*Année universitaire 2004-2005*

UNIVERSITÉ D'ORLÉANS

**Olivier GARET**

# Probabilités Approfondies

(Master 1ère année semestre 2)

# Table des matières

**Table des matières** **i**

**1 Théorème de Levy** **1**
1.1 Rappels sur la convergence en loi . . . . . . . . . . . . . . . . . . 1
1.2 Tension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
1.3 Théorèmes de Levy . . . . . . . . . . . . . . . . . . . . . . . . . . 6
1.4 Exercices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

**2 Martingales** **9**
2.1 Définitions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
2.2 Premières inégalités . . . . . . . . . . . . . . . . . . . . . . . . . 10
2.2.1 Martingales et fonctions convexes . . . . . . . . . . . . . 10
2.2.2 Inégalité de Kolmogorov . . . . . . . . . . . . . . . . . . 11
2.3 Convergence des martingales de carré intégrable . . . . . . . . . 12
2.4 Temps d'arrêts . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
2.5 Convergence $L^1$ des martingales . . . . . . . . . . . . . . . . . . 17
2.5.1 Théorème des traversées montantes . . . . . . . . . . . . 17
2.5.2 Le théorème de convergence de Doob . . . . . . . . . . . 18
2.6 Décomposition de Doob (*) . . . . . . . . . . . . . . . . . . . . . 19
2.7 Exercices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

**3 Loi d'un processus** **25**
3.1 Loi d'un processus . . . . . . . . . . . . . . . . . . . . . . . . . . 25
3.2 Théorème d'existence de Kolmogorov (admis) . . . . . . . . . . 27
3.3 Processus réels stationnaires . . . . . . . . . . . . . . . . . . . . 28
3.4 Processus gaussiens . . . . . . . . . . . . . . . . . . . . . . . . . 30
3.4.1 Caractérisation . . . . . . . . . . . . . . . . . . . . . . . 30
3.4.2 Condition d'existence . . . . . . . . . . . . . . . . . . . . 31
3.4.3 Processus gaussiens stationnaires . . . . . . . . . . . . . 32
3.5 Exercices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

**4 Chaînes de Markov** **35**
4.1 Dynamique markovienne . . . . . . . . . . . . . . . . . . . . . . . . . 35
4.2 Matrice stochastique . . . . . . . . . . . . . . . . . . . . . . . . . 36
4.2.1 Existence des chaînes de Markov . . . . . . . . . . . . . 37
4.2.2 Puissances des matrices stochastiques . . . . . . . . . . 37
4.3 Graphe associé à une matrice stochastique . . . . . . . . . . . 38
4.4 Exercices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40

**5 Récurrence et mesures invariantes** **45**
5.1 Temps d'arrêt et propriété de Markov forte . . . . . . . . . . . 45
5.2 Classification des états . . . . . . . . . . . . . . . . . . . . . . . . 46
5.3 Mesures invariantes . . . . . . . . . . . . . . . . . . . . . . . . . . 49
5.4 Théorème de la probabilité stationnaire . . . . . . . . . . . . . . 50
5.5 Théorème ergodique des chaînes de Markov . . . . . . . . . . . 53
5.6 Exercices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54

# Chapitre 1

# Théorème de Levy

## 1.1 Rappels sur la convergence en loi

On dit qu'une suite $(\mu_n)_{n\geq 1}$ de mesures de probabilités sur $\mathbb{R}^d$ converge faiblement vers la mesure de probabilité $\mu$ lorsque pour toute fonction $f$ continue bornée de $\mathbb{R}^d$ dans $\mathbb{R}$, on a

$$\lim_{n\to+\infty} \int_{\mathbb{R}^d} f \; d\mu_n = \int_{\mathbb{R}^d} f \; d\mu.$$

Par extension, on dit qu'une suite de variables aléatoires $X_n$ converge en loi vers la variable aléatoire $X$ (ou vers la loi $\mu$) si la suite de mesures $P_{X_n}$ converge faiblement vers $P_X$ (ou vers la loi $\mu$).

Ainsi, dire que $X_n$ converge en loi vers $X$ signifie que pour toute fonction continue bornée, $\mathbb{E}f(X_n)$ converge vers $\mathbb{E}f(X)$.

Rappel : si $\mu$ et $\nu$ sont deux mesures telles que pour toute fonction $f$ continue bornée de $\mathbb{R}^d$ dans $\mathbb{R}$, on a $\int_{\mathbb{R}^d} f \; d\mu = \int_{\mathbb{R}^d} f \; d\nu$, alors $\mu = \nu$.

On rappelle deux théorèmes très utiles dont la preuve peut être trouvée dans le cours de licence.

**Théorème 1 (Théorème du porte-manteau).** *Les propositions suivantes sont équivalentes*

1. *$\mu_n$ converge faiblement vers $\mu$.*
2. *Pour toute fonction $f$ uniformément continue bornée de $\mathbb{R}^d$ dans $\mathbb{R}$, on a*
$$\lim_{n\to+\infty} \int_{\mathbb{R}^d} f \; d\mu_n = \int_{\mathbb{R}^d} f \; d\mu.$$
3. *Pour tout fermé $F$, $\mu(F) \geq \overline{\lim_{n\to+\infty}} \mu_n(F)$.*

4. *Pour tout ouvert $O$, $\mu(O) \leq \varliminf_{n\to+\infty} \mu_n(O)$.*
5. *Pour tout borélien $A$ dont la frontière $\partial A$ vérifie $\mu(\partial A) = 0$, on a $\lim_{n\to+\infty} \mu_n(A) = \mu(A)$.*

**Théorème 2.** *Si une suite $(\mu_n)_{n\geq 1}$ de mesures de probabilités sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ est telle que pour toute fonction $f$ continue positive à support compact de $\mathbb{R}^d$ dans $\mathbb{R}$, on a*

$$\lim_{n\to+\infty} \int_{\mathbb{R}^d} f\ d\mu_n = \int_{\mathbb{R}^d} f\ d\mu,$$

*alors $\mu_n$ converge faiblement vers $\mu$.*

**Théorème 1.** *Soit $X_n$ une suite de variables aléatoires réelles, $X$ une autre variable aléatoire. Alors $X_n$ converge en loi vers $X$ si et seulement si pour tout point $x$ où $F_X$ est continue, $F_{X_n}(x)$ tend vers $F(x)$ lorsque $n$ tend vers l'infini.*

## 1.2 Tension

**Définition:** On dit qu'une famille $\mathcal{M}$ de mesures de probabilités sur $\mathbb{R}^d$ est tendue si pour tout $\varepsilon > 0$, il existe un compact $K$ de $\mathbb{R}^d$ tel que pour tout $\mu \in \mathcal{M}$, on a $\mu(K^c) \leq \varepsilon$.
**Exemple:** Une famille constitué d'une unique mesure $\mu$ sur $\mathbb{R}^d$ est tendue.

*Démonstration.* Considérer la suite d'ensemble $A_n = B(0, n)$ : La suite $A_n^c$ est décroissante et son intersection est $\varnothing$, donc d'après le théorème de continuité séquentielle décroisante $\mu(A_n)^c$ tend vers $\mu(\mathbb{R}^d)$, ce qui montre que bien que pour $n$ assez grand $\mu(A_n^c)$ ne dépasse pas $\varepsilon$. □

**Lemme 1.** *La réunion de deux familles tendues est une famille tendue.*

*Démonstration.* Soit $\mathcal{M}$ et $\mathcal{N}$ deux familles tendues. Soit $\varepsilon > 0$. Comme $\mathcal{M}$ est tendue, il existe un compact $K_1$ tel que $\forall \mu \in \mathcal{M} \mu(K_1^c) \leq \varepsilon$. De même, il existe un compact $K_1$ tel que $\forall \mu \in \mathcal{N} \mu(K_2^c) \leq \varepsilon$. Maintenant, si l'on pose $K = K_1 \cup K_2$, $K$ est compact et il est facile de voir que

$$\forall \mu \in \mathcal{M} \cup \mathcal{N} \mu(K^c) \leq \varepsilon.$$

Ainsi $\mathcal{M} \cup \mathcal{N}$. □

**Corollaire 1.** *Toute famille finie de mesures sur $\mathbb{R}^d$ est tendue.*

Le lemme suivant peut également être utile

**Lemme 2.** *Pour $k$ entre 1 et $d$ et $x$ on note $\pi_k(x)$ la $k$-ième composante de $x$. Soit $\mathcal{M}$ une famille de mesures sur $\mathbb{R}^d$. $\mathcal{M}$ est tendue si et seulement si pour tout $k$ entre 1 et $d$ la famille $\pi_k^{-1}\mathcal{M}$ est tendue.*

*Démonstration.*

- Supposons que $\mathcal{M}$ est tendue. Soit $k$ entre 1 et $d$ et $\varepsilon > 0$. Il existe un compact $K$ tel que pour tout $\mu \in \mathcal{M}$ $\mu(K) \geq 1 - \varepsilon$. Alors, il est clair que $\pi_k K$ est compact est que pour tout $\mu \in \mathcal{M}$, on a

$$\pi_k^{-1}\mu(\pi_k K) = \mu(x \in \mathbb{R}^d : \pi_k(x) \in \pi_k K) \geq \mu(K) \geq 1 - \varepsilon.$$

- Réciproquement, supposons que $pi_k^{-1}\mathcal{M}$ est tendue. Soit $\varepsilon > 0$. Pour tout $k$ entre 1 et $d$ il existe un compact $K_d$ tel que pour tout $\mu \in \mathcal{M}$ $\pi_k^{-1}\mu(K_k) \geq 1 - \varepsilon/d$. Posons $K = K_1 \times K_2 \times \ldots K_d$. On a $K^c = \cup_{i=1}^d \pi_k^{-1}(K_k^c)$, donc

$$\begin{aligned}
\mu(K^c) &\leq \sum_{i=1}^d \mu(\pi_k^{-1}(K_k^c)) \\
&= \sum_{i=1}^d \pi_k^{-1}\mu(K_k^c) \\
&\leq \sum_{i=1}^d \varepsilon/d \\
&\leq \varepsilon
\end{aligned}$$

□

**Définition:** On dit qu'une famille $\mathcal{M}$ de mesures de probabilités est relativement compacte si et seulement toute suite d'éléments de $\mathcal{M}$ admet une sous-suite convergente.

**Théorème 2 (Théorème de Prohorov).** *Toute famille tendue est relativement compacte.*

Afin de ne pas se perdre dans des détails techniques, on va donner la preuve uniquement dans le cas où $d = 1$. On va s'appuyer sur le lemme suivant

**Lemme 3 (Théorème de Helly).** *De toute suite $(F_n)_{n\geq 1}$ de fonctions de répartition on peut extraire une sous-suite $(F_{n_k})_{k\geq 1}$ telle qu'il existe une fonction $F$ croissante continue à droite avec*

$$F_{n_k}(x) \to F(x)$$

*en chaque point de continuité de $F$.*

*Démonstration.* À l'aide du procédé diagonal d'extraction, on commence par extraire une suite $(n_k)_{k\geq 1}$ telle que $F_{n_k}(x)$ converge en tout point $x$ rationnel. On note $G(x)$ la limite obtenue. C'est une fonction croissante. On définit alors

$$F(x) = \inf\{G(r); r \in \mathbb{Q} \cap ]x, +\infty[\}.$$

Il est encore clair que $F$ est croissante. Montrons que $F$ est continue à droite. Soit $x \in \mathbb{R}$ et $\varepsilon > 0$. Par définition de $F$, il existe $r > x$, avec $r$ rationnel tel que $G(r) < F(x) + \varepsilon$. Maintenant, on a

$$\forall y \in [x, r[ \quad F(x) \leq F(y) \leq G(r) < F(x) + \varepsilon,$$

ce qui montre bien que $F$ est continue à droite. Reste à montrer que $F_{n_k}$ converge vers $F$ en chaque point de continuité de $F$. Soit $x$ un point de continuité de $x$ et $\varepsilon > 0$. On peut trouver $\eta$ tel que $|F(x) - F(y)| \leq \varepsilon$ en tout $y$ de $[x - \eta, x + \eta]$. Comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, on peut trouver des rationnels $r$ et $s$ tels que $x - \eta \leq r \leq s \leq x + \eta$. On a pour tout $k \geq 1$ :

$$F_{n_k}(r) \leq F_{n_k}(x).$$

On en déduit

$$F(r) = \lim_{k\to+\infty} F_{n_k}(r) = \varliminf_{k\to+\infty} F_{n_k}(r) \leq \varliminf_{k\to+\infty} F_{n_k}(x),$$

ce qui implique que pout tout $\varepsilon > 0$

$$F(x) - \varepsilon \leq \varliminf_{k\to+\infty} F_{n_k}(x).$$

On en déduit finalement que

$$F(x) \leq \varliminf_{k\to+\infty} F_{n_k}(x).$$

De la même manière, on montre que

$$\varlimsup_{k\to+\infty} F_{n_k}(x) \leq F(x).$$

Finalement, on a

$$F(x) \leq \varliminf_{k\to+\infty} F_{n_k}(x) \leq \varlimsup_{k\to+\infty} F_{n_k}(x) \leq F(x),$$

ce qui montre bien que

$$\lim_{k\to+\infty} F_{n_k}(x) = F(x).$$

□

On peut maintenant prouver le théorème de Prohorov.

*Démonstration.* Soit $\mu_n$ une suite formées de mesures appartenant à une famille $\mathcal{M}$ tendue. On note $F_n$ la fonction de répartion de $\mu_n$. D'après le théorème de Helly, on peut trouver une suite strictement croissante $n_k$ et une fonction continue à droite croissante telle que $F_{n_k}(x)$ converge vers $F(x)$ en chaque point de continuité de $F$. Comme $F$ est croissante continue à droite, on sait (voir un cours de théorie de la mesure) qu'il existe une mesure $\mu$ telle que pour tous $a$ et $b$ $F(b) - F(a) = \mu(]a, b])$. Par construction, $F$ prend ses valeurs entre 0 et 1, donc $\mu(\Omega) = \lim_{n \to +\infty} \mu(] - n, n]) \leq 1$. Il s'agit maintenant de voir que cette mesure est une mesure de probabilité. Soit $\varepsilon > 0$. Comme $\mathcal{M}$ est tendue, il existe un compact $K$ tel que pour tout $\mu \in \mathcal{M}$, on ait $\mu(K) \geq 1 - \varepsilon$. Posons $M = \sup\{|x|; x \in K\}$. Comme l'ensemble des points de discontinuités d'une fonction croissante est au plus dénombrable, il existe $a < -M$ et $b > M$ tels que $F$ soit continue en $a$ et en $b$. Ainsi pour tout $k \geq 1$, on a

$$\begin{aligned} F_{n_k}(b) - F_{n_k}(a) &= \mu_{n_k}(]a, b]) \\ &\geq \mu_{n_k}([-m, m]) \\ &\geq \mu_{n_k}(k) \\ &\geq 1 - \varepsilon \end{aligned}$$

Comme $a$ et $b$ sont des points de continuité de $F$, on obtient, en faisant tendre $k$ vers $+\infty$ : $F(b) - F(a) \geq 1 - \varepsilon$, ce qui entraîne $\mu(\omega) \geq \mu(]a, b]) \geq 1 - \varepsilon$. Comme $\varepsilon$ peut être pris aussi petit que l'on veut, cela entraîne $\mu(\omega) \geq 1$, d'où $\mu(\Omega) = 1$. $\mu$ est donc bien une mesure de probabilité dont $F$ est la fonction de répartition. Comme $F_{n_k}(x)$ converge vers $F(x)$ en chaque point de continuité de $F$, on peut conclure que $\mu_{n_k}$ converge en loi vers $\mu$.

□

**Corollaire 2.** *Soit $(\mu_n)_{n \geq 1}$ telle que*
- $\{\mu_n; n \geq 1\}$ *est tendue*
- *Toute sous-suite de $(\mu_n)_{n \geq 1}$ qui converge en loi converge vers $\mu$.*

*Alors $(\mu_n)_{n \geq 1}$ converge en loi vers $\mu$.*

*Démonstration.* Soit $f$ une fonction continue bornée. On doit montrer que la suite $(\int f(x)\ d\mu_n(x))_{n \geq 1}$ converge vers $\int f(x)\ d\mu(x)$. Soit $(n_k)_{k \geq 1}$ une suite srictement croissante d'entiers quelconque. La suite $(\mu_{n_k})_{k \geq 1}$ est une suite d'éléments d'une famille tendue. Je peux donc en extraire une sous-suite convergente $(\mu_{n_{k_i}})_{i \geq 1}$. Mais d'après la deuxième hypothèse la limite ne peut être que $\mu$. Il s'ensuit que

$$\lim_{i \to +\infty} \int f(x)\ d\mu_{n_{k_i}} = \int f(x)\ d\mu(x).$$

Ainsi, de chaque sous-suite de la suite $(\int f(x)\ d\mu_n(x))_{n\geq 1}$, on peut extraire une sous-suite qui tend vers $\int f(x)\ d\mu(x)$. Cela signifie que la suite $(\int f(x)\ d\mu_n(x))_{n\geq 1}$ converge vers $\int f(x)\ d\mu(x)$, ce qui achève la preuve. □

## 1.3 Théorèmes de Levy

On rappelle que la fonction caractéristique $\varphi_\mu$ d'une mesure $\mu$ est définie par

$$\varphi_\mu(t) = \int_{\mathbb{R}^d} e^{i\langle t,x\rangle}\ d\mu(x).$$

On rappelle aussi que la fonction caractérique caractérise la loi : si deux mesures de probabilités $\mu$ et $\nu$ vérifient en tout point $t$ de $\mathbb{R}^d$ : $\varphi_\mu(t) = \varphi_\nu(t)$, alors $\mu = \nu$.

On rappelle enfin que la fonction caractéristique d'une loi est uniformément continue sur $\mathbb{R}^d$.

**Théorème 3.** *Soit $\mu_n$ une suite de mesures de probabilités sur $\mathbb{R}^d$. On a les résultats suivants :*

- *Si $\mu_n$ converge en loi vers une probabilité $\mu$, alors la suite $\varphi_{\mu_n}$ converge ponctuellement vers $\varphi_\mu$.*
- *Si la suite $\varphi_{\mu_n}$ converge ponctuellement vers $\varphi_\mu$, où $\varphi_\mu$ est la loi de la probabilité $\mu$, alors $\mu_n$ converge en loi vers $\mu$.*
- *Si la suite $\varphi_{\mu_n}$ converge ponctuellement vers une fonction $\varphi$ continue en l'origine, alors il existe une mesure de probabilité $\mu$ dont la fonction caractéristique est $\mu$ et $\mu_n$ converge en loi vers $\mu$.*

*Démonstration.*

Le premier point résulte de la définition de la convergence en loi et du fait que pour tout $t$, la fonction $x \mapsto e^{i\langle t,x\rangle}$ est continue bornée.

Le deuxième point est une conséquence du troisième, en utilisant le fait que la fonction caractéristique $\varphi_\mu$ est bien continue en l'origine.

Montrons donc le troisième point (sans utiliser le deuxième, bien sûr !). Ce qui nous manque maintenant, c'est la tension de la suite $(\mu_n)_{n\geq 1}$, car si $(\mu_n)_{n\geq 1}$ est tendue, elle admet une valheur d'adhérence $\mu$ et il ressort alors clairement du premier point que

- $\varphi_\mu = \varphi$.
- Si une mesure $\mu_*$ est valeur d'adhérence de la suite $(\mu_n)_{n\geq 1}$, alors $\varphi_{\mu_*} = \varphi$, donc $\mu_* = \mu$.

Grâce au corollaire précédent, cela implique alors la convergence en loi de $\mu_n$ vers $\mu$. Il reste donc à montrer que $(\mu_n)_{n\geq 1}$ est tendue. Pour cela, il suffit de montrer que pour tout $k$ entre 1 et $d$, la suite de mesures marginales $(\pi_k^{-1}\mu_n)_{n\geq 1}$ est tendue. Notons qu'on a simplement $\varphi_{\pi_k^{-1}\mu_n)}(x) = \varphi_{\mu_n}(0,\dots,0,x,0,\dots,0)$ et la suite de fonctions $(\varphi_{\pi_k^{-1}\mu_n)}(x))_{n\geq 1}$ converge simplement vers la fonction $x \mapsto \varphi((0,\dots,0,x,0,\dots,0)$, qui est continue en 0. Ainsi, on voit qu'on est ramené à démontrer qu'une famille $(\nu_n)_{n\geq 1}$ de mesures de probabilités sur $\mathbb{R}$ dont les fonctions caractéristiques $\psi_n$ convergent simplement vers une fonction $\psi$ continue en 0 est tendue. Soit $n \geq 1$ et $a > 0$

$$\begin{aligned}
\frac{1}{2a}\int_{-a}^{a}(1-\psi_n(t))\ dt &= \frac{1}{2a}\int_{-a}^{a}(1-\psi_n(t))\ dt \\
&= \frac{1}{2a}\int_{-a}^{a}\int(1-e^{itx})\ d\nu_n(x)\ dt \\
&= \int\frac{1}{2a}\int_{-a}^{a}(1-e^{itx})\ dt\ d\nu_n(x) \\
&= \int 1-\frac{\sin ax}{ax}\ d\nu_n(x)
\end{aligned}$$

La fonction $\theta \mapsto 1 - \frac{\sin\theta}{\theta}$ est positive et pour $|\theta| \geq \pi/2$, on a $\frac{\sin\theta}{\theta} \leq \frac{2}{\pi}$, d'où

$$\begin{aligned}
\frac{1}{2a}\int_{-a}^{a}(1-\psi_n(t))\ dt &= \int 1-\frac{\sin ax}{ax}\ d\nu_n(x) \\
&\geq \int \mathbb{1}_{|ax|\geq\pi/2}(1-\frac{2}{\pi}\ d\nu_n(x) \\
&\geq (1-\frac{2}{\pi})\nu_n(\{x\in\mathbb{R} : |x| \geq \frac{\pi}{2a}\})
\end{aligned}$$

Soit maintenant $\varepsilon > 0$ : on va montrer qu'il existe un compact qui porte à $\varepsilon$ près la charge de chaque $\nu_n$. Comme $\psi$ est continue en 0, il existe $a > 0$ tel que

$$\sup_{t\in[-a,a]}|1-\psi(t)| \leq \frac{\varepsilon}{2}(1-\frac{2}{\pi}),$$

ce qui entraîne

$$\frac{1}{2a}\int_{-a}^{a}(1-\psi(t))\ dt \leq \frac{\varepsilon}{2}(1-\frac{2}{\pi}).$$

D'autre part, d'après le théorème de convergence dominée, $\frac{1}{2a}\int_{-a}^{a}(1-\psi_n(t))\ dt$ converge vers $\frac{1}{2a}\int_{-a}^{a}(1-\psi(t))\ dt$, donc il existe $n_0$ tel que pour $n > n_0$, on ait

$$\frac{1}{2a}\int_{-a}^{a}(1-\psi(t))\ dt \leq \varepsilon(1-\frac{2}{\pi}),$$

ce qui entraîne $\nu_n(\{x\in\mathbb{R} : |x| \geq \frac{\pi}{2a}\}) \leq \varepsilon$. La famille $(\nu_n)_{n\leq n_0}$ est finie, donc elle est tendue : il existe un compact $K$ tel que pour tout $n \leq n_0$, on ait $\nu_n(K) \geq 1-\varepsilon$. Enfin $K' = K \cup [-\frac{\pi}{2a}, \frac{\pi}{2a}]$ est compact et pour tout $n \geq 1$, on a $\nu_n(K') \geq 1-\varepsilon$, ce qui achève la preuve. $\square$

## 1.4 Exercices

1. *Existence des lois stables d'indice $\alpha$*
   Soient $Y_{n,k}$ une suite de variables aléatoires indépendantes suivant la loi de Poisson $\mathcal{P}(cn^\alpha/|k|^{1+\alpha})$, où $c$ est une constante positive et $\alpha$ un réel vérifiant $0 < \alpha < 2$. On pose
$$Z_n = \frac{1}{n} \sum_{k=-n^2}^{n^2} kY_{n,k}.$$
   (a) Montrer que la fonction caractéristique de $Z_n$ peut s'écrire $\varphi_{Z_n}(\theta) = \exp(-2c\theta^\alpha u_n(\theta))$, avec
$$u_n(\theta) = \int_0^{n|\theta|} f(\frac{\theta}{n} \ulcorner \frac{nx}{\theta} \urcorner)\, dx,$$
   où $f(x) = \frac{1-\cos x}{x^{1+\alpha}}$.
   (b) Soit $A > 0$ et $\theta > 0$. À l'aide du théorème de convergence dominée, démontrer la convergence de
$$\int_A^{n\theta} f(\frac{\theta}{n} \ulcorner \frac{nx}{\theta} \urcorner)\, dx$$
   lorsque $n$ tend vers l'infini.
   (c) Montrer que pour tout $x > 0$, on a
$$f'(x) = \sum_{k=1}^{+\infty} (-1)^{k-1} \frac{2k-\alpha-1}{(2k)!} x^{2k-\alpha-2}.$$
   En déduire l'existence d'un réel $A$ tel que $f$ est monotone sur $]0, A[$.
   (d) Montrer que pour un choix approprié de $c$, la suite de fonctions $\varphi_{Z_n}$ converge vers $\varphi(\theta) = \exp(-|\theta|^\alpha)$.
   (e) Soit $\alpha$ avec $0 < \alpha \leq 2$. Montrer qu'il existe une mesure $m_\alpha$ dont la fonction caractéristique soit la fonction $\theta \mapsto \exp(-|\theta|^\alpha)$.
   (f) Soient $X_1, \dots X_n$ des variables aléatoires indépendantes identiquement distribuées suivant la loi $m_\alpha$. Montrer qu'il existe $\lambda_n$ tel que $\lambda_n(X_1 + X_2 + \dots + X_n)$ suive la loi $m_\alpha$.
2. Déterminer un réel $\lambda_n$ tel que la mesure $\mu_n$ dont le support est $\{-n, -(n-1), \dots, n-1, n\}$ et vérifiant
$$\forall k \in \{-n, -(n-1), \dots, n-1, n\} \quad \mu(k) = \lambda_n(2n + 1 - 2|k|)$$
   soit une mesure de probabilité. Soit $X_n$ suivant la loi $\mu_n$. Montrer que $X_n/n$ converge en loi. (Indication : on pourra déterminer $\nu_n$ telle que $\mu_n = \nu_n * \nu_n$.

# Chapitre 2

# Martingales

## 2.1 Définitions

Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé.

**Définition:** On appelle filtration toute suite croissante $(\mathcal{F}_n)_{n\geq 0}$ de sous-tribus de $\mathcal{F}$.

**Définition:** Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires et $(\mathcal{F}_n)_{n\geq 0}$ une filtration. On dit que la suite $(X_n)_{n\geq 0}$ est $(\mathcal{F}_n)_{n\geq 0}$ adaptée si pour tout $n$, $X_n$ est $\mathcal{F}_n$-mesurable.

**Définition:** Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires. On appelle filtration naturelle adaptée à la suite $(X_n)_{n\geq 0}$ la filtration définie par $\mathcal{F}_n = \sigma(X_0, X_1, \ldots X_n)$.

**Définition:** Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une suite de variables aléatoires. On dit que la suite $(X_n)_{n\geq 0}$ est une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$ si

1. la suite $(X_n)_{n\geq 0}$ est $(\mathcal{F}_n)_{n\geq 0}$ adaptée
2. Pour tout $n$, $X_n$ est intégrable.
3. Pour tout $n$, $X_n = \mathbb{E}[X_{n+1}|\mathcal{F}_n]$.

**Exemples:**

1. Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration, $X$ une variable aléatoire intégrable. La suite définie par $X_n = \mathbb{E}[X|\mathcal{F}_n]$ est une martingale
2. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes centrées. On pose pour tout $n \geq 1$ : $\mathcal{F}_n = \sigma(X_1, \ldots X_n)$ et $S_n = X_1 + X_2 + \ldots X_n$. Alors, $(S_n)_{n\geq 1}$ est une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 1}$.

**Remarque:** Une martingale est toujours adaptée à sa filtration naturelle.

**Définition:** Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une suite de variables aléatoires. On dit que la suite $(X_n)_{n\geq 0}$ est une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$ si

1. la suite $(X_n)_{n\geq 0}$ est $(\mathcal{F}_n)_{n\geq 0}$ adaptée
2. Pour tout $n$, $X_n$ est intégrable.
3. Pour tout $n$, $X_n \leq \mathbb{E}[X_{n+1}|\mathcal{F}_n]$.

**Définition:** Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une suite de variables aléatoires. On dit que la suite $(X_n)_{n\geq 0}$ est une surmartingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$ si

1. la suite $(X_n)_{n\geq 0}$ est $(\mathcal{F}_n)_{n\geq 0}$ adaptée
2. Pour tout $n$, $X_n$ est intégrable.
3. Pour tout $n$, $X_n \geq \mathbb{E}[X_{n+1}|\mathcal{F}_n]$.

**Proposition 1.** *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires intégrables. La suite $(\mathbb{E}X_n)_{n\geq 0}$ est*
- *décroissante si la suite $(X_n)_{n\geq 0}$ est une surmartingale.*
- *croissante si la suite $(X_n)_{n\geq 0}$ est une sous-martingale.*
- *constante si la suite $(X_n)_{n\geq 0}$ est une martingale.*

*Démonstration.* On va juste prouver la première assertion. Pour tout $n$, on a $X_n \geq \mathbb{E}[X_{n+1}|\mathcal{F}_n]$.
En prenant l'espérance, on a $\mathbb{E}[X_n] \geq \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n]] = \mathbb{E}[X_{n+1}]$. □

## 2.2 Premières inégalités

### 2.2.1 Martingales et fonctions convexes

**Théorème 4.** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $\varphi$ une fonction convexe.*
- *Si la suite $(X_n)_{n\geq 0}$ est une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$ et que les $(\varphi(X_n))_{n\geq 0}$ sont intégrables, alors la suite la suite $(\varphi(X_n))_{n\geq 0}$ est une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*
- *Si la suite $(X_n)_{n\geq 0}$ est une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$, que $\varphi$ est croissante et que les $(\varphi(X_n))_{n\geq 0}$ sont intégrables, alors la suite la suite $(\varphi(X_n))_{n\geq 0}$ est une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

*Démonstration.* – Comme $X_n = \mathbb{E}[X_{n+1}|\mathcal{F}_n]$, on a $\varphi(X_n) = \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) \leq \mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n]$, d'après l'inégalité de Jensen conditionnelle.

– $X_n \leq \mathbb{E}[X_{n+1}|\mathcal{F}_n]$ entraîne, avec l'hypothèse de croissance $\varphi(X_n) \leq \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n])$ et on conclut comme précédemment avec l'inégalité de Jensen conditionnelle.

□

**Exemple:** En particulier, si une suite $(X_n)_{n\geq 0}$ de variables aléatoires positives est une sous-martingale de carré intégrable, alors la suite $(X_n^2)_{n\geq 0}$ est une sous-martingale.

### 2.2.2 Inégalité de Kolmogorov

**Théorème 5.** *Soit $(X_n)_{n\geq 0}$ une sous-martingale. Pour tout $\alpha > 0$, on a*

$$P(\max_{1\leq i\leq n} X_i \geq \alpha) \leq \frac{1}{\alpha}\mathbb{E}|X_n|.$$

*Démonstration.* Notons $\tau = \inf\{i \geq 1; X_i \geq \alpha\}$. Il est clair que

$$\{\max_{1\leq i\leq n} X_i \geq \alpha\} = \{\tau \leq n\}.$$

Soit $k$ entre 1 et $n$ : l'événement $\tau = k$ est $\mathcal{F}_k$-mesurable, donc d'après la propriété de sous-martingale, on a

$$\mathbb{E}[X_n \mathbb{1}_{\{\tau=k\}}] = \mathbb{E}\ \mathbb{E}[X_n \mathbb{1}_{\{\tau=k\}}|\mathcal{F}_k] = \mathbb{E}\ \mathbb{1}_{\{\tau=k\}}\mathbb{E}[X_n|\mathcal{F}_k] \geq \mathbb{E}\ \mathbb{1}_{\{\tau=k\}}X_k.$$

Mais $\mathbb{1}_{\{\tau=k\}}X_k \geq \alpha\mathbb{1}_{\{\tau=k\}}$. Ainsi, en intégrant, on obtient

$$\mathbb{E}[X_n \mathbb{1}_{\{\tau=k\}}] \geq \alpha P(\tau = k).$$

En faisant la somme pour $k$ variant de 1 à $n$, on obtient

$$\mathbb{E}[X_n \mathbb{1}_{\{\tau\leq n\}}] \geq \alpha P(\tau \leq n).$$

Si $(X_n)_{n\geq 1}$ est une sous-martingale positive, on a fini, car alors

$$\mathbb{E}X_n \geq \mathbb{E}[X_n \mathbb{1}_{\{\tau\leq n\}}] \geq \alpha P(\tau \leq n).$$

Sinon, comme $(X_n^+)_{n\geq 1}$ est une sous-martingale positive, on peut lui appliquer le résultat que l'on vient de démontrer, et l'on a

$$P(\max_{1\leq i\leq n} X_i \geq \alpha) = P(\max_{1\leq i\leq n} X_i^+ \geq \alpha) \leq \frac{1}{\alpha}\mathbb{E}X_n^+ \leq \frac{1}{\alpha}\mathbb{E}|X_n|,$$

ce qui donne le résultat voulu. □

## 2.3 Convergence des martingales de carré intégrable

**Théorème 6.** *Soit $(X_n)_{n\geq 0}$ une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 1}$ telle que*
$$\sup_{n\geq 1} \mathbb{E}X_n^2 < +\infty.$$
*Alors $(X_n)_{n\geq 0}$ converge presque sûrement et dans $L^2$ vers une variable $X_\infty$ de carré intégrable.*

*Démonstration.* La convergence quadratique s'obtient par des méthodes hilbertiennes classiques : comme $L^2$ est complet, il suffit en effet de montrer que la suite $(X_n)_{n\geq 0}$ est de Cauchy. Soit $p < n$ entiers. Comme $X_n$ est dans $L^2$, on sait que $\mathbb{E}[X_n|\mathcal{F}_p] = X_p$ est le projeté orthogonal de $X_n$ sur le sous-espace des variables $\mathcal{F}_p$-mesurables. Ainsi, on peut écrire l'identité de Pythagore :
$$\|X_n\|_2^2 = \|X_p\|_2^2 + \|X_n - X_p\|_2^2,$$
ou encore
$$\mathbb{E}X_n^2 = \mathbb{E}X_p^2 + \mathbb{E}(X_n - X_p)^2.$$
Il est alors clair que la suite $(\mathbb{E}X_n^2)_{n\geq 1}$ est croissante : elle converge donc vers $\alpha = \sup_{n\geq 1} \mathbb{E}X_n^2$ que nous avons supposé fini. Soit $\varepsilon > 0$ et $N$ tel que $\alpha \geq \mathbb{E}X_n^2 \geq \alpha - \varepsilon^2$ pour $n \geq N$. Alors, pour $n, p \geq N$, on a $\|X_n - X_p\|_2 \leq \varepsilon$, ce qui contre bien que la suite est de Cauchy, et donc convergente.

On va maintenant montrer la convergence presque sûre. Pour cela, on va montrer que la suite $(X_n)_{n\geq 1}$ est presque sûrement de Cauchy, c'est à dire que $R_n = \sup_{i,j\geq n} |X_i - X_j|$ tend presque sûrement vers 0. Comme la suite $(R_n)_{n\geq}$ est monotone décroissante, il suffit de montrer qu'elle admet une sous-suite qui converge presque sûrement vers 0. Pour démontrer que $(R_n)_{n\geq}$ admet une sous-suite qui converge presque sûrement vers zéro, il suffit (voir le cours de licence) de montrer que $R_n$ converge en probabilité vers 0.

Soit donc $\varepsilon > 0$. On a
$$\{R_n > \varepsilon\} \subset \cup_{i\geq n}\{|X_n - X_i| > \varepsilon/2\}.$$
Ainsi
$$\begin{aligned} P(R_n > \varepsilon) &\leq P(\sup_{i\geq n} |X_n - X_i| \geq \varepsilon/2) \\ &\leq P(\sup_{i\geq n} |X_n - X_i| > \varepsilon/3). \end{aligned}$$
D'après le théorème de continuité séquentielle croissante, on a
$$P(\sup_{i\geq n} |X_n - X_i| > \varepsilon/3) = \sup_{N\geq n} P(\sup_{n\leq i\leq N} |X_n - X_i| > \varepsilon/3).$$

La suite $(X_n - X_i)_{n\geq i}$ est une martingale, donc la suite $((X_n - X_i)^2)_{n\geq i}$ est une sous-martingale positive : on a donc

$$P(\sup_{n\leq i\leq N} |X_n - X_i|^2 > \varepsilon^2/9) \leq \frac{9}{\varepsilon^2}\mathbb{E}(X_n - X_N)^2.$$

Ainsi

$$P(R_n > \varepsilon) \leq \frac{9}{\varepsilon^2} \sup_{N\geq n} \mathbb{E}(X_n - X_N)^2,$$

cette dernière suite tend bien vers 0, puisque $(X_n)_{n\geq 1}$ converge en moyenne quadratique.

□

## 2.4 Temps d'arrêts

On dit que variable aléatoire $T$ à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est un *temps d'arrêt* adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$ si pour tout $n \in \mathbb{N}$, l'événement $T \leq n$ est $\mathcal{F}_n$-mesurable.

Comme $\{T = n\} = \{T \leq n\}\backslash\{T \leq n-1\}$, il s'ensuit que $T = n$ est également $\mathcal{F}_n$-mesurable.

**Exemple:** Toute constante est un temps d'arrêt adapté à toute filtration.

*Démonstration.* Si $T$ est constant, $\{T \leq n\}$ ne peut valoir que $\Omega$ ou $\varnothing$, et est donc toujours dans $\mathcal{F}_n$. □

**Exemple:** Si $X_0, \dots, X_n$ est une suite de variables aléatoires à valeurs dans $S$ et $A$ un borélien de $S$, alors

$$T_A = \inf\{n \geq 1; X_n \in A\}$$

est un temps d'arrêt.

Preuve : $\{T_A \leq n\} = \cup_{k=1}^n \{X_k \in A\}$.

**Définition:** On dit qu'un événement $A$ se produit avant $T$ si pour tout $n$, l'événement $A \cap \{T \leq n\}$ est $\mathcal{F}_n$-mesurable.

**Remarque:** Si deux temps d'arrêts $S$ et $T$ adaptés à la filtration $(\mathcal{F}_n)_{n\geq 0}$ vérifient $S \leq T$, alors tout événement qui se produit avant $S$ se produit avant $T$.

*Démonstration.* Soit $A$ se produisant avant $S$. Comme $S \leq T$, on a $\{T \leq n\} \cap A = (A \cap \{S \leq n\}) \cap \{T \leq n\}$. Comme $A$ se produit avant $S$, $A \cap \{S \leq n\} \in \mathcal{F}_n$. Par définition d'un temps d'arrêt, $\{T \leq n\} \in \mathcal{F}_n$. On en déduit que $\{T \leq n\} \cap A \in \mathcal{F}_n$. Comme $n$ est quelconque, $A$ se produit avant $T$. □

**Proposition 2.** *Soit $T$ un temps d'arrêt adapté à une filtration $(\mathcal{F}_n)_{n\geq 0}$. L'ensemble $\mathcal{F}_T$ des événements qui se produisent avant $T$ forme une tribu.*

*Démonstration.* – Montrons que $\varnothing \in \mathcal{F}_T$. Pour tout $n$, on a $\varnothing \cap \{T \leq n\} = \varnothing \in \mathcal{F}_n$, car toute tribu contient $\varnothing$ donc on a bien $\varnothing \in \mathcal{F}_T$.
- Soit $A \in \mathcal{F}_T$. Montrons que $A^c \in \mathcal{F}_T$. Soit $n$ entier. On a $A^c \cap \{T \leq n\} = \{T \leq n\} \backslash (A \cap \{T \leq n\})$. Les événements $\{T \leq n\}$ et $A \cap \{T \leq n\}$ sont tous deux $\mathcal{F}_n$-mesurables, donc $A^c \cap \{T \leq n\}$ est bien dans $\mathcal{F}_n$. Comme $n$ est quelconque, $A^c \in \mathcal{F}_T$.
- Soit $(A_p)_{p\geq 1}$ une suite d'éléments de $\mathcal{F}_T$. Il faut montrer que $A = \sup_{p\geq 1} A_p \in \mathcal{F}_T$ Soit $n$ entier. On a

$$A \cap \{T \leq n\} = \cup_{p\geq 1}(A_p \cap \{T \leq n\}),$$

  $A$ est réunion dénombrable d'éléments de $\mathcal{F}_n$, donc $A \in \mathcal{F}_n$, et finalement $A \in \mathcal{F}_T$.

□

**Théorème 7 (Théorème de Hunt).** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

*Soient $S$ et $T$ deux temps d'arrêts $(\mathcal{F}_n)_{n\geq 0}$-adaptés bornés avec $S \leq T$. Alors*

$$E[X_T|\mathcal{F}_S] \geq X_S.$$

*Démonstration.* Il s'agit de montrer que pour tout événement $A$ $\mathcal{F}_S$-mesurable, on a $\mathbb{E}X_T 1\!\!1_A \geq \mathbb{E}X_S 1\!\!1_A$. Soit $M$ un entier déterministe tel que l'on ait $S \leq T \leq M$. Posons $\Delta_k = X_k - X_{k-1}$, avec $X_{-1} = 0$. Notons, que comme $(X_n)$ est une sous-martingale $\mathbb{E}1\!\!1_B \Delta_k$ est positif pour tout ensemble $B$ $\mathcal{F}_{k-1}$-mesurable. On a

$$\begin{aligned} \mathbb{E}((X_T - X_S)1\!\!1_A) &= \mathbb{E}(\sum_{k=0}^{M} \Delta_k 1\!\!1_{S<k\leq T} 1\!\!1_A) \\ &= \sum_{k=0}^{M} \mathbb{E}\Delta_k 1\!\!1_{\{S<k\leq T\}\cap A} \end{aligned}$$

Pour conclure, il reste donc à voir que $\{S < k \leq T\} \cap A$ est $\mathcal{F}_{k-1}$-mesurable, ce qui vient de l'identité

$$\{S < k \leq T\} \cap A = (\{S \leq k-1\} \cap A) \cap (T \leq k-1)^c.$$

□

On en déduit facilement les deux résultats suivants :

**Corollaire 3.** *Soit* $(\mathcal{F}_n)_{n\geq 0}$ *une filtration et* $(X_n)_{n\geq 0}$ *une surmartingale relative à la filtration* $(\mathcal{F}_n)_{n\geq 0}$.

*Soient* $S$ *et* $T$ *deux temps d'arrêts* $(\mathcal{F}_n)_{n\geq 0}$*-adaptés bornés avec* $S \leq T$. *Alors*

$$E[X_T|\mathcal{F}_S] \leq X_S.$$

**Corollaire 4.** *Soit* $(\mathcal{F}_n)_{n\geq 0}$ *une filtration et* $(X_n)_{n\geq 0}$ *une martingale relative à la filtration* $(\mathcal{F}_n)_{n\geq 0}$.

*Soient* $S$ *et* $T$ *deux temps d'arrêts* $(\mathcal{F}_n)_{n\geq 0}$*-adaptés bornés avec* $S \leq T$. *Alors*

$$E[X_T|\mathcal{F}_S] = X_S.$$

**Théorème 8 (Théorème d'arrêt).** *Soit* $(\mathcal{F}_n)_{n\geq 0}$ *une filtration et* $(X_n)_{n\geq 0}$ *une sous-martingale relative à la filtration* $(\mathcal{F}_n)_{n\geq 0}$. *Soit* $T$ *un temps d'arrêt adapté à la filtration* $(\mathcal{F}_n)_{n\geq 0}$. *Alors la suite* $(X_{n\wedge T})_{n\geq 0}$ *est une sous-martingale adaptée à la filtration* $(\mathcal{F}_n)_{n\geq 0}$.

*Démonstration.* Soit $A$ un événement $\mathcal{F}_n$-mesurable et $n$ un entier. On a

$$\begin{aligned}\mathbb{E}1\!\!1_A X_{n+1\wedge T} &= \mathbb{E}1\!\!1_A X_{n+1\wedge T}1\!\!1_{\{T\leq n\}} + \mathbb{E}1\!\!1_A X_{n+1\wedge T}1\!\!1_{\{T>n\}} \\ &= \mathbb{E}X_{n\wedge T}1\!\!1_A 1\!\!1_{\{T\leq n\}} + \mathbb{E}X_{(n+1)\wedge T}1\!\!1_A 1\!\!1_{\{T>n\}}\end{aligned}$$

Pour l'instant, on a juste utilisé que quand $T \leq n$, on a $(n+1)\wedge T = T = n \wedge T$.

Montrons que $A \cap \{T > n\}$ est $\mathcal{F}_{n\wedge T}$-mesurable : soit $p$ un entier ; on doit montrer que $A \cap \{T > n\} \cap \{n \wedge T \leq p\}$ est dans $\mathcal{F}_p$. Si $n > p$, l'intersection est l'ensemble vide, donc est $\mathcal{F}_p$-mesurable. Si $n \leq p$, alors

$$A \cap \{T > n\} \cap \{n \wedge T \leq p\} = A \cap \{T > n\} \in \mathcal{F}_n \subset \mathcal{F}_p.$$

Ainsi, en appliquant le théorème de Hunt aux temps d'arrêts $(n+1)\wedge T$ et $n \wedge T$, on a

$$\mathbb{E}X_{(n+1)\wedge T}1\!\!1_A 1\!\!1_{\{T>n\}} \geq \mathbb{E}X_{n\wedge T}1\!\!1_A 1\!\!1_{\{T>n\}}$$

D'où

$$\begin{aligned}A \cap \{T \leq n\} \cap \{n \wedge T \leq p\} &= \mathbb{E}X_{n\wedge T}1\!\!1_A 1\!\!1_{\{T\leq n\}} + \mathbb{E}X_{(n+1)\wedge T}1\!\!1_A 1\!\!1_{\{T>n\}} \\ &\geq \mathbb{E}X_{n\wedge T}1\!\!1_A 1\!\!1_{\{T\leq n\}} + \mathbb{E}X_{n\wedge T}1\!\!1_A 1\!\!1_{\{T>n\}} \\ &\geq \mathbb{E}X_{n\wedge T}1\!\!1_A\end{aligned}$$

ce qui achève la preuve.

□

On en déduit facilement les deux résultats suivants :

**Corollaire 5.** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une surmartingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Soit $T$ un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Alors la suite $(X_{n\wedge T})_{n\geq 0}$ est une surmartingale adaptée à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

**Corollaire 6.** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Soit $T$ un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Alors la suite $(X_{n\wedge T})_{n\geq 0}$ est une martingale adaptée à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

**Remarque:** Certains auteurs appellent « théorème d'arrêt »ce que nous avons appelé « théorème de Hunt ». En fait, ces deux résultats sont équivalents. Ici, nous avons choisi de démontrer le théorème de Hunt et d'en déduire le théorème d'arrêt, tandis que d'autres auteurs (par exemple Baldi, Mazliak et Priouret) font le choix inverse.

Les deux lemmes suivants seront utiles par la suite. Leur preuve relève des méthodes classiques.

**Lemme 4.** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Soit $T$ un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Alors, la variable aléatoire $1\!\!1_{\{T<+\infty\}}X_T$ est $\mathcal{F}_T$-mesurable.*

*Démonstration.* Il suffit de voir que pour tout borélien de $\mathbb{R}$, l'événement $\{1\!\!1_{\{T<+\infty\}}X_T \in A\}$ est $\mathcal{F}_T$-mesurable. Soit $n$ un entier : on a

$$\{1\!\!1_{\{T<+\infty\}}X_T \in A\} \cap \{T \leq n\} = \cup_{i\leq n}\{X_i \in A; T = i\},$$

qui est bien $\mathcal{F}_n$-mesurable. □

**Lemme 5.** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(A_n)_{n\geq 0}$ une suite d'événement adaptée à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Soit $T$ un temps d'arrêt adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Alors, la variable aléatoire $S$ définie par*

$$S(\omega) = \inf\{n \geq T(\omega); \omega \in A_n\}.$$

*est un temps d'arrêt.*

*Démonstration.* Il suffit de remarquer que pour tout $n$

$$S \leq n = \cup_{i\leq n} A_i \cap \{T \leq i\},$$

qui est clairement $\mathcal{F}_n$-mesurable. □

## 2.5 Convergence $L^1$ des martingales

### 2.5.1 Théorème des traversées montantes

**Théorème 9.** *Soit $(X_n)_{n\geq 1}$ une sous-martingale $a, b$ teux réels avec $a < b$. On note $U_n^{[a,b]}$ le nombre de nombre de traversées montantes de $a$ à $b$ entre les instants 1 et $n$. Alors, on a*

$$\mathbb{E}U_n^{[a,b]} \leq \frac{\mathbb{E}(X_n - a)^+}{b - a}$$

*Démonstration.* Posons pour $n \geq 1 : Y_n = \frac{\mathbb{E}(X_n - a)^+}{b-a}$ : comme $|Y_n| \leq \frac{1}{b-a}(|X_n| + a)$ et que $\varphi(x) = \frac{(x-a)^+}{b-a}$ est croissante et convexe, $(Y_n)_{n\geq 1}$ est une sous-martingale.

Posons $\tau_0 = 0$, et pour $k \geq 1$

$$\sigma_k = \inf\{n \geq \tau_{k-1}; X_n \leq a\}$$

et

$$\tau_k = \inf\{n \geq \sigma_k; X_n \geq b\}.$$

En utilisant le lemme 5, on voit facilement par récurrence que les $(\sigma_k)$ et les $(\tau_k)$ sont des temps d'arrêt. D'autre part, par construction, on a $\sigma_0 \leq \tau_0 \leq \sigma_1 \leq \tau_1 \leq \sigma_2 \leq \cdots \leq \sigma_n \leq \tau_n$.

On a

$$U_n^{[a,b]} = \sum_{k=1}^{n} \mathbb{1}_{\{\tau_k \leq n\}}.$$

Montrons que $\mathbb{1}_{\{\tau_k \leq n\}} \leq Y_{n\wedge\tau_k} - Y_{n\wedge\sigma_k}$ : il y a trois cas possibles

– si $\sigma_k > n$, alors $\tau_k > n$ : on a

$$\mathbb{1}_{\{\tau_k \leq n\}} = 0 = Y_n - Y - n =\leq Y_{n\wedge\tau_k} - Y_{n\wedge\sigma_k}.$$

– si $\sigma_k \leq \tau_k \leq n$, alors on a $Y_{n\wedge\tau_k} = Y_{\tau_k} \geq 1$ tandis que $Y_{n\wedge\sigma_k} = Y_{\sigma_k} = 0$, de sorte que

$$\mathbb{1}_{\{\tau_k \leq n\}} = 1 \leq Y_{n\wedge\tau_k} - Y_{n\wedge\sigma_k}.$$

– si $\sigma_k \leq n < \tau_k$, alors $Y_{n\wedge\tau_k} = Y_n \geq 0$ tandis que $Y_{n\wedge\sigma_k} = Y_{\sigma_k} = 0$, de sorte que

$$\mathbb{1}_{\{\tau_k \leq n\}} = 0 \leq Y_{n\wedge\tau_k} - Y_{n\wedge\sigma_k}.$$

On a donc

$$U_n^{[a,b]} \leq \sum_{k=1}^{n} Y_{n\wedge\tau_k} - Y_{n\wedge\sigma_k}.$$

On peut écrire

$$\begin{aligned} Y_n - Y_0 &= Y_{\tau_n \wedge n} - Y_{\tau_0 \wedge n} \\ &= \sum_{k=1}^{n} Y_{\tau_k \inf n} - Y_{\tau_{k-1} \inf n} \\ &= \sum_{k=1}^{n} (Y_{\tau_k \wedge n} - Y_{\sigma_k \wedge n}) + (Y_{\sigma_k \wedge n} - Y_{\tau_{k-1} \wedge n}) \end{aligned}$$

Il s'ensuit que

$$\begin{aligned} U_n^{[a,b]} &\le Y_n - Y_0 - \sum_{k=1}^{n} Y_{\sigma_k \wedge n} - Y_{\tau_{k-1} \wedge n} \\ &\le Y_n - \sum_{k=1}^{n} Y_{\sigma_k \wedge n} - Y_{\tau_{k-1} \wedge n}, \end{aligned}$$

soit, en prenant l'espérance

$$\mathbb{E} U_n^{[a,b]} \le \mathbb{E} Y_n - \sum_{k=1}^{n} \mathbb{E} Y_{\sigma_k \wedge n} - Y_{\tau_{k-1} \wedge n}.$$

Mais $\sigma_k \wedge n$ et $\tau_{k-1} \wedge n$ sont des temps d'arrêts bornés avec $\tau_{k-1} \le \sigma_k$. D'après le théorème de Hunt, on a donc, en réintégrant, $\mathbb{E} Y_{\sigma_k \wedge n} - Y_{\tau_{k-1} \wedge n} \ge 0$, d'où

$$\mathbb{E} U_n^{[a,b]} \le \mathbb{E} Y_n,$$

ce qui était le résultat voulu. □

### 2.5.2 Le théorème de convergence de Doob

**Théorème 10.** *Soit $(X_n)_{n\ge1}$ une sous-martingale. Si $K = \sup_{n\ge1} \mathbb{E}|X_n| < +\infty$, alors la suite $(X_n)_{n\ge1}$ converge cresque surement vers une variable aléatoire $X$ qui est telle que $\mathbb{E}|X| \le K$.*

*Démonstration.* Faisons d'abord une remarque d'analyse : si une suite $(x_n)_{n\ge1}$ ne converge pas dans $\overline{\mathbb{R}}$, alors il existe deux rationnels $a$ et $b$ tels que la suite $(x_n)_{n\ge1}$ traverse une infinité de fois l'intervalle $[a,b]$ de bas en haut : pour cela, il suffit de considérer deux rationnels $a$ et $b$ vérifiant $\underline{\lim}_{n\to+\infty} x_n < a < b < \overline{\lim}_{n\to+\infty} x_n$. Ainsi, si l'on note $U^{[a,b]}$ le nombre de traversées de l'intervalle $[a,b]$, on a

$$\{X_n \text{ ne converge pas dans } \overline{\mathbb{R}}\} \subset \bigcup_{a<b,(a,b)\in\mathbb{Q}^2} \{U^{[a,b]} = +\infty\}.$$

Cette réunion est dénombrable. Pour achever la preuve, il suffit donc de montrer que pour tout couple $(a,b)$ avec $a<b$, on a

$$P(U^{[a,b]}=+\infty)=0.$$

Cependant on a par convergence monotone $\mathbb{E}U^{[a,b]} = \lim_{n\to+\infty} EU^{[a,b]} \le \frac{K+|a|}{b-a}$ d'après le théorème des traversées montantes et la définition de $K$. Comme $U^{[a,b]}$ est intégrable, elle est presque sûrement finie.

Ainsi $(X_n)_{n\ge 1}$ converge presque sûrement vers un élément $X$ de $\overline{\mathbb{R}}$. D'après le lemme de Fatou, on a $\mathbb{E}|X| = \mathbb{E} \varliminf_{n\to+\infty} |X_n| \le \varliminf_{n\to+\infty} \mathbb{E}|X_n| \le K$. Ainsi, $|X|$ est intégrable, ce qui implique qu'elle prend presque surement ses valeurs dans $\mathbb{R}$.

□

## 2.6 Décomposition de Doob (*)

**Définition:** On dit qu'un processus $(\mathcal{F}_n)_{n\ge 0}$- adapté $(C_n)_{n\ge 0}$ est un processus croissant prévisible si $C_0=0$, $C_n\le C_{n+1}$ et si $C_{n+1}$ est $\mathcal{F}_n$-mesurable.

**Théorème 11.** *Toute sous-martingale $(X_n)_{ge0}$ s'écrit de manière unique comme somme d'une martingale $(M_n)_{n\ge 0}$ et d'un processus croissant prévisible intégrable $(C_n)_{n\ge 0}$.*

*Démonstration.* Supposons qu'une telle décomposition existe : on a alors

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

car $(M_n)_{n\ge 0}$ est une martingale et $C_{n+1}-C_n$ est $\mathcal{F}_n$-mesurable. Comme $C_0=0$, on doit nécessairement avoir

$$C_n=\sum_{i<n}\mathbb{E}[X_{i+1}-X_i|\mathcal{F}_\rangle].$$

Chaque terme de la somme est bien dans $\mathcal{F}_{n-1}$ et est positif, car $(X_n)_{n\ge 0}$ est une sous-martingale, donc $C_n$ est bien croissant prévisible. □

On présente maintenant une jolie application de la Décomposition de Doob. On va donner une nouvelle preuve du théorème de Hunt, basée sur son corollaire relatif aux martingales (lequel peut bien sûr se montrer sans utiliser la version sous-martingale du théorème de Hunt, voir par exemple Stroock).

**Corollaire 7 (Théorème de Hunt, version 2).** *Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration et $(X_n)_{n\geq 0}$ une sous-martingale relative à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

*Soient $S$ et $T$ deux temps d'arrêts $(\mathcal{F}_n)_{n\geq 0}$-adaptés bornés avec $S \leq T$. Alors*

$$E[X_T|\mathcal{F}_S] \geq X_S.$$

*Démonstration.* On écrit la décomposition de Doob $X_n = M_n + A_n$, avec la martingale $(M_n)_{n\geq 0}$ et le processus croissant prévisible intégrable $(C_n)_{n\geq 0}$. $X_T = M_T + C_T \geq M_T + C_S$, car $(C_n)_{n\geq 0}$ est croissant. Donc $\mathbb{E}[X_T|\mathcal{F}_S] \geq \mathbb{E}[M_T|\mathcal{F}_S] + \mathbb{E}[C_S|\mathcal{F}_S]$. D'après le théorème de Hunt sur les martingales, $\mathbb{E}[M_T|\mathcal{F}_S] = M_S$. D'autre part , le processus $(C_n)_{n\geq 0}$ est adapté, donc d'après le lemme 4, $C_S$ est $\mathcal{F}_S$-mesurable et $\mathbb{E}[C_S|\mathcal{F}_S] = C_S$. Finalement, $\mathbb{E}[X_T|\mathcal{F}_S] \geq M_S + C_S = X_S$. □

## 2.7 Exercices

1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées admettant un moment d'ordre 3 et vérifiant $\mathbb{E}X_1 = \mathbb{E}X_1^3 = 0$ et $\mathbb{E}X_1^2 = 1$. On note $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ et on pose

$$S_n = \sum_{k=1}^{n} X_i.$$

   (a) On pose $Y_n = S_n^3 - 3nS_n$. Montrer que $(Y_n)_{n\geq 1}$ est une martingale par rapport à la filtration $(\mathcal{F}_n)_{n\geq 1}$.

   (b) Soient $a, b, c, d$ des réels. On pose

$$Q(x,t) = x^2 + axt + bx + ct + d.$$

   Pour quelles valeurs du quadruplet $(a, b, c, d)$ la suite $Z_n = Q(S_n, n)$ forme-t-elle une martingale rapport à la filtration $(\mathcal{F}_n)_{n\geq 1}$ ?

2. Soit $(X_n)_{n\geq 1}$ une surmartingale $(\mathcal{F}_n)_{n\geq 1}$-adaptée. On suppose que les $(X_n)_{n\geq 1}$ ont toutes la même loi.

   (a) Montrer que $(X_n)_{n\geq 1}$ est une martingale.

   (b) Montrer que pour tout réel $a$ les suites $(X_n \wedge a)_{n\geq 1}$ et $(X_n \vee a)_{n\geq 1}$ sont des martingales.

   (c) En déduire que pour $n > p \geq 1$, $X_n$ est presque sûrement supérieur ou égal à $a$ sur l'événement $\{X_p \geq a\}$.

   (d) En déduire que $(X_n)_{n\geq 1}$ est presque sûrement constante.

3. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées dont la loi commune est non dégénérée et à support compact. On pose $S_n = X_1 + \cdots + X_n$, $\varphi(t) = \log \mathbb{E}e^{tX_1}$ et
$$Y_n^t = e^{tS_n - n\varphi(t)}.$$
   (a) Montrer que $(Y_n^t)_{n\geq 1}$ est une martingale par rapport à la filtration naturelle associée aux $(X_n)_{n\geq 1}$.
   (b) On suppose désormais que $t$ est non-nul. Montrer que $\varphi(t/2) < \varphi(t)/2$.
   (c) En déduire que $(Y_n^t)_{n\geq 1}$ converge presque sûrement vers 0.
   (d) Retrouver ce résultat à partir de la loi forte des grands nombres.
4. Soit $(\mathcal{F}_n)_{n\geq 1}$ une filtration de l'espace $(\Omega, \mathcal{F}, P)$. On note $\mathcal{F}_\infty = \sigma(\cup_{n\geq 1}\mathcal{F}_n))$. Soit $Z$ une variable aléatoire possédant un moment d'ordre 1. On pose $Y_n = \mathbb{E}[Z|\mathcal{F}_n]$.
   (a) Montrer que $(Y_n)_{n\geq 1}$ est une martingale par rapport à la filtration $(\mathcal{F}_n)_{n\geq 1}$.
   (b) Démontrer qu'il existe une variable aléatoire $Y$ intégrable vers laquelle $(Y_n)_{n\geq 1}$ converge presque sûrement.
   (c) Démontrer que la suite $(Y_n)$ est équiintégrable sous $P$, c'est à dire que
$$\lim_{a\to+\infty} \sup_{n\geq 1} \int_{\mathbb{R}\setminus[-a,a]} |x| dP_{Y_n}(x) = 0.$$
   (d) En déduire que $Y_n$ converge en norme 1 vers $Y$.
   (e) Montrer que $Y = \mathbb{E}[Z|\mathcal{F}_\infty]$.
5. Soit $a \in [0, 1/2]$. Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$. On définit par récurrence une suite $X_n$ par $X_0 = a$ et $X_{n+1} = U_{n+1} \cos X_n$. Montrer que $(X_n)_{n\geq 1}$ est bornée, puis que la suite $2^n X_n$ est une sous-martingale.
6. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, et soit $(X_n)_{n\in `N}$ une suite de variables aléatoires intégrables. On note $\mathcal{F}_n$ la tribu engendrée par $X_0, ..., X_n$.

   Dans un pays (fictif), seule la naissance d'un garçon parvient à combler de joie les parents. Si bien qu'après leur premier garçon, ils ne font plus d'enfants. Les lois de ce pays restreignent en outre le nombre d'enfants par famille : pas plus de quatre. Les couples les plus malheureux sont évidemment ceux qui ont eu quatre filles ! Les sociologues de ce pays se disputent. Tous affirment que l'équilibre hommes/femmes est menacé.

Mais ils ne sont pas d'accord sur les conséquences de cette menace. Deux écoles s'affrontent : les "masculo-tendanciels" et les "fémino-tendanciels". Les masculo-tendanciels notent qu'il n'y a presque pas de famille sans garçon et que par conséquent, il va y avoir plus de garçons que de filles. De leur côté, les fémino-tendanciels rétorquent que c'est le contraire qui va se produire : Il n'y a jamais plus d'un garçon par famille alors qu'il peut y avoir une, deux, trois ou quatre filles. Ce sont donc les filles qui vont finir par être trop nombreuses.

Les hypothèses suivantes sont faites :

-Chaque naissance a autant de chance d'être celle d'un garçon ou d'une fille

-le sexe d'un enfant d'une famille ne dépend pas du sexe de l'enfant précédent

-enfin, on se limite aux familles sans jumeaux, triplés, etc ...

En classe de seconde, on tranche la controverse entre masculo-tendanciels et fémino-tendanciels à l'aide d'une simulation sur ordinateur. Ici on va examiner une solution théorique pour ce problème, et d'autres problèmes du même type.

**i)** On considère qu'une famille a en général le potentiel d'avoir un nombre d'enfants $N$ assez important ($N = 4$ p.e.), et qu'un sexe est attribué à chacun de ces enfants potentiels. Pour $0 \leq n \leq N$, on note $F_n$ (resp. $G_n$) le nombre potentiel de filles (resp. de garçons) dans la famille après la n$^{\text{ème}}$ naissance, et on pose $X_n = F_n - G_n$. Montrer que $X_n$ est une martingale.

**ii)** Cette famille décide de cesser de s'agrandir losqu'elle compte un nombre $T$ d'enfants, $T \leq N$. Sa décision ne dépends que du sexe des enfants qui sont déja nés au moment de la prise de décision, ce qu'on peut formaliser par :

$$\{T = n\} \in \sigma(F_1, G_1, F_2, G_2, \ldots, G_n, F_n).$$

Montrer que

$$\mathbb{E}[F_T] = \mathbb{E}[G_T].$$

Vérifier ! que la règle d'arrêt définie en introduction est du type "$\{T = n\} \in \sigma(F_1, G_1, F_2, G_2, \ldots, G_n, F_n)$".

**iii)** Dans le cas où les deux sexes auraient des probabilités différentes $p_F$ et $p_G = 1 - p_F$, vérifier de la même manière que

$$\frac{\mathbb{E}[F_T]}{\mathbb{E}[G_T]} = \frac{p_F}{p_G}.$$

7. *Arrêt optimal pour une marche aléatoire.*
   On considère le jeu suivant. J'ai un pion qui se déplace sur les entiers compris entre 0 et $n$. A chaque point $i \in [1..n-1]$ est associé un somme $f(i)$ strictement positive (on la prolonge par $f(0) = f(n) = 0$). Mon pion part d'un point $i_0 \in [1..n-1]$ ; à chaque étape, je peux décider de partir avec le gain correspondant à ma position actuelle, ou alors lancer une pièce équilibrée qui me donnera ma position suivante (juste à droite si 'pile', juste à gauche si 'face'). Si je touche 0 ou $n$ je ne gagne rien et je suis éliminé. Quelle stratégie adopter ?

   Ce problème revient à trouver un temps d'arrêt $T$ optimal pour la marche aléatoire. Notons $X_n$ la position de mon pion à l'instant $n$, et $\mathcal{F}_n$ la tribu engendrée par $X_0, ..., X_n$.

   (a) Soit $g$ une fonction concave supérieure à $f$. Montrer que $(g(X_n))_{n\in`N}$ est une sous-martingale

   (b) Soit $T$ un temps d'arrêt fini p.s. Montrer que
   $$\mathbb{E}_{i_o}(f(X_T)) \leq \mathbb{E}_{i_o}(g(X_T)) \leq g(i_o).$$

   (c) Notons $\psi$ l'enveloppe concave de $f$. Montrer que
   $$\mathbb{E}_{i_o}(f(X_T)) \leq \psi(i_o).$$

   (d) Soit $T_{opt} = \min\{n \in \mathbb{N}, f(X_n) = \psi(X_n)\}$, $A = \min\{j \geq i_o, f(j) = \psi(j)\}$ et $B = \max\{j \leq i_o, f(j) = \psi(j)\}$. Calculer $\mathbb{E}_{i_o}(f(X_{T_{opt}}))$ en fonction de $f(A)$ et $f(B)$, et en déduire que
   $$\mathbb{E}_{i_o}(f(X_{T_{opt}})) = \psi(i_o).$$

# Chapitre 3

# Loi d'un processus

## 3.1 Loi d'un processus

**Définition:** Un processus stochastique est une famille $(X_t)_{t\in T}$ définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Le plus souvent, $T$ est un ensemble ordonné qui joue le rôle du temps, par exemple $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}$. Le cas $T = \mathbb{Z}^d$, qui évoque plutôt une structure spatiale est également intéressant. Dans ce cours, nous ne considérerons que des processus indexés par un ensemble dénombrable. Aussi tous les théorèmes que nous donnerons seront-ils énoncés dans le où $T = \mathbb{N}$, à charge pour le lecteur sérieux de faire les modifications qui s'imposent...et surtout de consulter des livres plus à même de satisfaire sa curiosité.

Concrètemement pour nous, un processus strochastique ne sera rien d'autre qu'un famille $X = (X_n, n \in \mathbb{N})$ de variables aléatoires définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$

$$\forall n \in \mathbb{N}, X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

**Définition:** On appelle trajectoire de $X$ tout élément $X(\omega) = (X_n(\omega), n \in \mathbb{N}), \omega \in \Omega$.
**Définition:** On définit la tribu borélienne sur $\mathbb{R}^{\mathbb{N}}$, notée $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$, comme étant la plus petite tribu qui rend mesurable les projections $\Pi_i : \mathbb{R}^{\mathbb{N}} \longrightarrow \mathbb{R}, \omega = (\omega_n)_{n\in\mathbb{N}} \longmapsto \omega_i$.
**Définition:** Pour toute partie non vide $S$ et $S'$ de $\mathbb{N}$ telle que $S \supseteq S'$, on appelle projection de $\mathbb{R}^S$ sur $\mathbb{R}^{S'}$ la fonction $\Pi^S_{S'} : \mathbb{R}^S \longrightarrow \mathbb{R}^{S'}$ définie par

$$\forall (x_s, s \in S) \in \mathbb{R}^S, \Pi^S_{S'}(x_s, s \in S) = (x_s, s \in S').$$

**Définition:** Étant donné un processus $X = (X_n, n \in \mathbb{N})$ où

$$\forall n \in \mathbb{N}, X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

$X : \omega \longmapsto X(\omega)$ est une application mesurable de $\Omega$ dans l'espace des trajectoires $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. L'image de $\mathbb{P}$ par $X$ s'appelle loi du processus $X$, notée $P_X$.

Notation : $\mathcal{P}^*(\mathbb{N}), \mathcal{F}(\mathbb{N}), \mathcal{D}(\mathbb{N})$ désignent respectivement l'ensemble des parties non vides, l'ensemble des parties finies non vides et l'ensemble des parties dénombrables non vides de $\mathbb{N}$.

Il est clair que $\mathcal{F}(\mathbb{N}) \subseteq \mathcal{D}(\mathbb{N}) \subseteq \mathcal{P}^*(\mathbb{N})$.

**Définition:** On appelle processus extrait d'un processus $X = (X_n, n \in \mathbb{N})$ tout processus $X_S = (X_n, n \in S), S \in \mathcal{P}^*(\mathbb{N})$.

**Définition:** On appelle loi de dimension finie d'un processus $X = (X_n, n \in \mathbb{N})$ la loi de tout processus extrait $X_S, S \in \mathcal{F}(\mathbb{N})$.

**Proposition 3.** *Les lois de dimension finie d'un processus $X = (X_n, n \in \mathbb{N})$ où*

$$\forall n \in \mathbb{N}, X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

*sont les images de la loi $P_X$ de $X$ par les projections de $\mathbb{R}^{\mathbb{N}}$ sur les espaces-produits finis $\mathbb{R}^S, S \in \mathcal{F}(\mathbb{N})$.*

**Théorème 12.** *Deux processus stochastiques $X = (X_n, n \in \mathbb{N})$ et $X' = (X'_n, n \in \mathbb{N})$ où*

$$\forall n \in \mathbb{N}, X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

*et*

$$X'_n : (\Omega', \mathcal{F}', \mathbb{P}') \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

*ont la même loi si et seulement s'ils ont les mêmes lois de dimension finie.*

*Démonstration.* La condition nécessaire est évidente, d'après la proposition précédente. Réciproquement, si $X$ et $X'$ ont les mêmes lois de dimension finie, d'après la proposition précédente, $P_X$ et $P'_{X'}$ ont les mêmes images par projection sur les espaces-produits de dimension finie $(\mathbb{R}^S, \mathcal{B}(\mathbb{R}^S))$. On considère

$$\mathcal{C} = \{\prod_{n \in \mathbb{N}} A_n, \forall n, A_n \in \mathcal{B}(\mathbb{R}), \text{ et } \exists N, \forall n \geq N A_n = \mathbb{R}\}$$

$P_X$ et $P'_{X'}$ coïncident sur $\mathcal{C}$ , c'est-à-dire $X$ et $X'$ ont même loi sur $\mathcal{C}$. $\mathcal{C}$ est un $\Pi$-système qui engendre $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$, donc $X$ et $X'$ ont la même loi. $\square$

**Définition:** On appelle processus canonique associé à un processus stochastique $X = (X_n, n \in \mathbb{N})$ où $\forall n \in \mathbb{N}, X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ le processus $\Pi = (\Pi_n, n \in \mathbb{N})$ formé par les projections $\Pi_n$ de $\mathbb{R}^{\mathbb{N}}$ sur $\mathbb{R}$, qui à $X = (X_k, k \in \mathbb{N})$ associe $X_n$, sa n-ième composante.

**Théorème 13.** *Tout processus stochastique a même loi que son processus canonique associé, quand on munit l'espace de ses trajectoires de la loi $P_X$.*

*Démonstration.* $\Pi : (X_n)_{n\in\mathbb{N}} \longmapsto (X_n(\omega), \omega \in \Omega)_{n\in\mathbb{N}}$ est l'application identité de $\mathbb{R}^{\mathbb{N}}$. Donc l'image de $P_X$ par $\Pi$ est $P_X$. $\square$

## 3.2 Théorème d'existence de Kolmogorov (admis)

**Définition:** On se place sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. On dit que $(Q_S, S \in \mathcal{F}(\mathbb{N}))$ où pour tout $S$, $Q_S$ désigne une probabilité sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$, est un système projectif de lois si pour tout S, S' de $\mathcal{F}(\mathbb{N})$ tels que $S \supseteq S'$, $Q_{S'}$ est l'image de $Q_S$ par $\Pi^S_{S'}$

**Théorème 14.** *Théorème d'existence de Kolmogorov (admis).*
*On se place sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. Pour toute partie $S \in \mathcal{F}(\mathbb{N})$, soit $Q_S$ une mesure de probabilité sur $(\mathbb{R}^S, \mathcal{B}(\mathbb{R}^S))$.*

1. *pour qu'il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et pour tout $n \in \mathbb{N}$ une variable aléatoire $X_n : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $(Q_S, S \in \mathcal{F}(\mathbb{N}))$ soit l'ensemble des lois de dimension finie du processus $X = (X_n, n \in \mathbb{N})$*
2. *ou pour qu'il existe une mesure de probabilité $Q$ sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ dont l'image par $\Pi_S$ soit $Q_S$ quelle que soit la partie finie, non vide $S$ de $\mathbb{N}$,*
3. *il faut et il suffit que $(Q_S, S \in \mathcal{F}(\mathbb{N}))$ soit un système projectif de lois.*

**Exemple:** Pour tout entier $n \geq 1$, soit $P_n$ une mesure de probabilité sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Pour qu'il existe une suite $X = (X_n, n \geq 1)$ de variables aléatoires réelles simultanées telle que $P_n$ soit la loi de $(X_1, X_2, ..., X_n)$ pour tout $n \geq 1$ il faut et il suffit que

$$\forall B \in \mathcal{B}(\mathbb{R}^n), P_{n+1}(B \times \mathbb{R}) = P_n(B)(*).$$

En effet, s'il existe une telle suite $X$ de variables aléatoires réelles simultanées, on a

$$\forall B \in \mathcal{B}(\mathbb{R}^n), P_{n+1}(B\times\mathbb{R}) = \mathbb{P}((X_1, ..., X_{n+1}) \in B\times\mathbb{R}) = \mathbb{P}((X_1, ..., X_n) \in B) = P_n(B).$$

Réciproquement, on suppose la condition $(*)$ satisfaite. Pour toute partie non vide $S$ de $\mathbb{N}$ de cardinal $p$, notée $S = \{s_1, \ldots, s_p\}$, on note $P_S$ l'image de $P_{\max(S)}$ par la projection $\Pi_S^{\{1,\ldots,\max(S)\}}$. D'après le théorème d'existence de Kolmogorov, il suffit de vérifier que $(P_S, S \in \mathcal{F}(T))$ est un système projectif

de loi. Pour cela, on vérifie que pour tout $S$ de $\mathcal{F}(T)$, et tout $n \geq \max(S)$, $P_S$ est l'image de $P_n$ par $\Pi_S^{\{1,\dots,n\}}$, ce qui se fait facilement par récurrence.
**Exemple:** Étant donnée une suite $(\mu_n, n \geq 1)$ de mesures de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ il existe toujours une suite de variables aléatoires réelles simultanées indépendantes $(X_n, n \geq 1)$ telle que $\forall n \geq 1, P_{X_n} = \mu_n$. En effet, si pour tout $n \geq 1$, on pose $P_n = \mu_1 \otimes \cdots \otimes \mu_n$, la condition $(*)$ est satisfaite puisque $\forall n \geq 1, \forall B \in \mathcal{B}(\mathbb{R}^n)$,

$$P_{n+1}(B \times \mathbb{R}) = \mu_1 \otimes \cdots \otimes \mu_n \otimes \mu_{n+1}(B \times \mathbb{R}) = \mu_1 \otimes \cdots \otimes \mu_n(B)\mu_{n+1}(\mathbb{R}) = P_n(B)$$

D'après l'exemple précédent, il existe une suite $(X_n, n \geq 1)$ de variables aléatoires simultanées telle que

$$\forall n \geq 1, P_{(X_1,\dots,X_n)} = \mu_1 \otimes \cdots \otimes \mu_n$$

On en déduit que $\forall n \geq 1, P_{X_1} = \mu_1, \dots, P_{X_n} = \mu_n$ et que les variables aléatoires réelles $X_1, \dots, X_n$ sont indépendantes.

## 3.3 Processus réels stationnaires

**Définition:** Un processus stochastique réel $(X_n, n \in \mathbb{N})$ est dit stationnaire si quels que soient les entiers $d \geq 1, n_1, \dots, n_d$ choisis dans $\mathbb{N}$, tels que $n_1 < \cdots < n_d$, les vecteurs aléatoires réels d-dimensionnels $(X_{n_1}, \dots, X_{n_d})$ et $(X_{n_1+1}, \dots, X_{n_d+1})$ suivent la même loi.

Il en résulte évidemment que $(X_{n_1}, \dots, X_{n_d})$ et $(X_{n_1+h}, \dots, X_{n_d+h})$ suivent la même loi quel que soit l'entier $h \geq 0$.
**Définition:** Étant donné un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on dit qu'une application $T : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{F})$ conserve la mesure $\mathbb{P}$ si $\forall A \in \mathcal{F}, \mathbb{P}(A) = \mathbb{P}(T^{-1}(A))$ ou si $\mathbb{P}$ est sa propre image par $T$.
**Définition:** On dit que $T : \Omega \longrightarrow \Omega$ est une bijection bimesurable si c'est une bijection $(\mathcal{F}, \mathcal{F})$-mesurable et si l'application réciproque $T^{-1}$ est $(\mathcal{F}, \mathcal{F})$-mesurable .

**Théorème 15.** *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et une application*

$$T : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{F}).$$

- *Si $T$ conserve la mesure $\mathbb{P}$, pour tout entier $n \geq 1$, $T^n$ conserve $\mathbb{P}$*
- *Si $T$ est une bijection bimesurable qui conserve $\mathbb{P}$, $T^{-1}$ conserve $\mathbb{P}$.*

*Démonstration.* – On raisonne par récurrence. Au rang initial, $\forall A \in \mathcal{F}, \mathbb{P}(A) = \mathbb{P}(T^{-1}(A))$ parce que $T$ conserve $\mathbb{P}$. Si pour un entier $n \geq 1$, on a démontré que

$$\forall A \in \mathcal{F}, \mathbb{P}(A) = \mathbb{P}(T^{-n}(A)),$$

alors comme

$$\mathbb{P}(T^{-n}(A)) = \mathbb{P}(T^{-1}(T^{-n}(A))) = \mathbb{P}(T^{-(n+1)}(A)),$$

et il vient

$$\forall A \in \mathcal{F}, \mathbb{P}(A) = \mathbb{P}(T^{-(n+1)}(A)).$$

D'où la conclusion par récurrence : pour tout $n \geq 1$, $T^n$ conserve $\mathbb{P}$.

– $\forall A \in \mathcal{F}, \mathbb{P}(T(A)) = \mathbb{P}(T^{-1}(T(A))) = \mathbb{P}(A)$.

□

**Théorème 16.** *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité et* $\mathcal{C}$ *un* $\Pi$*-système d'événements engendrant* $\mathcal{F}$*. Alors une application* $T : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{F})$ *conserve la mesure* $\mathbb{P}$ *si et seulement si*

$$\forall A \in \mathcal{C}, \mathbb{P}(A) = \mathbb{P}(T^{-1}(A)).$$

*Démonstration.* Le sens direct est évident.
Réciproquement, on suppose que $\mathbb{P}$ et son image par $T$ coïncident sur $\mathcal{C}$, donc sur $\mathcal{F}$. □

**Définition:** Pour $E = \mathbb{N}$ ou $E = \mathbb{Z}$, on notera $\theta$ l'application de $\mathbb{R}^E$ dans $\mathbb{R}^E$ appelée opérateur de translation définie par

$$\theta(x_n, n \in E) = (y_n, n \in E),$$

où $\forall n \in E, y_n = x_{n+1}$.

**Théorème 17.** *Pour qu'un processus réel* $X = (X_n, n \in \mathbb{N})$ *soit stationnaire, il faut et il suffit que l'opérateur de translation* $\theta$ *conserve sa loi* $P_X$*.*

*Démonstration.* D'après le théorème précédent, $\theta$ conserve $P_X$ si et seulement si pour tout cylindre mesurable à base finie $C$ de $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$, $P_X(\theta^{-1}(C)) = P_X(C)$. Un tel cylindre s'écrivant sous la forme $\Pi^{-1}_{(s_1,\ldots,s_n)}(B)$,où $n \in \mathbb{N}^*, s_1, \ldots, s_n \in \mathbb{N}$ tels que $s_1 < \cdots < s_n$, et $B \in \mathcal{B}(\mathbb{R}^n)$, la condition nécessaire et suffisante s'écrit : $\forall n \in \mathbb{N}^*, \forall s_1 < \cdots < s_n \in \mathbb{N}, \forall B \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned}
P_{(X_{s_1+1},\ldots,X_{s_n+1})}(B) &= P_X(\Pi^{-1}_{(s_1+1,\ldots,s_n+1)}(B)) \\
&= P_X(\theta^{-1}(\Pi^{-1}_{(s_1,\ldots,s_n)}(B))) \\
&= P_X(\Pi^{-1}_{(s_1,\ldots,s_n)}(B)) \\
&= P_{(X_{s_1},\ldots,X_{s_n})}(B).
\end{aligned}$$

Elle équivaut à l'égalité des lois $P_{(X_{s_1+1},\ldots,X_{s_n+1})}$ et $P_{(X_{s_1},\ldots,X_{s_n})}$, quels que soient $n \in \mathbb{N}^*$, et $s_1, \ldots, s_n \in \mathbb{N}$ avec $s_1 < \cdots < s_n$ ce qui est la définition de la stationnarité du processus $X$. □

Les processus stationnaires fournissent ainsi un exemple fondamental de transformation conservant la mesure.

**Théorème 18.** *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité et une application*

$$T : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{F})$$

*qui conserve la mesure* $\mathbb{P}$. *Alors*

- *quelle que soit la variable aléatoire* $\xi$ *sur* $(\Omega, \mathcal{F}, \mathbb{P})$, $(\xi \circ T^n, n \geq 0)$ *est un processus stationnaire.*
- *si* $T$ *est une bijection bimesurable de* $(\Omega, \mathcal{F})$, $(\xi \circ T^n, n \in \mathbb{Z})$ *est également un processus stationnaire.*

*Démonstration.* 1. Soit $n \in \mathbb{N}^*$, $s_1 < \cdots < s_n$ dans $\mathbb{N}$. On a :

$$\begin{aligned} T : (\Omega, \mathcal{F}, \mathbb{P}) &\rightarrow (\Omega, \mathcal{F}, \mathbb{P}) \\ (\xi \circ T^{s_1}, \ldots, \xi \circ T^{s_n}) : (\Omega, \mathcal{F}, \mathbb{P}) &\rightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \end{aligned}$$

puis

$$(\xi \circ T^{s_1+1}, \ldots, \xi \circ T^{s_n+1}) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

par composition. Alors la loi de $(\xi \circ T^{s_1+1}, \ldots, \xi \circ T^{s_n+1})$ est l'image de $\mathbb{P}$ par $(\xi \circ T^{s_1}, \ldots, \xi \circ T^{s_n})$ (car $T$ conserve la mesure), qui est la loi de $(\xi \circ T^{s_1}, \ldots, \xi \circ T^{s_n})$. $(\xi \circ T^n, n \geq 1)$ est donc stationnaire par définition.

2. Même démonstration en remplaçant $s_1, \ldots, s_n \in \mathbb{N}$ par $s_1, \ldots, s_n \in \mathbb{Z}$ . $\square$

## 3.4 Processus gaussiens

### 3.4.1 Caractérisation

**Définition:** On dit qu'un processus $(X_t)_{t \in T}$ est gaussien si pour tout $S \in \mathcal{F}(T)$, le vecteur $(X_s)_{s \in S}$ est gaussien.

A tout processus gaussien, on peut associer son espérance $(\mathbb{E}X_t)_{t \in T}$ est sa fonction de covariance

$$C_X : (s, t) \mapsto \mathbb{E}(X_t - \mathbb{E}X_t)(X_s - \mathbb{E}X_s)).$$

**Proposition 4.** *Deux processus gaussiens ont même espérance et même fonction de covariance si et seulement si ils ont même loi.*

*Démonstration.* Notos $(X_s)$ et $(Y_s)$ les deux processus considérés.

– Le sens « même loi implique même espérance, même covariance »est "presque" évident. Arrétons nous y tout de même quelques instants. On a
$$\mathbb{E}X_s = \int_{\mathbb{R}^T} \omega_s \; dP_X(\omega)$$
et
$$\mathbb{E}Y_s = \int_{\mathbb{R}^T} \omega_s \; dP_Y(\omega).$$
Dire que $(X_s)$ et $(Y_s)$ ont même loi, c'est précisément dire que $P_X = P_Y$. Cela implique donc qu'ils ont les mêmes espérances. Les identités
$$\mathbb{E}X_s X_t = \int_{\mathbb{R}^T} \omega_s \omega_t \; dP_X(\omega)$$
et
$$\mathbb{E}Y_s Y_t = \int_{\mathbb{R}^T} \omega_s \omega_t \; dP_Y(\omega)$$
permettent alors de compléter la preuve.
– Soit $F \subset T$, $F$ fini. Les vecteurs $(X_s)_{s \in S}$ et $(Y_s)_{Y \in S}$ sont gaussiens. Par hypothèse, ils ont même espérance et même matrice de covariance. Des vecteurs gaussiens qui ont même espérance et même matrice de covariance ont même loi. Ainsi $(X_s)$ et $(Y_s)$ ont mêmes lois de dimension finie. Ils ont donc la même loi.

□

### 3.4.2 Condition d'existence

**Théorème 19.** *Soit $(m_t)_{t \in T}$ et $(c_{s,t})_{(s,t) \in T \times T}$ des réels. Il existe un processus gaussien de moyenne $(m_t)_{t \in T}$ et de covariance $(c_{s,t})_{(s,t) \in T \times T}$ si et seulement si*
- *Pour tous $s, t \in T$, on a $c_{s,t} = c_{t,s}$.*
- *Four tous $S$ fini inclus dans $T$ et tout $x \in \mathbb{R}^T$, on a*

$$\sum_{(s,t) \in S \times S} c_{s,t}(x_s - m_s)(x_t - m_t) \geq 0.$$

*Démonstration.* La nécessité des deux conditions provient du fait que $(c_{s,t})_{(s,t) \in S \times S}$ doit être la matrice de covariance du vecteur $(X_s)_{s \in S}$ Pour voir que ces conditions sont suffisantes, il suffit d'appliquer le théorème de Kolmogorov à la famille de mesures $\mathcal{N} m_S, C_S$ où $m_S = (m_t)_{t \in S}$ et $C_S = (c_{s,t})_{(s,t) \in S \times S}$ qui est compatible. □

### 3.4.3 Processus gaussiens stationnaires

**Théorème 20.** *Soit $(X_n)_{n\in\mathbb{Z}}$ un processus gaussien. $(X_n)_{n\in\mathbb{Z}}$ est stationnaire si et seulement si il existe une constante $m$ et une fonction $\varphi$ telle que*

- *Pour tout $n$ $\mathbb{E}X_n = m$*
- *Pour tous $n, p$ entiers on a $\mathbb{E}(X_n - m)(X_p - m) = \varphi(n-p)$.*

*$\varphi$ est appelée fonction d'autocovariance du processus.*

*Démonstration.* Supposons que le processus est stationnaire et posons $m = \mathbb{E}X_0$ et $\varphi(n) = \mathbb{E}(X_n - m)(X_0 - m)$. Pour tout $n$ $X_0$ et $X_n$ ont même loi, donc $\mathbb{E}X_n = \mathbb{E}X_0 = m$. D'autre part, le couple $(X_n, X_p)$ a même loi que le couple $(X_{n-p}, X_0)$ : on a donc $\mathbb{E}(X_n - m)(X_p - m) = \mathbb{E}(X_{n-p})(X_0 - m)) = \varphi(n-p)$. Réciproquement, supposons que pour tout $n$ $\mathbb{E}X_n = m$ et que pour tous $n, p$ entiers on a $\mathbb{E}(X_n - m)(X_p - m) = \varphi(n-p)$. Il faut démontrer que le processus $(X_{n+1})_{n\in\mathbb{Z}}$ a même loi que le processus $(X_n)_{n\in\mathbb{Z}}$. Ces deux processus étant gaussiens, ils suffit de montrer qu'ils ont même espérance et même covariance. Or on a pour tout $n$ :$\mathbb{E}X_{n+1} = m = \mathbb{E}X_n$ et pour tous $n, p$

$$
\begin{aligned}
\mathbb{E}(X_{n+1} - \mathbb{E}X_{n+1})(X_{p+1} - \mathbb{E}X_{p+1}) &= \mathbb{E}(X_{n+1} - m)(X_{p+1} - m) \\
&= \varphi((n+1) - (p+1)) \\
&= \varphi(n-p) \\
&= \mathbb{E}(X_n - m)(X_p - m) \\
&) \ \mathbb{E}(X_n - \mathbb{E}X_n)(X_p - \mathbb{E}X_p),
\end{aligned}
$$

ce qui achève la preuve. □

## 3.5 Exercices

1. Soient $(X_n)_{n\geq 1}$, $(Y_n)_{n\geq 1}$ deux processus stationnaires indépendants. Montrer que pour toute application mesurable $\varphi$ de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$, le processus $\varphi(X_n, Y_n)$ est stationnaire. Donner un exemple de processus $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$ stationnaires tels que $X_n + Y_n$ ne soit pas stationnaire.
2. Soit $(X_n)_{n\geq 1}$ un processus stationnaire. Montrer que le processus $(Y_n)_{n\geq 1}$ défini par $Y_n = X_n + 2X_{n+1}$ est stationnaire.
3. Soit $\varphi$ une application mesurable de $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ et $(X_n)_{n\geq 1}$ un processus stationnaire. Démontrer que le processus $(Y_n)_{n\geq 1}$ défini par $Y_n = \varphi(X_n, X_{n+1}, \dots) = \varphi(\theta^{n-1} \circ X)$ est stationnaire.
4. Soit $(X_n)_{n\geq 0}$ une chaîne de Markov homogène dont l'espace d'état est fini. Montrer que $(X_n)_{n\geq 0}$ est stationnaire si et seulement si $X_0$ et $X_1$ ont même loi.

5. Soit $X_0$ une variable aléatoire suivant la loi uniforme sur $\mathbb{Z}/7\mathbb{Z}$. Soit $(T_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi uniforme sur l'ensemble $\{-1,1\}$. On définit par récurrence une suite $(T_n)_{n\geq 1}$ par $X_{n+1} = X_n + T_{n+1}$.On pose enfin $Z_n = \inf\{k \geq 0; X_{n+k} = 0\}$. Montrer que $(Z_n)_{n\geq 0}$ est un processus stationnaire.
6. Soit $\alpha$ un réel. Montrer que l'application de $[0,1]$ dans lui-même qui à $x$ associe la partie fractionnaire de $x+\alpha$ laisse invariante la mesure de Lebesgue sur $[0,1]$. Comment interpréter ce résultat si l'on identifie $[0,1[$ au cercle unité par l'application $x \mapsto e^{2i\pi x}$ ?
7. Montrer que l'application de $[0,1]$ dans lui-même qui à $x$ associe la partie fractionnaire de $2x$ laisse invariante la mesure de Lebesgue sur $[0,1]$.
8. On appelle bruit blanc une suite $(Z_n)_{n\in\mathbb{Z}}$ de variables aléatoires indépendantes suivant la loi $\mathcal{N}(0,1)$. Soit $(Z_n)_{n\in\mathbb{Z}}$ un bruit blanc et $\beta = (\beta_0,\ldots,\beta_q) \in \mathbb{R}^{q+1}$ avec $\beta_0 \neq 0$ et $\beta_q \neq 0$. On considère la moyenne mobile :

$$X_n = \sum_{k=0}^{q} \beta_k Z_{n-q}.$$

   Démontrer que $(Z_n)_{n\in\mathbb{Z}}$ est un processus stationnaire dont on calculera la fonction d'autocovariance.

# Chapitre 4

# Chaînes de Markov

## 4.1 Dynamique markovienne

**Définition:** Soit $S$ un ensemble fini ou dénombrable, $\nu$ une mesure de probabilité sur $S$ et $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice à coefficients positifs. Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires définies sur un espace $(\Omega, \mathcal{F}, P)$. On dit que la suite $(X_n)_{n\geq 0}$ est une *chaîne de Markov* de loi initiale $\nu$ et de matrice de passage $M$ si l'on a, pour tout entier $n \geq 1$ et toute suite $x_0, \dots x_n$ d'éléments de $S$ :

$$P(X_0 = x_0, X_1 = x_1, \dots X_n = x_n) = \nu(x_0) \prod_{i=0}^{n-1} p_{x_i, x_{i+1}}.$$

Exemple : une suite $(X_n)_{n\geq 0}$ de variables aléatoires indépendantes de même loi $\nu$ à valeurs dans $S$ dénombrable est une chaîne de Markov. En effet, il suffit de poser pour $(i,j) \in S \times S$ $p_{i,j} = \nu(j)$.

Propriétés :

1. Si $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de passage $M$ et que $x_0, \dots, x_{n-1}$ sont tels que $P(X_0 = x_0, X_1 = x_1, \dots X_{n-1} = x_{n-1})$, alors $P(X_n = x_n)|X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = p_{X_{n-1},x_n}$. Autrement dit,
$$P(X_n = x_n | X_0, \dots, X_{n-1}) = p_{X_{n-1},x_n}. \tag{4.1}$$
Cela signifie que toute l'information que $X_0, \dots, X_{n-1}$ peuvent nous apporter sur $X_n$ est comprise dans $X_n$.
2. (4.1) implique que $P(X_n = x_n | X_{n-1}) = p_{X_{n-1},x_n}$

Qu'est ce concrètement, qu'une chaîne de Markov ? On va voir que c'est une suite de réalisations, au cours du temps, des états d'un système soumis à des transformations aléatoires, la suites des transformations est une suite de transformations indépendantes, de même loi. Évidemment, le résultat de la transformation dépend de la transformation choisie et de l'état du système avant la transformation.

**Lemme 6.** *Soit $S$ un ensemble fini ou dénombrable, $\nu$ une loi sur $S$ et $\theta$ une mesure sur $S^S = \mathcal{F}(S,S)$.*

*Soit $(f_n)_{n\geq 1}$ une suite de v.a.i.i.d. de loi $\theta$ et $X_0$ une variable aléatoire de loi $\mu$ indépendante de $(f_n)_{n\geq 1}$. On définit $(X_n)_{n\geq 1}$ par*

$$\forall n \geq 0 \quad X_{n+1} = f_{n+1}(X_n)$$

*Alors $(X_n)_{n\geq 0}$ est une chaîne de Markov de loi initiale $\nu$ et de matrice de transition $M$, où $M$ est définie par*

$$\forall (i,j) \in S \times S \quad m_{i,j} = \theta(\{f \in S^S; f(i) = j\}).$$

*Démonstration.* Soit $A \subset S^{\{0,\dots,n\}}$.

$$\begin{aligned}
& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\} \cap \{X_{n+1} = j\}) \\
= \;& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\} \cap \{f_{n+1}(i) = j\}) \\
= \;& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\}) P(f_{n+1}(i) = j) \\
= \;& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\}) P(f_{n+1} \in S \times \dots \{j\} \times \dots S) \\
= \;& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\}) \theta(S \times \dots \{j\} \times \dots S) \\
= \;& P(\{(X_0,\dots,X_n) \in A\} \cap \{X_n = i\}) m_{i,j}
\end{aligned}$$

□

Exemple : la marche de l'ivrogne (ou marche aléatoire sur $\mathbb{Z}$)
Un ivrogne sort du café passablement éméché. À chaque pas, il prend une décision (enfin, si tant est que cela lui soit possible...) : aller à gauche, ou aller à droite. Si on repère par $X_n$ sa position dans la rue au temps $n$, on a $S = \mathbb{Z}$, $X_{n+1} = f_{n+1}(X_n)$, où $f_n$ est une suite de translations indépendantes : $P(f_n = (x \mapsto x+1)) = P(f_n = (x \mapsto x-1)) = 1/2$.

Comme on va le voir, ce procédé permet de fabriquer toutes les chaînes de Markov.

## 4.2 Matrice stochastique

**Définition:** Soit $S$ un ensemble dénombrable et $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice à coefficients positifs. On dit que $P$ est une matrice stochastique si on a

$$\forall i \in S \quad \sum_{j\in S} p_{i,j} = 1.$$

### 4.2.1 Existence des chaînes de Markov

**Théorème 21.** *Soit $S$ un ensemble dénombrable, $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique et $\nu$ une mesure de probabilité sur $S$. Alors, on peut construire une chaîne de Markov de loi initiale $\nu$ et de matrice de passage $P$.*

*Démonstration.* Définissons une mesure $\theta_P$ sur $S^S$ par

$$\theta_P = \bigotimes_{i\in S} \mu_i,$$

où $\mu_i$ est la mesure sur $S$ définie par $\mu_i(j) = p_{i,j}$. Alors $\theta_P$ vérifie $\theta_P(S \times \dots \{j\} \times \dots S) = p_{i,j}$ et il suffit d'appliquer le lemme précédent. □

Lorsque la matrice $P$ est fixée, on note souvent $\mathbb{P}_\nu$ une probabilité sous laquelle $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $P$ telle que la loi de $X_0$ sous $P_\nu$ est $\nu$. De même, on note $\mathbb{E}_\nu$ l'espérance correspondante. Dans le cas où la loi initiale est une masse de Dirac, on écrit simplement $\mathbb{P}_i$ (resp. $\mathbb{E}_i$) au lieu de $\mathbb{P}_{\delta_i}$ (resp. $\mathbb{E}_{\delta_i}$).

### 4.2.2 Puissances des matrices stochastiques

**Théorème 22.** *Soit $(X_n)$ une chaîne de Markov de matrice de transition $P$ et de loi initiale $\mathbb{P}_{X_0} = \nu$. Alors, la loi $\mu_n$ de la chaîne au temps $n$ s'écrit $\mu_n = \nu P^n$, où on a écrit $\nu$ et $\mu_n$ comme des vecteurs lignes.*

*Démonstration.* Il suffit de montrer que $\mu_{n+1} = \mu_n P$, puis procéder par récurrence sur $n$. D'après le principe de partition, on a

$$\begin{aligned}
\mu_{n+1}(j) &= \mathbb{P}_\nu(X_{n+1} = j) \\
&= \sum_{i\in S} \mathbb{P}_\nu(X_n = i, X_{n+1} = j) \\
&= \sum_{i\in S} \mathbb{P}_\nu(X_n = i, )p_{i,j} \\
&= \sum_{i\in S} \mu_n(i)p_{i,j} \\
&= (\mu_n M)(j)
\end{aligned}$$

□

## 4.3 Graphe associé à une matrice stochastique

Soit $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique. On peut associer à la matrice $P$ (où aux chaînes de Markov correspondantes) un graphe orienté $G = (S, A)$ avec

$$A = \{(x, y) \in S \times S; p_{i,j} > 0\}.$$

Considérons une chaîne de Markov associée à la matrice stochastique $P$ avec la condition initiale déterministe $x_0$, autrement dit $\nu = \delta_{x_0}$ et notons $\mathbb{P}_{x_0}$ la mesure de probabilité correspondante Alors, comme

$$\mathbb{P}_{x_0}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \prod_{i=0}^{n-1} p_{x_i,x_{i+1}},$$

il est clair que $\mathbb{P}_{x_0}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n)$ est non nul si et seulement si $(x_0, x_1, \dots, x_n)$ constitue un chemin dans le graphe $G$.

D'après le principe de partition, on a pour une chaîne de Markov avec une loi initiale $\nu$

$$P(\nu)(X_1 = x_1, X_n = x_n) = \sum_{(x_0,\dots x_{n-1})\in S^n} \mathbb{P}_i(X_0 = x_0, X_2 = X_2, \dots X_{n-1} = x_{n-1}, X_n = x_n). \tag{4.2}$$

En particulier, si l'on pose $p_{i,j}^{(n)} = \mathbb{P}_i(X_n = j)$, on a

$$p_{i,j}^{(n)} = \sum_{x\in S^{n-1}} \mathbb{P}_i(X_1 = x_1, X_2 = X_2, \dots X_{n-1} = x_{n-1}, X_n = j).$$

Donc $p_{i,j}^{(n)} > 0$, autrement dit il est possible d'aller en $n$ étapes de l'état $i$ à l'état $j$ si et seulement si on peut trouver dans le graphe $G$ un chemin de longueur $n$ allant de $i$ à $j$.

On en déduit que

$$\mathbb{P}_i(\exists n > 0 X_n = j) = \mathbb{P}_i(\bigcup_{i\geq 1} \{X_n = j\}),$$

qui représente la probabilté que, partant de $i$, on puisse arriver à $j$, est non nulle si et seulement si il existe dans le graphe $G$ un chemin allant de $i$ à $j$.

Si il y a à la fois un chemin de $i$ vers $j$ et un chemin de $j$ vers $i$, on dit que les états $i$ et $j$ communiquent et on écrit $i \leftrightarrow j$.

Si tous les états communiquent, on dit que la chaîne de Markov est *irréductible*.

On appelle *période* d'un état $x$ d'une chaîne de Markov et on note $d(x)$ le pgcd (plus grand commun diviseur) des longueurs des circuits du graphe $G$ contenant $x$. Lorsque la période est 1, on dit que l'état $x$ est *apériodique*.

**Lemme 7.** *Si deux états communiquent, alors ils ont même période.*

*Démonstration.* Soient $i, j$ avec $i \leftrightarrow j$. Soit $\gamma$ un chemin de $i$ à $j$, $\gamma'$ un chemin de $j$ à $i$. Soit $\mathcal{C}$ un circuit quelconque (éventuellement vide) contenant $j$ . $\gamma - \gamma'$ et $\gamma - \mathcal{C} - \gamma'$ sont deux circuits contenant $i$. Donc $d(i)$ divise leurs longueurs ainsi que la différence de leurs longueurs, soit la longueur de $\mathcal{C}$. Ainsi $d(i)$ divise les longueurs de tous les circuits contenant $j$, donc divise leur pgcd, soit $d(j)$. De la même manière, on montre que $d(j)$ divise $d(i)$, d'où $d(i) = d(j)$. □

**Définition:** Si une chaîne irréductible a ses états de période 1, on dit qu'elle est apériodique.

Les lemme suivant se révélera très utile par la suite

**Lemme 8.** *Soit $x$ un état de période 1. Il existe un entier $N(x)$ tel que pour tout $n \geq N(x)$ le graphe associé à la chaîne de Markov possède un circuit de longueur $n$ contenant $x$*

Soit $A$ l'ensemble des valeurs de $n$ telles que le graphe associé à la chaîne de Markov possède un circuit de longueur $n$ contenant $x$. Il est clair que $A$ est stable par addition (concaténation des circuits). Il existe $p \geq 1$ et $n_1, n_2, \ldots, n_p$ tels que le pgcd de $n_1, n_2, \ldots, n_p$ soit 1. D'apès le lemme de Bezout, il existe des relatifs $a_1, \ldots a_p$ tels que $1 = \sum_{k=1}^{p} a_k n_k$. Posons $P = \sum_{p:a_p>0} a_p n_p$ et $N = \sum_{p:a_p<0} (-a_p) n_p$. On a $P \in A, N \in A$ et $1 = P - N$. Soit $n \geq N(N-1)$. On peut écrire $n = bN + r$, avec $b \geq N - 1$ et $r \in \{0, 1, \ldots N - 1\}$. On a $n = bN + r = bN + r(P - N) = rP + (b - r)N \in A$ car $b - r \in \mathbb{N}$ et $A$ est stable par addition.

**Corollaire 8.** *Si $x$ est un état de période 1 et qu'il existe un chemin de longueur $d(x, y)$ allant de $x$ à $y$, alors pour tout $n \geq N(x, y) = N(x) + d(x, y)$, il existe un chemin de longueur $n$ allant de $x$ à $y$.Ainsi, si $P$ est la matrice associée, $P^n(x, y) > 0$.*

*Démonstration.* Il suffit de concaténer le chemin allant de $x$ à $x$ avec un chemin allant de $x$ à $y$. □

**Corollaire 9.** *Si une chaîne de Markov est irréductible, apériodique, à valeurs dans un ensemble fini $S$, alors il existe un entier $N$ tel que pour tout $n \geq N$ et tout couple $(i, j)$, il existe un chemin de longueur $n$ allant de $i$ à $j$. Ainsi, si $P$ est la matrice associée, $P^n$ est à coefficients strictement positifs.*

*Démonstration.* Il suffit de prendre $N = \max(N(x), x \in S) + \mathrm{diam}(G)$. □

## 4.4 Exercices

1. *Chaîne à deux états.* Soit $\{X_n : n \geq 0\}$ une chaîne de Markov à valeurs dans $\{0, 1\}$ et de probabilité de transition :

$$P := \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \ 0 \leq \alpha, \beta \leq 1.$$

   (a) Montrer que pour $(\alpha, \beta) \neq (0, 0)$ :

$$P^n = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta} \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}.$$

   Que se passe-t-il lorsque $\alpha = 0$ ou $\beta = 0$ ou $\alpha = \beta = 0$ ? On supposera pour la suite de l'exercice que $(\alpha, \beta) \neq (0, 0)$.

   (b) Vérifier (par récurrence) que, pour toute loi initiale $\mu$ :

$$\mathrm{P}_\mu(X_n = 0) = \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n \left(\mu(0) - \frac{\beta}{\alpha+\beta}\right).$$

   (c) Si $(\alpha, \beta) \neq (1, 1)$, montrer que $\{X_n : n \geq 0\}$ converge en loi vers une variable aléatoire de loi $\nu$. Que vaut $\nu$ ? On supposera pour la suite de l'exercice que $(\alpha, \beta) \neq (1, 1)$.

   (d) (*Mesure stationnaire*) Prouver que, pour tout $n \in \mathbb{N}$,

$$\mathrm{P}_\nu(X_n \in A) = \nu(A).$$

2. *Représentation canonique et simulation des chaînes de Markov.*

   (a) Soit $(Z_n)_{n\geq 1}$ une suite de vaiid à valeurs dans $F$, soit $g : E \times F \to E$ et soit $X_0$ une variable aléatoire à valeurs dans $E$ indépendante de $(Z_n)_{n\geq 1}$. Montrer que la suite $(X_n)_{n\geq 0}$ définie par $X_{n+1} = g(X_n, Z_{n+1})$ est une chaîne de Markov homogène. Donner sa matrice de transition.

   (b) On suppose qu'on dispose d'un générateur de nombres aléatoire de loi uniforme sur $[0, 1]$, noté 'rand'. Soit $\mu$ une mesure de probabilité sur $\mathbb{N}$. Donner un algorithme pour générer des nombres aléatoires suivant la loi $\mu$.

   (c) Soit $P = (p_{i,j})$ une matrice de transition sur $\mathbb{N}$. On note $s_{i,k} = \sum_{j=0}^{k} p_{i,j}$. Soit $(Z_n)_{n\geq 1}$ une suite de vaiid de loi uniforme sur $[0, 1]$ et $X_0$ une variable aléatoire à valeurs dans $\mathbb{N}$ indépendante de

$(Z_n)_{n\geq 1}$. On construit la suite $(X_n)_{n\geq 0}$ par récurrence de la façon suivante :

$$\text{si } X_n(\omega) = i \text{ et } Z_{n+1}(\omega) \in ]s_{i,j}, s_{i,j+1}] \text{ alors } X_{n+1} = j.$$

Montrer que la suite $(X_n)_{n\geq 0}$ ainsi définie est une chaîne de Markov homogène. Donner sa matrice de transition.

(d) Application. Comment simuler une chaîne de Markov homogène de matrice de transition $P = (p_{i,j})$ ? Ecrire un algorithme explicite si

$$P = \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}.$$

3. *La ruine du joueur* Un joueur possédant une fortune de $a$ unités joue à pile ou face jusqu'à ce qu'il ait fait sauter la banque ou qu'il soit ruiné. Les réserves de la banque sont de $b$ unités. Chaque victoire rapporte une unité et chaque défaite en coûte une. On suppose que les lancers sont indépendants et que la probabilité de gain de la banque est $p = 1 - q$. On veut déterminer la probabilité $p_g$ que la banque résiste.
On note $(X_n)_{n\geq 1}$ une suite de v.a.i.i.d. de loi $p\delta_1 + q\delta_{-1}$, puis $S_n = \sum_{k=1}^n X_k$ et $T = \inf\{n \geq 0; S_n = -b \text{ ou } S_n = a\}$. Si l'on pose $S'_n = S_{n\wedge T}$, il est aisé de constater que $S'_n$ représente la suite des gains relatifs de la banque.

(a) Montrer que $S'_n$ est une chaîne de Markov homogène à espace d'états $E = \{-b, \ldots, a\}$ dont on déterminera la loi initiale et la matrice de transition.

(b) Considérons les chaînes de Markov ayant la même matrice de transition que $(S'_n)_{n\geq 0}$ Montrer que la suite $(u_n)_{-b\leq n\leq a}$ définie par

$$u_n = \mathbb{P}_n(\{\text{la banque résiste}\})$$

vérifie la récurrence linéaire

$$pu_{n+1} - u_n + qu_{n-1} = 0.$$

Que valent $u_a$ et $u_{-b}$ ?

(c) Résoudre l'équation de récurrence et en déduire

$$p_g = \frac{(\frac{q}{p})^b - 1}{(\frac{q}{p})^{a+b} - 1}. \tag{4.3}$$

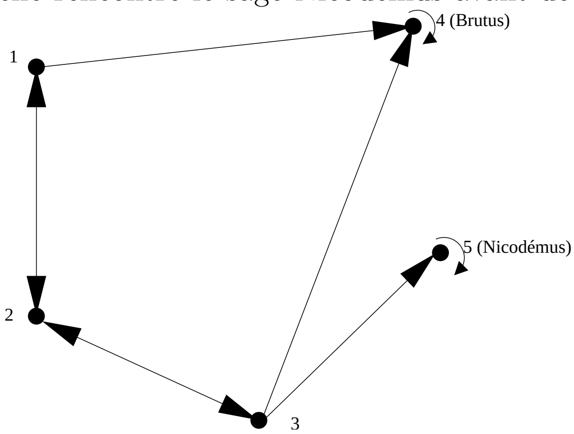
4. *Le joueur inruinable* Le problème est le même que le précédent, a ceci près que l'on suppose maintenant que le joueur est infiniment riche. On cherche toujours la probabilité que la banque résiste (ce qui ne signifie pas ici que le joueur est ruiné).

   Intuitivement, il suffit de faire tendre $a$ vers $+\infty$ dans la formule (4.3), le tout étant de le justifier...

   Indications : poser $T' = \inf\{n; S_n \le -b\}$ et, pour tout $a > 0$, $U_a = \inf\{n; S_n \ge a\}$ et $G^a = \{U_a \le T'\}$, puis montrer que

$$\lim_{a \to +\infty} \mathbb{1}_{G^a} = \mathbb{1}_{\{T' = +\infty\}}.$$

5. Madame Brisby dans le labyrinthe Madame Brisby s'est perdue dans le labyrinthe que forment les galeries où vivent les rats de Nim. Quelle est la probabilité qu'elle rencontre le sage Nicodémus avant de croiser le belliqueux Rufus ?




6. Soit $M$ la matrice d'une chaîne de Markov. Montrer que si $m_{i,i} > 0$, alors l'état $i$ est apériodique. Qu'en déduire pour une chaîne irréductible ?
7. Soit $a$ un entier supérieur ou égal à 2, $(D_n)_{n \ge 1}$ une suite de variables aléatoires i.i.d. à valeurs dans $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/a\mathbb{Z}$ vérifiant $P(D_1 = (0,1)) = P(D_1 = (1,0)) = \frac{1}{2}$. On pose

$$S_n = S_0 + \sum_{k=1}^{n} D_k.$$

   Montrer que $(S_n)$ est une chaîne de Markov . Est-elle irréductible, apériodique ?
8. *Madame Brisby II*

   Soit $f : E \to \mathbb{C}$ telle que $\forall x \in A \quad f(x) = 0$.

   On pose $F = \sum_{k=1}^{+\infty} f(X_k)$.

Montrer que $\mathbb{E}|F| \leq (\mathbb{E}T - 1)\|f\|_\infty$.
Montrer l'identité

$$(I - N)\begin{pmatrix} \mathbb{E}_1 F \\ \mathbb{E}_2 F \\ \mathbb{E}_3 F \end{pmatrix} = \begin{pmatrix} \mathbb{E}_1 f(X_1) \\ \mathbb{E}_2 f(X_1) \\ \mathbb{E}_3 f(X_1) \end{pmatrix},$$

En déduire $\mathbb{E}_1 T, \mathbb{E}_2 T, \mathbb{E}_3 T$.

9. *Évolution d'un genotype avec fixation*
Nous travaillons ici sur une population de taille fixe formée de $2N$ gènes. Il y a deux types de gènes possibles : le type "a" et le type "A". Chacun des gènes au temps $n+1$ est engendré par deux des $2N$ gènes présents au temps $N$. Son type est celui d'un de ses deux parents (choisi au hasard).
On considère la variable aléatoire $X_n$ égale au nombre d'individus de type "A" dans la population à l'étape $n$.
   (a) Montrer que $X_n$ est une chaîne de Markov à valeurs dans $E = \{0, \ldots, 2N\}$.
   (b) Montrer que la loi de $X_{n+1}$ sachant $X_n = k$ est une loi binomiale de paramètre $2N$ et $(k/2N)$.
   (c) Montrer que $(X_n)_{n\geq 0}$ converge presque sûrement vers une variable aléatoire $X_\infty$.
   (d) Déterminer la loi de $X_\infty$ en fonction de la loi de $X_0$. (On pourra remarquer que la suite $(\mathbb{E}X_n)_{n\geq 1}$ est constante.)
10. *L'image d'une chaîne de Markov n'est pas (toujours) une chaîne de Markov.* On considère la chaîne de Markov $(X_n)$ sur $E = \{0, 1, 2\}$ de transition $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ et de loi initiale $\pi_0 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Soit $f : E \to \{0, 1\}$ telle que $f(0) = f(1) = 0$, $f(2) = 1$. Montrer que $(f(X_n))$ n'est pas une chaîne de Markov.
11. *L'image d'une chaîne de Markov peut être une chaîne de Markov.* Soit $(X_n)$ une chaîne de Markov sur un ensemble dénombrable $E$ de matrice de transition $P$. Soit $\psi$ une application surjective de $E$ dans un ensemble $F$ telle que

$$\forall z \in F\ ,\ \forall x, y \in E \quad \psi(x) = \psi(y) \Rightarrow P(x, \psi^{-1}(z)) = P(y, \psi^{-1}(z))\ .$$

Montrer que la suite $(Y_n)$ définie par $Y_n = \psi(X_n)$ est une chaîne de Markov et déterminer sa matrice de transition. Montrer que si $\pi$ est une probabilité stationnaire pour la chaîne $(X_n)$ alors l'image de $\pi$ par $\psi$ est stationnaire pour $(Y_n)$.

# Chapitre 5

# Récurrence et mesures invariantes

## 5.1 Temps d'arrêt et propriété de Markov forte

**Théorème 23.** *Soit $(X_k)_{k\ge 0}$ une chaîne de Markov de matrice de passage $P$ et $T$ un temps d'arrêt adapté à cette suite. Soit $A$ un événement se produisant avant le temps $T$. Conditionnellement à l'événement $\{T < +\infty\} \cap A$, la suite $(X_{T+k})_{k\ge 0}$ est une chaîne de Markov de matrice de passage $P$.*

*Démonstration.* Comme être une chaîne de Markov est une propriété de la loi, on peut supposer que $(X_n)_{n\ge 0}$ est obtenue par le procédé décrit plus haut : $X_{n+1} = f_{n+1}(X_n)$ où $(f_n)_{n\ge 1}$ est une suite de variables aléatoires indépendantes de loi $\theta_M$ indépendante de $X_0$.

Posons $Y_k = X_{T+k}$ si $T < +\infty$ et $Y_k = X_k$ sinon. De même posons $g_n = f_{n+k}$ si $T < +\infty$ et $g_k = f_k$ sinon. Il est facile de voir que $(Y_k)_{k\ge 0}$ vérifie la récurrence $Y_{n+1} = g_{n+1}(Y_n)$

Soit $p$ un entier et $B$ un borélien de $\mathcal{F}(S,S)^{\mathbb{N}_*}$. On a

$$\begin{aligned} P(T=p, A, g \in B) &= P(T=p, A, f_{p+.} \in B) \\ &= P(T=p, A)P(f_{p+.} \in B) \\ &= P(T=p, A)P(f_{p+.} \in B) \end{aligned}$$

car comme l'événement $\{T = p\} \cap A$ est $\sigma(X_0, f_1, \dots, f_p)$-mesurable, il est indépendant de l'événement $\{f_{p+.} \in B\}$ qui est $\sigma(f_{p+1}, f_{p+2}, )$-mesurable. Maintenant la loi de $f_{p+.}$ est la même loi que celle de $f$ : c'est $\Theta_M^{\otimes \mathbb{N}^*}$. On a donc $P(T=p, A, g \in B) = P(T=p, A)P(f \in B)$. En faisant la somme pour

$p$ variant de 1 à $+\infty$, on obtient

$$P(T < +\infty, A, g \in B) = P(T < +\infty, A)P(f \in B)$$

Autrement dit , sachant $\{T < +\infty\} \cap A$ $(g_n)_{n\geq 1}$ $(g_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes de loi $\theta_M$ indépendante de $A$, donc $(X_{T+k})_{k\geq 0}$ est une chaîne de Markov de matrice de passage $P$. □

Remarque : une constante étant un temps d'arrêt, le théorème précédent s'applique lorsque $T$ est une constante. Dans ce cas, le résultat porte simplement le nom de *Propriété de Markov*.

## 5.2 Classification des états

**Définition:** Soit $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique. Pour $i \in S$, on considère une chaîne de Markov $(X_n)_{n\geq 0}$ partant de l'état $i$ et de matrice de passage $P$. On pose $T_i = \inf\{n \geq 1; X_n = i\}$. Si $\mathbb{P}_i(T_i < +\infty) = 1$, on dit que l'état $i$ est *récurrent*. Inversement, si $\mathbb{P}_i(T_i < +\infty) < 1$, on dit que l'état $i$ est *transient*.

**Théorème 24.** *Soit $P = (p_{i,j})_{(i,j)\in S\times S}$ une matrice stochastique. Pour $i \in S$, on considère une chaîne de Markov $(X_n)_{n\geq 0}$ partant de l'état $i$ et de matrice de passage $P$. On pose $T_i = \inf\{n \geq 1; X_n = i\}$ et $N_i = \sum_{k=1}^{+\infty} 1\!\!1_{\{i\}}(X_i)$. $N_i$ représente le nombre de passage de la chaîne en $i$ à partir de l'instant 1.*

- *Si $i$ est transient, alors $1 + N_i$ suit la loi géométrique de paramètre $1 - \mathbb{P}_i(T_i < +\infty)$. En particulier $N_i$ est presque sûrement fini et intégrable.*
- *Si $i$ est récurrent, alors $N_i$ est presque sûrement infinie. En particulier $\mathbb{E}N_i = +\infty$.*

*Démonstration.* Si $T_i < +\infty$ (ou de manière équivalente si $N_i > 0$, on a

$$N_i = 1 + \sum_{k=1}^{+\infty} 1\!\!1_{\{i\}}(X_{T+i}).$$

Soit $k$ un entier positif ou nul. On a

$$\mathbb{P}_i(N_i \geq k+1) = \mathbb{P}_i(N_i \geq k+1, T_i < +\infty) = \mathbb{P}_i(T_i < +\infty)\mathbb{P}_i(\sum_{k=1}^{+\infty} 1\!\!1_{\{i\}}(X_{T+i} \geq k|T_i < +\infty).$$

Or $T_i$ est un temps d'arrêt. Donc, d'après la propriété de Markov forte, sachant $T_i < +\infty$, $(X_{T+i})_{i\geq 0}$ a la loi d'une chaîne de Markov commençant en $i$ et de matrice de transition $P$, c'est à dire la même loi que $(X_i)_{i\geq 0}$. On en déduit

$$\mathbb{P}_i(N_i \geq k+1) = \mathbb{P}_i(T_i < +\infty)\mathbb{P}_i(N_i \geq k).$$

Par récurrence, on en déduit

$$\mathbb{P}_i(N_i \geq k) = \mathbb{P}_i(T_i < +\infty)^k$$

D'après le théorème de continuité séquencielle décroissante, on a $P(N_i = +\infty) = \lim_{k\to+\infty} \mathbb{P}_i(N_i \geq k)$. Cette limite vaut donc 0 si $i$ est transient, 1 si $i$ est récurrent. Pour $k \geq 1$, on a

$$\begin{aligned} \mathbb{P}_i(1+N_i = k) &= \mathbb{P}_i(1+N_i \geq k) - P(N_i \geq k) \\ &= \mathbb{P}_i(T_i < +\infty)^{k-1} - \mathbb{P}_i(T_i < +\infty)^k \\ &= \mathbb{P}_i(T_i < +\infty)^{k-1}(1 - \mathbb{P}_i(T_i < +\infty)), \end{aligned}$$

ce qui montre que $1 + N_i$ suit bien une loi géométrique de paramètre $1 - \mathbb{P}_i(T_i < +\infty)$. De plus

$$\mathbb{E}_i N_i = \sum_{k=1}^{+\infty} \mathbb{P}_i(N_i \geq k) = \sum_{k=1}^{+\infty} = \mathbb{P}_i(T_i < +\infty)^k = \frac{\mathbb{P}_i(T_i < +\infty)}{1 - \mathbb{P}_i(T_i < +\infty)} < +\infty.$$

□

**Corollaire 10.** *Un état $i$ est transient si et seulement si*

$$\sum_{k=1}^{+\infty} \mathbb{P}_i(X_k = i) < +\infty.$$

*Démonstration.* D'après le théorème précédent, $i$ est transient si et seulement si $N_i$ est intégrable sous $\mathbb{P}_i$. Or

$$\begin{aligned} \mathbb{E}_i N_i &= \mathbb{E}_i \sum_{k=1}^{+\infty} \mathbb{1}_{\{i\}}(X_k) \\ &= \sum_{k=1}^{+\infty} \mathbb{E}_i \mathbb{1}_{\{i\}}(X_k) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}_i(X_k = i) \end{aligned}$$

(On utilise le théorème de Tonelli pour échanger la somme et l'espérance) Ceci achève la preuve. □

**Corollaire 11.** *Soient $i$ et $j$ deux états d'une chaîne de Markov de matrice de transition $P = (p_{i,j})_{(i,j)\in S\times S}$. On suppose que $i$ et $j$ communiquent. Alors $i$ et $j$ sont tous les deux transients ou tous les deux récurrents*

*Démonstration.* Soit $n, p$ tels que $p_{i,j}^{(n)} > 0$ et $p_{j,i}^{(p)} > 0$. Pour tout $k \geq 0$, on a

$$p_{j,j}^{(n+p+k)} \geq p_{j,i}^{(p)} p_{i,i}^{(k)} p_{i,j}^{(n)}.$$

Ainsi, si la série de terme général $p_{i,i}^{(k)}$ diverge, la série de terme général $p_{j,j}^{(k)}$ aussi. Comme les rôles de $i$ et $j$ sont symétriques, les deux séries sont de même nature. Comme $p_{i,i}^{(k)} = \mathbb{P}_i(X_k = i)$, le résultat découle du corollaire précédent. □

**Corollaire 12.** *Considérons une chaîne de Markov irréductible de matrice de transition $P = (p_{i,j})_{(i,j)\in S\times S}$ et pour tous les $i \in S$ , notons $\mathbb{P}_i$ les lois markoviennes correspondantes. Les propriétés suivantes sont équivalentes*

1. $\exists i, j \in S \ \mathbb{P}_j(N_i = +\infty) > 0$.
2. $\exists i \in S$, *$i$ est récurrent*
3. $\forall i \in S$, *$i$ est récurrent*
4. $\forall i, j \in S \ \mathbb{P}_j(N_i = +\infty) = 1$.

*Démonstration.*
- (1) $\Longrightarrow$ (2). Soit $l$ tel que $\mathbb{P}_i(X_l = j) > 0$. On a $\mathbb{P}_i(N = i = +\infty) \geq \mathbb{P}_i(X_l = j, \sum_{k=1}^{+\infty} X_{k+l} = i) \geq \mathbb{P}_i(X_l = j)\mathbb{P}_j(N_i = +\infty) > 0$, donc $i$ est récurrent.
- (2) $\Longrightarrow$ (3). C'est une conséquence du corollaire précédent.
- (3) $\Longrightarrow$ (4). Considérons $\mathbb{P}_i(T_j < +\infty, \forall k > T_y X_k \neq x)$. Comme $i$ et $j$ communiquent $\mathbb{P}_i(T_j < +\infty) > 0)$. D'après la propriété de Markov forte, on a

  $$\mathbb{P}_i(T_j < +\infty, \forall k > T_j X_k \neq i) = \mathbb{P}_i(T_j < +\infty)\mathbb{P}_j(\forall k > 0 X_k \neq i) = \mathbb{P}_i(T_j < +\infty)\mathbb{P}_j(T_i =$$

  Mais $\{T_j < +\infty, \forall k > T_j X_k \neq i\} \subset \{N_j < +\infty\}$, donc comme $i$ est récurrent, $\mathbb{P}_i(N_i < +\infty) = 0$, donc $\mathbb{P}_j(T_i = +\infty) = 0$. Mais

  $$N_i = 1 + \sum_{k=1}^{+\infty} \mathbb{1}_{\{i\}}(X_{k+T_i}),$$

  Donc d'après la propriété de Markov forte

  $$\mathbb{P}_j(N_i = +\infty) = \mathbb{P}_i(\sum_{k=1}^{+\infty} \mathbb{1}_{\{i\}}(X_k) = +\infty) = \mathbb{P}_i(N_i = +\infty) = 1.$$

- (4) $\Longrightarrow$ (1). Évident.

□

**Définition:** Si une chaîne de Markov vérifie une des 4 propriétés équivalentes ci-dessus, on dit que c'est une *chaîne récurrente*.

## 5.3 Mesures invariantes

**Définition:** On dit qu'une mesure $\mu$ est *invariante* sous l'action de la matrice de transition markovienne $M$ si $\mu M = \mu$, c'est à dire.

$$\forall j \in S \quad \sum_{i \in S} \mu(i) m_{i,j} = \mu(j).$$

Si $\mu$ est *invariante* sous l'action de $M$, une récurrence immédiate donne $\forall n \geq 0 \mu M^n = \mu$. Ainsi, si $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $M$ et de mesure initiale $\mathbb{P}_{X_0} = \mu$, alors pour tout $n$, la loi de $X_n$ est $\mathbb{P}_{X_n} = \mu$.
**Définition:** On dit qu'une mesure $\mu$ est réversible sous l'action de la matrice de transition markovienne $M$ si

$$\forall i, j \in S \quad \mu(i) m_{i,j} = \mu(j) m_{j,i}.$$

**Théorème 25.** *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de loi initiale $\nu$ réversible sous l'action de $M$. Alors*

$$\forall n \geq 1 (X_0, X_1, \dots X_n) \textit{ et } (X_n, X_{n-1}, \dots, X_0) \textit{ ont même loi sous} \mathbb{P}_\nu.$$

*Démonstration.* Il suffit de procéder par récurrence sur $n$ que

$$\forall (X_0, \dots X_n) \in S^{n+1} \mathbb{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}_\nu(X_0 = x_n, X_1 = x_{n-1}, \dots, X_n = x_0).$$

Pour $n = 1$, il suffit de voir que

$$\mathbb{P}_\nu(X_0 = x_0, X_1 = x_1) = \nu(x_0) m_{x_0,x_1} = \nu(x_1) m_{x_1,x_0} = \mathbb{P}_\nu(X_0 = x_1, X_1 = x_0).$$

Ensuite

$$\begin{aligned}
\mathbb{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) &= \mathbb{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) m_{x_{n-1},x_n} \\
&= m_{x_{n-1},x_n} \mathbb{P}_\nu(X_0 = x_{n-1}, X_1 = x_{n-2}, \dots, X_{n-1} = x_0) \\
&= m_{x_{n-1},x_n} \nu(x_{n-1}) \prod_{i=1}^{n-1} m_{x_{n-i},x_{x-i-1}} \\
&= \nu(x_{n-1}) m_{x_{n-1},x_n} \prod_{i=1}^{n-1} m_{x_{n-i},x_{x-i-1}} \\
&= \nu(x_n) m_{x_n,x_{n-1}} \prod_{i=1}^{n-1} m_{x_{n-i},x_{x-i-1}} \\
&= \nu(x_n) \prod_{i=0}^{n-1} m_{x_{n-i},x_{x-i-1}} \\
&= \mathbb{P}_\nu(X_0 = x_n, X_1 = x_{n-1}, \dots, X_n = x_0).
\end{aligned}$$

□

Il est facile de voir que toute mesure réversible est invariante.

**Théorème 26.** *Si la matrice de transition $M$ est irréductible et admet une probabilité invariante, alors les chaînes de Markov associées à $M$ sont récurrentes.*

*Démonstration.* Soit $\mu$ une probabilité invariante Pour tout $n \geq 0$, on a $\mu M^n = \mu$, soit

$$\forall j \in S \quad \forall n \geq 0 \quad \sum_{i \in S} \mu(i) m_{i,j}^{(n)} = \mu(j)$$

Si une chaîne de Markov irréductible n'est pas récurrente, les états sont tous transitoires et $\lim_{n \to +\infty} \mu(i) m_{i,j}^{(n)} = 0$ quels que soient $i$ et $j$. D'après le théorème de convergence dominée, on a alors

$$\forall j \in S \quad 0 = \mu(j),$$

ce qui est impossible. □

**Théorème 27.** *Toute chaîne de Markov sur un espace d'états $S$ fini admet une probabilité invariante.*

*Démonstration.* L'ensemble $\mathcal{M}(S)$ des mesures de probabilité sur $S$ s'identifie au compact $K = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n ; \sum_{k=1}^n x_k = 1\}$, avec $n = |S|$. $\mathcal{M}(S)$ est un convexe stable par $\mu \mapsto \mu M$. Ainsi, si $\mu$ est une mesure quelconque sur $S$, la suite $(\mu_n)_{n \geq 0}$ définie par

$$\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \mu M^k$$

est à valeurs dans $\mathcal{M}(S)$. On a $\mu_n(I - M) = \frac{\mu(I - M^n)}{n}$. Comme la suite $(\mu I - M^n))_{n \geq 0}$, est bornée, il s'ensuit que toute valeur d'adhérence de $(\mu_n)_{n \geq 0}$ est laissée fixe par $M$. Comme $\mathcal{M}(S)$ est compacte, $(\mu_n)_{n \geq 0}$ a au moins une valeur d'adhérence donc $M$ au moins une mesure invariante. □

**Corollaire 13.** *Une chaîne de Markov irréductible dont l'espace d'états est fini est récurrente.*

## 5.4 Théorème de la probabilité stationnaire

**Théorème 28.** *Soit $M$ la matrice de transition d'une chaîne de Markov irréductible apériodique admettant $\mu$ comme loi stationnaire. Alors pour toute loi $\nu$ sur $S$, la chaîne de Markov de matrice de transition $M$ et de loi initiale $\nu$ converge vers $\mu$.*

*Démonstration.* Soit $X_0$, $X'_0$ deux variables aléatoires indépendantes, $X_0$ suivant la loi $\mu$, $X'_0$ la loi $\nu$. On note également $Y_0 = X'_0$. Soit également $(f_n)_{n\geq 1}$ et $(f'_n)_{n\geq 1}$ deux suites de variables aléatoires i.i.i.d. de loi $\theta_M$ définie au lemme 1, ces deux suites étant indépendantes de $X_0$ et $X'_0$. On définit par récurrence les suites $(g_n)_{n\geq 1}, (X_n)_{n\geq 1}$ et $(X'_n)_{n\geq 1}, (Y_n)_{n\geq 1}$ par

$$\begin{cases} X_{n+1} &= f_{n+1}(X_n) \\ Y_{n+1} &= f'_{n+1}(Y_n) \\ g_{n+1} &= \begin{cases} f_{n+1} & \text{si } X_n = X'_n \\ f'_{n+1} & \text{sinon} \end{cases} \\ X'_{n+1} &= g_{n+1}(X'_n) \end{cases}$$

Il n'est pas difficile de voir qu'en tout point $\omega$ on a

$$(X_n(\omega) = X'_n(\omega)) \Longrightarrow (f_{n+1}(\omega) = g_{n+1}(\omega)) \Longrightarrow (X_{n+1}(\omega) = X'_{n+1}(\omega))$$

Ainsi, les processus $X_n$ et $X'_n$ évoluent de manière indépendante jusqu'au moment où ils se rencontrent. À partir de là, $X'_n$ demeure scotché à $X_n$.

**Lemme 9.** *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition $M$ et de loi initiale $\mu$ , $(Y_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition $N$ et de loi initiale $\nu$. On suppose en outre que les suites $(X_n)_{n\geq 0}$ et $(Y_n)_{n\geq 0}$ sont indépendantes sous $P$. Alors la suite $(Z_n)_{n\geq 0}$ définie par $Z_n = (X_n, Y_n)$ est une chaîne de markov de matrice de transition $M \otimes N$.*

*Démonstration.* Soient $(x_0, \dots, x_n) \in S^{n+1}$ et $(y_0, \dots, y_n) \in S^{n+1}$.

$$\begin{aligned} P(\forall i \in \{0,n\}(X_i, Y_i) = (x_i, y_i)) &= P(\{\forall i \in \{0,n\} X_i = x_i\} \cap \{\forall i \in \{0,n\} Y_i = y_i\}) \\ &= P(\forall i \in \{0,n\} X_i = x_i) P(\forall i \in \{0,n\} Y_i = y_i) \\ &= \mu(\{x_0\}) \prod_{i=0}^{n-1} m_{x_i,x_{i+1}} \times \nu(\{y_0\}) \prod_{i=0}^{n-1} n_{y_i,y_{i+1}} \\ &= \mu(\{x_0\})\nu(\{y_0\}) \prod_{i=0}^{n-1} m_{x_i,x_{i+1}} n_{y_i,y_{i+1}} \\ &= (\mu \otimes \nu)(\{x_0, y_0\}) \prod_{i=0}^{n-1} (M \otimes N)((x_i, y_i), (x_{i+1}, y_{i+1})) \end{aligned}$$

□

**Lemme 10.** *Soit $U,V$ deux variables aléatoires de loi $\theta$. On suppose que sous $P$, $U$ et $V$ sont indépendantes de la tribu $\mathcal{A}$. Soit $A$ un événement*

*$\mathcal{A}-mesurable$. On définit $W$ par*

$$W(\omega) = \begin{cases} U(\omega) & si\ \omega \in A \\ V(\omega) & si\ \omega \notin A \end{cases}$$

*Alors, sous $P$, $W$ suit la loi $\theta$ et $W$ est indépendante de $\mathcal{A}$.*

*Démonstration.* Soit $A'$ un événement $\mathcal{A}-mesurable$ et $B$ un borélien

$$\begin{aligned} P(A' \cap \{W \in B\}) &= P(A \cap A' \cap \{W \in B\}) + P(A^c \cap A' \cap \{W \in B\}) \\ &= P(A \cap A' \cap \{U \in B\}) + P(A^c \cap A' \cap \{V \in B\}) \\ &= P(A \cap A')P(U \in B) + P(A^c \cap A')P(V \in B) \\ &= P(A \cap A')\theta(B) + P(A^c \cap A')\theta(B) \\ &= (P(A \cap A') + P(A^c \cap A'))\theta(B) \\ &= P(A')\theta(B) \end{aligned}$$

En prenant $A' = \Omega$, on en déduit d'abord que $P(W \in B) = \theta(B)$ pour tout borélien $B$. $\theta$ est donc la loi de $W$ sous P. En réinsérant dans la formule précédente, on a pour tout événement $\mathcal{A}-mesurable$ $A'$ et pour tout borélien $B$ :

$$P(A' \cap \{W \in B\}) = P(A')P(W \in B),$$

ce qui veut dire que $W$ est indépendante de $\mathcal{A}$. □

En appliquant le lemme précédent à $\mathcal{A} = \sigma(X_0, X'_0, f_1, \ldots, f_n, f'_1, \ldots, f'_n)$, $A = \{X_n = X'_n\}$, $U = f_{n+1}$, $V = f'_{n+1}$ et $W = g_{n+1}$ on voit que $g_{n+1}$ suit la loi $\theta_M$ et que $g_{n+1}$ est indépendante de $\sigma(X_0, X'_0, f_1, \ldots, f_n, f'_1, \ldots, f'_n)$. Comme $(g_1, \ldots, g_n)$ est $\sigma(X_0, X'_0, f_1, \ldots, f_n, f'_1, \ldots, f'_n)$-mesurable, il s'ensuit que $(g_n)_{n \geq 1}$ est une suite de v.a.i.i.d de loi $\theta_M$.

D'après le lemme 6, $(X_n)$ est une chaîne de Markov de matrice de transition $M$ et de loi initiale $\mu$ tandis que $(X'_n)$ est une chaîne de Markov de matrice de transition $M$ et de loi initiale $\nu$.

On va maintenant montrer que $\tau = \inf\{n; X_n = X'_n\}$ est presque sûrement fini. Il est facile de voir que $\tau = \inf\{n; X_n = Y_n\}$. Ce qui est intéressant, c'est que $(X_n)_{n \geq 0}$ et $(Y_n)_{n \geq 0}$ sont indépendants.

Ainsi, d'après le lemme 9, $(X_n, Y_n)$ est une chaîne de Markov de loi initiale $\nu \otimes \mu$ et de matrice de transition $N = M \otimes M$. Soient $(x, y, z, t) \in S^4$. Comme $M$ est la matrice d'une chaîne de Markov irréductible et apériodique, on peut, d'après le lemme 8, trouver un entier $n_0$ tel que $M^{n_0}$ soit à coefficients strictement positifs. Or $N^{n_0} = (M \otimes M)^{n_0} = M^{n_0} \otimes M^{n_0}$ : on a

$$N^{n_0}((x, y), (z, t)) = M^{n_0}(x, z)M^{n_0}(y, t) > 0.$$

Ainsi $((Z_n)_{n\ge 0} = ((X_n, Y_n))_{n\ge 0}$ est une chaîne de Markov irréductible. Comme $M \otimes M$ admet $\mu \otimes \mu$ comme mesure invariante, la dynamique est donc récurrente : $(Z_n)_{n\ge 0}$ passe donc presque sûrement en tout point de $S \times S$. En particulier, elle passe presque sûrement sur sa diagonale, ce qui implique que $P(\tau < +\infty) = 1$.

Soit $f$ une fonction bornée de $S$ dans $\mathbb{R}$. Pour $n \ge \tau$, on a $f(X_n) = f(X'_n)$. Donc $f(X_n) - f(X'_n)$ converge presque sûrement vers 0. D'après le théorème de convergence dominée, on en déduit que $\mathbb{E}(f(X_n) - f(X'_n))$ converge vers 0. Comme $\mu$ est invariante $\mathbb{E}(f(X_n) - f(X'_n)) = \int f\ d\mu - \mathbb{E}f(X'_n)$. Ainsi pour toute fonction $f$, $\mathbb{E}f(X'_n)$ converge vers $\int f\ d\mu$, ce qui veut dire que $X'_n$ converge en loi vers $\mu$.

□

**Corollaire 14.** *Une chaîne de Markov irréductible apériodique a au plus une loi stationnaire.*

**Remarque-exercice :** l'hypothèse d'apériodicité est importante. En effet, on peut construire deux chaînes de Markov indépendantes $(X_n)_{n\ge 0}$ et $(Y_n)_{n\ge 0}$ ayant la même matrice de transition irréductibles, telles que $(X_n, Y_n)_{n\ge 0}$ ne soit pas irréductible et que $(X_n, Y_n)_{n\ge 0}$ ne coupe jamais la diagonale. Donner deux exemples d'un tel phénomène, l'un avec $S$ fini, l'autre avec $S$ infini.

## 5.5 Théorème ergodique des chaînes de Markov

**Théorème 29.** *Soit $(X_n)_{n\ge 0}$ une chaîne de Markov irréductible et récurrente. Pour tout $x \in S$, on a*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} 1\!\!1_{\{x\}}(X_k) \to \frac{1}{\mathbb{E}_x T_x}.$$

*Démonstration.* Pour tout $k \ge 1$, posons $T^k = \inf\{n > 0, \sum_{k=1}^n 1\!\!1_{\{x\}}(X_k) \ge k$. Comme la chaîne est récurrente, les $T^k$ sont presque sûrement finis. Il est aisé de constater que la suite $(T^k)_{k\ge 1}$ est croissante. Soit $k \ge 1$ et $A \in \sigma(T^1, \dots, T^k)$ : il est clair que $A$ se produit avant $T^k$. Donc

$$\begin{aligned}
\mathbb{P}_x(A, T^{k+1} - T_k > n) &= \mathbb{P}_x(A, \cap_{j=T^k+1}^{T^k+n} 1\!\!1_{\{x\}}(X_j) = 0) \\
&= \mathbb{P}_x(A)\mathbb{P}_x(\cap_{j=1}^{n} 1\!\!1_{\{x\}}(X_j) = 0) \\
&= \mathbb{P}_x(A)\mathbb{P}_x(T^1 > n)
\end{aligned}$$

On en déduit que, sous la loi $\mathbb{P}_x$, les variables aléatoires $T^1, T^2 - T^1, T^3 - T^2, \dots$ forment une suite de variables aléatoires positives indépendantes ayant même loi que $T^1 = T_x$. D'après la loi forte des grands nombres on en déduit que $\frac{T^n}{n} = \frac{1}{n}\big(T^1 + (T^2 - T^1) + (T^3 - T^2) + \dots (T^n - T^{n-1}\big)$ converge presque sûrement vers $\mathbb{E}_x T^x$. Lorsque la loi initiale n'est pas la masse de Dirac en $x$, il suffit de remarquer que la suite $\frac{T^n}{n}$ a le même comportement asymtotique que la suite $\frac{T^n - T^1}{n-1}$ et appliquer à nouveau la propriété de Markov forte : la loi sous $P$ de $(\frac{T^n - T^1}{n-1})_{n\ge 2}$ est la loi sous $\mathbb{P}_x$ de $(\frac{T^n}{n})_{n \ge 1}$, d'où le résultat. Posons $S_n = \sum_{k=0}^{n-1} \mathbb{1}_{\{x\}}(X_k)$. Un instant de réflexion montre que $T^{S_n} \le n < T^{S_n+1}$. On en déduit

$$\frac{T^{S_n}}{S_n} \le \frac{n}{S_n} < \frac{T^{S_n+1}}{S_n+1}\frac{S_n+1}{S_n}$$

Comme $x$ est récurrent $\lim_{n\to+\infty} S_n = +\infty$, donc $\lim_{n\to+\infty} \frac{n}{S_n} = \mathbb{E}_x T_x$, d'où le résultat. □

**Théorème 30.** *Soit $(X_n)_{n\ge 0}$ une chaîne de Markov irréductible admettant une probabilité invariante $\mu$. Pour tout $x \in S$, on a*

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

*Démonstration.* Une chaîne de Markov irréductible admettant une probabilité invariante est toujours récurrente. Le théorème précédent s'applique donc. Comme $|\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{\{x\}}(X_k)| \le 1$, le théorème de convergence dominée s'applique et on a

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{P}_\nu(X_k = x) \to \frac{1}{\mathbb{E}_x T_x}.$$

Si l'on prend pour *nu* la mesure invariante $\mu$, on a pour tout $k \ge 0$ $\mathbb{P}_\nu(X_k = x) = \mu(x)$. On en déduit que $\frac{1}{\mathbb{E}_x T_x} = \mu(x)$, ce qui achève la preuve. □

## 5.6 Exercices

1. *Temps d'atteinte d'un état absorbant.*

   Soit $(X_n)$ une chaîne de Markov sur un ensemble dénombrable $E$ de noyau de transition $Q$, et $a \in E$ un état absorbant (*ie* : $Q(a,a) = 1$). Montrer que $P(X_n = a) = P(T_a \le n)$.

2. *Temps d'entrée : une propriété d'invariance.*
   Soit $(X_n)$ une chaîne de Markov sur un ensemble dénombrable $E$ de noyau de transition $Q$. Pour $f : E \to \mathbb{R}_+$, soit $Qf$ la fonction définie par $Qf(x) = \sum_{y\in E} Q(x,y)f(y)$. Pour $B \subset E$ on note $D_B = \inf\{n \geq 0; X_n \in B\}$ le temps d'entrée dans $B$. Montrer que la fonction $f$ définie par $f(x) = \mathbb{P}_x(T_B < +\infty)$ vérifie $(I-Q)f = 0$ sur $E \setminus B$, et $f = 1$ sur $B$.
3. *Chaîne de Markov arrêtée.*
   Soit $(X_n)$ une chaîne de Markov sur un ensemble dénombrable $E$ de noyau de transition $Q$. Etant donné un ensemble $B \subset E$, on pose $Y_n = X_{n\wedge D_B}$. Montrer que $(Y_n)$ est une chaîne de Markov sur $E$ de noyau de transition $Q'$ défini par $Q'(x,y) = \delta_x(\{y\})$ si $x \in B$ et $Q'(x,y) = Q(x,y)$ si $x \notin B$.
4. *Chaîne observée quand elle bouge.* Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène sur l'espace d'états $E$, de matrice de transition $P$. On définit la suite $(T_k)_{k\in\mathbb{N}}$ de la façon suivante : $T_0 = 0$ et
   $$T_{k+1} = \min\{n \geq T_k, X_n \neq X_{T_k}\}.$$
   On suppose que la chaîne n'admet aucun état absorbant.
   (a) Montrer que les $(T_k)_{k\in\mathbb{N}}$ sont des temps d'arrêt pour $(X_n)_{n\in\mathbb{N}}$, finis presque sûrement.
   (b) On définit $Y_k = X_{T_k}$. Montrer que $(Y_k)_{K\in\mathbb{N}}$ est une chaîne de Markov homogène, donner son espace d'états et sa matrice de transition.
5. *Chaîne restreinte.* Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène d'espace d'états $E$ dénombrable et de matrice de transition $P = (p_{i,j})_{i,j\in E}$. Soit $J$ une partie de $E$. On observe cette chaîne de Markov seulement lors de ses passages par $J$, et on note $Y_m$ la $m^{\text{ième}}$ observation. Plus formellement, pour $m \geq 1$, on note
   $$T_m = \inf\left\{n \geq 1 + T_{m-1} \,\middle|\, X_n \in J\right\} \text{ et } T_0 = \inf\left\{n \geq 0 \,\middle|\, X_n \in J\right\}.$$
   On suppose que $\forall m \geq 0$, $T_m < +\infty$, et on pose $Y_m = X_{T_m}$.
   (a) Montrer que $T_m$ est un temps d'arrêt, que $X_{T_m}$ est mesurable pour $\mathcal{F}_{T_m}$ et que pour $k \leq m$, $T_k$ et $X_{T_k}$ sont $\mathcal{F}_{T_m}$-mesurables.
   (b) Montrer que $(Y_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène.
   *Indication.* On pourra montrer, en décomposant sur les valeurs possibles de $T_m - T_{m-1}$, que, si $(k_0,\dots,k_m) \in J^{m+1}$ alors $P(Y_0 =$

$k_0, \dots Y_m = k_m) = P(Y_0 = k_0, \dots Y_{m-1} = k_{m-1})\mathbb{P}_{k_{m-1}}(X_S = k_m)$, où $S = \inf\left\{n \geq 1 \,\middle|\, X_n \in J\right\}$.

(c) On note $Q = (q_{i,j})_{i,j\in J}$ la matrice de transition de $(Y_n)_{n\in\mathbb{N}}$. Montrer que

$$q_{i,j} = p_{i,j} + \sum_{l\notin J} p_{i,l}\mathbb{P}_l(X_{T_0} = j)$$

et donner une caractérisation des $(\mathbb{P}_l(X_{T_0} = j))_{l\notin J, j\in J}$.

(d) Exprimer, en fonction de la loi initiale $\mu_0$ de $(X_n)_{n\in\mathbb{N}}$, la loi initiale $\nu_0$ de $(Y_n)_{n\in\mathbb{N}}$.

(e) Exemple. On considère la marche aléatoire symétrique aux plus proches voisins sur l'hypercube $\{0,1\}^d$, et on choisit $J = \{(x_i)_{1\leq i\leq d}, \sum x_i \text{ impaire}\}$. Déterminer les caractéristiques de $(Y_n)_{n\in\mathbb{N}}$ dans ce cas.

6. *Marche aléatoire sur $\mathbb{Z}/d\mathbb{Z}$.* On considère $X = \{X_n : n \geq 0\}$ la marche aléatoire sur $\mathbb{Z}/d\mathbb{Z}$ dont les pas sont indépendants de même loi $p\delta_1 + (1-p)\delta_{-1}$.

   (a) Quelle est sa matrice de transition $P$ ? Dessiner son graphe.

   (b) Calculer $P^n$. (Indication : écrire $P = pA + (1-p)B$ et trouver une relation entre $A$ et $B$.)

   (c) Calculer la probabilité stationnaire $\pi$.

   (d) (calculatoire) Calculer la loi $\mu_n$ de $X_n$. Discuter l'existence de $\lim_n \mu_n$, $\lim_n P^n$.

7. *Modèle d'Ehrenfest.* Etant donné deux enceintes séparées par une paroi poreuse et contenant ensemble $N$ particules diffusant à travers cette paroi, on décrit le nombre aléatoire $X_n$ de particules se trouvant dans la première enceinte ($0 \leq X_n \leq N$) aux instants successifs $n \in \mathbb{N}$ de transition des particules par une chaîne Markov de probabilité de transition :

$$p(x, x-1) := \frac{x}{N},\ p(x, x+1) := \frac{N-x}{N},\ 0 \leq x \leq N.$$

   À chaque instant $n$, les probabilités que la transition se fasse de la première enceinte vers la seconde ou de la seconde enceinte vers la première sont donc proportionnelles aux nombres de particules en présence dans la première et la deuxième enceinte :

$$\frac{p(x, x-1)}{x} = \frac{p(x, x+1)}{N-x}.$$

   Montrer que cette chaîne est récurrente irréductible et trouver sa probabilité invariante.

8. *Chaîne de Markov avec décision.*

   Le $n^{\text{ème}}$ Lundi de l'année, une petite entreprise reçoit $A_n$ propositions de travail de type A, et $B_n$ propositions de travail de type B. Un travail de type A mobilise toute la capacité de travail de l'entreprise durant une semaine et lui rapporte 200 euros, alors qu'un travail de type B l'occupe deux semaines pour un rapport de 360 euros. Une semaine d'inactivité coûte 100 euros, un travail non traité pendant la semaine où il arrive est perdu. On suppose $A_n$,$B_n$ indépendants, les couples $(A_n, B_n)_{n\geq 1}$ indépendants, et

   $$\mathbb{P}(A_n = 1) = 1-\mathbb{P}(A_n = 0) = 0,5, \qquad \mathbb{P}(B_n = 1) = 1-\mathbb{P}(B_n = 0) = 0,6.$$

   Modéliser la situation par une chaîne de Markov , avec si possible un nombre d'états minimal. Quelle est la meilleure stratégie, quand on reçoit simultanément une offre de chaque type : donner la préférence à celle de type A ou à celle de type B ? On pourra faire appel au Théorème ergodique pour départager les deux politiques.

9. *Un modèle de prédiction météo (!)* On suppose que le temps qu'il fera demain depend des deux jours précédents. On suppose que :

   $$\begin{aligned}
   &\mathbb{P}(\text{ il pleut demain} \mid \text{il a plu hier et aujourd'hui}) = 0,7\\
   &\mathbb{P}(\text{ il pleut demain} \mid \text{il a plu aujourd'hui mais pas hier}) = 0,5\\
   &\mathbb{P}(\text{ il pleut demain} \mid \text{il a plu hier mais pas aujourd'hui}) = 0,4\\
   &\mathbb{P}(\text{ il pleut demain} \mid \text{il n'a pas plu ni hier, ni aujourd'hui}) = 0,2
   \end{aligned}$$

   Montrer qu'on peut modéliser ceci par une chaîne de Markov. Quelle est la probabilité, sachant qu'il a plu lundi et mardi qu'il pleuve jeudi ? Sur le long terme, quelle proportion de jours de pluie observe-t-on ?

10. *Chaîne de Markov réversible*

    (a) Soit $P$ une matrice de transition sur un espace d'états $E$ dénombrable. On suppose qu'il existe une probabilité $\pi$ telle que

    $$\pi_i p_{i,j} = \pi_j p_{j,i}.$$

    Montrer que $\pi$ est stationnaire pour la $P$.

    (b) Trouver rapidement la probabilité stationnaire de la marche aléatoire symétrique sur les sommets de l'hypercube de dimension $d$.

    (c) Marche aléatoire symétrique sur un damier. Calculer les temps de retours moyens des différents points de l'échiquier.

11. *Modèle de Laplace-Bernoulli.* $N$ boules noires et $N$ boules blanches sont placées dans deux urnes de sorte que chacune contienne $N$ boules. Après chaque unité de temps on choisit au hasard une boule de chaque urne ; les deux boules ainsi choisies changent d'urne. On note par $X_n$ le nombre de boules noires dans la première urne. Montrer que $\{X_n : n \geq 0\}$ est une chaîne de Markov irréductible réversible et trouver sa mesure stationnaire.