AGRÉGATION OPTION A

CHRISTOPHE POQUET, MARIELLE SIMON

Contents

1. Martingales

Les martingales sont utilisées pour modéliser les jeux de hasard équitables : si, à un instant t, on a perdu ou gagné une certaine somme S, alors on n'a pas plus de chance dans le futur d'augmenter ou de diminuer le gain. Un théorème dit que si un joueur début le jeu avec une fortune initiale finie, il n'existe pas de stratégie pour gagner à coup sûr (et donc les casinos s'enrichissent).

Les martingales ont d'importantes propriétés qui permettent d'obtenir des résultats de convergence forts.

1.1. **Définitions.**

Définition 1.1 (Filtration)**.** *Une* filtration *d'un espace de probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$ *est une suite croissante* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *de sous-tribus de* $\mathcal{F}$*. On dit que* $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathbb{P})$ *est un espace de probabilité* filtré.

Remarque 1.2. *Souvent on utilise la convention* $\mathcal{F}_{-1} := \{\emptyset, \Omega\}$.

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

Exemple 1.4. *Un processus* (X_n) *est toujours adapté pour sa* filtration canonique *définie par*

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

Définition 1.5 (Martingale). *Le processus* $(X_n)_{n\in\mathbb{N}}$ *est une* martingale pour la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *si les trois conditions suivantes sont vérifiées :*

(i) pour tout $n \in \mathbb{N}$, $\mathbb{E}[|X_n|] < \infty$ *(*i.e. X_n *est dans* L^1*)*
(ii) $(X_n)_{n\in\mathbb{N}}$ *est adapté à* $(\mathcal{F}_n)_{n\in\mathbb{N}}$
(iii) pour tout $n \in \mathbb{N}$,

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

On dit que (X_n) *est une* sur-martingale *(resp.* sous-martingale*) si (i) et (ii) sont vérifiées et si (iii) est remplacé par*

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leqslant X_n \;\; p.s. \quad resp. \;\; \mathbb{E}[X_{n+1}|\mathcal{F}_n] \geqslant X_n \;\; p.s.$$

Exemple 1.6 (Quelques exemples). *Voici quelques exemples à connaître :*

(1) Soit $(\mathcal{F}_n)$ *une filtration, et* $X \in L^1(\Omega)$ *une variable aléatoire intégrable. Alors* $M_n = \mathbb{E}[X|\mathcal{F}_n]$ *est une martingale pour* $(\mathcal{F}_n)$*.*
(2) Soient $(Z_\ell)_{\ell\in\mathbb{N}}$ *des v.a.i.i.d. intégrables, d'espérance nulle. Alors*

$$M_n = Z_0 + \cdots + Z_n$$

est une martingale pour sa filtration canonique. En effet:

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \underbrace{\mathbb{E}[M_n|\mathcal{F}_n]}_{=M_n \text{ car mesurable}} + \underbrace{\mathbb{E}[X_{n+1}|\mathcal{F}_n]}_{=\mathbb{E}[X_{n+1}]=0 \text{ car } \perp}$$

Proposition 1.7. *(1) Si* (X_n) *est une martingale pour la filtration* $(\mathcal{F}_n)$ *alors* $\mathbb{E}[X_n] = \mathbb{E}[X_0]$*.*
De plus, $\mathbb{E}[X_{n+k}|\mathcal{F}_n] = X_n$ *pour tout* $k \geqslant 0$*.*
(2) Si (X_n) *est une sur-martingale (resp. sous-martingale) alors* $\mathbb{E}[X_n] \leqslant \mathbb{E}[X_0]$ *(resp.* $\mathbb{E}[X_n] \geqslant \mathbb{E}[X_0]$*).*

Proposition 1.8. *Si* (X_n) *est une martingale, si* $\varphi : \mathbb{R} \to \mathbb{R}$ *est convexe, et si, pour tout* n, $\varphi(X_n)$ *est intégrable, alors* $(\varphi(X_n))_{n\in\mathbb{N}}$ *est une sous-martingale.*

Exemple 1.9. *Par exemple, si* (X_n) *est une martingale, alors* $(|X_n|)$, (X_n^2), *et* $(X_n^+) = \max(X_n, 0)$ *sont des sous-martingales.*

1.2. **Temps d'arrêt.** On souhaite arrêter un processus à des temps aléatoires, dépendant uniquement du passé et du présent (pas du futur). On rappelle ici la définition :

Définition 1.10 (Temps d'arrêt). *Une application mesurable* $T : \Omega \to \mathbb{N} \cup \{\infty\}$ *est un* temps d'arrêt pour la filtration $(\mathcal{F}_n)$ *si, pour tout* $n \in \mathbb{N}$,

$$\{T = n\} \in \mathcal{F}_n \tag{1}$$

Remarque 1.11. *L'événement* $\{T = \infty\}$ *appartient à* $\mathcal{F}_\infty := \sigma(\cup_{n\in\mathbb{N}}\mathcal{F}_n)$*.*

Remarque 1.12. *Il est important de remarquer que la condition* (1) *est équivalente à* $\{T \leqslant n\} \in \mathcal{F}_n$ *pour tout* n*. Remarquons aussi :* $\{T < n\} = \{T \leqslant n-1\} \in \mathcal{F}_{n-1}$*.*

Exemple 1.13. *Voici quelques exemples :*

(1) Un temps déterministe : $T(\omega) \equiv n_0$ *est un temps d'arrêt.*
(2) Les temps d'atteinte *sont des temps d'arrêt pour la filtration naturelle des processus,* i.e.

$$T_B = \inf\{n \in \mathbb{N} \; ; \; X_n \in B\}$$

car

$$\{T_B = n\} = \bigcup_{k=0}^{n-1} \{X_k \notin B\} \cup \{X_n \in B\}$$

(3) *Un temps défini comme* $\tau_B = \sup\{n \in \mathbb{N}\ ;\ X_n \in B\}$ *n'est* pas *un temps d'arrêt, car il dépend de tout le futur.*

Proposition 1.14. *Si* S *et* T *sont deux temps d'arrêt pour la même filtration* $(\mathcal{F}_n)$*, alors* $S+T$*,* $S \wedge T$ *et* $S \vee T$ *sont des temps d'arrêt pour* $(\mathcal{F}_n)$*.*

Définition 1.15 (Tribu des événements antérieurs)**.** *Si* T *est un temps d'arrêt, on définit*

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty\ ;\ \textit{pour tout } n \in \mathbb{N},\ A \cap \{T = n\} \in \mathcal{F}_n\}$$

la tribu des événements antérieurs à T*.*

Exemple 1.16. *Soit* $(X_n)_{n\in\mathbb{N}}$ *un processus à valeurs dans* $\mathbb{Z}$ *et soit* $(\mathcal{F}_n)$ *sa filtration canonique. Notons* T *le temps d'atteinte de* 0*. On considère* A *l'événement :*

$$A = \{\ \textit{la trajectoire } (X_n) \textit{ passe par un nombre pair}\ \}$$

Alors on peut récrire $A = \{\tau < \infty\}$ *où* τ *est le temps d'atteinte de l'ensemble des entiers pairs. Alors*

$$A \in \mathcal{F}_T$$

En effet : comme $\tau \leqslant T$ *p.s., on a*

$$\{\tau < \infty\} \cap \{T = n\} = \{\tau \leqslant n\} \cap \{T = n\} \in \mathcal{F}_n$$

Proposition 1.17. *Si* S *et* T *sont deux temps d'arrêt pour la même filtration* $(\mathcal{F}_n)$ *et si* $S \leqslant T$ *alors* $\mathcal{F}_S \subset \mathcal{F}_T$*.*

Proposition 1.18. *La variable aléatoire* $\mathbf{1}_{T<\infty} X_T$ *est* $\mathcal{F}_T$*-mesurable.*

Proof.

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

□

On parle maintenant du processus arrêté.

Proposition 1.19. *Si* $(X_n)_{n\in\mathbb{N}}$ *est une martingale (resp. sur-martingale, resp. sous-martingale) pour* $(\mathcal{F}_n)$ *et si* T *est un temps d'arrêt pour* $(\mathcal{F}_n)$ *alors* $(X_{n\wedge T})_{n\in\mathbb{N}}$ *est une martingale (resp. sur-martingale, resp. sous-martingale) pour* $(\mathcal{F}_n)$*.*

Proof. On a

$$\begin{aligned}
\mathbb{E}[X_{(n+1)\wedge T} - X_{n\wedge T} | \mathcal{F}_n] &= \mathbb{E}[(X_{(n+1)\wedge T} - X_{n\wedge T})\mathbf{1}_{T>n} + 0 \times \mathbf{1}_{T\leqslant n} | \mathcal{F}_n] \\
&= \mathbb{E}[(X_{n+1} - X_n)\ \underbrace{\mathbf{1}_{T>n}}_{\text{mesurable}}\ |\mathcal{F}_n] \\
&= \mathbf{1}_{T>n}\ \mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] = 0.
\end{aligned}$$

□

1.3. **Théorème d'arrêt.** D'après le résultat précédent, si T est un temps d'arrêt, et si (X_n) est une martingale (pour $(\mathcal{F}_n)$) alors

$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0]$$

De plus, si $T < \infty$ p.s., alors $X_{T\wedge n} \xrightarrow[n\to\infty]{p.s.} X_T$. Si on pouvait passer à la limite sous l'espérance, alors on aurait $\mathbb{E}[X_T] = \mathbb{E}[X_0]$, ce qui a d'importantes conséquences pratiques. On va énoncer le résultat général :

Théorème 1.20 (Théorème d'arrêt)**.** *Si l'une des trois conditions suivantes est satisfaite :*

(1) T est borné (autrement dit, il existe $N \in \mathbb{N}$ tel que $T \leqslant N$ p.s.)
*(2) T est intégrable (*i.e. $\mathbb{E}[T] < \infty$*) et il existe $M > 0$ tel que, pour tout $n \geqslant 0$, $|X_{n+1} - X_n| \leqslant M$ p.s.*
(3) $T < \infty$ p.s. et $(X_{T\wedge n})$ est uniformément borné (il existe $C > 0$ tel que, pour tout n, $X_{T\wedge n} \leqslant C$ p.s.),

alors $(X_{T\wedge n})$ est intégrable, et on a l'égalité : $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

Remarque 1.21. *Si l'une des trois conditions est satisfaite et si (X_n) est une sur-(resp. sous-) martingale, alors $\mathbb{E}[X_T] \leqslant \mathbb{E}[X_0]$ (resp. $\mathbb{E}[X_T] \geqslant \mathbb{E}[X_0]$).*

Exemple 1.22. *Donnons un contre-exemple au théorème d'arrêt. Supposons $X_0 = 1$ p.s., et $X_{n+1} = X_n \times Z_{n+1}$ où les (Z_i) sont i.i.d. de même loi $\frac{1}{2}\delta_2 + \frac{1}{2}\delta_0$. On note T le temps d'atteinte de 0. Alors T est fini presque sûrement (il suit la loi géométrique de paramètre $\frac{1}{2}$). Mais on a $\mathbb{E}[X_0] = 1$ et $\mathbb{E}[X_T] = 0$.*

1.4. **Inégalité maximale de Doob.**

Théorème 1.23 (Inégalité maximale de Doob)**.** *Si (X_n) est une sous-martingale positive, alors, pour tout $\lambda > 0$,*

$$\mathbb{P}(\max_{0\leqslant k\leqslant n} X_k > \lambda) \leqslant \frac{\mathbb{E}[X_n]}{\lambda}$$

En particulier, si (X_n) est une martingale et si $X_n \in L^p$ pour un certain $p \geqslant 1$, alors on peut appliquer ce résultat à $(|X_n|^p)$ qui est une sous-martingale positive, et on obtient:

$$\mathbb{P}(\max_{0\leqslant k\leqslant n} |X_k| > \lambda) = \mathbb{P}(\max_{0\leqslant k\leqslant n} |X_k|^p > \lambda^p) \leqslant \frac{\mathbb{E}[|X_n|^p]}{\lambda^p}$$

Exemple 1.24. *On applique ce résultat à la marche aléatoire simple sur $\mathbb{Z}$: $X_n = Z_1 + \cdots + Z_n$ où les Z_i sont i.i.d. de même loi $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Alors, $\mathbb{E}[X_n^2] = \mathrm{Var}(X_n) = n\mathrm{Var}(Z_1) = n$ et*

$$\mathbb{P}(\max_{0\leqslant k\leqslant n} |X_k| \geqslant n^{\frac{1}{2}+\varepsilon}) \leqslant \frac{\mathbb{E}[X_n^2]}{n^{1+2\varepsilon}} = n^{-2\varepsilon}$$

1.5. **Convergences.**

Théorème 1.25 (Martingales bornées dans L^1)**.** *Si (X_n) est une martingale (resp. sous- ou sur-martingale) vérifiant $\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] < \infty$, alors*

$$X_n \xrightarrow[n\to\infty]{p.s.} X_\infty, \quad \text{où } X_\infty \text{ est une v.a. intégrable.}$$

Remarque 1.26. *La convergence a lieu p.s. mais pas dans L^1. E, effet, prenons le même exemple que tout à l'heure : $X_0 = 1$ p.s., et $X_{n+1} = X_n \times Z_{n+1}$ où les (Z_i) sont i.i.d. de même loi $\frac{1}{2}\delta_2 + \frac{1}{2}\delta_0$. Alors $\mathbb{E}[X_n] = \mathbb{E}[X_0] = 1$, mais $X_n \to 0$ p.s.*

Théorème 1.27 (Martingales bornées dans L^2)**.** *Si (X_n) est une martingale uniformément bornée dans L^2 i.e. vérifiant $\sup_{n\in\mathbb{N}} \mathbb{E}[X_n^2] < \infty$ alors*

$$X_n \xrightarrow[n\to\infty]{p.s. \text{ et dans } L^2} X_\infty, \quad \text{où } X_\infty \text{ est une v.a. intégrable.}$$

De plus : pour tout $n \in \mathbb{N}$

$$X_n = \mathbb{E}[X_\infty|\mathcal{F}_n].$$

Les martingales dans L^2 ont des propriétés très intéressantes, par exemple :

Théorème 1.28 (Pythagore pour les martingales)**.** *Soit* (X_n) *une martingale telle que* $\mathbb{E}[X_n^2] < \infty$ *pour tout* n*. alors, pour tous* $m \leqslant n \leqslant p$ *entiers,*

$$\mathbb{E}[(X_n - X_m)(X_p - X_n)] = 0$$

et en particulier

$$\mathbb{E}[X_n^2] = \mathbb{E}[X_0^2] + \sum_{k=0}^{n-1} \mathbb{E}[(X_{k+1} - X_k)^2]$$

Proof. On a

$$\begin{aligned}\mathbb{E}[(X_n - X_m)(X_p - X_n)] &= \mathbb{E}\big[\mathbb{E}[(X_n - X_m)(X_p - X_n)|\mathcal{F}_n]\big] \\ &= \mathbb{E}\big[(X_n - X_m)\mathbb{E}[X_p - X_n|\mathcal{F}_n]\big] \\ &= \mathbb{E}[(X_n - X_m) \times 0] = 0.\end{aligned}$$

□

Théorème 1.29 (Convergence dans L^1)**.** *Si* (X_n) *est une martingale, alors les conditions suivantes sont équivalentes :*

(i) (X_n) *converge vers une v.a.* X_∞ *p.s. et dans* L^1*.*
(ii) il existe une v.a. $Y \in L^1$ *intégrable telle que* $X_n = \mathbb{E}[Y|\mathcal{F}_n]$ *pour tout* $n \in \mathbb{N}$*.*
(iii) la suite (X_n) *est* uniformément intégrable*, au sens suivant :*

$$\lim_{c\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\big[|X_n|\mathbf{1}_{|X_n|>c}\big] = 0.$$

Exemple 1.30. *Si* (X_n) *vérifie* $\sup_{n\in\mathbb{N}} |X_n| \leqslant M$ *p.s., alors* (X_n) *est uniformément intégrable. C'est aussi vrai si*

$$\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^{1+\alpha}] < \infty$$

pour un $\alpha > 0$*.*

2. Chaînes de Markov

2.1. Définitions.

Définition 2.1. *Soit* E *un ensemble fini ou dénombrable.*

Une suite $X = (X_n)$ *de variables aléatoires à valeurs dans* E*, définie sur un espace de probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$ *est une* chaîne de Markov de loi initiale μ_0 *si la loi de* X_0 *est* μ_0 *et si, pour tout* $n \geqslant 0$*, pour tous* $x_1, x_2, \ldots, x_n, x_{n+1} \in E$*,*

$$\mathbb{P}(X_{n+1} = x_{n+1}|X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n)$$

dès lors que $\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) > 0$*.*

La chaîne est dite homogène *si la probabilité de transition* $\mathbb{P}(X_{n+1} = y|X_n = x) =: P(x, y)$ *ne dépend pas de* n*. On appelle* matrice de transition *(ou* noyau de transition*) l'application* $P : E \times E \to [0, 1]$*, qui vérifie*

$$\sum_{y\in E} P(x, y) = 1$$

À partir de maintenant on ne considère que des chaînes de Markov homogènes.

Proposition 2.2. *Soit* X_0 *une v.a. à valeurs dans* $(E, \mathcal{P}(E))$*, et soit* (U_n) *une suite de v.a. i.i.d. à valeurs dans un espace mesurable* $(G, \mathcal{G})$ *et indépendantes de* X_0*. Enfin, soit* $f : E \times F \to E$

une application mesurable. La suite récurrente aléatoire

$$X_{n+1} = f(X_n, U_{n+1})$$

est une chaîne de Markov homogène à valeurs dans E.

Proof. On vérifie la propriété :

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) \\
&\quad = \mathbb{P}(f(X_n, U_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) \\
&\quad = \frac{\mathbb{P}(f(x_n, U_{n+1}) = x_{n+1}, X_0 = x_0, \dots, X_n = x_n)}{\mathbb{P}(X_0 = x_0, \dots, X_n = x_n)} \\
&\quad = \mathbb{P}(f(x_n, U_{n+1}) = x_{n+1}) \qquad \text{car } U_{n+1} \text{ est indépendante des } (X_i) \\
&\quad = \mathbb{P}(f(x_n, U_1) = x_{n+1}) \qquad \text{car les } U_i \text{ sont de même loi.}
\end{aligned}$$

□

La loi d'une chaine de Markov est entièrement donnée par sa loi initiale μ_0, et sa matrice de transition P. On définit par récurrence les *itérés* de P comme suit : $P^0 := P$ et

$$P^{n+1}(x, y) = \sum_{z \in E} P(x, z) P^n(z, y) \qquad \text{(produit standard de matrices dans le cas } E \text{ fini)}$$

Il est facile de voir que pour tout $n \geqslant 1$, P^n est encore un noyau de transition.

On identifie les *mesures* ν sur E à des vecteurs lignes $(\nu(x_1), \dots, \nu(x_n), \dots)$ et les *fonctions* sur E à des vecteurs colonnes ${}^{\mathrm{T}}(f(x_1), \dots, f(x_n), \dots)$. Les produits entre "matrice" et "vecteur" s'étendent aisément au cas E dénombrable. En particulier, si ν est une mesure sur E est $f : E \to \mathbb{R}$ est une fonction définie sur E, alors

$$\int_E f \mathrm{d}\nu = \nu f = \sum_{x \in E} \nu(x) f(x), \qquad \text{i.e. un produit scalaire entre deux vecteurs.}$$

Proposition 2.3. *Soit (X_n) une chaîne de Markov homogène de probabilité de transition P et de loi initiale μ_0. Alors, pour tout $x_0, x_1, \dots, x_n \in E$,*

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \mu(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n)$$

De plus, la loi de X_n est $\mu_n = \mu_0 P^n$. En particulier,

$$\mathbb{P}(X_n = y | X_0 = x) = P^n(x, y) \qquad \textit{et} \qquad \mathbb{E}[f(X_n)] = \mu_0 P^n f.$$

Proposition 2.4 (Relation de Chapman-Kolmogorov)**.** *Pour tout $n, m \geqslant 1$, pour tout $x, y \in E$*

$$\begin{aligned}
\mathbb{P}(X_{n+m} = y | X_0 = x) &= \sum_{z \in E} \mathbb{P}(X_m = y | X_0 = z) \mathbb{P}(X_n = z | X_0 = x) \\
&= \sum_{z \in E} P^m(z, y) P^n(x, z)
\end{aligned}$$

2.2. **Exemples.**

2.2.1. *Cas E fini.*

- **Graphe des transitions**

- **Marche aléatoire sur les sommets d'un graphe**. Soit V un ensemble de sommets d'un graphe supposé connexe. On note $x \sim y$ (ou de manière équivalente $y \sim x$) s'il existe une arête (non orientée) reliant x à y. Pour tout $x \in V$, on note le nombre de ses voisins $d(x) = \#\{y \in V \ ; \ y \sim x\} \geqslant 1$ que l'on suppose fini. L'espace d'états est V, et la matrice de transition est donnée par

$$P(x,y) = \begin{cases} \dfrac{1}{d(x)} & \text{si } y \sim x \\ 0 & \text{sinon.} \end{cases}$$

- **Urne d'Ehrenfest**. L'espace d'états est $\{0, 1, \ldots, d\}$ et la matrice de transition est donnée par

$$P(x,y) = \begin{cases} x/d & \text{si } y = x - 1 \\ (d-x)/d & \text{si } y = x + 1 \\ 0 & \text{sinon.} \end{cases}$$

- **Ruine du joueur**. L'espace d'états est $\{0, 1, \ldots, K\}$ et la matrice de transition est donnée pour $p \in (0,1)$ par

$$\begin{aligned} P(0,0) &= 1 \\ P(K,K) &= 1 \\ P(x, x+1) &= p \qquad \text{et } P(x, x-1) = 1 - p \qquad \text{pour tout } x \in \{1, \ldots, K-1\} \end{aligned}$$

2.2.2. *Cas E dénombrable.*

- **Marche aléatoire simple symétrique sur $\mathbb{Z}^d$**. Soit $(e_i)_{i=1,\ldots,d}$ est une base de $\mathbb{Z}^d$. Un marcheur fait un pas de longueur 1 toutes les secondes, il choisit sa direction au hasard uniformément sur $\{e_i, -e_i \ ; \ i = 1, \ldots, d\}$. Il part de l'origine 0 et on note X_n sa position au temps n. La suite (X_n) est une chaîne de Markov sur $\mathbb{Z}^d$ de matrice de transition

$$P(x,y) = \begin{cases} 1/(2d) & \text{si } |x - y| = 1 \\ 0 & \text{sinon.} \end{cases}$$

- **Chaîne de vie et de mort**. L'espace d'états est $\mathbb{N}$ et la matrice de transition est définie par

$$P(x, x-1) = q_x, \quad P(x, x+1) = p_x, \quad P(x,x) = r_x,$$

  où les p_x, q_x, r_x vérifient

$$q_0 = 0, \quad q_x + p_x + r_x = 1, \quad \begin{cases} p_x > 0 & \text{pour tout } x \geqslant 0 \\ q_x > 0 & \text{pour tout } x \geqslant 1 \end{cases}$$

2.3. **Propriété de Markov.** Soit $X = (X_n)$ une chaîne de Markov sur E de matrice de transition P et de loi initiale μ. On peut voir X comme une variable aléatoire à valeur dans *l'espace des trajectoires* $E^{\mathbb{N}}$. Pour cela, on munit $E^{\mathbb{N}}$ de la tribu $\mathcal{B}$ engendrée par les *cylindres* $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$. Le résultat suivant est fondamental (et peut être admis) :

Théorème 2.5. *Il existe une unique mesure de probabilité $\mathbb{P}_\mu$ sur $(E^{\mathbb{N}}, \mathcal{B})$ telle que, sous $\mathbb{P}_\mu$, (X_n) est une chaîne de Markov sur E de loi initiale μ et de noyau de transition P.*

À partir de maintenant on supposera que la chaîne de Markov est toujours définie sur l'espace de probabilité $(E^{\mathbb{N}}, \mathcal{B}, \mathbb{P}_\mu)$, au lieu de $(\Omega, \mathcal{F}, \mathbb{P})$ comme initialement, et sa matrice de transition est P. On l'appelle alors la *chaîne de Markov canonique.* Lorsque $\mu = \delta_x$ où $x \in E$ on note plus simplement $\mathbb{P}_{\delta_x} = \mathbb{P}_x$. On note $\mathbb{E}_\mu$ l'espérance par rapport à $\mathbb{P}_\mu$.

On introduit à présent la chaîne X^{+k} *décalée d'un certain rang* k, comme suit :

$$X_n^{+k} := X_{k+n}, \quad \text{pour tout } n \geqslant 0.$$

On la note aussi parfois $(\theta_k X)_n$, avec θ_k l'opérateur de translation.

Théorème 2.6 (Propriété de Markov faible)**.** *Pour tout $A \in \mathcal{B}$, pour tout $k \in \mathbb{N}$*

$$\mathbb{P}_\mu(X^{+k} \in A | X_k = x_k, \ldots, X_0 = x_0) = \mathbb{P}_{x_k}(X \in A)$$

dès lors que $\mathbb{P}_\mu(X_k = x_k, \ldots, X_0 = x_0) > 0$. En particulier,

$$\begin{aligned}\mathbb{P}_\mu(X_{k+1} = y_1, \ldots X_{k+r} = y_r | X_k = x_k, \ldots, X_0 = x_0) &= \mathbb{P}_{x_k}(X_1 = y_1, \ldots, X_r = y_r) \\ &= P(x_k, y_1)P(y_1, y_2) \cdots P(y_{r-1}, y_r).\end{aligned}$$

De manière encore plus générale, on énonce la propriété de Markov faible de la manière suivante :

Théorème 2.7 (Propriété de Markov faible, énoncé plus général)**.** *Soit $F : (E^{\mathbb{N}}, \mathcal{B}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ une fonction positive mesurable. Alors, quelque soit $\psi : E^{k+1} \to \mathbb{R}_+$ on a*

$$\mathbb{E}_\mu[\psi(X_0, \ldots, X_k)\ F(X^{+k})] = \mathbb{E}_\mu[\psi(X_0, \ldots, X_k)\ \mathbb{E}_{X_k}[F(X)]]$$

ce qui est équivalent à

$$\mathbb{E}_\mu[F(X^{+k}) | \mathcal{F}_k] = \mathbb{E}_{X_k}[F] \quad \mathbb{P}_\mu\text{–}a.s.$$

Cette propriété s'étend aux *temps aléatoires* qui possèdent certaines propriétés.

Définition 2.8 (Temps d'arrêt)**.** *Un* temps d'arrêt T *est une variable aléatoire à valeurs dans* $\mathbb{N} \cup \{\infty\}$ *telle que, pour tout $n \in \mathbb{N}$*

$$\{T = n\} \in \mathcal{F}_n = \sigma(X_0, \ldots, X_n)$$

Théorème 2.9 (Propriété de Markov forte)**.** *Soit T un temps d'arrêt. Alors, pour tout $k \in \mathbb{N}$, pour tout $A \in \mathcal{B}$,*

$$P_\mu(X^{+T} \in A | T = k, X_k = x_k, \ldots, X_0 = x_0) = \mathbb{P}_{x_k}(X \in A)$$

ou encore : pour toute fonction $F : (E^{\mathbb{N}}, \mathcal{B}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ positive mesurable, et toute fonction $\psi : E^{k+1} \to \mathbb{R}_+$,

$$\mathbb{E}_\mu[\mathbf{1}_{T=k}\ \psi(X_0, \ldots, X_k)\ F(X^{+T})] = \mathbb{E}_\mu[\mathbf{1}_{T=k}\ \psi(X_0, \ldots, X_k)\ \mathbb{E}_{X_T}[F(X)]]$$

En particulier, définissons la tribu engendrée par le temps d'arrêt :

$$\mathcal{F}_T := \{A \in \mathcal{B}\ ;\ A \cap \{T = k\} \in \mathcal{F}_k \text{ pour tout } k \in \mathbb{N}\}$$

Alors, si $x, y \in E$ et T est un temps d'arrêt tel que, $T < \infty$ et $X_T = y$ $\mathbb{P}_x$–p.s., on peut dire que la chaîne translatée X^{+T} est une chaîne de Markov de matrice de transition P, de loi initiale δ_y, et est indépendante de $\mathcal{F}_T$.

2.4. Classification des états.

Définition 2.10. *Pour $x, y \in E$, on note $x \to y$ s'il existe $n \in \mathbb{N}$ tel que $P^n(x, y) > 0$. On dit que deux états $x, y \in E$ communiquent si $x \to y$ et $y \to x$. On note alors $x \leftrightarrow y$. On dit que x est absorbant si $P(x, x) = 1$.*

La relation $\leftrightarrow$ *est une relation d'équivalence sur* E *dont les classes sont appelées les* classes de communication *de la chaîne de Markov.*

Une classe de communication C *est* fermée *si, pour tout* $x \in C$*, il n'existe aucun* $y \notin C$ *tel que* $x \to y$*. S'il n'y a qu'une seule classe de communication* $C = E$*, la chaîne de Markov est dite* irréductible*.*

Exemple 2.11. *La marche aléatoire simple sur* $\mathbb{Z}^d$ *est irréductible. La ruine du joueur possède deux états absorbants, et une classe de communication ouverte.*

On s'intéresse aux états après une longue trajectoire de la chaîne de Markov : en particulier, quels états ne seront plus visités après un certain temps, et quels sont ceux qui au contraire seront revisités perpetuellement ?

2.4.1. *Récurrence et transience.* Soit $x \in E$ et (X_n) une chaîne de Markov sur E de noyau P. On définit

$$T_x := \inf \{n > 0\ ;\ X_n = x\}$$

à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$.

Définition 2.12. *On dit que* x *est transient si* $\mathbb{P}_x(T_x < \infty) < 1$ *et on dit que* x *est récurrent si* $\mathbb{P}_x(T_x < \infty) = 1$*.*

Un état récurrent est dit récurrent nul *si* $\mathbb{E}_x[T_x] = \infty$ *et* récurrent positif *si* $\mathbb{E}_x[T_x] < \infty$*.*

Ainsi, si x est récurent, alors partant de x, la chaîne de Markov revient presque sûrement en x. Par la propriété de Markov forte au temps T_x, un second retour en x est aussi presque sûr, etc. Pour étudier cela, on introduit le nombre de visites en x :

$$N_x = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = x\}}$$

Proposition 2.13. *Soit* $x \in E$*.*

(1) L'état x *est récurrent si et seulement si* $\mathbb{P}_x(N_x = \infty) = 1$*.*

(2) L'état x *est transient si et seulement si* $\mathbb{P}_x(N_x < \infty) = 1$*. De plus* N_x *suit une loi géométrique de paramètre* $p = \mathbb{P}_x(T_x = \infty)$ *et*

$$\mathbb{E}_x[N_x] = \frac{1}{\mathbb{P}_x(T_x = \infty)}.$$

Idées de preuve. Pour tout $k \geqslant 1$ on calcule $\mathbb{P}_x(N_x \geqslant k+1)$ en utilisant la propriété de Markov fort :

$$\mathbb{P}_x(N_x \geqslant k+1) = \mathbb{E}_x\left[\mathbb{E}_{X_{T_x}}(\mathbf{1}_{\{N_x \geqslant k\}})\mathbf{1}_{\{T_x < \infty\}}\right] = \mathbb{P}_x(N_x \geqslant k)\mathbb{P}_x(T_x < \infty).$$

Comme $\mathbb{P}_x(N_x \geqslant 1) = 1$ on en déduit par induction

$$\mathbb{P}_x(N_x \geqslant k) = \left(\mathbb{P}_x(T_x < \infty)\right)^{k-1}.$$

Si $\mathbb{P}_x(T_x < \infty) = 1$ alors $\mathbb{P}_x(N_x = \infty) = 1$. Sinon,

$$\mathbb{E}_x[N_x] = \sum_{k \geqslant 1} \mathbb{P}_x(N_x \geqslant k) = \frac{1}{1 - \mathbb{P}_x(T_x < \infty)}$$

□

2.4.2. *Propriétés de classe.*

Proposition 2.14. *Soit* $x, y \in E$*. On suppose que* x *est récurrent et que* $x \to y$*.*

(1) *Alors y est récurrent. Par suite, les éléments d'une même classe sont tous récurrents, ou tous transients. On parle de* classe récurrente *ou* classe transiente.
(2) *De plus $y \to x$. Par suite, si une classe n'est pas fermée, elle est nécessairement transiente. Enfin,*
$$\mathbb{P}_x(T_y < \infty) = \mathbb{P}_y(T_x < \infty) = 1.$$
(3) *(Corollaire) Une classe fermée et de cardinal fini est récurrente.*

Cette proposition permet de déterminer la nature de tous les états d'une chaîne de Markov sur un espace d'états E *fini* : les classes fermées sont récurrentes, les autres transientes. En particulier :

Corollaire 2.15. *Une chaîne de Markov irréductible sur un espace fini est récurrente.*

Remarque 2.16. *Supposons la chaîne de Markov irréductible sur E (de cardinal fini ou infini).*
Dans le cas récurrent, tous les points *de E sont visités infiniment souvent : pour $x, y \in E$,*
$$\mathbb{P}_x(X_n = y \text{ pour une infinité de } n) = 1$$
Dans le cas transient, tous les sous-ensembles finis de *E sont visités un nombre fini de fois : pour $A \subset E$ de cardinal fini,*
$$\mathbb{P}_x(X_n \in A \text{ pour une infinité de } n) = 0.$$

Il est alors parfois utile, pour mener à bien les calculs, de considérer la *fonction de Green* associée à la chaîne de Markov, définie de la manière suivante :

Définition 2.17. *Pour $x, y \in E$ on définit*
$$G(x,y) = \mathbb{E}_x[N_y] = \mathbb{E}_x\Big[\sum_{n \geqslant 0} \mathbf{1}_{\{X_n = y\}}\Big]$$

Proposition 2.18. *Pour tout $x \neq y$ dans E*
$$G(x,x) = \frac{1}{\mathbb{P}_x(T_x = \infty)}, \qquad G(x,y) = \mathbb{P}_x(T_y < \infty) G(y,y)$$

En particulier : x est récurrent si et seulement si $G(x,x) = +\infty$. De plus, on a toujours : pour tout $x \neq y$, $G(x,y) \leqslant G(y,y)$. Enfin $x \to y$ si et seulement si $G(x,y) > 0$.

2.4.3. *Instants de retours.* Soit $x \in E$ et (X_n^x) une chaîne de Markov sur E de probabilité de transition P partant de $X_0 = x$. On introduit la suite des instants successifs de retours en x, par récurrence :
$$T_x^1 = \inf\{k > 0 \,;\, X_k^x = x\}, \qquad T_x^{n+1} = \inf\{k > T_x^n \,;\, X_k^x = x\}$$
avec la convention $\inf \emptyset = \infty$. Le nombres de visites en x avant le temps n est noté N_x^n et est défini par
$$N_x^n = \sum_{k=0}^{n} \mathbf{1}_{\{X_k^x = x\}}.$$
Ces nombres de visites sont reliés aux temps de passage par les relations suivantes :
$$N_x^n \geqslant k+1 \;\Leftrightarrow\; T_x^k \leqslant n.$$
La propriété de Markov fort nous permet également de montrer :

Proposition 2.19. *Soit $x \in E$. Sur l'évènement $T_x^n < \infty$, les v.a. $T_x, T_x^2 - T_x^1, \ldots, T_x^n - T_x^{n-1}$ sont i.i.d.*

2.5. **Mesures invariantes.** À partir de maintenant on supposera la chaîne de Markov (X_n) irréductible. Dans le cas général, on pourra néanmoins appliquer les résultats suivantes à la chaîne de Markov restreinte à une classe fermée.

Les questions de récurrence ou transience s'intéressent plutôt au comportement asymptotique de la chaîne (les états sont-ils visités infiniment souvent ?). On va maintenant regarder des aspects plus quantitatifs, par exemple :

- combien de temps (X_n) passe-t-elle dans les différents états (en proportion) ?
- combien de temps (X_n) met-elle pour revenir à son point de départ (en espérance, i.e. calculer $\mathbb{E}_x[T_x]$) ?
- la loi de (X_n) admet-elle une limite quand n est grand ?

Pour tout n, on note μ_n la loi de (X_n) sur E (que l'on peut considérer comme un vecteur ligne, possiblement infini). On rappelle : pour tout n, $\mu_{n+1} = \mu_n P$. Ainsi, si $\mu_n \to \mu$ on a de bonnes raisons de penser que μ devrait vérifier $\mu = \mu P$.

Définition 2.20. *Une mesure μ sur E (non nécessairement de masse finie) est dite* invariante *pour P si $\mu = \mu P$, i.e.: pour tout $y \in E$*

$$\mu(y) = \sum_{x \in E} \mu(x) P(x, y).$$

Si $\mu(E) = 1$, alors μ est une loi de probabilité invariante*.*

On s'intéresse maintenant à l'existence de telles mesures. La première façon d'en trouver est de rechercher parmi une certaine famille de mesures :

Définition 2.21. *Une mesure μ sur E est* réversible *pour P si, pour tout $x, y \in E$*

$$\mu(x)P(x, y) = \mu(y)P(y, x)$$

Une mesure réversible est invariante, mais l'inverse n'est pas vrai. Ces mesures sont souvent plus faciles à trouver s'il en existe.

2.5.1. *Cas E fini.* Dans ce cas, la relation $\mu = \mu P$ signifie que μ est un vecteur propre à gauche de P pour la valeur propre 1 (ou encore que ${}^T\mu$ est un vecteur propre à droite de ${}^TP$). On sait déjà que 1 est une valeur propre car la somme des lignes vaut 1, donc on sait qu'il existe $\mu \neq 0$ tel que $\mu P = \mu$. Il faut donc comprendre si l'on peut choisir μ à coefficients *positifs*. C'est la conséquence (par exemple) d'un résultat d'algèbre (le théorème de Perron-Frobenius) mais il peut aussi être démontré par un argument probabiliste (cf. cas général).

Proposition 2.22. *Si E est de cardinal fini, alors il existe une mesure invariante μ pour P. De plus :*

- *Si la chaîne est irréductible alors il existe une unique loi de probabilité invariante π et elle vérifie $\pi(x) > 0$ pour tout $x \in E$.*
- *Si la chaîne n'est pas irréductible, toute mesure invariante pour P s'écrit comme combinaison convexe des mesures invariantes des chaîne restreintes à chaque composante* récurrente.

2.5.2. *Cas général.* Dans ce cas il n'existe pas toujours de loi de probabilité invariante.

Proposition 2.23. *Supposons la chaîne (X_n) récurrente irréductible. Alors, il existe une mesure invariante (non nulle), unique à un facteur près.*

Soit $x \in E$. La mesure invariante notée μ_x est donnée par

$$\mu_x(y) = \mathbb{E}_x\left[\sum_{n=0}^{T_x-1} \mathbf{1}_{\{X_n=y\}}\right]$$

et c'est l'unique mesure invariante qui vérifie $\mu_x(x) = 1$.

Dans le cas transient, on peut juste dire :

Proposition 2.24. *Si π est une loi de probabilité invariante et si $x \in E$ est transient, alors $\pi(x) = 0$. En particulier, si la chaîne est transiente et irréductible, il ne peut pas exister de loi de probabilité invariante non nulle.*

2.6. **Récurrence positive.** Grâce à la formule explicite des mesures invariante μ_x dans le cas récurrent, on peut donner un critère d'existence de loi de probabilité invariante. Notons que sa masse totale est égale à

$$\mu_x(E) = \sum_{y\in E} \mu_x(y) = \mathbb{E}_x\left[\sum_{n=0}^{T_x-1} \sum_{y\in E} \mathbf{1}_{\{X_n=y\}}\right] = \mathbb{E}_x[T_x]$$

Ceci nous mène au résultat suivant :

Proposition 2.25. *Supposons la chaîne de Markov (X_n) irréductible récurrente. Alors il n'existe que deux situations possibles :*

(1) Toutes les mesures invariantes ont une masse infinie. Dans ce cas : pour tout $x \in E$, $\mathbb{E}_x[T_x] = \infty$. La chaîne est dite récurrente nulle.

(2) Toutes les mesures invariantes ont une masse totale finie. Il existe alors une unique loi de probabilité invariante π, qui vérifie : pour tout $x \in E$,

$$\pi(x) = \frac{1}{\mathbb{E}_x[T_x]} > 0,$$

et la chaîne est dite récurrente positive.

On en déduit alors, d'après ce qui précède :

Corollaire 2.26. *Toute chaîne de Markov irréductible sur un espace d'états fini est récurrente positive.*

Dans le cas général, l'ensemble des mesures de probabilité invariantes d'une chaîne est formé des combinaisons convexes des meures invariantes des chaînes restreintes aux classes de communication récurrentes positives.

Exemple 2.27. *Dans l'exemple de la ruine du joueur, les lois de probabilité invariantes sont les combinaisons convexes $p\delta_0 + (1-p)\delta_K$ pour $p \in [0,1]$.*

2.7. **Théorèmes limites.**

Théorème 2.28 (Théorème ergodique)**.** *On suppose la chaîne (X_n) irréductible récurrente.*

(1) Si la chaîne est récurrente positive, *notons π son unique loi de probabilité invariante. Alors, pour tout $x \in E$, pour toute loi initiale μ sur E,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=x\}} \xrightarrow[n\to\infty]{\mathbb{P}_\mu\text{-}p.s.} \pi(x).$$

Plus généralement, pour toute fonction $f : E \to \mathbb{R}$ intégrable par rapport à π

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[n\to\infty]{\mathbb{P}_\mu\text{-}p.s.} \int_E f\,\mathrm{d}\pi = \sum_{x\in E} \pi(x) f(x)$$

(2) *Si la chaîne est* récurrente nulle*, alors, pour tout état $x \in E$, pour toute loi initiale μ sur E,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=x\}} \xrightarrow[n\to\infty]{\mathbb{P}_\mu\text{-}p.s.} 0.$$

Lorsque l'on prend l'espérance dans les convergences ci-dessus, le théorème de convergence dominée nous permet d'en déduire, dans le cas irréductible récurrent :

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{P}_\mu(X_k = x) \xrightarrow[n\to\infty]{} \frac{1}{\mathbb{E}_x[T_x]}.$$

Ce n'est pas suffisant pour en déduire que la suite $(\mathbb{P}_\mu(X_n = x))_{n\geqslant 0}$ converge vers $\pi(x)$. Un contre-exemple typique est celui de la matrice de transition donnée par

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

qui a pour loi invariante $\pi = (\frac{1}{2}, \frac{1}{2})$ mais pour laquelle $\mathbb{P}_0(X_n = 0) = 0$ dès que n est pair. On a donc besoin d'une notion supplémentaire :

Définition 2.29. *La* période *d'un état $x \in E$ est*

$$d(x) := \mathrm{PGCD}\{n \geqslant 1 \; ; \; P^n(x,x) > 0\}$$

Si $d(x) = 1$, x est dit apériodique*.*

Dans l'exemple précédent, $d(x) = 2$ pour tout état x. Notons que si $P(x,x) > 0$ alors immédiatement $d(x) = 1$.

Proposition 2.30. *Si x et y communiquent, alors $d(x) = d(y)$: la période est la même pour tous les états d'une même classe.*

Théorème 2.31 (Convergence en loi)**.** *On suppose que la chaîne (X_n) est irréductible, récurrente positive, et apériodique. Alors, pour tout $x \in E$ et toute loi initiale μ sur E,*

$$\mathbb{P}_\mu(X_n = x) \xrightarrow[n\to\infty]{} \pi(x).$$

Si la chaîne est irréductible, récurrente nulle et apériodique, ou transiente, alors

$$\mathbb{P}_\mu(X_n = x) \xrightarrow[n\to\infty]{} 0.$$

2.8. **Quelques exemples.**

2.8.1. *La marche aléatoire simple symétrique sur $\mathbb{Z}^d$.* La marche aléatoire simple sur $\mathbb{Z}^d$ peut être définie de manière équivalente de la manière suivante : $X_0 \in \mathbb{Z}^d$ aléatoire, et pour tout $n \geqslant 0$,

$$X_{n+1} = X_n + U_{n+1}$$

où les $(U_i)_{i\geqslant 1}$ sont des v.a.i.i.d. uniformes dans $\{-e_i, e_i \; ; \; i = 1, \ldots, d\}$, avec (e_i) base de $\mathbb{Z}^d$.

Théorème 2.32 (Polya, 1921)**.** *Si $d = 1$ ou 2 alors la marche aléatoire est récurrente. Si $d \geqslant 3$ elle est transiente.*

Proof. Le caractère transient ou récurrent d'une chaîne irréductible ne dépend pas de son point de départ. On suppose donc $X_0 = 0$ p.s.

(1) On suppose $d = 1$. Pour n impair, on a $P^n(0,0) = 0$, tandis que pour $n = 2k$ on a

$$P^{2k}(0,0) = \mathbb{P}(\text{autant de pas à droite qu'à gauche}) = \binom{2k}{k}\frac{1}{2^{2k}} \sim_{k\to\infty} \frac{1}{\sqrt{\pi k}}.$$

On calcule alors

$$\mathbb{E}_0[N_0] = \mathbb{E}_0\Big[\sum_{n\geqslant 0} \mathbf{1}_{\{X_n=0\}}\Big] = \sum_{k=0}^{\infty} \mathbb{P}_0(X_{2k} = 0) = \sum_{k=0}^{\infty}\binom{2k}{k}\frac{1}{2^{2k}}$$

qui est une série divergente, donc $\mathbb{E}_0[N_0] = +\infty$ et la chaîne est récurrente.

(2) On suppose $d = 2$. On note $X_n = (X_n^1, X_n^2)$ et $U_n = (U_n^1, U_n^2)$. Les variables

$$S_n = U_n^1 + U_n^2, \qquad D_n = U_n^1 - U_n^2$$

sont indépendantes et de même loi $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. On a alors

$$\begin{aligned} P^{2k}(0,0) &= \mathbb{P}_0(X_{2k}^1 + X_{2k}^2 = X_{2k}^1 - X_{2k}^2 = 0) \\ &= \mathbb{P}(S_1 + S_2 + \cdots + S_{2k} = 0)\mathbb{P}(D_1 + D_2 + \cdots + D_{2k} = 0) \sim \frac{1}{\pi k} \end{aligned}$$

d'où la récurrence de la chaîne.

(3) Pour $d \geqslant 3$, l'intuition nous dit que la chaîne va devenir transiente (la série devient convergente). Pour une preuve rigoureuse (à l'aide des fonctions caractéristiques) est disponible dans [1, Section 2.3.2].

□

2.8.2. *La marche aléatoire simple asymétrique sur* $\mathbb{Z}$. On suppose maintenant $d = 1$ et les v.a. (U_i) vérifient $\mathbb{P}(U_i = 1) = 1 - \mathbb{P}(U_i = -1) = p \neq \frac{1}{2}$. Dans ce cas, la marche aléatoire est transiente.

En effet, dans ce cas

$$X_n = \sum_{k=1}^{n} U_k$$

qui est une somme de v.a.i.i.d. et d'après la loi des grands nombres, on a

$$\frac{1}{n}\sum_{k=1}^{n} U_k \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[U_1] = 2p - 1 \neq 0$$

Ainsi,

$$X_n \xrightarrow[n\to\infty]{p.s.} \begin{cases} +\infty & \text{si } p > \frac{1}{2} \\ -\infty & \text{si } p < \frac{1}{2} \end{cases}$$

et la chaîne est transiente.

2.8.3. *Probabilités d'absorption.* Dans le cas où la chaîne de Markov n'est pas irréductible, et où il y a plusieurs classes fermées, on se demande dans quelle classe la chaîne de Markov va se retrouver "bloquée" au bout d'un certain temps. On définit alors :

Définition 2.33. *Soit C une classe fermée (pour la chaîne de Markov X). On note τ_C son temps d'atteinte :*

$$\tau_C(X) = \inf\{n \geqslant 1\ ;\ X_n \in C\}.$$

Pour tout $x \in E$, la probabilité d'absorption par C partant de x est donnée par

$$q_C(x) = \mathbb{P}_x(\tau_C(X) < \infty)$$

On a évidemment : $q_C(x) = 1$ si $x \in C$, et aussi $q_C(x) = 0$ si x appartient à une autre classe fermée différente de C.

Si x appartient à une classe non fermée, alors, pour atteindre C depuis X, il faut l'atteindre à partir de la position X_1. On peut donc décaler la chaîne d'un rang, et utiliser la propriété de Markov :

$$q_C(x) = \mathbb{P}_x(\tau_C(X^{+1}) < \infty) = \sum_{y \in E} \mathbb{P}_x(X_1 = y, \tau_C(X^{+1}) < \infty) = \sum_{y \in E} \mathbb{P}_x(X_1 = y)\mathbb{P}_y(\tau_C(X) < \infty).$$

Ainsi, si on note $\mathcal{T}$ la réunion des classes non fermées, on obtient

$$q_C(x) = \sum_{y \in \mathcal{T}} P(x, y) q_C(y) + \sum_{y \in C} P(x, y).$$

C'est un système linéaire d'équations, d'inconnues $q_C(x)$ pour $x \in \mathcal{T}$. On peut montrer, si $\mathcal{T}$ est fini, qu'il admet une unique solution.

References

[1] M. Benaïm, N. El Karoui, *Promenade aléatoire*, Ed. École Polytechnique. 2007.