a)) \Longrightarrow \phi'(a)X$

Remarque - En pratique on l'utilise avec le TCL, cela fournit la vitesse d'estimateurs qui sont composés (par ϕ) d'un autre estimateur de vitesse connue.

Calculs de sommes - L'application du TCL et de la LGN aux lois de Poisson fournissent des résultats asymptotiques sur certaines sommes telles que :

- $\lim_{n\to+\infty} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} = \frac{1}{2}$ ou $\forall f : \mathbb{R} \to \mathbb{R}$ continue bornée
- $\lim_{n\to+\infty} \sum_{k=0}^{+\infty} e^{-n\lambda}\frac{(n\lambda)^k}{k!} f\left(\frac{k}{n}\right) = f(\lambda)$

Référence - Ouvrard - Probabilité 1 et 2 (surtout 2).

<u>rem</u> Thm Levy demo inutilement compliquée

question : que se passe-t-il si on enlève une des deux hyp ?

(en gen : si on enlève croissante, monotone suffit)

• Théorème (de Lévy, une réciproque partielle) : Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de va réelles indépendantes, X une va réelle, définies sur $(\Omega, \mathcal{A}, P)$. Alors, $\sum_{i=0}^{n} X_n \xrightarrow[n]{P} X \implies \sum_{i=0}^{n} X_n \xrightarrow[n]{p.s} X$

• Lemme (Inégalité d'Ottaviani) : Soient $n \in \mathbb{N}^*$ et $X_1, \dots, X_n$ des va réelles indépendantes, $\varepsilon > 0$, alors on a :

$$\min_{k\in[\![1,n-1]\!]} P\left(\left|\sum_{i=k+1}^{n} X_i\right| \leq \varepsilon\right) P\left(\max_{k\in[\![1,n]\!]} \left|\sum_{i=1}^{k} X_i\right| > 2\varepsilon\right)$$

$$\leq P\left(\left|\sum_{i=1}^{n} X_i\right| > \varepsilon\right)$$

• Preuve du théorème : On suppose $\sum_{i=0}^{n} X_n \xrightarrow[n]{P} X$

On pose $n \in \mathbb{N}^*$, $S_n = \sum_{i=1}^{n} X_i$, $A_n = \sup_{k\in\mathbb{N}^*} |S_{n+k} - S_n|$

et $A = \inf_{n\in\mathbb{N}^*} A_n$

On remarque $\{\sum_n X_n \text{ converge}\} = \{A = 0\}$ (critère de Cauchy)

On va montrer que $P(A \neq 0) = 0$

On a $\{A \neq 0\} = \bigcup_{\varepsilon\in\mathbb{Q}_+^*} \{A > 2\varepsilon\}$

$$\{A > 2\varepsilon\} \subset \bigcap_{n\in\mathbb{N}^*} \{A_n > 2\varepsilon\}$$

D'où $\{A \neq 0\} \subset \bigcup_{\varepsilon\in\mathbb{Q}_+^*} \bigcap_{n\in\mathbb{N}^*} \{A_n > 2\varepsilon\}$

Mais encore, pour $\varepsilon \in \mathbb{R}_+^*$, $m \in \mathbb{N}^*$, on a

$$\{ A_m > 2\varepsilon \} = \bigcup_{r \in \mathbb{N}^*} \{ \max_{k \in \{1,..,r\}} |S_{m+k} - S_m| > 2\varepsilon \}$$

On remarque que $S_{m+k} - S_m = \sum_{i=1}^{k} X_{m+i}$, on applique Ottaviani avec $(X_{m+i})_{i \in \{1,..,r\}}$, pour $r \in \mathbb{N}^*$, on a

$$\min_{k \in \{1,..,r-1\}} P(|S_{m+r} - S_{m+k}| \leq \varepsilon) P(\max_{k \in \{1,..,r\}} |S_{m+k} - S_m| > 2\varepsilon)$$

$$\leq P(|S_{m+r} - S_m| > \varepsilon) \quad (*)$$

On remarque, pour $k \in \{0,..,r\}$, pour $\delta \in]0;1[$, on a

$$P(|S_{m+r} - S_{m+k}| > \varepsilon)$$

$$\leq P(|S_{m+r} - X| > \frac{\varepsilon}{2}) + P(|S_{m+k} - X| > \frac{\varepsilon}{2})$$

Comme $\sum_{i=1}^{n} X_n \xrightarrow[n]{P} X$, on peut poser $N(\varepsilon, \delta)$, ne dépendant pas de r et k, tel que dès que $m > N(\varepsilon, \delta)$, on a

$P(|S_{m+r} - S_{m+k}| > \varepsilon) < \delta$ et donc aussi que

$$P(|S_{m+r} - S_{m+k}| \leq \varepsilon) \geq 1 - \delta$$

En particulier $\min_{k \in \{1,..,r-1\}} P(|S_{m+r} - S_{m+k}| \leq \varepsilon) \geq 1 - \delta$

Donc on a $P(\max_{k \in \{1,..,r\}} |S_{m+k} - S_m| > 2\varepsilon) \leq \frac{\delta}{1-\delta}$

(En appliquant $(*)$.)

$\boxed{2/8}$

Les $\{\max_{k\in\{1,..,n\}} |S_{m+k} - S_m| > 2\varepsilon\}$ formant une suite croissante en n,

pour $m > N(\varepsilon,\delta)$, $N(\varepsilon,\delta)$ ne dépendant pas de n et k, on a :

$$P(A_m > 2\varepsilon) = \lim_n P(\max_{k\in\{1,..,n\}} |S_{m+k} - S_m| > 2\varepsilon) \leq \frac{\delta}{1-\delta}$$

Donc, pour $\delta \in]0,1[$, pour $m > N(\varepsilon,\delta)$, on a

$$P(\bigcap_{p\in\mathbb{N}^*} (A_p > 2\varepsilon)) \leq P(A_m > 2\varepsilon) \leq \frac{\delta}{1-\delta}$$

Ceci étant vrai pour tout $\delta \in]0,1[$, on a

$$P(\bigcap_{p\in\mathbb{N}^*} (A_p > 2\varepsilon)) = 0$$

Donc $P(A \neq 0) \leq P(\bigcup_{\varepsilon\in\mathbb{Q}_+^*} \bigcap_{p\in\mathbb{N}^*} (A_m > 2\varepsilon)) = 0$

Donc $P(A\neq 0) = 0$, donc $P(\sum_n X_n \text{ converge}) = P(A=0) = 1$

Donc $\sum_{i=0}^{n} X_i \xrightarrow[n]{P\text{-}ps} X$ car la limite est nécessairement

égale P-ps.

• Preuve de l'inégalité d'Ottaviani : Pour $k \in \{1,..,n\}$,

On pose $S_k = \sum_{i=1}^{k} X_i$; $M_k = \max_{i\in\{1,..,k\}} |S_i|$

$S_{n,k} = \sum_{i=k+1}^{n} X_i$ avec $S_{n,n} = 0$

Pour $\varepsilon > 0$, on pose $E_\varepsilon = (M_n > 2\varepsilon)$ $\boxed{3/8}$

$E_\varepsilon^1 = (|S_1| > 2\varepsilon)$ et pour $k \geq 2$,

$E_\varepsilon^k = (|S_k| > 2\varepsilon) \cap \left[\bigcap_{i=1}^{k-1} (|S_i| \leq 2\varepsilon)\right]$

On remarque $E_\varepsilon = \dot{\bigcup}_{k=1}^{m} E_\varepsilon^k$ (càd les $(E_\varepsilon^k)_k$ forment une partition de E_ε)

On a $(|S_m| > \varepsilon) \supset (|S_m| > \varepsilon) \cap E_\varepsilon \overset{(*)}{=} \dot{\bigcup}_{k=1}^{m} [(|S_m| > \varepsilon) \cap E_\varepsilon^k]$

On a $S_m = S_k + S_{m,k}$, donc $(|S_k| > 2\varepsilon$ et $|S_{k,m}| \leq \varepsilon \Rightarrow |S_m| > \varepsilon)$

(Sinon on aurait $|S_k| \leq |S_m| + |S_{k,m}| \leq 2\varepsilon$, contradiction avec $|S_k| > 2\varepsilon$.)

On a donc $(|S_m| > \varepsilon) \cap E_\varepsilon^k \supset (|S_{k,m}| \leq \varepsilon) \cap E_\varepsilon^k$

Donc $P(|S_m| > \varepsilon) \geq \sum_{k=1}^{m} P\left[(|S_{k,m}| \leq \varepsilon) \cap E_\varepsilon^k\right]$ d'après $(*)$

$(|S_{k,m}| \leq \varepsilon)$ et E_ε^k étant toujours indépendants, on a

$$P(|S_m| > \varepsilon) \geq \sum_{k=1}^{m} P(|S_{k,m}| \leq \varepsilon) \times P(E_\varepsilon^k)$$

$$\geq \min_{k \in \{1,\dots,m\}} P(|S_{k,m}| \leq \varepsilon) \times \sum_{k=1}^{m} P(E_\varepsilon^k)$$

$$= \min_{k \in \{1,\dots,m\}} P(|S_{k,m}| \leq \varepsilon) \times P(E_\varepsilon)$$

$$= \min_{k \in \{1,\dots,m-1\}} P(|S_{k,m}| \leq \varepsilon) \times P(E_\varepsilon)$$

cqfd.

$\boxed{4/8}$

• Théorème de Kintchine (loi faible des grands nombres) : Soit $(X_n)_n$ une suite de va réelles définies sur $(\Omega, \mathcal{A}, P)$, deux à deux indépendantes, de même loi μ, L^1. Alors,

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n]{P} EX_1 \text{, leur moyenne commune}$$

• Loi des grands nombres "basiques" utilisée : Soit $(X_n)_n$ une suite de va réelles définies sur $(\Omega, \mathcal{A}, P)$, L^2, deux à deux non corrélées (càd de coefficient de corrélation nul, càd de covariance nulle). On suppose, pour un certain $m \in \mathbb{R}$:

$$\frac{1}{n}\sum_{i=1}^{n} EX_i \xrightarrow[n]{} m \quad \text{et} \quad \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2_{X_i} \xrightarrow[n]{} 0$$

Alors $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n]{P} m$

• Lemme de Kronecker : Soient une série convergente de terme général réel x_n, $n \in \mathbb{N}^*$, et une suite croissante $(b_n)_{n\in\mathbb{N}^*}$ de réels, tendant vers infini avec n, alors : $\frac{1}{b_n}\sum_{i=1}^{n} b_i x_i \xrightarrow[n]{} 0$

• Preuve du théorème : On va se ramener à la LGN "basique". Pour $n \in \mathbb{N}^*$, on pose $Y_n = \mathbb{1}_{(|X_n| \leq n)} X_n$ et $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$

Les va Y_n, $n \in \mathbb{N}^*$, sont indépendantes deux à deux, donc non corrélées, et bornées, donc L^2.

$\boxed{5/8}$

On a $\sum_{i=1}^{n} EY_i = \sum_{i=1}^{n} \int_{(|x| \leq i)} x \, d\mu(x)$

l'intégrale étant nulle sur $\{0\}$

$$\overset{\curvearrowleft}{=} \sum_{i=1}^{n} \sum_{k=0}^{i} \int_{k < |x| \leq k+1} x \, d\mu(x)$$

$$= \sum_{k=1}^{n} (n-k) \int_{k < |x| \leq k+1} x \, d\mu(x)$$

D'où $\frac{1}{n} \sum_{i=1}^{n} EY_i = \int_{|x| \leq n} x \, d\mu(x) - \frac{1}{n} \sum_{i=1}^{n} k \int_{(k < |x| \leq k+1)} x \, d\mu(x)$

X_1 étant L^1, on a $\int_{|x| \leq n} x \, d\mu(x) \xrightarrow[n]{} EX_1$

De plus la série de terme général $\int_{(k < |x| \leq k+1)} x \, d\mu(x)$, pour $k \in \mathbb{N}^*$

est convergente ~~ent~~, en appliquant le lemme de Kronecker, on a :

$$\frac{1}{n} \sum_{k=0}^{n-1} k \int_{(k < |x| \leq k+1} x \, d\mu(x) \xrightarrow[n]{} 0$$

Donc $\frac{1}{n} \sum_{i=1}^{n} EY_i = EX_1$

Par ailleurs, pour $n \geq i \geq 0$, on a

$$\sigma_{Y_i}^2 \leq EY_i^2 = \int_{(|x| \leq i)} x^2 \, d\mu(x) \leq \int_{(|x| \leq n)} x^2 \, d\mu(x)$$

Et donc,

$\boxed{6/8}$

$$0 \leq \frac{1}{n^2} \sum_{i=1}^{n} \sigma_{Y_i}^2 \leq \frac{1}{n} \int_{(|x| \leq n)} x \, d\mu(x)$$

$$\leq \frac{1}{n} \left[\int_{(|x| \leq \sqrt{n})} x^2 d\mu(x) + \int_{(\sqrt{n} < |x| \leq n)} x^2 d\mu(x) \right]$$

$$\leq \frac{1}{\sqrt{n}} \int_{(|x| \leq \sqrt{n})} |x| d\mu(x) + \int_{(\sqrt{n} < |x|)} |x| \, d\mu(x)$$

$$\xrightarrow[n]{} 0 \quad \text{car} \int_{\mathbb{R}} |x| \, d\mu(x) < +\infty$$

On a bien $\frac{1}{n^2} \sum_{i=1}^{n} \sigma_{Y_i}^2 \xrightarrow[n]{} 0$

D'après la LGN "basique" $\overline{Y}_n \xrightarrow[n]{P} EX_1$

Pour $n > r \in \mathbb{N}$, on note $\overline{Y}_{n,r} = \frac{1}{n} \left(\sum_{i=1}^{r} X_i + \sum_{i=r+1}^{n} Y_i \right)$.

On a $P(\overline{Y}_{n,r} \neq \overline{X}_n) \leq P\left(\bigcup_{i=r+1}^{n} (X_i \neq Y_i)\right)$

$$\leq \sum_{i=r+1}^{n} P(X_i \neq Y_i) = \sum_{i=r+1}^{n} P(|X_1| > i)$$

$$= \sum_{i=r+1}^{n} \int_{\mathbb{R}} \mathbb{1}_{(|x| > i)}(x) \, d\mu(x)$$

$$= \int_{\mathbb{R}} \left(\sum_{i=r+1}^{n} \mathbb{1}_{(|x| > i)}(x) \right) d\mu(x)$$

$\boxed{7/8}$

$$\leqslant \int_{\mathbb{R}} \mathbb{1}_{(|x|>n)}(x) \times |x| \, d\mu(x)$$

Donc pour $\varepsilon > 0$ quelconque, on peut poser $n > 0$ tel que pour tout $m > n$, on ait $P(\overline{Y_{m,n}} \neq \overline{X}_m) \leqslant \frac{\varepsilon}{2}$, alors pour $\delta > 0$, on a :

$P(|\overline{X}_m - EX_1| > \delta)$

$= P[(|\overline{X}_m - EX_1| > \delta) \cap (\overline{Y_{m,n}} \neq \overline{X}_m)] + P[(|\overline{X}_m - EX_1| > \delta) \cap (\overline{Y}_{m,n} = \overline{X}_m)]$

$\leqslant P(\overline{Y_{m,n}} \neq \overline{X}_m) + P(|\overline{Y_{m,n}} - EX_1| > \delta)$

Or $\overline{Y}_m \xrightarrow[m]{P} EX_1$ donc $\overline{Y_{m,n}} = \frac{1}{m}\left(\sum_{i=1}^{n} X_i + \sum_{i=n+1}^{m} Y_i\right)$

$$= \overline{Y}_m - \frac{1}{m}\sum_{i=1}^{n}(X_i - Y_i)$$

$$\xrightarrow[m]{P} EX_1 + 0$$

Donc on peut poser $N > n$ tel que dès que $m \geqslant N$, on ait $P(|\overline{Y}_{m,n} - EX_1| > \delta) \leqslant \frac{\varepsilon}{2}$ et $P(|\overline{X}_m - EX_1| < \delta) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$

ε et δ étant arbitraires, on a bien $\overline{X}_m \xrightarrow[m]{P} EX_1$

Référence : Probabilité 2, Jean-Yves Ouvrard
Pour le lemme de Kronecker, voir p 105 (parce que ça fait déjà 8 pages j'écris trop gros !)