Probabilités et statistique

Cours de Antti Knowles

Notes de Ekin Arikök

Printemps 2022

Table des matières

Chapitre 0

Prélude : Conditionnement

Cours 1

Intuition : On prend un espace probabilisé $(\Omega, \mathcal{G}, \mathbb{P})$, où $\mathcal{G}$ est une tribu finie, et $\mathbb{P}[A] > 0$ pour tout $A \in \mathcal{G}$. Soit X une variable aléatoire dans L^1. L'espérence de X sur A est

$$\mathbb{E}[X|A] := \frac{\mathbb{E}[X \cdot \mathbb{1}_A]}{\mathbb{P}[A]}$$

Image 1

On définit *l'espérence conditionnelle* de X sachant $\mathcal{G}$ par

$$\mathbb{E}[X|\mathcal{G}] := \sum_{A \in \mathcal{G}} \mathbb{E}[X|A] \cdot \mathbb{1}_A$$

(À corriger)

Proposition 0.1. *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Pour $\mathcal{G} \subset \mathcal{F}$ une sous-tribu de $\mathcal{F}$ et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, il existe une unique variable aléatoire $Z \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ telle que*

$$\mathbb{E}[X \cdot \mathbb{1}_B] = \mathbb{E}[Z \cdot \mathbb{1}_B], \forall B \in \mathcal{G}$$

On appelle cette variable aléatoire ***l'espérance conditionnelle*** *de X sachant $\mathcal{G}$, et on note $\mathbb{E}[X|\mathcal{G}] := Z$.*

Proposition 0.2.

1. *$\mathbb{E}[X|\mathcal{G}]$ est $\mathcal{G}$-mesurable*
2. *$\forall B \in \mathcal{G}$, on a*
$$\mathbb{E}[X \cdot \mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot \mathbb{1}_B]$$
3. *Si Y est $\mathcal{G}$-mesurable, alors*
$$\mathbb{E}[XY|\mathcal{G}] = Y \cdot \mathbb{E}[X|\mathcal{G}]$$
4. *Si $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, alors*
$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$$
5. *Si X est indépendant de $\mathcal{G}$, alors*
$$\mathbb{E}[X|G] = \mathbb{E}[X]$$
6. *L'inégalité de Hölder :*
$$\mathbb{E}[XY|\mathcal{G}] \leq \mathbb{E}[X^p|\mathcal{G}]^{1/p}\, \mathbb{E}[Y^q|\mathcal{G}]^{1/q}$$
pour tout $1 \leq p, q \leq +\infty$ satisfaisant $\frac{1}{p} + \frac{1}{q} = 1$.
7. *Jensen : Pour $\varphi : \mathbb{R} \to \mathbb{R}$ convexe tel que $\varphi(X) \in L^1$*
$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}])$$

À compléterles p.s.

Chapitre 1

Martingales

Notation. Dans ce cours, on a $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

1.1 Définition et exemples

Un **processus aléatoire** (PA) est une suite de variables alétoires à valeurs réelles :

$$(X_n)_{n\in\mathbb{N}}, \mathbb{N} = \{0, 1, 2, ...\}$$

Définition 1.1. Une **filtration** $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une suite croissante de sous-tribus de $\mathcal{F}$:

$$F_0 \subset F_1 \subset F_2 \subset ... \subset \mathcal{F}$$

Exemple 1.2.

1. $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire. On définit

$$\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$$

   la **fibration canonique** (ou **naturelle**) des processus aléatoires $(X_n)_{n\in\mathbb{N}}$

2. $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1))$. Posons

$$\mathcal{F}_n := \sigma([\frac{i-1}{2^n}, \frac{i}{2^n}) : i = 1, ..., 2^n)$$

   la **tribu dyadique** de $[0, 1)$.

Définition 1.3. Un processus aléatoire $(X_n)_{n\in\mathbb{N}}$ est **adapté** à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si $\forall n \in \mathbb{N}$, X_n est $\mathcal{F}_n$-mesurable

Dorénavant, on sous-entend une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$.

Définition 1.4. Soit $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire adapté, avec $X_n \in L^1, \forall n \in \mathbb{N}$. Alors $(X_n)_{n\in\mathbb{N}}$ est :

— **une martingale** si

$$\forall n \in \mathbb{N}, \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \ p.s.$$

— **une sur-martingale** si

$$\forall n \in \mathbb{N}, \mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n \ p.s.$$

— **une sous-martingale** si

$$\forall n \in \mathbb{N}, \mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n \ p.s.$$

Remarque.

1. (X_n) sur-martingale $\iff$ $(-X_n)$ sous-martingale.
2. Pour une martingale (X_n) on a, pour $0 \leq n \leq m$, $\mathbb{E}[X_m|\mathcal{F}_n] = X_n$. En effet, Pour $m \geq n+2$, on a $\mathcal{F}_n \subset F_{m-1}$, donc

$$\mathbb{E}[X_m|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X_m|\mathcal{F}_{m-1}]|\mathcal{F}_n] = \mathbb{E}[X_{m-1}|\mathcal{F}_n] = ... = \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$$

Interprétation : Un martingale est un jeu équitable, une sur-martingale est une jeu défavorable.

Exemple 1.5. $X \in L^1$, on pose $X_n := \mathbb{E}[X|\mathcal{F}_n]$. Alors $(X_n)_{n\in\mathbb{N}}$ est une martingale (exercice), appelée *martingale fermée* ou *martingale de Doob.*

Exemple 1.6 (marche aléatoire)**.** Pour $(X_n)_{n\geq 1}$ i.i.d. et $x \in \mathbb{R}$ déterministe. On pose

$$S_n := x + X_1 + ... + X_n, n \geq 0, \mathcal{F}_n := \sigma(X_1, ..., X_n)$$

Alors $(S_n)_{n\in\mathbb{N}}$ est adapté à $(\mathcal{F}_n)$ et

$$\begin{aligned}\mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}[S_n + X_{n+1}|\mathcal{F}_n] \\ &= \mathbb{E}[S_n|\mathcal{F}_n] + \mathbb{E}[X_{n+1}|\mathcal{F}_n] \\ &= S_n + \mathbb{E}[X_{n+1}]\end{aligned}$$

Donc $(S_n)_{n\in\mathbb{N}}$ est une martingale ssi $\mathbb{E}[X_1] = 0$ (sous-martingale ssi $\mathbb{E}[X_1] \geq 0$).

Exemple 1.7. (carré d'une marche aléatoire) Comme dans l'exemple précédent, $\mathbb{E}[X_1] = 0, \mathbb{E}[X_1^2] = \sigma^2 < +\infty$. Soit $M_n := S_n^2 - n\sigma^2$. Alors $(M_n)_{n\in\mathbb{N}}$ est une martingale :

$$\begin{aligned}\mathbb{E}[M_{n+1}|\mathcal{F}_n] - M_n &= \mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n] \\ &= \mathbb{E}[(S_n + X_n + 1)^2 - S_n^2 - \sigma^2|\mathcal{F}_n] \\ &= \mathbb{E}[2S_nX_{n+1} + X_{n+1}^2 - \sigma^2|\mathcal{F}_n] = 2S_n\,\mathbb{E}[X_{n+1}] + \mathbb{E}[X_{n+1}^2] - \sigma^2 = 0\end{aligned}$$

Cours 2

Exemple 1.8 (Marche aléatoire asymétrique)**.** $(X_n)_{n\geq 1}$ i.i.d., $\mathbb{P}[X_n = 1] = p, \mathbb{P}[X_n = -1] = 1-p, 0 < p < 1$. On pose

$$S_n := X_1 + X_2 + ... + X_n, (S_0 = 0)$$

Posons $M_n := \left(\frac{1-p}{p}\right)^{S_n}$. Alors $(M_n)_{n\in\mathbb{N}}$ est une martingale :

$$\begin{aligned}\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}\left[M_n \cdot \left(\frac{1-p}{p}\right)^{X_{n+1}} |\mathcal{F}_n\right] \\ &= M_n\,\mathbb{E}\left[\left(\frac{1-p}{p}\right)^{X_{n+1}} |\mathcal{F}_n\right] \\ &= M_n\underbrace{\mathbb{E}\left[\left(\frac{1-p}{p}\right)^{X_{n+1}}\right]}_{=1} = M_n\end{aligned}$$

On note que si $p \neq \frac{1}{2}$ on a que $M_n \to 0$ p.s. (par $(*)$) mais $\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1$

1.2 Propriétés de base

Proposition 1.9. *Soit $\varphi : \mathbb{R} \to \mathbb{R}$ convexe et $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire tel que $\varphi(X_n) \in L^1, \forall n$.*

1. *Si (X_n) est un martingale, alors $(\varphi(X_n))$ est une sous martingale.*
2. *Si (X_n) est une sous-martingale et φ est croissante, alors $(\varphi(X_n))$ est une sous-martingale.*

Démonstration. 1. $\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \overset{\text{Jensen}}{\geq} \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) = \varphi(X_n)$

2. $\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \overset{\text{Jensen}}{\geq} \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) \overset{\varphi \text{ croissante}}{\geq} \varphi(X_n)$

□

Exemple 1.10. Si (X_n) est une martingale, alors $(X_n \vee a)$ sous-martingale

Définition 1.11. $(X_n)_{n\geq 1}$ est **prévisible** si $\forall n \geq 1, X_n$ est $\mathcal{F}_{n-1}$-mesurable.

Proposition 1.12 (Décomposition de Doob)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire adapté tel que $X_n \in L^1, \forall n$. Alors $X_n = M_n + A_n$, où (M_n) est une martingale et (A_n) est un processus aléatoire prévisible dans L^1 avec $A_0 = 0$. Cette décomposition est unique.*

Démonstration. *Unicité :* Soient M_n et A_n comme dans l'énoncé. Alors :

$$\begin{aligned}\mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n &= \mathbb{E}[X_{n+1} - X_n|\mathcal{F}_n] \\ &= \underbrace{\mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n]}_{=0 \text{car martingale}} + \underbrace{\mathbb{E}[A_{n+1} - A_n|\mathcal{F}_n]}_{=A_{n+1}-A_n \text{ car prévisible}} = A_{n+1} - A_n\end{aligned}$$

On conclut que forcément

$$\begin{cases} A_{n+1} = A_n + \mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n, A_0 = 0 \\ M_n = X_n - A_n \end{cases} \qquad (*)$$

Existence : Définissons A_n et M_n par $(*)$. Alors $A_n, M_n \in L^1$, A_n est par construction prévisible, et M_n est une martingale, car

$$\begin{aligned}\mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n] &= \mathbb{E}[X_{n+1} - X_n - (A_{n+1} - A_n)|\mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n - (A_{n+1} - A_n) = 0\end{aligned}$$

□

Corollaire 1.13. *Si (X_n) est en plus une sous-martingale, alors (A_n) est p.s. croissante.*

Démonstration. Suit de $(*)$ □

Disons que je joue à un jeu de hasard. X_n est ma fortune au temps n. Si le jeu est équitable, (X_n) est une martingale. Dans la partie au temps $n-1$, je gagne $X_n - X_{n-1}$

Je peux modifier la mise en multipliant par H_n : je gagne $H_n(X_n - X_{n-1})$. H_n doit être $\mathcal{F}_{n-1}$-mesurable.

Définition 1.14. Pour des processus aléatoires $(X_n)_{n\geq 0}$ et $(H_n)_{n\geq 1}$, on définit un processus aléatoire $((H \cdot X)_n)_{n\geq 0}$ par

$$(H \cdot X)_n := \sum_{k=1}^{n} H_k(X_k - X_{k-1})$$

et $(H \cdot X)_0 = 0$.

(Ceci est l'analogue de l'intégrale stochastique $\int H dX$.)

Exemple 1.15 ("La martingale")**.** Soit $\xi_1, \xi_2, ...$ i.i.d. $\mathbb{P}[\xi_1 = 1] = \mathbb{P}[\xi_1 = -1] = \frac{1}{2}$, $X_n = \xi_1 + ... + \xi_n$.

Définissons $H_1 = 1$, $H_n = \begin{cases} 2H_{n-1} & \text{si } \xi_{n-1} = -1 \\ 1 & \text{si } \xi_{n-1} = 1 \end{cases}$

Stratégie de la martingale : $((H \cdot X)_n)$

(Je double ma mise tant que je pers, je recommence avec une mise de 1 quand je gagne.)

Proposition 1.16 (Un jeu est imbattable)**.** *Soit (H_n) prévisible et borné.*

1. *Si (X_n) est un martingale, alors $((H \cdot X)_n)$ est une martingale.*
2. *Si (X_n) est un sous-martingale et $H_n \geq 0, \forall n$, alors $((H \cdot X)_n)$ est une sous-martingale.*

Démonstration. $(H \cdot X)_n \in L^1$ par hypothèse, et $(H \cdot X)_n$ est $\mathcal{F}_n$-mesurable. On trouve :

$$\begin{aligned}\mathbb{E}[(H \cdot X)_{n+1} - (H \cdot X)_n|\mathcal{F}_n] &\overset{\text{Déf.13}}{=} \mathbb{E}[H_{n+1}(X_{n+1} - X_n)|\mathcal{F}_n] \\ &= H_{n+1}\,\mathbb{E}[X_{n+1} - X_n|\mathcal{F}_n]\end{aligned}$$

Les deux cas découlent immédiatement. □

1.3 Temps d'arrêt

Cours 3

Définition 1.17. Une variable aléatoire $T \in \bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ est appelé **temps d'arrêt** (TA) (de la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$) si $\forall n \in \mathbb{N}, \{T = n\} \in \mathcal{F}_n$

Remarque 1.18. 1. T est un temps d'arrêt si et seulement si $\forall n \in \mathbb{N}, \{T \leq n\} \in \mathcal{F}_n$

Démonstration. $\{T \leq n\} = \bigcup_{k\leq n}\{T = k\}$ et $\{T = n\} = \{T \leq n\} \backslash \{T \leq n-1\}$ □

2.

$$\{T = \infty\} = \{T < \infty\}^c = (\bigcup_{n\in\mathbb{N}}\{T = n\})^c \in \mathcal{F}_\infty := \sigma(\bigcup_{n\in\mathbb{N}}\mathcal{F}_n)$$

Interprétation : Un temps d'arrêt est un temps aléatoire auquel je peux décider d'arrêter de jouer, en me basant uniquement sur l'information que j'ai (sans voir dans le futur !). Décision d'arrêter de jouer au temps $\{T = n\}$, mesurable par rapport à $\mathcal{F}_n$, l'information dont je dispose à l'instant n.

Exemple 1.19. 1. $T = k$ pour un $k \in \mathbb{N}$ fixé est un temps d'arrêt

2. (X_n) adapté, $A \in \mathcal{B}(\mathbb{R})$. Alors

$$T_A := \inf\{n \in \mathbb{N} : X_n \in A\}$$

(première visite en A)

Démonstration.

$$\{T_A = n\} = \{X_0 \notin A, X_1 \notin A, ..., X_{n-1} \notin A, X_n \in A\} \in \mathcal{F}_n$$

□

3. Fixons $N > 0$ est posons

$$L_A := \sup\{n \leq N : X_n \in A\}$$

n'est en général par un temps d'arrêt. (dernière visite en A).

Démonstration.

$$\{L_A = n\} = \{X_n \in A, X_{n+1} \notin A, ..., X_N \notin A\} \overset{\text{Général}}{\notin} \mathcal{F}_n\}$$

□

Proposition 1.20. *1. Soient S et T des temps d'arrêt. Alors $S \vee T$ et $S \wedge T$ sont des temps d'arrêt.*

2. $(T_k)_{k\in\mathbb{N}}$ une suite de temps d'arrêt. Alors $\inf T_k, \sup T_k, \liminf T_k, \limsup T_k$ des temps d'arrêt.

Démonstration. En exercice. □

Exemple 1.21. Soit T un temps d'arrêt et $H_n := \mathbb{1}_{n\leq T}$. (On parie un CHF jusqu'à l'instant T). Alors (H_n) est prévisible, car $\{H_n = 0\} = \{n > T\} = \{T \leq n-1\} \in \mathcal{F}_{n-1}$.

Soit (X_n) une sous-martingale. Par la proposition 1.16, $(H \cdot X)_n$ est une sous-martingale. On a

$$\begin{aligned}(H \cdot X)_n &= \sum_{k=1}^{n} \mathbb{1}_{k\leq T}(X_k - X_{k-1}) \\ &= \sum_{k\geq 1} \mathbb{1}_{k\leq T\wedge n}(X_k - X_{k-1}) \\ &= X_{T\wedge n} - X_0\end{aligned}$$

On conclut :

Proposition 1.22. *Si (X_n) est une sous-martingale et T un temps d'arrêt, alors $(X_{T\wedge n})$ est aussi une sous-martingale.*

— En particulier, pour tout $n \in \mathbb{N}$, on a $\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0]$

— En prenant $n \to \infty$, si $T < \infty$ p.s., on pourrait donc penser que $\mathbb{E}[X_T] = \mathbb{E}[X_0]$. Ceci est faux en général !

Exemple 1.23. Soit $X_n = \xi_1 + ... + \xi_n$, $\xi_1, \xi_2, ...$ i.i.d. avec $\mathbb{P}[\xi_1 = +1] = \mathbb{P}[\xi_1 = -1] = \frac{1}{2}$ (marche aléatoire simple). (X_n) est une martingale, $T := \inf\{n : X_n = 1\}$.

Alors $T < \infty$ p.s. (Cf. Chapitre 2), mais $\mathbb{E}[X_T] = 1 \neq 0 = \mathbb{E}[X_0]$

Proposition 1.24. *Soit (X_n) une sur-martingale et T un temps d'arrêt. Alors $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$ si une des conditions suivantes est vraie pour une constante $c > 0$.*

1. *$T \leq c$ p.s*
2. *$T < \infty$ p.s. et $|X_n| \leq c, \forall n$*
3. *$\mathbb{E}[T] < \infty$ et $|X_n - X_{n-1}| \leq c, \forall n \geq 1$*

Démonstration. Par Prop 22 À compléter, $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0], \forall n \in \mathbb{N}$. Dans les trois cas, on a $T < \infty$ p.s. et donc $X_{T\wedge n} \overset{n\to\infty}{\longrightarrow} X_T$ p.s.

1. choisir $n \geq c$ fixé.
2. Convergence dominée.
3. Convergence dominée, par

$$|X_{T\wedge n} - X_0| \leq \sum_{k=1}^{T\wedge n} \underbrace{|X_k - X_{k-1}|}_{\leq c} \leq cT \in L^1$$

Donc $|X_{T\wedge n}| \leq |X_0| + cT \in L^1$

□

1.4 Convergence presque sûre des martingales

Définition 1.25. Soient $\alpha = (\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ et $a < b$.

DESSIN À compléter

On définit $(S_k(\alpha))_{k\geq 1}$ et $(T_k(\alpha))_{k\geq 1}$ dans $\bar{\mathbb{N}}$ comme suit :

$$\begin{aligned} S_1(\alpha) &:= \inf\{n \geq 0 : \alpha_n \leq a\} \\ T_k(\alpha) &:= \inf\{n \geq S_k(\alpha) : \alpha_n \geq b\} \\ S_{k+1}(\alpha) &:= \inf\{n \geq T_k(\alpha) : \alpha_n \leq a\} \end{aligned}$$

Pour $n \leq 0$, on définit

$$\begin{aligned} N_n([a,b],\alpha) &:= \sum_{k=1}^{\infty} \mathbb{1}_{T_k(\alpha)\leq n} \\ N_\infty([a,b],\alpha) &:= \sum_{k=1}^{\infty} \mathbb{1}_{T_k(\alpha)<\infty} \end{aligned}$$

Le nombre de montées le long de l'intervalle [a,b].

Cours 4 Si $X = (X_n)_{n\in\mathbb{N}}$ est un processus adapté, alors $S_k(X)$ et $T_k(X)$ sont des temps d'arrêt pour tout $k \leq 1$.

Démonstration. Par exemple,

$$\begin{aligned} \{T_k(X) \leq n\} &= \bigcup_{0\leq m_1<n_1<...<m_k<n_k\leq n} \{X_{m_1} \leq a, X_{n_1} \geq b, ..., X_{m_k} \leq a, X_{n_k} \geq b\} \\ &= \bigcup_{0\leq m_1<n_1<...<m_k<n_k\leq n} (\bigcap_{i=1}^{k}\{X_{m_i} \leq a\} \cap \bigcap_{i=1}^{k}\{X_{n_i} \geq b\}) \in \mathcal{F}_n \end{aligned}$$

□

En particulier, $N_n([a,b],X)$ est $\mathcal{F}_n$-mesurable.

Notation. $X^+ = X \vee 0$, $X^- = (-X)^+$, $X = X^+ - X^-$

Proposition 1.26 (Inégalité des nombres de montées de Doob)**.** *Soit X une sous-martingale. Alors pour tout $a < b$ et $n \in \mathbb{N}$ on a*

$$(b-a)\,\mathbb{E}[N_n([a,b],X)] \leq \mathbb{E}[(X_n - a)^+ - (X_0 - a)^+]$$

Démonstration. Soit $Y_n := (X_n - a)^+$. Par la proposition 1.9.2 (cf. Exemple 1.10), (Y_n) est une sous-martingale. On abrège $N_n \equiv N_n([a,b],X)$, $S_k \equiv S_k(X)$, $T_k \equiv T_k(X)$. La famille ("Je me trouve dans une montée.) :

$$H_n := \sum_{k=1}^{\infty} \mathbb{1}_{S_k < n \leq T_k}$$

est dans $\{0,1\}$ et prévisible, parce que

$$\{S_k < n \leq T_k\} = \{S_k \leq n-1\} \cap \{n \leq T_k\} = \{S_k \leq n-1\} \backslash \{T_k \leq n-1\} \in \mathcal{F}_{n-1}$$

Donc $H \cdot Y$ est une sous-martingale, par la proposition 16.2. On trouve

$$\begin{aligned}
(H \cdot Y)_n &= \sum_{m=1}^{n} H_m (Y_m - Y_{m-1}) \\
&= \sum_{k=1}^{N_n} (Y_{T_k} - Y_{S_k}) + \mathbb{1}_{S_{N_n+1} < n} (\underset{\geq 0}{Y_n} - \underset{=0}{Y_{S_{N_n+1}}}) \\
&\geq \sum_{k=1}^{N_n} (\underset{\geq b-a}{Y_{T_k}} - \underset{=0}{Y_{S_k}}) \geq (b-a) N_n
\end{aligned}$$

Avec $K_n := 1 - H_n$ (prévisible et positive), on a que $K \cdot Y$ est une sous-martingale par la proposition 1.16.. Donc $\mathbb{E}[(K \cdot Y)_n] \geq \mathbb{E}[(K \cdot Y)_0] = 0$

Puisque $(K \cdot Y)_n + (H \cdot Y)_n = ((K+H) \cdot Y)_n = Y_n - Y_0$, on conclut que

$$(b-a)\,\mathbb{E}[N_n] \leq \mathbb{E}[(H \cdot Y)_n] \leq \mathbb{E}[(H \cdot Y)_n + (K \cdot Y)_n] = \mathbb{E}[Y_n - Y_0]$$

□

Proposition 1.27. *Soit (X_n) une sous-martingale telle que* $\sup_{n \in \mathbb{N}} \mathbb{E}[(X_n)^+] < \infty$*. Alors $\exists X_\infty \in L^1$ telle que $X_n \to X_\infty$ p.s.*

Remarque. L'hypothèse est équivalente à $\sup_{n \in \mathbb{N}} \mathbb{E}[|X_n|] < \infty$.

Démonstration.

$$\begin{aligned}
\mathbb{E}[(X_n)^-] &= -\,\mathbb{E}[X_n] + \mathbb{E}[(X_n)^+] \\
&= -\,\mathbb{E}[X_0] + \mathbb{E}[(X_n)^+]
\end{aligned}$$

□

Preuve de la proposition. Soient $a, b \in \mathbb{Q}$ tels que $a < b$. Par la proposition 26,

$$(b-a)\,\mathbb{E}[N_n([a,b],X)] \leq \mathbb{E}[(X_n - a)^+] \leq |a| + \mathbb{E}[(X_n)^+] \leq |a| + \sup_{k \in \mathbb{N}} \mathbb{E}[(X_k)^+]$$

Donc, par convergence monotone, $(b-a)\,\mathbb{E}[N_\infty([a,b],X)] < \infty$ et donc $N_\infty([a,b],X) < \infty$ p.s.. En prenant l'union sur $a, b \in \mathbb{Q}, a < b$, on conclut que p.s. $\forall a < b$ dans $\mathbb{Q}$, $N_\infty([a,b],X) < \infty$. On conclut que

X_n converge dans $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. (Exercice d'Analyse I, Série 3). Soit X_∞ la limite p.s. de X. Ensuite,

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf_{n\to\infty} |X_n|] \leq \liminf_{n\to\infty} \mathbb{E}[|X_n|] \leq \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] \underset{\text{remarque}}{<} \infty$$

□

Corollaire 1.28. *Soit (X_n) une sur-martingale positive. Alors X_n converge p.s. vers une variable aléatoire X_∞ qui est dans L^1 et satisfait $X_n \geq \mathbb{E}[X_\infty|\mathcal{F}_n], \forall n \in \mathbb{N}$*

Démonstration. Applique la proposition 27 à $\tilde{X}_n = -X_n$. La dernière assertion suit de

$$X_n \underset{\text{sur-martingale}}{\geq} \liminf_{m\to\infty} \mathbb{E}[X_m|\mathcal{F}_n] \underset{\text{Fatou}}{\geq} \mathbb{E}[X_\infty|\mathcal{F}_n]$$

□

Exemple 1.29. $Y_n = 1 + \xi_1 + \xi_2 + ... + \xi_n$ marche aléatoire simple (ξ_n i.i.d., $\mathbb{P}[\xi_n = 1] = \mathbb{P}[\xi_n = -1] = \frac{1}{2}$). Soit le temps d'arrêt $T := \inf\{n \in \mathbb{N} : Y_n = 0\}$. Par la proposition 22, $X_n := Y_{n\wedge T}$ est une martingale positive.

Par le Corollaire 28, X_n converge p.s. vers X_∞ satisfaisant $X_\infty < \infty$ p.s. (parce que dans L^1).

Mais si $T = \infty$ on a $|X_{n+1} - X_n| = |Y_{n+1} - Y_n| = 1$ ce qui est impossible p.s. parce que $X_n \to X_\infty$ p.s. avec $X_\infty \in \mathbb{R}$. Donc $T < \infty$ p.s.

(C.f. lacune dans Exemple 23)

Dans ce cas on a $X_\infty = 0$ p.s. En particulier $X_n \to X_\infty \in L^1$ p.s., mais À compléter

Cours 5

Exemple 1.30 (Processus de branchement, arbre de Galton-Watson)**.** Soit μ une mesure de probabilité sur $\mathbb{N}$ telle que

$$m := \sum_{k\in\mathbb{N}} k\mu(k) < \infty$$

On exclut les cas triviaux $\mu = \delta_0$ et $\mu = \delta_1$. Soient $(\xi_{n,j})_{n,j\in\mathbb{N}}$ des variables aléatoires i.i.d. de loi μ. Pour un $l \in \mathbb{N}^*$ fixé on définit $(X_n)_{n\in\mathbb{N}} \in \mathbb{N}^\mathbb{N}$ par

$$X_0 := l$$
$$X_{n+1} := \sum_{j=1}^{X_n} \xi_{n,j}, \quad \forall n \in \mathbb{N}$$

Interprétation : X_n est le nombre d'individus d'une population en temps n, chaque individu produit ξ individus de loi$(\xi) = \mu$.

Alors $m^{-n}X_n$ est une martingale positive par rapport à

$$\mathcal{F}_0 := \{\varnothing, \Omega\}$$
$$\mathcal{F}_n := \sigma(\xi_{k,j} : k < n, j \in \mathbb{N}) \text{ si } n \geq 1$$

Démonstration. (X_n) est adapté et dans L^1 (par récurrence)

$$\begin{aligned}\mathbb{E}[X_{n+1}|\mathcal{F}_n] &= \mathbb{E}[\sum_{j\geq 1} \mathbb{1}_{j\leq X_n} \xi_{n,j}|\mathcal{F}_n] \\ &= \sum_{j\geq 1} \mathbb{1}_{j\leq X_n} \mathbb{E}[\xi_{n,j}|\mathcal{F}_n] \\ &= \sum_{j\geq 1} \mathbb{1}_{j\leq X_n} m = mX_n\end{aligned}$$

□

Par le Corollaire 28, $m^{-n}X_n \overset{p.s.}{\to} \mathbb{Z} \in L^1$

On distingue trois cas :

- $m < 1$ Puisque $Z < \infty$ p.s. et $X_n \in \mathbb{N}$, on déduit que p.s. $X_n = 0$ pour n assez grand (extinction presque sûre).
- $m = 1$ On a $X_n \to Z$ p.s., et donc

$$\begin{aligned}\mathbb{P}[Z > 0] &= \mathbb{P}[\exists N \in \mathbb{N}, p \geq 1 : \forall n \geq N, X_n = p] \\ &\leq \sum_{N \in \mathbb{N}} \sum_{p \geq 1} \mathbb{P}[\forall n \geq N, X_n = p]\end{aligned}$$

  On a

$$\begin{aligned}\mathbb{P}[\forall n \geq N, X_n = p] &\leq \mathbb{P}[X_N = p, X_{N+1}, ..., X_{N+k} = p] \\ &= \underbrace{\mathbb{P}[X_N = p]}_{\leq 1} \underbrace{\mathbb{P}[X_{N+1} = p | X_n = p]}_{\leq \beta < 1} ... \underbrace{\mathbb{P}[X_{N+k} = p | X_N = ... = X_{N+k-1} = p]}_{\leq \beta} \\ &\leq \beta^k \underset{k \to \infty}{\longrightarrow} 0\end{aligned}$$

  Donc $Z = 0$ p.s. (Extinction presque sûre).
- $m > 1$. On a $m^{-n} X_n \to Z$ p.s.. Si cette convergence était dans L^1, alors on aurait $\mathbb{E}[Z] = \mathbb{E}[X_0] = l$, et donc $\mathbb{P}[Z > 0] > 0$. Ceci est vrai si $\sum_{k \in \mathbb{N}} k \log(k) \mu(k) < \infty$ (Kesten - Stygum). Cf. aussi section 1.5. (à venir).

Proposition 1.31. *Soit (X_n) une martingale. Les assertions suivantes sont équivalentes :*

1. *$X_n \to X_\infty$ p.s. et dans L^1*
2. *$\exists$ une variable aléatoire $Z \in L^1$ telle que $X_n = \mathbb{E}[Z | \mathcal{F}_n], \forall n$*

Si ces conditions sont satisfaites, on peut prendre $Z = X_\infty$ (martingale fermée).

Démonstration. 1. $\Rightarrow$ 2. Pour $m > n$, on a $X_n = \mathbb{E}[X_m | \mathcal{F}_n]$.

$$X_m \xrightarrow{p.s \,\&\, L_1} X_\infty \implies \mathbb{E}[X_m | \mathcal{F}_n] \xrightarrow{p.s.} \mathbb{E}[X_\infty | \mathcal{F}_n], \forall n$$

En effet :

$$|\mathbb{E}[X_m | \mathcal{F}_n] - \mathbb{E}[X_\infty | \mathcal{F}_n]| \overset{\text{Jensen}}{\leq} \mathbb{E}[|X_m - X_\infty| | \mathcal{F}_n]$$

En prenant $\mathbb{E}[.]$, on trouve

$$\mathbb{E}[|\mathbb{E}[X_m | \mathcal{F}_n] - \mathbb{E}[X_\infty | \mathcal{F}_n]|] \leq \mathbb{E}[|X_m - X_\infty|] \underset{m \to \infty}{\longrightarrow} 0$$

donc en prenant $m \to \infty$ on obtient $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$.

2. $\Rightarrow$ 1. (X_n) est bornée dans L^1. ($\mathbb{E}[|X_n|] \leq \mathbb{E}[|Z|]$). Par la proposition 1.27, $\exists$ une variable aléatoire $X_\infty \in L^1$ telle que $X_n \xrightarrow{p.s.} X_\infty$.
Supposons d'abord que $|Z| \leq M$ p.s. pour un $M > 0$ fixé. Alors $|X_n| \leq M$ p.s. et par convergence dominée on a $X_n \xrightarrow{L^1} X_\infty$.
Dans le cas général, soit $\epsilon > 0$ et choisissons $M > 0$ tel que $\mathbb{E}[|Z| \mathbb{1}_{|Z| > M}] < \epsilon$. Alors $\forall n \in \mathbb{N}$,

$$\mathbb{E}\left[\left| \underbrace{X_n}_{= \mathbb{E}[Z | \mathcal{F}_n]} - \mathbb{E}[Z \mathbb{1}_{|Z| \leq M} | \mathcal{F}_n] \right|\right] = \mathbb{E}[|\mathbb{E}[Z \mathbb{1}_{|Z| > M} | \mathcal{F}_n]|] < \epsilon$$

De plus, par l'observation précédente, la martingale $\mathbb{E}[Z \mathbb{1}_{|Z| \leq M} | \mathcal{F}_n]$ converge dans L^1. Donc pour n_0 assez grand, $\forall n, m \geq n_0$,

$$\mathbb{E}\left[\left| \underbrace{\mathbb{E}[Z \mathbb{1}_{|Z| \leq M} | \mathcal{F}_n]}_{= Y_n} - \underbrace{\mathbb{E}[Z \mathbb{1}_{|Z| \leq M} | \mathcal{F}_m]}_{= Y_m} \right|\right] < \epsilon$$

On conclut que

$$\mathbb{E}[|X_n - X_m|] \leq \mathbb{E}[|X_n - Y_n|] + \mathbb{E}[|Y_n - Y_m|] + \mathbb{E}[|Y_m - X_m|] < 3\epsilon$$

Donc (X_n) est de Cauchy dans L^1

□

Corollaire 1.32. *Soit $Z \in L^1$. La martingale $X_n := \mathbb{E}[Z|\mathcal{F}_n]$ converge p.s. et dans L^1 vers $\mathbb{E}[Z|\mathcal{F}_\infty]$ où $\mathcal{F}_\infty = \sigma(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n)$*

Démonstration. Par la proposition 31, il reste seulement à montrer que $X_\infty := \lim_{n\to\infty} X_n = \mathbb{E}[Z|\mathcal{F}_\infty]$. Clairement, X_∞ est $\mathcal{F}_\infty$-mesurable car les X_n le sont. De plus, pour $n \leq m$ et $A \in \mathcal{F}_n$

$$\begin{aligned}
\mathbb{E}[Z\mathbb{1}_A] &= \mathbb{E}[\mathbb{E}[Z\mathbb{1}_A|\mathcal{F}_m]] \\
&= \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_m]\mathbb{1}_A] \\
&= \mathbb{E}[X_m\mathbb{1}_A] \overset{\text{Prop. 31}}{\underset{m\to\infty}{\longrightarrow}} \mathbb{E}[X_\infty\mathbb{1}_A] \\
\implies \mathbb{E}[Z\mathbb{1}_A] &= \mathbb{E}[X_\infty\mathbb{1}_A]
\end{aligned}$$

Donc $\forall A \in \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$, on a $\mathbb{E}[Z\mathbb{1}_A] = \mathbb{E}[X_\infty\mathbb{1}_A]$ $(*)$. Par le lemme des classes monotones, $(*)$ est vrai pour tout $A \in \mathcal{F}_\infty$. On conclut que $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty]$. □

Cours 6

1.5 Convergence dans L^p des martingales

Soit (X_n) une martingale avec $X_n \xrightarrow{p.s.} X_\infty$. On veut, pour arriver à $X_n \xrightarrow{L^p} X_\infty$, que $\mathbb{E}[|X_n - X_\infty|^p] \underset{n\to\infty}{\longrightarrow} 0$. On aimerait utiliser le théorème de convergence dominée. Cette section présente des inégalités permettant de borner nos martingales.

Proposition 1.33. *Soient (X_n) une sous-martingale et S, T deux temps d'arrêt bornés tels que $S \leq T$. Alors on a $\mathbb{E}[X_T] \geq \mathbb{E}[X_S]$.*

Démonstration. On pose

$$H_n := \mathbb{1}_{S<n} - \mathbb{1}_{T<n} = \mathbb{1}_{S<n\leq T} \geq 0$$

H_n est prévisible. Par la proposition 16, $((H \cdot X)_n)$ est une sous-martingale et on a $\mathbb{E}[(H \cdot X)_n] \geq 0$.

Pour $N \in \mathbb{N}$ tel que p.s. $S \leq T \leq N$, on a $(H \cdot X)_N = X_T - X_S$. Comme $\mathbb{E}[(H \cdot X)_n] \geq 0$, on obtient le résultat. □

Notation. On note $\tilde{X}_n := \sup_{0\leq k\leq n} X_k$.

Proposition 1.34 (Inégalité maximale de Doob)**.** *Soit (X_n) une sous-martingale. $\forall a > 0, \forall n \in \mathbb{N}$,*

$$a\,\mathbb{P}[\tilde{X}_n \geq a] \leq \mathbb{E}[X_n\mathbb{1}_{\tilde{X}_n\geq a}] \leq \mathbb{E}[(X_n)^+]$$

Démonstration. Posons l'évènement $A := \{\tilde{X}_n \geq a\}$ et le temps d'arrêt $T := \inf\{k : X_k \geq a\}$. On a $A = \{T \leq n\}$. En appliquant la proposition 33 à $T \wedge n$ et n, on obtient :

$$a\,\mathbb{P}[A] + \mathbb{E}[X_n\mathbb{1}_{A^c}] \leq \mathbb{E}[X_T\mathbb{1}_{T\leq n}] + \mathbb{E}[X_n\mathbb{1}_{T>n}] = \mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_n]$$

En soustrayant $\mathbb{E}[X_n\mathbb{1}_{A^c}]$ des deux cotés, on obtient la première inégalité.

La seconde inégalité est immédiate. □

Proposition 1.35. *Soit $p > 1$. Soit (X_n) une sous-martingale positive. ALors,*

$$\mathbb{E}[\tilde{X}_n^p] \leq (\frac{p}{p-1})^p\,\mathbb{E}[X_n^p]$$

Remarque. Si $(Y_n)_n$ est une martingale et $Y_n^* := \sup_{k\leq n} |Y_k|$, alors

$$\mathbb{E}[(Y_n^*)^p] \leq (\frac{p}{p-1})^p\,\mathbb{E}[|Y_n|^p]$$

En effet, $x \mapsto |x|$ est convexe et positive, donc par l'inégalité de Jensen, $(|Y_n|)$ est une sous-martingale positive. On peut lui appliquer la proposition 35.

Démonstration. C'est facile si $\mathbb{E}[X_n^p] = \infty$. On suppose donc $\mathbb{E}[X_n^p] < \infty$.

$$\forall k \leq n, \mathbb{E}[X_k^p] \overset{X_n \text{ sous-martingale}}{\leq} \mathbb{E}[\mathbb{E}[X_n|\mathcal{F}_k]^p] \overset{x\mapsto x^p \text{ convexe}}{\leq} \mathbb{E}[\mathbb{E}[X_n^p|\mathcal{F}_k]] = \mathbb{E}[X_n^p]$$

Donc, $\mathbb{E}[\tilde{X}_n^p] < \infty$.

Par la proposition 34, on a $\forall a > 0$,

$$a\,\mathbb{P}[\tilde{X}_n \geq a] \leq \mathbb{E}[X_n \mathbb{1}_{\tilde{X}_n \geq a}]$$

donc $\forall a > 0$,

$$a^{p-1}\,\mathbb{P}[\tilde{X}_n \geq a] \leq a^{p-2}\,\mathbb{E}[X_n \mathbb{1}_{\tilde{X}_n \geq a}]$$

On intègre cette inégalité :

$$\int_0^\infty a^{p-1}\,\mathbb{P}[\tilde{X}_n \geq a]\ da \leq \int_0^\infty a^{p-2}\,\mathbb{E}[X_n \mathbb{1}_{\tilde{X}_n \geq a}]\ da$$

Par Tonelli (car positif) :

$$\begin{aligned}\int_0^\infty a^{p-1}\,\mathbb{P}[\tilde{X}_n \geq a]\ da &= \mathbb{E}[\int_0^\infty a^{p-1}\mathbb{1}_{\tilde{X}_n \geq a}\ da] \\ &= \mathbb{E}[\int_0^{\tilde{X}_n} a^{p-1}\ da] \\ &= \frac{1}{p}\,\mathbb{E}[\tilde{X}_n^p]\end{aligned}$$

Par Tonelli (car positif) :

$$\begin{aligned}\int_0^\infty a^{p-2}\,\mathbb{E}[X_n \mathbb{1}_{\tilde{X}_n \geq a}]\ da &= \mathbb{E}[X_n \int_0^{\tilde{X}_n} a^{p-2}\ da] \\ &= \frac{1}{p-1}\,\mathbb{E}[X_n \tilde{X}_n^{p-1}]\end{aligned}$$

Donc, $\mathbb{E}[\tilde{X}_n^p] \leq (\frac{p}{p-1})\,\mathbb{E}[X_n \tilde{X}_n^{p-1}]$

On écrit $\frac{1}{p} + \frac{p-1}{p} = 1$. Donc par Hölder :

$$\mathbb{E}[X_n \tilde{X}_n^{p-1}] \leq \mathbb{E}[X_n^p]^{\frac{1}{p}}\,\mathbb{E}[(\tilde{X}_n^{p-1})^{\frac{p-1}{p}}]^{\frac{p-1}{p}} = \mathbb{E}[X_n^p]^{\frac{1}{p}}\,\mathbb{E}[\tilde{X}_n^p]^{1-\frac{1}{p}}$$

Ainsi, comme $\mathbb{E}[\tilde{X}_n^p]^{\frac{1}{p}} \leq \frac{p}{p-1}\,\mathbb{E}[X_n^p]$ ☐

Théorème 1.36 (Convergence L^p des martingales)**.** *Soit (X_n) une martingale. Supposons qu'il existe $p > 1$ tel que $\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^p] < \infty$. Alors il existe $X_\infty \in L^p$ telle que $X_n \overset{p.s\ et\ L^p}{\longrightarrow} X_\infty$, avec*

$$\mathbb{E}[|X_\infty|^p] = \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^p]$$

et pour $X_\infty^* := \sup_{k\geq 0} |X_k|$, on a

$$\mathbb{E}[(X_\infty^*)^p] \leq (\frac{p}{p-1})^p \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^p] = (\frac{p}{p-1})^p\,\mathbb{E}[|X_\infty|^p]$$

Démonstration. (X_n) est bornée dans L^p, donc bornée dans L^1. Par la proposition 27 on a alors

$$\exists X_\infty \in L^1, X_n \overset{p.s.}{\longrightarrow} X_\infty$$

Par la proposition 35, on a

$$\mathbb{E}[(X_n^*)^p] \leq (\frac{p}{p-1})^p \underbrace{\sup \mathbb{E}[|X_n|^p]}_{\leq c < \infty}$$

Par le théorème de convergence dominée,

$$\mathbb{E}[(X_\infty^*)^p] = \lim \mathbb{E}[(X_n^*)^p] \leq (\frac{p}{p-1})^p c < \infty$$

Donc $X_\infty^* \in L^p$.

On a $|X_n| \leq X_\infty^*, \forall n \in \mathbb{N}$. Donc par le théorème de convergence dominée, $X_n \xrightarrow{L^p} X_\infty$

On a $\mathbb{E}[|X_n|^p] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[|X_\infty|^p]$. Or, $(\mathbb{E}[|X_n|^p])_n$ est croissante car $(|X_n|^p)$ est une sous-martingale. Donc

$$\mathbb{E}[|X_n|^p] \overset{n\to\infty}{\longrightarrow} \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^p]$$

□

Exemple 1.30 (suite) : $(\xi_{n,j})_{n,j}$ i.i.d. de loi $p \neq \delta_0, \delta_1$, avec $m := \sum k\mu(k) < \infty$.

À compléter(1 page)

1.6 Intégrabilité uniforme et martingales

Définition 1.37. $(X_i)_{i\in I} \in (L^1)^I$. On dit que $(X_i)_{i\in I}$ est **uniformément intégrable** (UI) si

$$\lim_{n\to\infty} \sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i|\geq a}] = 0$$

Remarque. 1. Si $(X_i)_{i\in I}$ est uniformément intégrable, alors $(X_i)_{i\in I}$ est bornée dans L^1 : $\exists a, \sup \mathbb{E}[|X_i|\mathbb{1}_{|X_i|\geq a}] \leq 1$. Alors $\forall i \in I, \mathbb{E}[|X_i|] = \mathbb{E}[|X_i|\mathbb{1}_{|X_i|<a}] + \mathbb{E}[|X_i|\mathbb{1}_{|X_i|\geq a}] \leq a+1$

2. La réciproque est fausse. (Exemple : À compléter)

Exemple 1.38. 1. Si $X \in L^1$, $\{X\}$ est uniformément intégrable.

2. Toute famille finie dans L^1 est uniformément intégrable.
3. Si (X_i) et (Y_i) sont uniformément intégrables, alors $(X_i + Y_i)$ l'est.
4. Si $Z \geq 0$, $Z \in L^1$, alors $\{X \in L^1, |X| \leq Z \text{ p.s.}\}$ est uniformément intégrable.
5. Si $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ est telle que $\frac{\phi(x)}{x} \to \infty$, Alors $\forall c \geq 0, \{X : \mathbb{E}[\phi(X)] \leq c\} =: \mathcal{A}$ est uniformément intégrable. En effet :

$$X \in \mathcal{A}, \mathbb{E}[|X|\mathbb{1}_{|X|\geq a}] = \mathbb{E}[\frac{|X|}{\phi(|X|)}\phi(|X|)\mathbb{1}_{|X|\geq a}] \leq \sup_{x\geq a} \frac{x}{\phi(x)} \mathbb{E}[\varphi(|X|)] \leq \sup_{x\geq a} \frac{x}{\phi(x)} c \underset{a\to\infty}{\longrightarrow} 0 \text{ unif.}$$

6. $p > 1$. Toute famille bornée dans L^p est uniformément intégrable.

Proposition 1.39. *Soit $(X_i)_{i\in I}$ bornée dans L^1. Les assertions suivantes sont équivalentes :*

1. *$(X_i)_{i\in I}$ est uniformément intégrable.*
2. *$\forall \epsilon > 0, \exists \delta > 0, \forall A \in \mathcal{F}$ tel que $\mathbb{P}[A] < \delta$, on a $\forall i \in I$, $\mathbb{E}[|X_i|\mathbb{1}_A] < \epsilon$.*

Démonstration. À compléter □

Cours 7

Corollaire 1.40. *Soit $X \in \mathrm{L}^1$. Alors $(\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \subset \mathcal{F}$ sous-tribu$)$ est uniformément intégrable.*

Démonstration. Soit $\epsilon > 0$. On choisit $\delta > 0$ t.q. $\forall A \in \mathcal{F}$ avec $\mathbb{P}[A] < \delta, \mathbb{E}[|X|\mathbb{1}_A] < \epsilon$ (Par la proposition 39 car (X) est uniformément intégrable). Soit $\mathcal{G} \subset \mathcal{F}$. Alors

$$\mathbb{P}[|\mathbb{E}[X|\mathcal{G}]| > a] \leq \frac{\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|]}{a} \leq \frac{\mathbb{E}[|X|]}{a} < \delta$$

Donc

$$\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|\mathbb{1}_{|\mathbb{E}[X|\mathcal{G}]|>a}] \leq \mathbb{E}[\mathbb{E}[|X||\mathcal{G}]\mathbb{1}_{|\mathbb{E}[X|\mathcal{G}]|>}a] = \mathbb{E}[|X|\mathbb{1}_{|\mathbb{E}[X|\mathcal{G}]|>a}] \leq \epsilon$$

Car $|\mathbb{E}[X|\mathcal{G}] > a|$ est $\mathcal{G}$-mesurable et $\mathbb{P}[|\mathbb{E}[X|\mathcal{G}] > a|] < \delta$. □

Proposition 1.41. *(X_n) dans L^1, convergeant en probabilité vers X_∞. Sont équivalents :*

1. *(X_n) converge en L^1 vers X_∞.*
2. *(X_n) est uniformément intégrable.*

Remarque. Soit (X_n) dans L^1 t.q. $X_n \overset{p.s.}{\to} X_\infty$.

— Par convergence dominée : $\exists Z \in L^1$ t.q. $|X_n| \leq Z, \forall n \implies X_n \overset{L^1}{\to} X_\infty$.

— Proposition 41 : (X_n) est uniformément intégrable si et seulement si $X_n \overset{L^1}{\to} X_\infty$. C'est un critère optimal (cf. exemple 38.4)

Démonstration. 1. $\Rightarrow$ 2. Clairement, (X_n) est bornée dans L^1. Soit $\epsilon > 0$. Soit N tel que

$$\forall n \geq N, \mathbb{E}[|X_n - X_\infty|] < \frac{\epsilon}{2}$$

$(X_n)_{0 \leq n \leq N}$ est UI (exemple 38.2). Alors

$$\exists \delta > 0, \forall A \text{ avec } \mathbb{P}[A] < \delta, \forall n \leq N, \mathbb{E}[|X_n| \mathbb{1}_A] < \frac{\epsilon}{2} \textbf{ (proposition 39)}$$

Pour $n > N$ on a (pour $\mathbb{P}[A] < \delta$)

$$\mathbb{E}[|X_n|] \leq \mathbb{E}[|X_N| \mathbb{1}_A] + \mathbb{E}[]$$

À compléter

2. $\Rightarrow$ 1. Par la proposition 39.2. $(X_n - X_m)_{n,m} \in \mathbb{N}$ est UI. Par 39,

$$\forall \epsilon > 0, \exists a > 0, \forall n, m \in \mathbb{N}, \mathbb{E}[|X_n - X_m| \mathbb{1}_{|X_n - X_m| > a}] < \epsilon$$

Donc, $\forall n, m \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}[|X_n - X_m|] = \underbrace{\mathbb{E}[|X_n - X_m| \mathbb{1}_{|X_n - X_m| \leq \epsilon}]}_{\leq \epsilon} \\
+ \underbrace{\mathbb{E}[|X_n - X_m| \mathbb{1}_[\epsilon < |X_n - X_m| \leq a]}_{\leq < \mathbb{P}[|X_n - X_m| > \epsilon} \\
+ \underbrace{\mathbb{E}[|X_n - X_m| \mathbb{1}_{|X_n - X_m| > a}]}_{< \epsilon} & \qquad \leq 2\epsilon + a\,\mathbb{P}[|X_n - X_m| > \epsilon]
\end{aligned}$$

(X_n) converge en probabilité, donc

$$\mathbb{P}[|X_n - X_m| > \epsilon] \leq \mathbb{P}[|X_n - X_\infty| > \epsilon/2] + \mathbb{P}[|X_m - X_\infty| > \epsilon/2] \to 0$$

pour $n, m \to \infty$. On conclut : $\limsup_{n,m \to \infty} \mathbb{E}[|X_n - X_m|] \leq 2\epsilon$. Donc (X_n) est de Cauchy dans L^1.

□

Remarque 1.42. Si (X_n) converge en probabilité et est bornée dans L^p pour un $p > 1$, alors (X_n) converge en $L^q, \forall q < p$.

En effet, il suffit de vérifier que $|X_n - X_\infty|^q$ est UI.

$$\mathbb{E}[|X_n - X_\infty|^q \mathbb{1}_{|X_n - X_\infty|^q > a^q}] \leq \mathbb{E}[\underbrace{|X_n - X_\infty|^p}_{\leq C}]^{\frac{q}{p}} \underbrace{\mathbb{P}[|X_n - X_\infty| > a]}_{\leq \frac{\mathbb{E}[|X_n - X_\infty|]}{a} \leq \frac{C'}{a}}^{1 - \frac{q}{p}}$$

Retour aux martingales :

Corollaire 1.43. *Soit (X_n) une martingale. Sont équivalentes :*

1. *(X_n) converge p.s. et dans L^1 vers X_∞.*
2. *(X_n) est UI.*
3. *(X_n) est fermée, c.à.d. $\exists Z \in L^1$ t.q. $X_n = \mathbb{E}[Z|\mathcal{F}_n], \forall n$*

On a aussi $Z = X_\infty$ p.s.

Démonstration. 1. $\Leftrightarrow$ 3. Proposition 31.

1. $\Rightarrow$ 2. Proposition 41.

2. $\Rightarrow$ 1. Par la proposition 27, $X_n \overset{p.s.}{\to} X_\infty$ et par la proposition 41, $X_n \overset{L^1}{\to} X_\infty$.

□

1.7 Théorèmes d'arrêt

Soit T un temps d'arrêt. Comment définir $\mathcal{F}_T$?

Exemple : $T =$ 3ème jour de pluie de l'année. $A = \{$quantité de pluie dans les 3 premiers jours de pluie $\geq$ 20 mm$\}$

Définition 1.44. Soit T un temps d'arrêt. On définit

$$\mathcal{F}_T := \{A \in \mathcal{F} : \forall n, A \cap \{T = n\} \in \mathcal{F}_n\}$$

Remarque 1.45.
1. $\mathcal{F}_T$ est une tribu.
2. Si $T = n$ p.s., alors $\mathcal{F}_T = \mathcal{F}_n$.

Démonstration. En exercice. □

Cours 8

Proposition 1.46. *Soit T et S des temps d'arrêt avec $S \leq T$ p.s.. Alors $\mathcal{F}_S \subset \mathcal{F}_T$.*

Démonstration. Soit $A \in \mathcal{F}_S$. Alors $\forall n \in \mathbb{N}$,

$$A \cap \{T = N\} = \bigcup_{k=0}^{n} \underbrace{(A \cap \{S = k\})}_{\in \mathcal{F}_k \subset \mathcal{F}_n} \cup \underbrace{\{T = n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n$$

□

Proposition 1.47. *Soit (X_n) un processus aléatoire adapté et T un temps d'arrêt. Si $\mathbb{P}[T = \infty] > 0$ on suppose que $X_n \to X_\infty$ p.s. et on pose*

$$X_T := \sum_{n \in \mathbb{N}} \mathbb{1}_{T=n} X_n + \mathbb{1}_{T=\infty} X_\infty$$

Alors X_T est $\mathcal{F}_T$-mesurable.

Démonstration. Soit $B \in \mathcal{B}(\mathbb{R})$. Alors $\forall n \in \mathbb{N}$,

$$\{\mathbb{1}_{T<\infty} X_T \in B\} \cap \{T = n\} = \{\underbrace{X_n}_{\in \mathcal{F}_n} \in B\} \cap \underbrace{\{T = n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n$$

donc $\{\mathbb{1}_{T<\infty} X_T \in B\} \in \mathcal{F}_T$, donc $\mathbb{1}_{T<\infty} X_T$ est $\mathcal{F}_T$-mesurable.

De plus, $\{\mathbb{1}_{T=\infty} X_\infty \in B\} \cap \{T = n\} = \begin{cases} \{T = n\} & \text{si } 0 \in B \\ \varnothing & \text{si } 0 \notin B \end{cases} \in \mathcal{F}_n$. Donc $\mathbb{1}_{T=\infty} X_\infty$ est $\mathcal{F}_T$-mesurable. □

Proposition 1.48. *Soit (X_n) une martingale uniformément intégrable, avec $X_\infty = \lim_{n\to\infty} X_n$. Soient $S \leq T$ des temps d'arrêt. Alors $X_S = \mathbb{E}[X_T|\mathcal{F}_S]$. En particulier, $X_S = \mathbb{E}[X_\infty|\mathcal{F}_S]$ et $\mathbb{E}[X_S] = \mathbb{E}[X_\infty] = \mathbb{E}[X_n], \forall n \in \mathbb{N}$.*

Remarque. Cela généralise la proposition 1.24.

Démonstration. Il suffit de montrer que

$$X_T = \mathbb{E}[X_\infty|\mathcal{F}_T] \quad (*)$$

En effet,

$$X_S \overset{(*)}{=} \mathbb{E}[X_\infty|\mathcal{F}_S] \overset{1.46}{=} \mathbb{E}[\mathbb{E}[X_\infty|\mathcal{F}_T]|\mathcal{F}_S] \overset{(*)}{=} \mathbb{E}[X_T|X_S]$$

Montrons d'abord que $X_T \in L^1$.

$$\begin{aligned}
\mathbb{E}[|X_T|] &= \sum_{n \in \mathbb{N}} \mathbb{E}[\mathbb{1}_{T=n}|X_n|] + \mathbb{E}[\mathbb{1}_{T=\infty}|X_n|] \\
&\overset{4.2}{=} \mathbb{E}[\mathbb{1}_{T=n}|\mathbb{E}[X_\infty|\mathcal{F}_n]|] \\
&\leq \mathbb{E}[\mathbb{1}_{T=n}\, \mathbb{E}[|X_\infty||\mathcal{F}_n]] \\
&= \mathbb{E}[\mathbb{1}_{T=n}|X_\infty|] \\
&\leq \mathbb{E}[|X_\infty|] < \infty
\end{aligned}$$

Pour $A \in \mathcal{F}_T$ on a

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_A X_T] &\overset{\text{Fubini}, X_T \in L^1}{=} \sum_{n \in \bar{\mathbb{N}}} \mathbb{E}[\mathbb{1}_{A \cap \{T=n\}} X_T] \\
&= \sum_{n \in \bar{\mathbb{N}}} \mathbb{E}[\mathbb{1}_{\underbrace{A \cap \{T=n\}}_{\in \mathcal{F}_n}} \underbrace{X_n}_{= \mathbb{E}[X_\infty|\mathcal{F}_n]}] \\
&= \sum_{n \in \bar{\mathbb{N}}} \mathbb{E}[\mathbb{1}_{A \cap \{T=n\}} X_\infty] \\
&\overset{\text{Fubini}, X_\infty \in L^1}{=} \mathbb{E}[\mathbb{1}_A X_\infty]
\end{aligned}$$

$\square$

Chapitre 2

Chaînes de Markov

Dans ce chapitre, E est un ensemble fini ou dénombrable, muni de la tribu $\mathcal{P}(E)$.

2.1 Définitions et premières propriétés

Définition 2.1. Une **matrice stochastique** sur E est une famille $Q = (Q(x,y) : x, y \in E) \in \mathbb{R}^{E\times E}$ satisfaisant

1. $0 \leq Q(x,y) \leq 1, \forall x, y \in E$
2. $\forall x \in E, \sum_{y\in E} Q(x,y) = 1$

Notation matricielle :

— On définit Q^n par récurrence :

$$Q^0(x,y) = \delta_{xy}$$
$$Q^{n+1}(x,y) = \sum_{z\in\mathcal{E}} Q^n(x,z)Q(z,y)$$

— Pour $f : E \to \mathbb{R}$ une fonction, on écrit :

$$Qf(x) := \sum_{y\in E} Q(x,y)f(y)$$

— Pour $\mu : E \to [0,1]$ une mesure de probabilité, on écrit :

$$\mu Q(x) := \sum_{y\in E} \mu(y)Q(y,x)$$

Définition 2.2. Soit Q une matrice stochastique sur E. Soit $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire à valeurs dans E. (X_n) est une **chaîne de Markov** (CM) de **matrice de transition** (MT) Q si $\forall n \in \mathbb{N}, \forall y \in E$

$$\mathbb{P}[X_{n+1} = y|X_0, X_1, ..., X_n] \overset{\text{propriété de Markov}}{=} \mathbb{P}[X_{n+1} = y|X_n] \overset{\text{homogénéité en } n}{=} Q(X_n, y)$$

De façon explicite :

$$\mathbb{P}[X_{n+1} = y|X_0 = x_0, X_1 = x_1, ..., X_n = x_n] = Q(x_n, y)$$

Pour tout $x_0, x_1, ..., x_n \in E$ tels que $\mathbb{P}[X_0 = x_0, X_1 = x_1, ..., X_n = x_n] > 0$

Remarque. 1. Le futur X_{n+1} dépend du passé $(X_0, X_1, ..., X_n)$ qu'à travers le présent X_n.

2. La fonction $Q(x,y) = \mathbb{P}[X_{n+1} = y|X_n = x]$ ne dépend pas de n. On parle de chaîne de Markov **homogène**. On peut considérer des chaînes de Markov inhomogènes, où Q dépend de n.

Proposition 2.3. *Soit $(X_n) \in E^{\mathbb{N}}$ un processus aléatoire. Les deux assertions suivantes sont équivalentes :*

1. *(X_n) est une chaîne de Markov de matrice de transition Q*
2. *$\forall n \in \mathbb{N}, \forall x_0, ..., x_n \in E$*

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, ..., X_n = x_n] = \mathbb{P}[X_0 = x_0]Q(x_0, x_1)Q(x_1, x_2)...Q(x_{n-1}, x_n)$$

Cours 9 *En particulier, si* $\mathbb{P}[X_0 = x_0] > 0, \mathbb{P}[X_n = x_n | X_0 = X_0] = Q^n(x_0, x_n)$.

Démonstration. $\Rightarrow$ Soit (X_n) une chaîne de Markov de matrice de transition Q. Alors

$$\begin{aligned}\mathbb{P}[X_0 = x_0, ..., X_n = x_n] &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{X_n = x_n} | X_0, ..., X_{n-1}] \mathbb{1}_{X_0 = x_0} ... \mathbb{1}_{X_{n-1} = x_{n-1}}] \\ &= \mathbb{E}[Q(x_{n-1}, x_n) \mathbb{1}_{X_0 = x_0} ... \mathbb{1}_{X_{n-1} = x_{n-1}}] \\ &= Q(x_{n-1}, x_n)\, \mathbb{P}[X_0 = x_0, ..., X_{n-1} = x_{n-1}]\end{aligned}$$

et on conclut par récurrence

$\Leftarrow$ Si $\mathbb{P}[X_0 = x_0, ..., X_n = x_n] > 0$ alors

$$\begin{aligned}\mathbb{P}[X_{n+1} = y | X_0 = x_0, ..., X_n = x_n] &= \frac{\mathbb{P}[X_{n+1} = y, X_n = x_n, ..., X_0 = x_0]}{\mathbb{P}[X_n = x_n, ..., X_0 = x_0]} \\ &= \frac{\mathbb{P}[X_0 = x_0] Q(x_0, x_1) Q(x_1, x_2) ... Q(x_{n-1}, x_n) Q(x_n, y)}{\mathbb{P}[X_0 = x_0] Q(x_0, x_1) Q(x_1, x_2) ... Q(x_{n-1}, x_n)} \\ &= Q(x_n, y)\end{aligned}$$

ce qui est la condition de la définition.

□

Remarque 2.4. La proposition précédente montre que la loi de (X_n) déterminée par Q et la loi de X_0.

Avec $\mu(x) = \mathbb{P}[X_0 = x]$, on a $\mathbb{E}[f(X_n)] = \mu Q^n f$ et $\mathbb{E}[f(X_{n+1}) | X_0, ..., X_n] = \mathbb{E}[f(X_{n+1}) | X_n] = Qf(X_n)$

Quelques exemples :

Exemple 2.5. $E = \{A, B\}$.

$$Q = \begin{pmatrix} Q(A,A) & Q(A,B) \\ Q(B,A) & Q(B,B) \end{pmatrix} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix},\ 0 \leq p, q \leq 1$$

À compléter

Exemple 2.6. Soit (X_n) une suite de variables aléatoires de même loi μ. Alors (X_n) est une chaîne de Markov avec $Q(x, y) = \mu(y)$.

Exemple 2.7 (Marche aléatoire)**.** Soit $d \in \mathbb{N}^*$, $\eta, \xi_1, \xi_2, ...$ des variables aléatoires indépendantes dans $\mathbb{Z}^d$, avec $\xi_1, \xi_2, ...$ ayant la même loi ν. Soit $X_n := \eta + \xi_1 + ... + \xi_n$. Alors (X_n) est une chaîne de Markov avec $Q(x, y) = \nu(y - x)$

Démonstration.

$$\begin{aligned}\mathbb{P}[X_{n+1} = y | X_0, ..., X_n] &= \mathbb{P}[\xi_{n+1} = y - X_n | X_0, ..., X_n] \\ &= \mathbb{P}[\xi_{n+1} = y - X_n | X_n] \\ &= \nu(y - X_n)\end{aligned}$$

□

Si $\nu(x) = \frac{1}{2d} \mathbb{1}_{|x|=1}$, (X_n) est la **marche aléatoire simple** sur $\mathbb{Z}^d$.

Plus généralement, si (V, E) est une graphe sur V, tel que le *degré* $D_x := |\{e \in E : x \in e\}|$ est fini $\forall x \in V$. Soit $Q(x, y) := \frac{\mathbb{1}_{\{x,y\} \in A}}{D_x}$ Alors la chaîne de Markov de matrice de transition Q est la **marche aléatoire simple** sur (V, E).

2.2 La chaîne de Markov canonique

Proposition 2.8. *Soit Q une matrice stochastique. Alors il existe $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ un espace de probabilité sur lequel $\forall x \in E, \exists (X_n^x)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice de transition Q avec $X_0^x = x$ p.s.*

Démonstration. Soit

- $\tilde{\Omega} = [0, 1)^{\mathbb{N}}$
- $\tilde{\mathcal{F}} =$ tribu produit
- $\tilde{\mathbb{P}} = Leb([0, 1))^{\otimes \mathbb{N}}$
- $\omega = (\omega_i)_{i \in \mathbb{N}^*} \in \tilde{\Omega}$

Soit $y_1, y_2, ...$ une énumération de E. Soit $x \in E$. On pose $X_0^x = x$, puis $X_1^x = y_k$ si $\sum_{j=1}^{k-1} Q(x, y_j) < \omega_1 \leq \sum_{j=1}^{k} Q(x, y_j)$, alors $\tilde{\mathbb{P}}[X_1^x = y] = Q(x, y) \forall y \in E$.

On continue par récurrence : $X_{n+1}^x = y_k$ si $\sum_{j=1}^{k-1} Q(X_n^x, y_j) < \omega_{n+1} \leq \sum_{j=1}^{k} Q(X_n^x, y_j)$

On trouve

$$\begin{aligned}
\tilde{\mathbb{P}}[X_{n+1}^x = y | X_0^x, ..., X_n^x] &= \tilde{\mathbb{P}}[\sum_{j=1}^{k-1} Q(X_n^x, y_j) < \omega_{n+1} \leq \sum_{j=1}^{k} Q(X_n^x, y_j) | X_0^x, ..., X_n^x] \\
&= Q(X_n^x, y_k)
\end{aligned}$$

car les (ω_i) sont indépendants $\square$

Il est très souvent utile de faire un choix canonique de l'espace de probabilité d'une chaîne de Markov.

- $\Omega = E^{\mathbb{N}}$, $\omega = (\omega_n)_{n \in \mathbb{N}}$
- $X_n(\omega) := \omega_n$
- $\mathcal{F} = \sigma(X_0, X_1, X_2, ...)$ tribu cylindrique (ou produit).

En d'autres termes, $\mathcal{F}$ est engendrée par les *cylindres* $\{\omega \in \Omega : \omega_0 = x_0, \omega_1 = x_1, ..., \omega_n = x_n\}, n \in \mathbb{N}, x_0, ..., x_n \in E$

Proposition 2.9. *Soit Q une matrice stochastique. Pour tout $x \in E$, il existe une unique $\mathcal{P}_x$ sur $(\Omega, \mathcal{F})$, telle que sous $\mathcal{P}_x$ la famille (X_n) est une chaîne de Markov de matrice de transition Q avec $X_0 = x$ p.s.*

Démonstration. Existence : Pour $x \in E$, soit $(\tilde{\Omega}, \tilde{F}, \tilde{\mathbb{P}})$ et $(X_n^x)_n$ comme dans la proposition 2.8. On définit $\mathbb{P}_x . = \psi_* \tilde{\mathbb{P}}$, où

$$\begin{array}{rccc}
\psi : & \tilde{\Omega} & \to & \Omega \\
& \tilde{\omega} & \mapsto & (X_n^*(\tilde{\omega}))_{n \in \mathbb{N}}
\end{array}$$

Alors

$$\begin{aligned}
\mathbb{P}_x[X_0 = x_0, ..., X_n = x_n] &= \tilde{\mathbb{P}}[X_0^x = x_0, ..., X_n^x = x_n] \\
&= \delta_{x x_0} Q(x_0, x_1) ... Q(x_{n-1}, x_n)
\end{aligned}$$

Unicité : Si $\mathbb{P}'_x$ est une autre mesure comme dans la proposition alors $\mathbb{P}_x$ et $\mathbb{P}'_x$ coincident sur les cylindres. On conclut par le lemme des classes monotones que $\mathbb{P}_x = \mathbb{P}'_x$ $\square$

Cours 10

Définition 2.10. Si μ est une mesure de probabilité sur E, on note

$$\mathbb{P}_\mu := \sum_{x \in E} \mu(x) \, \mathbb{P}_x$$

Alors, sous $\mathbb{P}_\mu$, (X_n) est une chaîne de Markov associée à Q et $\mathbb{P}_\mu[X_0 = x] = \mu(x)$. On note $\mathbb{E}_x$ et $\mathbb{E}_\mu$ les espérances de ces mesures de probabilité.

Définition 2.11. Pour $k \in \mathbb{N}$ on définit **l'opérateur de transition**

$$\begin{array}{rccc}
\Theta_k : & \Omega & \to & \Omega \\
& (\omega_n)_{n \in \mathbb{N}} & \mapsto & (\omega_{n+k})_{n \in \mathbb{N}}
\end{array}$$

On abrège $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$ (filtration canonique)

Proposition 2.12 (Propriété de Markov simple). *Soit G un variable aléatoire positive sur Ω et μ une mesure de probabilité sur E. Alors*

$$\forall n \in \mathbb{N}, \mathbb{E}_\mu[G \circ \Theta_n | \mathcal{F}_n] = \mathbb{E}_{X_n}[G]$$

(La loi conditionnelle de $\Theta_n(\omega)$ sachant $(X_0, ..., X_n)$ est $\mathbb{P}_{X_n}$)

Démonstration. Soit $A = \{X_0 = x_0, ..., X_n = x_n\} \in \mathcal{F}_n$ un cylindre et soit $G = G(\omega_0, ..., \omega_m)$ $\mathcal{F}_m$-mesurable. Alors

$$\begin{aligned}
\mathbb{E}_\mu[\mathbb{1}_A \cdot G \circ \Theta_n] &= \mathbb{E}_\mu[\mathbb{1}_{X_0=x_0} ... \mathbb{1}_{X_n=x_n} G(X_n, X_{n+1}, ..., X_{n+m})] \\
&= \mu(x_0)Q(x_0, x_1)...Q(x_{n-1}, x_n) \sum_{x_{n+1}, ..., x_{n+m}} Q(x_n, x_{n+1})...Q(x_{n+m-1}, x_{n+m})G(x_n, ..., x_{n+m}) \\
&= \mu(x_0)Q(x_0, x_1)...Q(x_{n-1}, x_n)\, \mathbb{E}_{x_n}[G] \\
&= \mathbb{E}_\mu[\mathbb{1}_A\, \mathbb{E}_{X_n}[G]]
\end{aligned}$$

On conclut que, pour tout cylindres A, B,

$$\mathbb{E}_\mu[\mathbb{1}_A \cdot \mathbb{1}_B \circ \Theta_n] = \mathbb{E}_\mu[\mathbb{1}_A\, \mathbb{E}_{X_n}[\mathbb{1}_B]] \quad (*)$$

Par une application du lemme des classes monotones, on trouve que $(*)$ est vraie pour tout $A \in \mathcal{F}_n, B \in \mathcal{F}$. Donc

$$\mathbb{E}_\mu[\mathbb{1}_B \circ \Theta_n | \mathcal{F}_n] = \mathbb{E}_{X_n}[\mathbb{1}_B]$$

et on conclut. □

Proposition 2.13 (Propriété de Markov forte). *Soit T un temps d'arrêt, G une variable aléatoire positive sur Ω et μ une mesure de probabilité sur E. Alors*

$$\mathbb{E}_\mu[\mathbb{1}_{T<\infty} G \circ \Theta_T | \mathcal{F}_T] = \mathbb{1}_{T<\infty}\, \mathbb{E}_{X_T}[G]$$

Remarque 2.14.
— $\{T < \infty\} \in \mathcal{F}_T$
— X_T, défini sur $\{T < \infty\}$, est $\mathcal{F}_T$-mesurable :

$$\{X_T \in \mathcal{B}\} \cap \{T = n\} = \{\underbrace{X_n \in B}_{\in \mathcal{F}_n}\} \cap \{\underbrace{T = n}_{\in \mathcal{F}_n}\} \in \mathcal{F}_n$$

Démonstration. Soit $F \geq 0$ F_T-mesurable et $n \in \mathbb{N}$. Alors

$$\begin{aligned}
\mathbb{E}_\mu[F \cdot \mathbb{1}_{T=n} G \circ \Theta_T] &= \mathbb{E}_\mu[F \cdot \mathbb{1}_{T=n} G \circ \Theta_n] \\
&= \mathbb{E}_\mu[F \mathbb{1}_{T=n}\, \mathbb{E}_\mu[G \circ \Theta_n | \mathcal{F}_n]] \\
&= \mathbb{E}_\mu[F \mathbb{1}_{T=n}\, \mathbb{E}_{X_n}[G]]
\end{aligned}$$

En sommant sur $n \in \mathbb{N}$, on obtient

$$\mathbb{E}_\mu[F \mathbb{1}_{T<\infty} G \circ \Theta_T] = \mathbb{E}_\mu[F \mathbb{1}_{T<\infty}\, \mathbb{E}_{X_T}[G]]$$

□

2.3 La classification des états

Notation. Pour $x \in E$ on écrit
— $H_x := \inf\{n \geq 1 : X_n = x\}$ (temps de la première visite / du premier retour à x)
— $N_x := \sum_{n \geq 0} \mathbb{1}_{X_n = x}$ (nombre de visites à x)

Proposition-Définition 2.15. *Soit $x \in E$*
— *Si $\mathbb{P}_x[H_x < \infty] = 1$, alors $N_x = \infty$ $\mathbb{P}_x$-p.s. On dit que x est* ***récurrent****.*
— *Si $\mathbb{P}_x[H_x < \infty] < 1$, alors $N_x < \infty$ $\mathbb{P}_x$-p.s. et même $\mathbb{E}_x[N_x] = \frac{1}{\mathbb{P}_x[H_x = \infty]} < \infty$. On dit que x est* ***transitoire****.*

Démonstration. Si $X_0 = x$ et $H_x < \infty$, alors

$$N_x = \sum_{n\geq 0} \mathbb{1}_{X_n=x} = 1 + \sum_{n\geq 1} \mathbb{1}_{X_n=x} = 1 + N_x \circ \Theta_{H_x}$$

Donc pour $k \geq 1$, on a

$$\begin{aligned}
\mathbb{P}_x[N_x \geq k+1] &= \mathbb{E}_x[\mathbb{1}_{H_x<\infty}\mathbb{1}_{N_x\geq k+1}] \\
&= \mathbb{E}_x[\mathbb{1}_{H_x<\infty}\mathbb{1}_{N_x\circ\Theta_{H_x}\geq k}] \\
&= \mathbb{E}_x[\mathbb{1}_{H_x<\infty}\mathbb{1}_{N_x\geq k}\circ\Theta_{H_x}] \\
&= \mathbb{E}_x[\mathbb{1}_{H_x<\infty}\,\mathbb{E}_{X_{H_x}}[\mathbb{1}_{N_x\geq k}] \\
&= \mathbb{P}_x[H_x<\infty]\,\mathbb{P}_x[N_x \geq k]
\end{aligned}$$

Puisque $\mathbb{P}_x(N_n \geq 1) = 1$, on a

$$\mathbb{P}_x[N_x \geq k] = \mathbb{P}_x[H_x < \infty]^{k-1}$$

- Si $\mathbb{P}_x[H_x < \infty] = 1$, on a $\mathbb{P}_x[N_x = \infty] = 1$.
- Si $\mathbb{P}_x[H_x < \infty] < 1$, on a $\mathbb{E}[N_x] = \sum_{k\geq 1} \mathbb{P}[N_x \geq k] = \frac{1}{1-\mathbb{P}[H_x<\infty]} = \frac{1}{\mathbb{P}[H_x=\infty]} < \infty$

□

Définition 2.16. Le **noyau potentiel**, ou la **fonction de Green**, de la chaîne de Markov est la fonction $U : E \times E \to [0, \infty]$ définie par $U(x, y) := \mathbb{E}_x[N_y]$

Proposition 2.17. *1. $U(x,y) = \sum_{n\geq 0} Q^n(x,y)$, ($(I-Q)^{-1}(x,y)$ si 1 n'est pas dans le spectre)*

2. $U(x,x) = \infty \iff x$ est récurrent.

3. $x \neq y \implies U(x,y) = \mathbb{P}_x(H_y < \infty)U(y,y)$

Démonstration. 1. $U(x,y) = \mathbb{E}_x[\sum_{n\geq 0} \mathbb{1}_{X_n=y}] = \sum_{n\geq 0} \mathbb{P}[X_n = y] = \sum_{n\geq 0} Q^n(x,y)$

Cours 11 2. Immédiat par la proposition précédente.

3. Si $X_0 = x$ et $H_y < \infty$, alors $N_y = N_y \circ \Theta_{H_y}$. Donc

$$\mathbb{E}_x[N_y] = \mathbb{E}[\mathbb{1}_{H_y<\infty} N_y \circ \Theta_{H_y}] \overset{2.13}{=} \mathbb{E}_x[\mathbb{1}_{H_y<\infty}\,\mathbb{E}_y[N_y]] = \mathbb{P}_x[H_y < \infty]U(y,y)$$

□

Exemple 2.18 (Marche aléatoire dans $\mathbb{Z}^d$)**.** On considère la chaîne de Markov de matrice de transition

$$Q(x,y) = \frac{1}{2^d}\prod_{i=1}^{d} \mathbb{1}_{|x^i-y^i|=1}$$

(Intuition : une marche aléatoire où on se déplace sur les diagonales) Sous $\mathbb{P}_0$, cette chaîne de Markov a la même loi que $((Y_n^1, ..., Y_n^1))$, où $(Y_n^1), ..., (Y_n^d)$ sont des marches aléatoires simples sur $\mathbb{Z}$, indépendantes, partant de 0. Donc $Q^n(0,0) = \mathbb{P}[Y_n^1 = 0, ..., Y_n^d = 0] = \mathbb{P}[Y_n^1 = 0]^d$. On trouve facilement :

$$\mathbb{P}[Y_n^1 = 0] = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{1}{2^n}\binom{n}{n/2} & \text{si } n \text{ est pair} \end{cases}$$

Par Stirling, on a $(n = 2k)$

$$\frac{1}{2^{2k}}\binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$$

Donc

$$U(0,0) = \sum_{k=0}^{\infty} Q^{2k}(0,0) = \sum_{k=0}^{\infty} (\frac{1}{2^{2k}}\binom{2k}{k})^d$$

Par la proposition 2.17 :

- 0 est récurrent $\iff d \leq 2$
- 0 est transitoire $\iff d \geq 3$

Définition 2.19. $R := \{x \in E : x \text{ est récurrent}\}$

Proposition 2.20. *Soit $x \in R$ et $y \in E$ avec $U(x,y) > 0$. Alors $y \in R$ et $\mathbb{P}_y[H_x < \infty] = 1$. En particulier, $U(y,x) > 0$ (par la proposition 2.17 (3.))*

Démonstration. Montrons que $\mathbb{P}[H_x < \infty] = 1$

$$\begin{aligned} 0 &= \mathbb{P}_x[N_x < \infty] \\ &\geq \mathbb{P}_x[H_x = \infty] \\ &\geq \mathbb{P}_x[H_x = \infty, H_y < \infty] \\ &= \mathbb{P}_x[H_y < \infty, H_x \circ \Theta_{H_y} = \infty] \\ &= \mathbb{E}_x[\mathbb{1}_{H_y<\infty} \mathbb{1}_{H_x=\infty} \circ \Theta_{H_y}] \\ &= \mathbb{P}_x[H_y < \infty] \, \mathbb{P}_y[H_x = \infty] \end{aligned}$$

Comme $U(x,y) > 0$, on a $\mathbb{P}_x[H_y < \infty] > 0$. Donc $\mathbb{P}_y[H_x = \infty] = 0$.

Pour montrer que $y \in R$: il existe $n_1, n_2 \in \mathbb{N}$ tels que $Q^{n_1}(x,y) > 0$ et $Q^{n_2}(y,x) > 0$. Pour tout $p \in \mathbb{N}$, on a $Q^{n_2+p+n_1}(y,y) \geq Q^{n_2}(y,x)Q^p(x,x)Q^{n_1}(x,y)$. Donc

$$U(y,y) \overset{2.17}{\geq} \sum_{p\geq 0} Q^{n_2+p+n_1}(y,y) \geq Q^{n_2}(y,x)Q^{n_1}(x,y) \sum_{p\geq 0} Q^p(x,x) = \infty$$

□

Proposition 2.21 (Classification des états)**.** *Il existe une partition $R = \bigsqcup_{i\in I} R_i$ (les R_i sont appelées classes de récurrence) satisfaisant :*

1. *Si $x \in R_i$, alors $\mathbb{P}_x$-p.s. : $N_y = \infty, \forall y \in R_i$ et $N_y = 0, \forall y \in E\backslash R_i$*
2. *Si $x \in E\backslash R$ et $T := \inf\{n \in \mathbb{N} : X_n \in R\}$, alors $\mathbb{P}_x$-p.s. :*
   — *Soit $T = \infty$ et $N_y < \infty, \forall y \in E$.*
   — *Soit $T < \infty$ et $\exists i \in I$ aléatoire t.q. $X_n \in R_i, \forall n \geq T$.*

Démonstration. Pour $x, y \in R$, on écrit $x \sim y \iff U(x,y) > 0$. Alors $\sim$ est une relation d'équivalence sur R, par la proposition 2.20 (exercice). $\{R_i\}$ sont les classes d'équivalence de $\sim$.

1. Soit $x \in R_i$. Alors $U(x,y) = 0, \forall y \in E\backslash R_i$. (si $y \in E\backslash R$, on utilise la proposition 2.20). Donc $N_y = 0$ $\mathbb{P}_x$-p.s. $\forall y \in E\backslash R_i$. En revanche, si $y \in R_i$, on a $\mathbb{P}_x[H_y < \infty] = 1$ par la proposition 2.20, et donc $\mathbb{P}_x[N_y = \infty] = \mathbb{E}_x[\mathbb{1}_{H_y<\infty} \cdot \mathbb{1}_{N_y=\infty} \circ H_y] \overset{2.13}{=} \mathbb{P}_x[H_y < \infty] \, \mathbb{P}_y[N_y = \infty] = 1$.
2. Soit $x \in E\backslash R$
   Cas $T = \infty$:
   — Si $y \in E\backslash R$, alors
   $$U(x,y) \overset{2.17}{=} \mathbb{P}_x[H_y < \infty] U(y,y) \overset{2.17}{<} \infty$$
   donc $N_y < \infty$ $\mathbb{P}_x$-p.s..
   — Si $y \in R$, alors $T = \infty \implies N_y = 0$.
   Conclusion : $\forall y \in E, \mathbb{P}_x[N_y < \infty, T = \infty] = \mathbb{P}_x[T = \infty]$.
   Cas $T < \infty$: Soit i l'indice (aléatoire) tel que $X_T \in R_i$, i.e. $R_i = [X_T]$. Alors $\forall k \geq 0$,
   $$\begin{aligned} \mathbb{P}_x[X_{T+k} \in R_i, T < \infty] &= \mathbb{E}_x[\mathbb{1}_{T<\infty} \cdot \mathbb{1}_{X_k} \circ \Theta_T \in R_i] \\ &= \sum_{Z\in R} \mathbb{E}_x[\mathbb{1}_{T<\infty} \cdot \mathbb{1}_{X_T=Z} \mathbb{1}_{X_k\circ\Theta_T \sim Z}] \\ &\overset{2.13}{=} \sum_{z\in R} \mathbb{E}_x[\mathbb{1}_{T<\infty} \mathbb{1}_{X_T=Z} \underbrace{\mathbb{P}[X_k \in Z]}_{=1}] \\ &= \mathbb{P}_x[T < \infty] \end{aligned}$$

□

Cours 12

Définition 2.22. Une chaîne de Markov est **irréductible** si $U(x,y) > 0, \forall x, y \in E$.

2.4 Mesures invariantes

Soit μ une mesure positive sur E, t.q. $\mu(E) > 0$, et $\mu(x) < \infty, \forall x \in E$.

Définition 2.23. μ est **invariante** (pour Q) si $\forall y, \mu(y) = \sum_{x \in E} \mu(x) Q(x,y)$ (i.e. $\mu = \mu Q$)

Remarque 2.24. Si $\mu(E) = 1$ et μ est invariante, alors les chaînes de Markov $(X_n)_{n\in\mathbb{N}}$ et $(X_{n+1})_{n\in\mathbb{N}}$ ont la même loi sois $\mathbb{P}_\mu$.

Démonstration. $\forall F : \Omega \to \mathbb{R}_+$, on a

$$\begin{aligned}
\mathbb{E}_\mu[F \circ \Theta_1] &\overset{2.12}{=} \mathbb{E}_\mu[\mathbb{E}_{X_1}[F]] \\
&= \sum_x (\mu Q)(x)\, \mathbb{E}_x[F] \\
&= \mu(x)\, \mathbb{E}_x[F] \\
&= \mathbb{E}_\mu[F]
\end{aligned}$$

□

Exemple 2.25. Marche aléatoire homogène sur $\mathbb{Z}^d$. $Q(x,y) = \gamma(x-y)$. $\mu =$ mesure de comptage, invariante. On a

$$\sum_x Q(x,y) = \sum \gamma(x-y) = 1$$

Définition 2.26. μ est **réversible** (pour Q) si $\forall x, y \in E$, $\mu(x)Q(x,y) = \mu(y)Q(y,x)$. ("Detailed balance")

Remarque 2.27. μ réversible $\implies$ μ invariante.

Démonstration.

$$\sum_x \mu(x)Q(x,y) = \sum_x \mu(y)Q(y,x) = \mu(y)$$

□

L'inverse n'est pas vrai. l'exemple 2.25 avec $\gamma(x) \neq \gamma(-x)$

Exemple 2.28. $E = \mathbb{Z}, Q(x,x+1) = p, Q(x,x-1) = 1-p, 0 < p < 1$. $\mu(x) = \left(\frac{p}{1-p}\right)^x$ est réversible (donc invariante).

Proposition 2.29. *Soit $x \in R$. Alors*

$$\mu(y) := \mathbb{E}_x\left[\sum_{k=0}^{H_x - 1} \mathbb{1}_{X_k = y}\right]$$

est invariante. De plus, $\mu(y) > 0 \iff x \sim y$.

Démonstration. Si $x \not\sim y$, alors $\mathbb{E}_x[N_y] = U(x,y) = 0$, et donc $\mu(y) = 0$. On trouve

$$\mu(y) = \mathbb{E}_x\left[\sum_{k=1}^{H_x} \mathbb{1}_{X_k = y}\right]$$

car $X_0 = X_{H_x} = x$

$$\begin{aligned}
&= \sum_{z \in E} \mathbb{E}_x\left[\sum_{k=1}^{H_x} \mathbb{1}_{X_{k-1}=z} \mathbb{1}_{X_k = y}\right] \\
&= \sum_{z \in E} \sum_{k=1}^{\infty} \mathbb{E}_x[\mathbb{1}_{k \le H_x} \mathbb{1}_{X_{k-1}=z} \mathbb{1}_{X_k=y}] \\
&= \sum_{z \in E} \sum_{k=1}^{\infty} \mathbb{E}_x[\mathbb{1}_{k \le H_x} \mathbb{1}_{X_{k-1}=z} \underbrace{\mathbb{E}_x[\mathbb{1}_{X_k=y} | \mathcal{F}_{k-1}]}_{=Q(X_{k-1},y)}] \\
&= \sum_{z \in E} \mathbb{E}_x\left[\sum_{k=1}^{H_x} \mathbb{1}_{X_{k-1}=z}\right] Q(z,y) \\
&= \sum_{z \in E} \mu(z) Q(z,y)
\end{aligned}$$

Donc $\mu = \mu Q$ et μ est invariante.

Enfin, soit $x \sim y$ et montrons que $0 < \mu(y) < \infty$. $\exists n \geq 0$ t.q. $Q^n(y,x) > 0$, et $\mu(y) < \infty$ car

$$1 = \mu(x) = \sum_{z \in E} \mu(z) Q^n(z,x) \geq \mu(y) \underbrace{Q^n(y,x)}_{>0}$$

$\exists m \geq 0$ t.q. $Q^m(x,y) > 0$, et

$$\mu(y) = \sum_{z \in E} \mu(z) Q^m(z,y) \geq \mu(x) Q^m(x,y) > 0$$

□

Remarque 2.30. S'il y a plusieurs classes de récurrence $R_i, i \in I$, on peut choisir $x_i \in R_i$ et poser

$$\mu_i(y) := \mathbb{E}_{x_i}\left[\sum_{k=0}^{H_{x_i}-1} \mathbb{1}_{X_k = y}\right]$$

Alors $\mu_i, i \in I$, sont invariantes à supports disjoints.

Proposition 2.31. *Si la chaîne de Markov est récurrente irréductible alors la mesure invariante est unique (à une constante près)*

Démonstration. Cf. de Gall □

Proposition 2.32. *Soit la chaîne de Markov récurrente irréductible.*

1. *Ou bien il existe une mesure de probabilité invariante μ, et on a $\mathbb{E}_x[H_x] = \frac{1}{\mu(x)}, \forall x$. On dit que la chaîne de Markov est* ***récurrente positive****.*
2. *Ou bien toute mesure invariante μ satisfait $\mu(E) = \infty$ et on a $\mathbb{E}_x[H_x] = \infty, \forall x$. On dit que la chaîne de Markov est* ***récurrente nulle****.*

Démonstration. Par la proposition 2.31, la mesure invariante est unique. Soit $x \in E$ et ν_x la mesure de la proposition 2.29. Alors $\mu = C\nu_x$ pour un $C > 0$

Cas 1 : $1 = \mu(E) = C\nu_x(E)$, donc $\mu(x) = \frac{\nu_x(x)}{\nu_x(E)} = \frac{1}{\nu_x(E)}$. Or $\nu_x(E) = \sum_{y \in E} \mathbb{E}_x[\sum_{k=0}^{H_x - 1} \mathbb{1}_{X_k = y}] = \mathbb{E}_x[H_x]$.

Cas 2 : $\nu_x(E) = \infty$, donc $\mathbb{E}_x[H_x] = \infty$. □

Proposition 2.33. *Soit la chaîne de Markov irréductible. S'il existe une mesure invariante finie, la chaîne de Markov est récurrente (positive).*

Démonstration. γ mesure invariante finie, soit $y \in E$ t.q. $\gamma(y) > 0$. Alors

$$\gamma(E) U(y,y) = \sum_x \gamma(x) U(y,y) \overset{2.17}{\geq} \sum_x \gamma(x) U(x,y) = \sum_{n \geq 0} \sum_x \gamma(x) Q^n(x,y) \overset{inv.}{=} \sum_{n \geq 0} \gamma(y) = \infty$$

Pour $\gamma(E) < \infty$ on conclut $U(y,y) = \infty$, donc y est récurrent. La chaîne de Markov est donc récurrente par la proposition 2.20 □

2.5 Comportement asymptotique

Proposition 2.34. *Soit une chaîne de Markov irréductible, et soit μ une mesure invariante. Soient $f, g : E \to \mathbb{R}_+$ tels que $0 < \int g d\mu < \infty$, Alors $\forall x \in E, \mathbb{P}_x$-p.s.,*

$$\frac{\sum_{k=0}^n f(X_k)}{\sum_{k=0}^n g(X_k)} \xrightarrow{n \to \infty} \frac{\int f d\mu}{\int g d\mu}$$

Démonstration. Idée : Pour $k \geq 0$, on définit $Z_k(f) := \sum_{n=T_k}^{T_{k+1}-1} f(X_n)$, où $T_0 = 0, T_1 = H_x$, $T_2 = \inf\{n > T_1, X_n = x\}$. On utilise que $(Z_k(f))_{k \geq 0}$ sont *i.i.d.*, et on applique la loi des grands nombres (Cf. de Gall). □

Corollaire 2.35. *Soit une chaîne de Markov irréductible et récurrente positive, et μ sa mesure invariante de probabilité, on a $\mathbb{P}_x$-p.s.*

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \overset{n\to\infty}{\longrightarrow} \int f d\mu$$

Cours 13

Remarque 2.36. Par le corollaire 2.35, si la chaîne est récurrente positive, on a

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{P}_x[X_k = y] \to \mu(y)$$

$\forall x, y$, convergence des moyennes. Est-il vrai que

$$\mathbb{P}_x[X_n = y] \to \mu(y), \forall x, y \in E$$

En général non ! On a un problème de periodicité. (Voir la marche aléatoire sur $E = 2$, déterministe de période 2).

Définition 2.37. Soit x récurrent et

$$L_x := \{n \in \mathbb{N} : Q^n(x, x) > 0\}$$

La **période** de x, notée $d(x)$, est le gcd de L_x.

Proposition 2.38. *Soit une chaîne de Markov récurrente irréductible. Tous les états de $x \in E$ ont la même période $d(x)$, la* ***période de la chaîne****.*

Définition 2.39. Une chaîne de Markov irréductible, récurrente positive et **apériodique** (i.e. de période 1) est appelée **ergodique**.

Proposition 2.40. *Soit une chaîne de Markov ergodique avec mesure de probabilité invariante μ. Alors $\forall x \in E$,*

$$\sum_{y\in E} |\mathbb{P}_x[X_n = y] - \mu(y)| \underset{n\to\infty}{\longrightarrow} 0$$

Chapitre 3

Introduction à la statistique

— **Probabilités** (mathématiques) : On se donne $\mathbb{P}$ et une famille de variables aléatoires, et on étudie leurs propriétés.
— **Statistique** (science empirique) : On se donne une collection d'observations d'expériences aléatoires indépendantes (ou pas), et on veut trouver la loi des variables aléatoires sous-jacentes.

3.1 Estimateurs

Définition 3.1. Un n**-échantillon de loi** $\mathbb{P}$ est une famille $X_1, ..., X_n$ de variables aléatoires indépendantes de loi $\mathbb{P}$.

En **estimation paramétrique**, on suppose que $\mathbb{P} = \mathbb{P}_\theta$ dépend d'un paramètre $\theta \in \Theta$, et on cherche à estimer θ à partir d'un n-échantillon de loi $\mathbb{P}_\theta$ (l'observation).

Exemple 3.2.
1. Sondage avec question oui/non : $\mathbb{P}_\theta = Ber(\theta), \theta \in [0, 1] = \Theta$
2. Durée de vie d'un appareil (par hypothèse sans vieillissement) : $\mathbb{P}_\theta = Exp(\theta), \theta \in \mathbb{R}_+ = \Theta$
3. Taille des individus : $\mathbb{P}_\theta = \mathcal{N}(\mu, \sigma^2), \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ = \Theta$ (Note : on peut conditionner pour avoir des valeurs positives)

Définition 3.3.
1. Une **statistique** est une fonction mesurable d'un n-échantillon, $T = F(X_1, ..., X_n)$
2. Pour $f : \Theta \to E$, on appelle **estimateur de** $f(\theta)$ toute statistique T à valeurs dans $f(\Theta)$ (qui sera utilisé pour estimer $f(\theta)$).

Remarque. Un estimateur est une fonction de l'échantillon, ne dépend pas de θ (inconnu).

Définition 3.4. Une famille d'estimateurs $(T_n)_{n\geq 1}$ est **consistante** si

$$T_n \xrightarrow[n\to\infty]{\mathbb{P}_\theta} f(\theta), \forall \theta \in \Theta$$

..

Exemple 3.5 (Moyenne empirique)**.** $\bar{X}_n := \frac{1}{n}(X_1 + ... + X_n)$ est un estimateur consistant de l'espérence $f(\theta) := \mathbb{E}_\theta[X_1]$ (par rapport à $\mathbb{P}_\theta$).

Définition 3.6.
1. Le **biais** d'un estimateur T de $f(\theta)$ est $\mathbb{E}_\theta[T] - f(\theta)$. T est **sans biais** si son biais est 0, $\forall \theta \in \Theta$, sinon il est biaisé.
2. Une famille d'estimateurs $(T_n)_{n\geq 1}$ de $f(\theta)$ est **asymptotiquement sans biais** si $\lim_{n\to\infty}(\mathbb{E}_\theta[T_n] - f(\theta)) = 0, \forall \theta \in \Theta$.

Exemple 3.7 (Suite de l'exemple 3.5)**.** On cherche à estimer $f(\theta) := \mathbb{E}_\theta[X_1]$. Pour la moyenne empirique $\bar{X}_n := \frac{1}{n}(X_1 + ... + X_n)$, on a $\mathbb{E}[\bar{X}_n] = \mathbb{E}_\theta[X_1] = f(\theta), \forall \theta$, donc sans biais.

Exemple 3.8. Estimons la variance $\sigma^2 = f(\theta) = \mathrm{Var}_\theta(X_1) = \mathbb{E}_\theta[X_1^2] - \mathbb{E}_\theta[X_1]^2$. On a un estimateur naturel :

$$\tilde{\sigma}_n^2 := \frac{1}{n}(X_1^2 + ... + X_n^2) - (\frac{1}{n}(X_1 + ... + X_n))^2$$

Par la loi des grands nombres, $\tilde{\sigma}_n^2$ est consistant.

Calculons le biais :
$$\mathbb{E}_\theta[\tilde{\sigma}_n^2] = \mathbb{E}_\theta[X_1^2] - \frac{1}{n^2}\,\mathbb{E}[(X_1 + ... + X_n)^2] = \frac{n-1}{n}\sigma^2$$
$\tilde{\sigma}_n^2$ est donc biaisé. Un estimateur sans biais est la **variance empirique** $S_n^2 := \frac{n}{n-1}\tilde{\sigma}_n^2$

Construction d'estimateurs :

A. Méthode des moments : $\theta = \mathbb{E}_\theta[g(X_1)]$, alors on peut estimer θ avec $\hat{\theta} := \frac{1}{n}(g(X_1) + ... + g(X_n))$, qui est consistant et sans biais. Choix classique : $g(x) = x^r, r \in \mathbb{R}$ applicable au cas où $\theta = h(\mathbb{E}_\theta[X_1^r])$ pour une fonction h, avec estimateur $\hat{\theta} = h(\frac{1}{n}(X_1^r + ... + X_n^r))$, en général biaisé.

Exemple 3.9. $\mathbb{P}_\theta = Exp(\theta)$, avec $\mathbb{E}_\theta[X_1] = \frac{1}{\theta}$, on peut utiliser $\hat{\theta}_n = \frac{1}{\bar{X}_n}$ pour θ.

B. Maximum de vraisemblance : Étant donné une réalisation $x_1, ..., x_n$ d'un n-échantillon, on cherche le θ sous lequel il était le plus probable d'observer $x_1, ..., x_n$.

Définition 3.10 (Vraisemblance).

1. Soit $\mathbb{P}_\theta$ discret. La **vraisemblance** de $\mathbb{P}_\theta$ en $x_1, ..., x_n$ est la fonction
$$L(\theta; x_1, ..., x_n) := \prod_{i=1}^{n} \mathbb{P}_\theta[X_i = x_i]$$
(θ variable, $x_1, ..., x_n$ paramètres)
2. Soit $\mathbb{P}_\theta(dx) = f_\theta(x)dx$ continue avec densité f_θ. La **vraisemblance** de $\mathbb{P}_\theta$ en $x_1, ..., x_n$ est
$$L(\theta; x_1, ..., x_n) := \prod_{i=1}^{n} f_\theta(x_i)$$

Définition 3.11. L'estimateur de **maximum de vraisemblance** de θ est
$$\hat{\theta}(X_1, ..., X_n) := \arg\max_\theta L(\theta; X_1, ..., X_n)$$

Exemple 3.12. $\mathbb{P}_\theta = Exp(\theta)$, $x_1, ..., x_n > 0$.
$$L(\theta; x_1, ..., x_n) = \prod_{i=1}^{n} \theta e^{-\theta x_i}$$
$$\hat{\theta} = \frac{n}{X_1 + ... + X_n}$$
(Cf. exemple 9)

Exemple 3.13. $\mathbb{P}_\theta = Unif([0, \theta])$, $x_1, ..., x_n \geq 0$.
$$L(\theta; x_1, ..., x_n) = \prod_{i=1}^{n} \frac{1}{\theta}\mathbb{1}_{X_i \leq \theta} = \frac{1}{\theta^n}\mathbb{1}_{\max_{x_i} \leq \theta}$$
donc $\hat{\theta} = \max_i x_i$.

Comparaison d'estimateurs

Définition 3.14.

1. Le **risque quadratique** de l'estimateur $\hat{\theta}$ de θ est $R_{\hat{\theta}}(\theta) := \mathbb{E}_\theta[(\hat{\theta} - \theta)^2]$
2. Si $\hat{\theta}$ et $\tilde{\theta}$ sont deux estimateurs de θ, $\hat{\theta}$ est meilleur que $\tilde{\theta}$ si $R_{\hat{\theta}}(\theta) < R_{\tilde{\theta}}(\theta)$.

Exemple 3.13. (suite)

— Estimateur sans biais de $\theta = 2\,\mathbb{E}_\theta[X_1]$ est $\bar{\theta} = \frac{2}{n}(X_1 + ... + X_n)$, avec risque quadratique
$$R_{\hat{\theta}}(\theta) = \frac{4}{n}\,\mathrm{Var}_\theta(X_1) = \frac{\theta^2}{3n}$$

— Estimateur de maximum de vraisemblance $\hat{\theta} = \max\{X_1, ...X_n\}$ Clairement, $\hat{\theta}$ est biaisé, car $\mathbb{E}[\hat{\theta}] < \theta$. De plus, $\mathbb{P}_\theta[\hat{\theta} \leq x] = \mathbb{P}_\theta[X_1 \leq x, ..., X_n \leq x] = \mathbb{P}_\theta[X_1 \leq x]^n = \left(\frac{x}{\theta}\right)^n$ pour $x \leq \theta$, donc la densité de $\hat{\theta}$ est

$$f_{\hat{\theta}}(x) = \frac{n}{\theta^n} x^{n-1} \mathbb{1}_{0 \leq x \leq \theta}$$

On en déduit que $\mathbb{E}_\theta[\hat{\theta}] = \frac{n}{n+1}\theta$ (asymptote sans biais). De plus,

$$R_{\hat{\theta}}(\theta) = \frac{2\theta^2}{(n+1)(n+2)}$$

Donc $\hat{\theta}$ est meilleur que $\bar{\theta}$ dès que $n \geq 3$ (même si $\hat{\theta}$ est biaisé).

3.2 Intervalles de confiance

But : à partir d'un échantillon $X_1, ..., X_n$ de loi $\mathbb{P}_\theta$ avec $\theta \in \mathbb{R}$, trouver un intervalle $I \subset \mathbb{R}$ tel que $\theta \in I$ avec une certaine confiance.

Définition 3.15. Soit $0 < \alpha < 1$. Un intervalle $I = I(X_1, ..., X_n)$ est appelé **intervalle de confiance pour θ au niveau** $1 - \alpha$ si

$$\mathbb{P}_\theta[\theta \in I] = 1 - \alpha, \forall \theta \in \Theta$$

Si on a $\mathbb{P}_\theta[\theta \in I] \geq 1 - \alpha$, alors on dit que I est un **intervalle de confiance pour θ au niveau** $1 - \alpha$ **par excès**.

Exemple 3.16. $\mathbb{P}_\theta = \mathcal{N}(\theta, 1)$. Soit $I = [\bar{X}_n - a, \bar{X}_n + a]$. On sait que

$$\sqrt{n}(\bar{X}_n - \theta) \stackrel{d}{=} \mathcal{N}(0, 1) = Z$$

donc

$$1 - \alpha \stackrel{!}{=} \mathbb{P}_\theta[\theta \in I] = \mathbb{P}_\theta[|X_n - \theta| \leq a] = \mathbb{P}[|Z| \leq a\sqrt{n}]$$

Avec $\alpha = \frac{1}{10}$, on trouve $I = [\bar{X}_n - \frac{1.64}{\sqrt{n}}, \bar{X}_n + \frac{1.64}{\sqrt{n}}]$ pour l'intervalle de confiance au niveau $\frac{9}{10}$.

Exemple 3.17. $\mathbb{P}_\theta = Unif([0, \theta]), \hat{\theta} = \max\{X_1, ..., X_n\}$ (Exemple 3.13). Clairement, $\hat{\theta} \leq \theta$. Soit $I = [\hat{\theta}, C\hat{\theta}]$. Alors

$$1 - \alpha \stackrel{!}{=} \mathbb{P}_\theta[\theta \in I] = \mathbb{P}_\theta[\theta \leq C\hat{\theta}] = 1 - \mathbb{P}[\hat{\theta} < \frac{\theta}{C}] \stackrel{3.13}{=} 1 - \frac{1}{C^n}$$

Conclusion : $I = [\hat{\theta}, \alpha^{-\frac{1}{n}}\hat{\theta}]$