

Cours 3

Proposition : $(\mu_j)_{j \in \mathbb{N}^*}$ suite de probabilités sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$

Il existe une suite de v.a.r.i $(X_j)_{j \in \mathbb{N}^*}$

sur $([0,1], \mathcal{B}_{[0,1]}, \mathbb{P})$ t.q.
 \hookrightarrow Lebesgue

X_j est de loi μ_j

preuve : On a construit v.a.r.i.

$$Y_j = \sum_{k=1}^{\infty} \frac{1}{2^k} D_{\varphi_j(k)}$$

lemme $\left\{ \begin{array}{l} \forall j \\ Y_j \end{array} \right.$ est de loi uniforme sur $[0,1]$

preuve : On pose

$$Y_{j,n} = \sum_{k=1}^n \frac{1}{2^k} D_{\varphi_j(k)} \rightarrow D_k ?$$

On verra : la loi d'une somme de v.a.r.i ne dépend que de la loi de chaque v.a.
 donc la loi de $Y_{j,n}$ est la même

que la loi de

$$Z_n = \sum_{k=1}^n \frac{1}{2^k} D_k \quad \text{développement dyadique dans } [0,1]$$

On montre alors que $Z_n \rightarrow Z = \text{identité}$ sur $[0,1]$

Pour obtenir la loi de Y_j :

$$Y_j = \lim_{n \rightarrow \infty} Y_{j,n} \quad (\text{croissante})$$

$$\begin{aligned} \Rightarrow P(Y_j \leq y) &= \lim_{n \rightarrow \infty} P(Y_{j_1^n} \leq y) \quad \left. \begin{array}{l} \text{même} \\ \text{loi} \end{array} \right\} \\ &= \lim_{n \rightarrow \infty} P(Z_n \leq y) \\ &= P(Z \leq y) = y \\ &\quad (\text{car } Z = \text{identité}) \end{aligned}$$

□

Lemme: (simulation d'une v.a.)

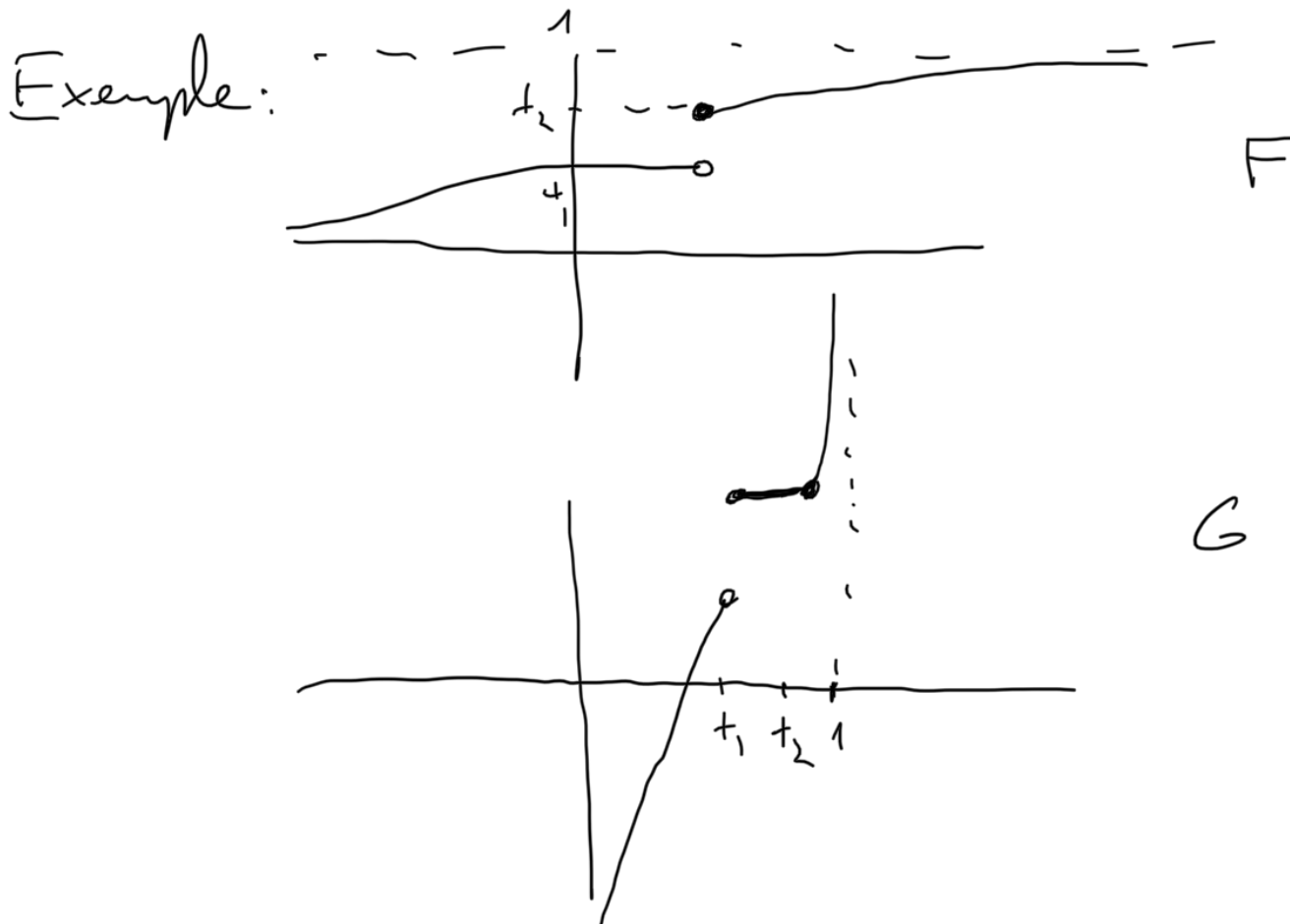
Soit Y v.a.r de loi uniforme sur $[0, 1]$

On fixe une fonction de répartition $F: \mathbb{R} \rightarrow [0, 1]$

On définit la pseudo-inverse

$$\forall t \quad \underline{G}(t) = \inf \{ x \mid F(x) \geq t \}$$

Alors $\underline{G}(Y)$ admet F comme fonction de répartition.



preuve: On observe que par définition

$$F(x) \geq t \iff x \geq \underline{G}(t)$$

$$\Rightarrow \forall x \in \mathbb{R} \quad P(G(Y) \leq x) = P(\underline{Y \leq F(x)}) = \underline{F(x)}$$

donc $F(x)$ est la fonction de répartition de $G(Y)$. □

A partir de Y_j on compose avec la fonction G_j obtenue par lemmes à partir des lois μ_j :

$G_j(Y_j)$ est la loi μ_j et sont indépendantes □

9.4 Sommes de variables aléatoires indépendantes.
Soient μ et ν mesures de probabilités sur \mathbb{R}^d

Definition (Produit de convolution de deux mesures)

$\mu * \nu$ est la mesure image de $\mu \otimes \nu$

l'application $(x, y) \xrightarrow{s} x + y$

Remarque : $\forall \varphi : \mathbb{R}^d \rightarrow \mathbb{R}_+$ mesurable

$$\int_{\mathbb{R}^d} \varphi(z) \mu * \nu(dz) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x+y) \mu \otimes \nu(dx dy)$$

définition

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x+y) \mu(dx) \nu(dy)$$

par Fubini-Tonelli = $\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x+y) \mu(dx) \right) \nu(dy)$.

Proposition Soient X, Y v.a.i à valeurs dans \mathbb{R}^d

i) La loi de $X+Y$ est $P_X * P_Y$

En particulier, si X est de densité p_X et Y est de densité p_Y alors $X+Y$ est de densité la convolution $p_X * p_Y$.

ii) La Fonction caractéristique de $X+Y$ est

$$\Phi_{X+Y}(z) = \Phi_X(z) \Phi_Y(z)$$

(de façon équivalente : si μ et ν sont deux mesures de probabilités sur \mathbb{R}^d)

iii) si X et Y sont dans L^2 alors

$$K_{X+Y} = K_X + K_Y$$

$$\text{où } K_X = (\text{cov}(X_i, X_j))$$

En particulier si $d=1$

$$(\text{cov}(X, X) = \text{var}(X))$$

$$\text{d'où } \text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$$

preuve: i) On suppose X, Y v.a.i

$$\Leftrightarrow P_{(X,Y)} = P_X \otimes P_Y$$

Rappel: supposons $f: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ mesurable

$$E[f] = \int f(x,y) P_{X,Y}(dx dy)$$

$$x, y \text{ indépendantes} = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x, y) P_x(dx) P_y(dy)$$

alors

$$E[e^{(x+y)}] = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{(x+y)} P_x(dx) P_y(dy)$$

par définition

$$= \int_{\mathbb{R}^d} e(z) P_x * P_y(dz)$$

si de plus X est de densité P_x
 Y " " P_y

$$E[e^{(x+y)}] = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{(x+y)} p_x(x) dx p_y(y) dy$$

changement de variable $x+y=z \Rightarrow x=z-y$

$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e(z) p_x(z-y) dz p_y(y) dy$$

F.-T

$$= \int_{\mathbb{R}^d} e(z) \left(\int_{\mathbb{R}^d} p_x(z-y) p_y(y) dy \right) dz$$

donc $P_x * P_y$ est

la densité de $X+Y$.

Observer que la convolution est bien définie car P_x et P_y sont dans L^1 .

ii) Par définition:

$$\Phi_{X+Y}(z) = E[e^{iz(x+y)}]$$

par F.-T. $= \mathbb{E}[e^{i3^x}] \mathbb{E}[e^{i3^y}]$
 $= \phi_x(z) \phi_y(z)$

iii) Matrice de covariance:

$$X = (X_1, \dots, X_d) \quad Y = (Y_1, \dots, Y_d)$$

indépendantes $\Rightarrow \underline{\text{cov}(X_i, Y_j) = 0 \quad \forall i, j}$

par définition: $(\text{cov}(X_i + Y_i, X_j + Y_j))$
 est la matrice de covariance de
 $X + Y$

$$= (\text{cov}(X_i, X_j) + \text{cov}(Y_i, Y_j))$$

d'où

$$K_{X+Y} = K_X + K_Y$$

Problème: Trouver la loi d'une somme
 d'un grand nombre de v.a.i.

On notera souvent $S_n = X_1 + \dots + X_n$

la somme de n variables aléatoires

Théorème (loi faible de grands nombres)

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r.i

de même loi. Si $\mathbb{E}[X_1^2] < \infty$

alors

$$\frac{1}{n} S_n \xrightarrow[n \rightarrow \infty]{L} \mathbb{E}[X_1]$$

preuve : $E\left[\frac{1}{n} S_n\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] = E[X_1]$

$$E\left[\left(\frac{S_n}{n} - E[X_1]\right)^2\right] \xrightarrow{?} 0$$

$$\frac{1}{n^2} \text{var}(S_n)$$

mais par indépendance $\text{var}(S_n) = \sum_{k=1}^n \text{var}(X_k)$
 $= n \text{var}(X_1)$

$$E\left[\left(\frac{S_n}{n} - E[X_1]\right)^2\right] = \frac{1}{n^2} n \text{var}(X_1)$$

$$\rightarrow 0$$

□

Remarques: (1) La preuve montre que le théorème est vraie avec les hypothèses:

- $E[X_i] = E[X_1] \quad \forall i$
- $E[X_i^2] < \infty$
- $\text{cov}(X_i, X_j) = 0 \quad \forall i, j$

(2) loi faible : convergence L^P
 loi forte : convergence p.s.

Proposition $(X_n)_{n \geq 1}$ v.a.r.i. de même loi
 t.q. $E[X_1^4] < \infty \leftarrow$ hypothèse forte.
 A.D.

4 x 1000

$$\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{P.S.} \mathbb{E}[X_1]$$

preuve: On suppose, sans perte de généralité,
que $\mathbb{E}[X_1] = 0$

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] \stackrel{\text{linéarité}}{=} \frac{1}{n^4} \sum_{\substack{i_1, i_2, i_3, i_4 \\ \in \{1, \dots, n\}}} \mathbb{E}[X_{i_1} X_{i_2} X_{i_3} X_{i_4}]$$

Par regroupement de paquets on obtient
dans la somme que les termes où
 $\{i_1, i_2, i_3, i_4\}$ est de cardinalité au plus 2

Exemple: On sait que X_1, X_2, X_3 sont
indépendants.

$\Rightarrow X_1, (X_2, X_3)$ indépendants

$\Rightarrow X_1, X_2^2 X_3$ indépendants

$$\begin{aligned} \Rightarrow \mathbb{E}[X_1 \cdot (X_2^2 X_3)] &= \mathbb{E}[X_1] \mathbb{E}[X_2^2 X_3] \\ &\stackrel{||}{=} 0 \\ &= 0 \end{aligned}$$

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] = \frac{1}{n^4} \left(n \mathbb{E}[X_1^4] + c_n \mathbb{E}[X_1^2 X_2^2] \right)$$

où $c_n = \frac{4!}{2!2!} \binom{n}{2} \rightarrow \frac{n(n-1)}{2}$

$\hookrightarrow X_1 X_1 X_2 X_2$

(le résultat de l'expérience numéro i est $\mathbb{1}_{A_i}$)

preuve: On vérifie les hypothèses du théorème précédent pour les v.a. $\mathbb{1}_{A_i}$:

$$E[\mathbb{1}_{A_i}] = P(A_i) = P(A_1)$$

$$E[\mathbb{1}_{A_i}^4] = E[\mathbb{1}_{A_i}] < \infty$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{A_i} \xrightarrow[n \rightarrow \infty]{\text{P.S.}} P(A_1) \quad \square$$

Application au développement dyadique :

On rappelle que

$$\forall \omega \in [0, 1[\quad \omega = \sum_{k=1}^{\infty} \frac{X_k(\omega)}{2^k} \quad \text{où } X_k \text{ v.a.i. de même loi}$$

Soit $p \geq 1$ fixé. On regroupe par

paquets: $Y_1 = (X_1, \dots, X_p)$

$$Y_2 = (X_{p+1}, \dots, X_{2p})$$

sont v.a.i. de même loi.

On considère les événements indépendants

$$A_n = \{ Y_n = (c_1, \dots, c_p) \}$$

Par le corollaire :

$\frac{1}{n} \sum_{i=1}^n \dots$

$$\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{A_i} \xrightarrow[n \rightarrow \infty]{P.S.} P(A_i) = \frac{1}{2^p}$$

(calculé la semaine dernière)

On utilise le même argument pour

les variables $Y_1 = (X_{l+1}, \dots, X_{l+p-1})$

$Y_2 = (X_{l+2p}, \dots, X_{l+3p-1})$

⋮

et on obtient $A_n = \{ Y_n = (c_1, \dots, c_p) \}$

$$\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{A_i} \xrightarrow[n \rightarrow \infty]{P.S.} P(A_i) = \frac{1}{2^p}$$

Autre tl

$$\frac{1}{n} \# \{ j \leq n \mid X_{jp+l}(\omega) = c_1, \dots, X_{(j+1)p+l-1}(\omega) = c_p \}$$

$$\longrightarrow \frac{1}{2^p}$$

conclusion:

$$\frac{1}{n} \# \{ k \leq n \mid X_{k+1}(\omega) = c_1, \dots, X_{k+p}(\omega) = c_p \}$$

$$\xrightarrow[n \rightarrow \infty]{P.S.} \downarrow$$

$$\frac{1}{2^p}$$

i.e. la fréquence d'apparition d'un bloc de longueur p de 0 et 1

est égale à $\frac{1}{2^p}$.

Semi-groupe de convolution

$$\mathbb{I} = \underline{\mathbb{N}} \text{ ou } \underline{\mathbb{R}^+}$$

Définition: Soit $(\mu_t)_{t \in \mathbb{I}}$ une famille de mesures de probabilités sur \mathbb{R}^d . On dit que $(\mu_t)_{t \in \mathbb{I}}$ est un semi-groupe de convolution si

$$i) \mu_0 = \delta_0$$

$$ii) \mu_{t+s} = \mu_t * \mu_s \quad \forall t, s \in \mathbb{I}$$

Remarque: si μ_t est la loi de X_t et $\forall t, s \in \mathbb{I}$ X_t et X_s sont indépendantes alors la loi de

$$X_t + X_s \text{ est } \mu_{t+s} = \mu_t * \mu_s.$$

lemme: $(\mu_t)_{t \in \mathbb{I}}$ est un semi-groupe de convolution si $\exists \varphi: \mathbb{R} \rightarrow \mathbb{C}$

$$t.g. (* | \hat{\mu}_t(\xi) = (\varphi(\xi))^t \quad \forall t \in \mathbb{I} = \mathbb{N}$$

$$\text{ou } (** | \hat{\mu}_t(\xi) = e^{-t\varphi(\xi)} \quad \forall t \in \mathbb{I} = \mathbb{R}_+.$$

preuve: On sait que

$$\widehat{\mu}_+ * \widehat{\mu}_{+'} = \widehat{\mu}_+ * \widehat{\mu}_{+'}$$

par définitions $(\mu_+)_{t \in I}$ est semi-groupe de convolution ssi

$$\widehat{\mu}_+ * \widehat{\mu}_{+'} = \widehat{\mu}_{++'} \quad \forall t, t' \in I$$

$$\text{ssi } \widehat{\mu}_+ * \widehat{\mu}_{+'} = \widehat{\mu}_{++'} \quad *$$

On vérifie maintenant que (*) et (**) satisfont la condition *

$$\mathcal{Q}(\xi)^+ \cdot \mathcal{Q}(\xi)^{+'} = \mathcal{Q}(\xi)^{++'}$$

□

Exemples (1) $\forall n, \mu_n$ de loi binomiale $B(n, p)$
 où $p \in [0, 1]$: $P(X_n = k) = C_n^k p^k (1-p)^{n-k}$

$$\begin{aligned} \widehat{\mu}_n(\xi) &= \int e^{i\xi x} dP_x(dx) \\ &= \sum_{k=0}^n e^{i\xi k} C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n C_n^k (e^{i\xi} p)^k (1-p)^{n-k} \\ &= (e^{i\xi} p + (1-p))^n \end{aligned}$$

$$\mathcal{Q}(\xi) = e^{i\xi} p + 1-p$$

(2) $\forall t > 0 \dots$

(2) $\forall t \neq 0$ μ_t loi de Poisson
de paramètre t :

$$P_t(X_t = k) = \frac{t^k e^{-t}}{k!}$$

Alors $\hat{\mu}_t^{\lambda}(z) = \int e^{izx} P_t(dx)$

$$= \sum_{k=0}^{\infty} e^{izk} \frac{t^k e^{-t}}{k!}$$
$$= e^{-t} \sum_{k=0}^{\infty} \frac{(te^{iz})^k}{k!}$$
$$= e^{-t} (1 - e^{iz})$$
$$e(z) = (1 - e^{iz})$$

(3) μ_t est de loi gaussienne

$$\mathcal{N}(0, t)$$

$$\hat{\mu}_t^{\lambda}(z) = e^{-\frac{t z^2}{2}} \quad \text{exercice}$$

- si X et Y v.a.i suivent des lois de Poisson de paramètre λ et λ' alors $X+Y$ suit une loi de Poisson de paramètre $\lambda + \lambda'$
- si X et Y ^{v.a.i} sont des lois gaussiennes $\mathcal{N}(m, \sigma^2)$ et $\mathcal{N}(m', \sigma'^2)$ alors $X+Y$ suit une loi $\mathcal{N}(m+m', \sigma^2 + \sigma'^2)$.
exercice.

Chapitre 10 : Convergence de variables aléatoires.

- loi de grands nombres (forte)
- théorème central limite.

Soient (Ω, \mathcal{F}, P) un espace de probabilité et $(X_n)_{n \geq 1}$ une suite de v.a. à valeurs dans \mathbb{R}^d . X v.a.

Définition (1) $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ si

$$P\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

$$(2) \quad \forall \epsilon > 0 \quad \lim_{n \rightarrow \infty} P\left(\left\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| \geq \epsilon\right\}\right) = 0$$

$$\left(\int \right) \wedge_n \xrightarrow{n \rightarrow \infty} \wedge \quad P \in \mathcal{L}, \sim \mathcal{L}$$

si $\lim_{n \rightarrow \infty} P(|X_n - X|^p) = 0$

(3) convergence en probabilité

$$X_n \xrightarrow[n \rightarrow \infty]{(P)} X \quad \text{si}$$

$$\forall \varepsilon > 0 \quad P(|X_n - X| > \varepsilon) \xrightarrow[n \rightarrow \infty]{} 0$$

Proposition: Si $X_n \xrightarrow[n \rightarrow \infty]{P.S.} X$ alors

$$X_n \xrightarrow[n \rightarrow \infty]{(P)} X$$

preuve: $\forall \varepsilon > 0$ on définit

$$\Omega_n^\varepsilon = \left\{ \omega \in \mathcal{X} \mid \sup_{k \geq n} |X_k(\omega) - X(\omega)| > \varepsilon \right\}$$

• On observe que $\Omega_{n+1} \subset \Omega_n$
(i.e. la suite est décroissante)

• Aussi $X_n \xrightarrow[n \rightarrow \infty]{P.S.} X \implies P\left(\bigcap_{n \geq 1} \Omega_n\right) = 0$

donc $\lim_{n \rightarrow \infty} P(\Omega_n) = 0$

• On observe que $\{\omega \mid |X_n(\omega) - X(\omega)| > \varepsilon\} \subset \Omega_n$
donc $P(|X_n(\omega) - X(\omega)| > \varepsilon) \leq P(\Omega_n)$

$$\Rightarrow \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$$

$$\Rightarrow X_n \xrightarrow[n \rightarrow \infty]{(P)} X$$

□

La réciproque est fautive

Exemple 1 : Soit $(X_n)_{n \geq 1}$ ^{v.a.i} i.i.d. f.g.

$$\bullet P(X_n = 0) = 1 - \frac{1}{\sqrt{n}} \quad P(X_n = n) = \frac{1}{\sqrt{n}}$$

$$\bullet X(\omega) = 0 \quad \forall \omega$$

On calcule :

$$P(|X_n - \underset{0}{X}| > \varepsilon) = P(X_n = n) = \frac{1}{\sqrt{n}} \rightarrow 0$$

$$X_n \xrightarrow[n \rightarrow \infty]{(P)} X$$

$$X_n \xrightarrow[n \rightarrow \infty]{P.S.} X \quad ? \quad \text{non}$$

$$\text{On a } i) \sum_{n \geq 1} P(X_n = 0) = \sum_{n \geq 1} \left(1 - \frac{1}{\sqrt{n}}\right) = \infty$$

$$ii) \sum_{n \geq 1} P(X_n = n) = \sum_{n \geq 1} \frac{1}{\sqrt{n}} = \infty$$

Borel - Cantelli :

i) P.S. $\#\{n \mid X_n(\omega) = 0\}$ est infini

ii) P.S. $\#\{n \mid X_n(\omega) = n\}$ est infini

$\Rightarrow X_n \not\xrightarrow[n \rightarrow \infty]{P.S.} X$ contradiction.

~~$n \rightarrow \infty$~~

Exemple ② $\Omega = [1, 2]$

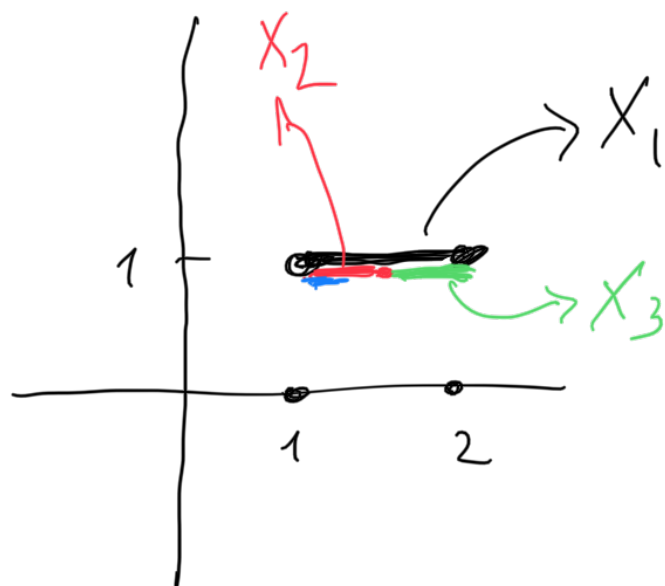
si $2^k \leq n < 2^{k+1}$ on définit

$$X_n(t) = \mathbb{1}_{\left[\frac{n}{2^k}, \frac{n+1}{2^k}\right]}$$

$$X_1(t) = \mathbb{1}_{[1, 2]}$$

$$X_2(t) = \mathbb{1}_{\left[1, \frac{3}{2}\right]}$$

$$X_3(t) = \mathbb{1}_{\left[\frac{3}{2}, 2\right]}$$



\vdots

$$X(\omega) = 0$$

$$P(|X_n - X| > \varepsilon) = P(X_n > \varepsilon) = \frac{1}{2^k}$$

si $2^k \leq n < 2^{k+1}$

$$\Rightarrow X_n \xrightarrow{(P)} 0$$

mais

$$\cancel{X_n \xrightarrow[n \rightarrow \infty]{P.S.} 0}$$

Proposition: Si $X_n \xrightarrow[n \rightarrow \infty]{(P)} X$

Alors il existe une suite extraite

$$X_{n_k} \xrightarrow[n \rightarrow \infty]{P.S.} X$$

Remarque: $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ si et seulement

$$\xrightarrow{\text{si}} \forall \varepsilon > 0 \quad P\left(\bigcup_{N \geq 0} \bigcap_{k \geq N} \{|X_k - X| \leq \varepsilon\}\right) = 1$$

preuve: Par hypothèse

$$P(|X - X_n| > \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$$

i.e. $\forall k \exists n_k$ t.q. si $n \geq n_k$

$$P(|X - X_{n_k}| > \frac{1}{k}) \leq \frac{1}{2^k}$$

$\varepsilon \qquad \delta$

On obtient que $\sum_{k=1}^{\infty} P(|X - X_{n_k}| > \frac{1}{k})$ converge

Borel-Cantelli :

$$P\left(\bigcap_{N \geq 1} \bigcup_{k \geq N} \{|X - X_{n_k}| > \frac{1}{k}\}\right) = 0$$

i.e.

$$P\left(\bigcup_{N \geq 1} \bigcap_{k \geq N} \{|X - X_{n_k}| < \frac{1}{k}\}\right) = 1$$

$$\Rightarrow \lim_{n \rightarrow \infty} X_{n_k} = X$$

~~Q~~