Advanced Probability

Lectures by Grégory Miermont
Notes by Alexis Marchand

ENS de Lyon
S1 2018-2019
M1 course

Contents

Notation 0.0.1. *Throughout these notes,* $(\Omega, \mathcal{F}, \mathbb{P})$ *will be a probability space.*

1 Conditional expectation

1.1 An elementary case

Example 1.1.1.

(i) *If* $A \in \mathcal{F}$*, with* $\mathbb{P}(A) > 0$*, then one can define a new probability measure* $\mathbb{P}(\cdot \mid A)$ *on* $(\Omega, \mathcal{F})$ *by* $\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$*. Now, if* $X : \Omega \to \mathbb{R}$ *is an integrable or nonnegative random variable, measurable w.r.t.* $\mathcal{F}$*, one can consider its integral w.r.t.* $\mathbb{P}(\cdot \mid A)$*:*

$$\mathbb{E}[X \mid A] = \int_\Omega X(\omega)\mathbb{P}(\mathrm{d}\omega \mid A) = \frac{\mathbb{E}[X\mathbb{1}_A]}{\mathbb{P}(A)}.$$

(ii) *Let* $(A_i)_{i \in I}$ *be a partition of* Ω *indexed by a countable set* I*. Let* X *be an integrable random variable. By convention, we set* $\mathbb{E}[X \mid A_i] = 0$ *if* $\mathbb{P}(A_i) = 0$*. Now, define:*

$$X' = \sum_{i \in I} \mathbb{E}[X \mid A_i] \mathbb{1}_{A_i}.$$

We note that (1) X' *is a random variable, measurable w.r.t.* $\mathcal{G} = \sigma(A_i,\ i \in I)$*, and* (2) *for all* $B \in \mathcal{G}$*,* $\mathbb{E}[X'\mathbb{1}_B] = \mathbb{E}[X\mathbb{1}_B]$*. Properties* (1) *and* (2) *will be the fundamental properties of conditional expectation.*

1.2 The general case

Theorem 1.2.1 (Kolmogorov's Theorem)**.** *Let* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-algebra of* $\mathcal{F}$*. Let* $X : \Omega \to \mathbb{R}$ *be an integrable random variable. Then there exists a random variable* X'*, integrable and s.t.*

(i) X' *is* $\mathcal{G}$*-measurable.*

(ii) $\forall B \in \mathcal{G},\ \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[X'\mathbb{1}_B]$.

The random variable X' *is unique up to almost sure equality; its equivalence class will be called the* expectation of X given $\mathcal{G}$ *and denoted by* $\mathbb{E}[X \mid \mathcal{G}]$*. It is an element of* $L^1(\Omega, \mathcal{G}, \mathbb{P})$*.*

Proof. *Uniqueness.* Let X' and X'' be two integrable random variables satisfying (i) and (ii). Consider $B = (X' > X'') \in \mathcal{G}$. We have $X'\mathbb{1}_B \geq X''\mathbb{1}_B$; therefore:

$$\mathbb{E}[X''\mathbb{1}_B] \leq \mathbb{E}[X'\mathbb{1}_B] = \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[X''\mathbb{1}_B].$$

Hence, equality must hold throughout, which implies that $X'\mathbb{1}_B = X''\mathbb{1}_B$ almost surely, i.e. $\mathbb{P}(B) = 0$. Thus, $X' \leq X''$ a.s., and by symmetry, $X' \geq X''$ a.s., so $X' = X''$ a.s. *Existence. First step*: we assume that $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Note that $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space, and $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace, so there exists a unique orthogonal projection $\pi : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to L^2(\Omega, \mathcal{G}, \mathbb{P})$, and π is characterised by:

$$\forall Z \in L^2(\Omega, \mathcal{G}, \mathbb{P}),\ \langle X - \pi(X) \mid Z \rangle = 0,$$

or in other words: $\forall Z \in L^2(\Omega, \mathcal{G}, \mathbb{P}),\ \mathbb{E}[XZ] = \mathbb{E}[\pi(X)Z]$. In particular, $\pi(X)$ satisfies (i) and (ii), which proves the theorem in the special case where X is in L^2. *Second step*: note that $\mathbb{E}[\cdot \mid \mathcal{G}] : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to L^2(\Omega, \mathcal{G}, \mathbb{P})$ is a positive linear map. Linearity is clear (it is a projection). For positivity, take $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ s.t. $X \geq 0$ a.s. and consider the event $B = (\mathbb{E}[X \mid \mathcal{G}] < 0) \in \mathcal{G}$. We have $\mathbb{E}[X \mid \mathcal{G}] \mathbb{1}_B \leq 0$, therefore:

$$0 \leq \mathbb{E}[X \mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mathbb{1}_B] \leq 0.$$

Hence, equality holds throughout and $\mathbb{P}(B) = 0$. *Third step*: we assume that X is a nonnegative random variable, measurable w.r.t. $\mathcal{F}$ (but not necessarily integrable). For $n \in \mathbb{N}$, the random variable $X \wedge n = \min(X, n)$ is bounded, so it is in L^2 and has a conditional expectation given $\mathcal{G}$ (according to the first step). Since $X \wedge n \leq X \wedge (n+1)$ for all $n \in \mathbb{N}$, and $\mathbb{E}[\cdot \mid \mathcal{G}]$ is positive linear on L^2 (according to the second step), we have $\mathbb{E}[X \wedge n \mid \mathcal{G}] \leq \mathbb{E}[X \wedge (n+1) \mid \mathcal{G}]$ for all $n \in \mathbb{N}$. Therefore, there exists a $\mathcal{G}$-measurable random variable X' with values in $[0, +\infty]$ s.t. $\mathbb{E}[X \wedge n \mid \mathcal{G}] \xrightarrow[n \to +\infty]{} X'$ a.s. But by monotone convergence, for all $B \in \mathcal{G}$, $\mathbb{E}[(X \wedge n) \mathbb{1}_B] \xrightarrow[n \to +\infty]{} \mathbb{E}[X \mathbb{1}_B]$ and $\mathbb{E}[(X \wedge n) \mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X \wedge n \mid \mathcal{G}] \mathbb{1}_B] \xrightarrow[n \to +\infty]{} \mathbb{E}[X' \mathbb{1}_B]$. This proves that $\forall B \in \mathcal{G}$, $\mathbb{E}[X \mathbb{1}_B] = \mathbb{E}[X' \mathbb{1}_B]$. Moreover, note that if X is integrable, then so is X' because $\mathbb{E}[X'] = \mathbb{E}[X' \mathbb{1}_\Omega] = \mathbb{E}[X \mathbb{1}_\Omega] < +\infty$. This proves the result in the special case where X is nonnegative. *Fourth step*: we assume that X is any integrable random variable in $L^1(\Omega, \mathcal{F}, \mathbb{P})$. We write $X = X^+ - X^-$, with $X^+ = X \vee 0$ and $X^- = (-X) \vee 0$, and apply the third step to X^+ and X^-. $\square$

Remark 1.2.2.

(i) *One could also prove Kolmogorov's Theorem by applying the Radon-Nikodym Theorem and setting* $X' = \frac{(X \cdot \mathrm{d}\mathbb{P})_{|\mathcal{G}}}{\mathrm{d}\mathbb{P}_{|\mathcal{G}}}$.

(ii) *Note that, by a standard approximation argument, one could replace property* (ii) *in the theorem by* $\mathbb{E}[XZ] = \mathbb{E}[X'Z]$ *for every bounded and* $\mathcal{G}$*-measurable random variable* Z.

(iii) *In the course of the proof, we have also showed the existence of* $\mathbb{E}[X \mid \mathcal{G}]$ *if* X *is any random variable with values in* $[0, +\infty]$. *It is actually possible to adapt the proof of the uniqueness to this setting.*

1.3 Elementary properties of conditional expectation

Proposition 1.3.1. *Let* $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *and consider a sub-*σ*-algebra* $\mathcal{G}$ *of* $\mathcal{F}$.

(i) $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$.

(ii) *If* X *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}[X \mid \mathcal{G}] = X$ *a.s.*

(iii) *If* X *is independent of* $\mathcal{G}$*, then* $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ *a.s.*

(iv) *If* $a, b \in \mathbb{R}$*, then* $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$ *a.s.*

(v) *If* $X \geq 0$ *a.s., then* $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ *a.s.*

(vi) $|\mathbb{E}[X \mid \mathcal{G}]| \leq \mathbb{E}[|X| \mid \mathcal{G}]$ *a.s., and in particular* $\mathbb{E}[|\mathbb{E}[X \mid \mathcal{G}]|] \leq \mathbb{E}[|X|]$.

Proposition 1.3.2 (Conditioned basic integration theorems)**.** *Let* $(X_n)_{n \in \mathbb{N}} \in L^1(\Omega, \mathcal{F}, \mathbb{P})^{\mathbb{N}}$ *and* $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

(i) Monotone Convergence. *Assume that* $\forall n \in \mathbb{N}$, $X_n \geq 0$ *a.s., and that* $(X_n)_{n \in \mathbb{N}}$ *is a.s. nondecreasing and that* $X_n \xrightarrow[n \to +\infty]{} X$ *a.s. Then* $\mathbb{E}[X_n \mid \mathcal{G}] \xrightarrow[n \to +\infty]{} \mathbb{E}[X \mid \mathcal{G}]$ *a.s.*

(ii) Fatou's Lemma. *Assume that* $\forall n \in \mathbb{N}$, $X_n \geq 0$ *a.s. Then:*

$$\mathbb{E}\left[\liminf_{n \to +\infty} X_n \mid \mathcal{G}\right] \leq \liminf_{n \to +\infty} \mathbb{E}[X_n \mid \mathcal{G}] \text{ a.s.}$$

(iii) Dominated Convergence. *If* $X_n \xrightarrow[n\to+\infty]{} X$ *a.s. and if there exists an integrable random variable* φ *s.t.* $\forall n \in \mathbb{N}$, $|X_n| \leq \varphi$ *a.s., then* $\mathbb{E}[X_n \mid \mathcal{G}] \xrightarrow[n\to+\infty]{} \mathbb{E}[X \mid \mathcal{G}]$ *a.s.*

(iv) Jensen's Inequality. *If* $\varphi : \mathbb{R} \to]-\infty, +\infty]$ *is a convex function and if* $\varphi(X)$ *is integrable or nonnegative, then:*

$$\mathbb{E}[\varphi(X) \mid \mathcal{G}] \geq \varphi(\mathbb{E}[X \mid \mathcal{G}]) \text{ a.s.}$$

(v) *For all* $p \in [1, +\infty]$, $\|\mathbb{E}[X \mid \mathcal{G}]\|_p \leq \|X\|_p$, *so* $\mathbb{E}[\cdot \mid \mathcal{G}] : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to L^p(\Omega, \mathcal{G}, \mathbb{P})$ *is a continuous projection.*

1.4 Specific properties of conditional expectation

Proposition 1.4.1. *Let* $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. *If* $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *is bounded and* $\mathcal{G}$*-measurable, then:*

$$\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}].$$

Proposition 1.4.2. *Let* X *be an integrable or nonnegative random variable. Consider* $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ *sub-*σ*-algebras of* $\mathcal{F}$. *We have:*

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}]$$

Notation 1.4.3. *If* $(\mathcal{G}_i)_{i\in I}$ *is a family of sub-*σ*-algebras of* $\mathcal{F}$, *we shall denote by* $\bigvee_{i\in I} \mathcal{G}_i$ *the* σ*-algebra generated by* $\bigcup_{i\in I} \mathcal{G}_i$.

Proposition 1.4.4. *Let* X *be an integrable or nonnegative random variable. Consider sub-*σ*-algebras* $\mathcal{G}$ *and* $\mathcal{H}$ *of* $\mathcal{F}$ *and assume that* $\mathcal{H}$ *is independent of* $\sigma(X) \vee \mathcal{G}$. *Then:*

$$\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}].$$

Proof. Let $(A, B) \in \mathcal{G} \times \mathcal{H}$. Since $\mathbb{1}_{A\cap B}$ is $(\mathcal{G} \vee \mathcal{H})$-measurable, we have:

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] \mathbb{1}_{A\cap B}] &= \mathbb{E}[X\mathbb{1}_{A\cap B}] = \mathbb{E}[(X\mathbb{1}_A)\mathbb{1}_B] = \mathbb{E}[X\mathbb{1}_A]\mathbb{E}[\mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]\mathbb{1}_A]\mathbb{E}[\mathbb{1}_B] \\ &= \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]\mathbb{1}_A\mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]\mathbb{1}_{A\cap B}].\end{aligned}$$

But note that $\{A \cap B,\ (A, B) \in \mathcal{G} \times \mathcal{H}\}$ is stable under finite intersections and generates $\mathcal{G} \vee \mathcal{H}$. By the Monotone Class Theorem, we have $\forall C \in \mathcal{G} \vee \mathcal{H}$, $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}]\mathbb{1}_C] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]\mathbb{1}_C]$, therefore $\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}]$. □

Notation 1.4.5. *If* X *and* Y *are two random variables s.t.* X *is integrable or nonnegative, we shall write* $\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)]$.

Proposition 1.4.6. *Let* X *and* Y *be random variables taking values in measurables spaces* $(E, \mathcal{E})$ *and* $(E', \mathcal{E}')$ *respectively. Assume that* X *and* Y *are independent and consider a function* $f : E \times E' \to \mathbb{R}_+$ *that is measurable w.r.t.* $\mathcal{E} \otimes \mathcal{E}'$. *Then:*

$$\mathbb{E}[f(X, Y) \mid Y] = \int_E f(x, Y)\mathbb{P}(X \in \mathrm{d}x).$$

Note that, according to Fubini's Theorem, the function $y \mapsto \int_E f(x, y)\mathbb{P}(X \in \mathrm{d}x)$ *is measurable, so it makes sense to talk about the random variable* $\int_E f(x, Y)\mathbb{P}(X \in \mathrm{d}x)$.

1.5 Notion of conditional distribution

Definition 1.5.1 (Conditional distribution). *If $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$, we define $\mathbb{P}(A \mid \mathcal{G}) = \mathbb{E}[\mathbb{1}_A \mid \mathcal{G}]$ for $A \in \mathcal{F}$.*

Remark 1.5.2. *The latter definition is very dangerous because $\mathbb{P}(A \mid \mathcal{G})$ is a random variable that is only defined $\mathbb{P}$-almost everywhere. For any fixed family $(A_n)_{n\in\mathbb{N}}$ of pairwise disjoint $\mathcal{F}$-measurable sets, it will be $\mathbb{P}$-almost surely the case that $\mathbb{P}(\bigsqcup_{n\in\mathbb{N}} A_n \mid \mathcal{G}) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n \mid \mathcal{G})$. However, this needs not be true for all choices of $(A_n)_{n\in\mathbb{N}}$ $\mathbb{P}$-almost surely.*

Definition 1.5.3 (Kernel). *Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. A* kernel *from E to F is a function $K : E \times \mathcal{F} \to \mathbb{R}_+$ s.t.*

(i) *For all $A \in \mathcal{F}$, $K(\cdot, A) : E \to \mathbb{R}_+$ is measurable.*

(ii) *For all $x \in E$, $K(x, \cdot) : \mathcal{F} \to \mathbb{R}_+$ is a probability measure.*

Definition 1.5.4 (Regular version of a conditional distribution). *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. We say that the kernel Q from $(\Omega, \mathcal{G})$ to $(\Omega, \mathcal{F})$ is a* regular version *of $\mathbb{P}(\cdot \mid \mathcal{G})$ if for all $A \in \mathcal{F}$, $Q(\omega, A) = \mathbb{P}(A \mid \mathcal{G})(\omega)$ $\mathbb{P}$-almost surely.*

Theorem 1.5.5. *Let $(E, \mathcal{E})$ be a Borel space (i.e. $(E, \mathcal{E})$ is isomorphic as a measurable space to a Borel subset of $\mathbb{R}$). If X is a random variable with values in $(E, \mathcal{E})$, then the conditional distribution of X given $\mathcal{G}$ admits a regular version, i.e. there exists a kernel Q from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ s.t. for all $A \in \mathcal{E}$, $Q(\omega, A) = \mathbb{P}(X \in A \mid \mathcal{G})(\omega)$ $\mathbb{P}$-almost surely.*

Example 1.5.6. *Assume that (X, Y) is a random variable in $\mathbb{R}^2$ whose law has a density f w.r.t. Lebesgue's Measure:*

$$\forall A \in \operatorname{Bor}\left(\mathbb{R}^2\right),\ \mathbb{P}((X, Y) \in A) = \int_A f(x, y)\, \mathrm{d}x\, \mathrm{d}y.$$

Then Y has a density given by $f_Y(y) = \int_{\mathbb{R}} f(x, y)\, \mathrm{d}x$. Therefore, for all nonnegative measurable function h, we have:

$$\mathbb{E}[h(X) \mid Y] = \int_{\mathbb{R}} \left(\mathrm{d}x \cdot \frac{f(x, y)}{f_Y(y)}\right) h(x).$$

Hence the kernel $Q(y, A) = \int_A \mathrm{d}x \cdot \frac{f(x,y)}{f_Y(y)}$ is a regular version of X given Y.

1.6 The Gaussian case

Example 1.6.1. *Let (X, Y) be a Gaussian vector in $\mathbb{R}^2$, i.e. $(sX + tY)$ has a Gaussian law for all $(s, t) \in \mathbb{R}^2$. We assume that $\mathbb{E}[X] = \mathbb{E}[Y] = 0$. Recall that for Gaussian variables, independence is equivalent to orthogonality. This leads to:*

$$\mathbb{E}[X \mid Y] = \frac{\operatorname{Cov}(X, Y)}{\operatorname{Var}(Y)} Y.$$

More precisely, the conditional law of X given Y is $\mathcal{N}\left(\frac{\operatorname{Cov}(X,Y)}{\operatorname{Var}(Y)} Y, \frac{\operatorname{Var}(X)\operatorname{Var}(Y) - \operatorname{Cov}(X,Y)^2}{\operatorname{Var}(Y)}\right)$.

2 Martingales

2.1 Filtrations and martingales

Definition 2.1.1 (Filtration). *Given $(\Omega, \mathcal{F}, \mathbb{P})$, a* filtration *is a nondecreasing sequence $(\mathcal{F}_n)_{n\in\mathbb{N}}$ of sub-σ-algebras of $\mathcal{F}$:*

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \cdots.$$

One should think of $\mathcal{F}_n$ as the information available at time n. Given a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, the space $\left(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P}\right)$ is called a filtered probability space. *A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is said to be* adapted *to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if X_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$.*

Example 2.1.2 (Canonical filtration of a sequence of random variables)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. For $n \in \mathbb{N}$, define $\mathcal{F}_n^X = \sigma(X_0, \ldots, X_n)$. Then $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$ is called the* canonical filtration *of $(X_n)_{n\in\mathbb{N}}$. The sequence $(X_n)_{n\in\mathbb{N}}$ is adapted to $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$.*

Definition 2.1.3 (Martingale, supermartingale, submartingale)**.** *Consider a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and a sequence $(X_n)_{n\in\mathbb{N}}$ of integrable random variables.*

(i) *We say that $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-*martingale *if $\forall n \in \mathbb{N}$, $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$. Equivalently, $\forall (m,n) \in \mathbb{N}^2$, $m \leq n \Longrightarrow \mathbb{E}[X_n \mid \mathcal{F}_m] = X_m$.*

(ii) *We say that $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-*supermartingale *if $\forall n \in \mathbb{N}, \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n$. Equivalently, $\forall (m,n) \in \mathbb{N}^2$, $m \leq n \Longrightarrow \mathbb{E}[X_n \mid \mathcal{F}_m] \leq X_m$.*

(iii) *We say that $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-*submartingale *if $\forall n \in \mathbb{N}$, $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n$. Equivalently, $\forall (m,n) \in \mathbb{N}^2$, $m \leq n \Longrightarrow \mathbb{E}[X_n \mid \mathcal{F}_m] \geq X_m$.*

Remark 2.1.4. *If $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale), then the sequence $(\mathbb{E}[X_n])_{n\in\mathbb{N}}$ is constant (resp. nonincreasing, nondecreasing).*

Proposition 2.1.5. *If $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale and $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function s.t. $\varphi(X_n)$ is integrable for all $n \in \mathbb{N}$, then $(\varphi(X_n))_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-submartingale.*

Proof. Use Jensen's Inequality (Proposition 1.3.2). □

Remark 2.1.6. *If $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale), then it is also a $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale). Therefore, when we say that a sequence $(X_n)_{n\in\mathbb{N}}$ is a martingale (resp. supermartingale, submartingale) without mentioning a filtration, we mean that it is a $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale).*

Example 2.1.7.

(i) Random walks. *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables s.t. $\xi_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Define $S_n = \sum_{i=1}^n \xi_i$ for $n \in \mathbb{N}$. Then $(S_n)_{n\in\mathbb{N}}$ is a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-martingale if $\mathbb{E}[\xi_1] = 0$, a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-supermartingale if $\mathbb{E}[\xi_1] \leq 0$, and a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-submartingale if $\mathbb{E}[\xi_1] \geq 0$.*

(ii) Random products. *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. nonnegative random variables s.t. $\xi_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Define $P_n = \prod_{i=1}^n \xi_i$ for $n \in \mathbb{N}$. Then $(P_n)_{n\in\mathbb{N}}$ is a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-martingale if $\mathbb{E}[\xi_1] = 1$, a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-supermartingale if $\mathbb{E}[\xi_1] \leq 1$, and a $\left(\mathcal{F}_n^\xi\right)_{n\in\mathbb{N}}$-submartingale if $\mathbb{E}[\xi_1] \geq 1$.*

(iii) Closed martingales. *Fix a random variable $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Then $(\mathbb{E}[Z \mid \mathcal{F}_n])_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale.*

(iv) Martingales in L^2. *Consider a sequence $(X_n)_{n\in\mathbb{N}} \in L^2(\Omega, \mathcal{F}, \mathbb{P})^{\mathbb{N}}$ that is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale. Then the increments $(X_{n+1} - X_n)_{n\in\mathbb{N}}$ are orthogonal to each other in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.*

2.2 Building new martingales from old ones

Definition 2.2.1 (Previsible process). *Given a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, we say that a random process $(C_n)_{n\in\mathbb{N}^*}$ is* previsible *if C_n is $\mathcal{F}_{n-1}$-measurable for all $n \in \mathbb{N}^*$.*

Proposition 2.2.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale) and let $(C_n)_{n\in\mathbb{N}^*}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-previsible process that is almost surely bounded. Then the process $C \cdot X$ defined by:*

$$\forall n \in \mathbb{N},\ (C \cdot X)_n = \sum_{k=1}^{n} C_k \left(X_k - X_{k-1}\right),$$

is also a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale).

Remark 2.2.3. *One can interpret Proposition 2.2.2 in terms of gambling games: $X_n - X_{n-1}$ represents the outcomes of a game, C_n represents the bet placed on the nth-outcome, so that $(C \cdot X)_n$ represents the fortune of the gambler after n steps. Hence, Proposition 2.2.2 means that one cannot turn an unfair game (a supermartingale) into a fair one (a martingale or submartingale).*

2.3 Stopping times and stopping theorems

Definition 2.3.1 (Stopping time). *Given a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, a* stopping time *is a random variable $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ s.t. one of the three following equivalent assertions is verified:*

(i) $\forall n \in \mathbb{N},\ (\tau \leq n) \in \mathcal{F}_n$.

(ii) $\forall n \in \mathbb{N},\ (\tau > n) \in \mathcal{F}_n$.

(iii) $\forall n \in \mathbb{N},\ (\tau = n) \in \mathcal{F}_n$.

Remark 2.3.2. *Intuitively, a stopping time is a random time at which a decision can be taken given the information available at that time.*

Example 2.3.3.

(i) *Constant random variables are stopping times.*

(ii) *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with law* $\mathrm{Be}\left(\frac{1}{2}\right)$. *Then the random variable $\tau = \inf\{n \in \mathbb{N},\ \xi_n = 1\}$ is a $\left(\mathcal{F}_n^{\xi}\right)_{n\in\mathbb{N}}$-stopping time.*

(iii) *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted sequence of random variables. If A is a measurable subset of Ω, then $T_A = \inf\{n \in \mathbb{N},\ X_n \in A\}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time. Moreover, if $A \in \mathcal{F}_n$ for some $n \in \mathbb{N}$, then $\tau = n\mathbb{1}_A + (n+1)\mathbb{1}_{(\complement A)}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time.*

Proposition 2.3.4. *Let σ, τ, $(\tau_k)_{k\in\mathbb{N}}$ be $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times. Then $(\sigma + \tau)$, $\inf_{k\in\mathbb{N}} \tau_k$, $\sup_{k\in\mathbb{N}} \tau_k$, $\liminf_{k\to+\infty} \tau_k$ and $\limsup_{k\to+\infty} \tau_k$ are $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times.*

Definition 2.3.5 (Events measurable before τ). *Given a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and a stopping time τ, the σ-algebra of* events measurable before τ *is defined by:*

$$\mathcal{F}_\tau = \{A \in \mathcal{F},\ \forall n \in \mathbb{N},\ A \cap (\tau \leq n) \in \mathcal{F}_n\}.$$

Proposition 2.3.6. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted process and consider τ a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time. Then the random variable $X_\tau \mathbb{1}_{(\tau<+\infty)}$ is $\mathcal{F}_\tau$-measurable.*

Proposition 2.3.7. *Let σ and τ be two $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times s.t. $\sigma \leq \tau$. Then $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.*

Theorem 2.3.8 (Doob's Stopping Theorem, first version)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale) and σ, τ be two almost surely bounded $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times s.t. $\sigma \le \tau$, then:*

$$\mathbb{E}\left[X_\tau \mid \mathcal{F}_\sigma\right] = X_\sigma \quad \textit{(resp. } \le, \ge\textit{).}$$

In particular, if τ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time that is bounded, then:

$$\mathbb{E}\left[X_\tau\right] = \mathbb{E}\left[X_0\right] \quad \textit{(resp. } \le, \ge\textit{).}$$

Proof. Let $N \in \mathbb{N}$ be s.t. $\sigma \le N$ almost surely. We shall compute $\mathbb{E}\left[X_N \mid \mathcal{F}_\sigma\right]$. For $A \in \mathcal{F}_\sigma$, we have:

$$\begin{aligned}\mathbb{E}\left[X_N \mathbb{1}_A\right] &= \mathbb{E}\left[X_N \sum_{n=0}^{N} \mathbb{1}_{A\cap(\sigma=n)}\right] = \sum_{n=0}^{N} \mathbb{E}\left[X_N \mathbb{1}_{A\cap(\sigma=n)}\right] = \sum_{n=0}^{N} \mathbb{E}\left[\mathbb{E}\left[X_N \mid \mathcal{F}_n\right] \mathbb{1}_{A\cap(\sigma=n)}\right] \\ &= \sum_{n=0}^{N} \mathbb{E}\left[X_n \mathbb{1}_{A\cap(\sigma=n)}\right] = \sum_{n=0}^{N} \mathbb{E}\left[X_\sigma \mathbb{1}_{A\cap(\sigma=n)}\right] = \mathbb{E}\left[X_\sigma \mathbb{1}_A\right].\end{aligned}$$

Since X_σ is $\mathcal{F}_\sigma$-measurable, we get $\mathbb{E}\left[X_N \mid \mathcal{F}_\sigma\right] = X_\sigma$. Likewise, $\mathbb{E}\left[X_N \mid \mathcal{F}_\tau\right] = X_\tau$. Therefore:

$$\mathbb{E}\left[X_\tau \mid \mathcal{F}_\sigma\right] = \mathbb{E}\left[\mathbb{E}\left[X_N \mid \mathcal{F}_\tau\right] \mid \mathcal{F}_\sigma\right] = \mathbb{E}\left[X_N \mid \mathcal{F}_\sigma\right] = X_\sigma.$$

The proof is similar for supermartingales and submartingales, using the fact that if X' is a $\mathcal{G}$-measurable random variable s.t. $\forall A \in \mathcal{G}$, $\mathbb{E}\left[X \mathbb{1}_A\right] \ge \mathbb{E}\left[X' \mathbb{1}_A\right]$, then $\mathbb{E}\left[X \mid \mathcal{G}\right] \ge X'$. ☐

Proposition 2.3.9. *If $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale) and τ is (any) $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time, then $X^\tau = (X_{n\wedge\tau})_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale), called the* martingale stopped at τ *(resp. supermartingale, submartingale stopped at τ).*

Proof. For $n \in \mathbb{N}^*$, set $C_n = \mathbb{1}_{(n\le\tau)}$. Hence, $(C_n)_{n\in\mathbb{N}^*}$ is a bounded previsible process. Therefore, according to Proposition 2.2.2, $X^\tau = C \cdot X + X_0$ is a martingale (resp. supermartingale, submartingale). ☐

Proposition 2.3.10. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale (resp. supermartingale, submartingale) and τ be an almost surely finite $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time. Assume that $\mathbb{E}[\tau] < +\infty$ and that $\exists M \in \mathbb{R}_+$, $\forall n \in \mathbb{N}$, $|X_{n+1} - X_n| \le M$ almost surely. Then $\mathbb{E}\left[X_\tau\right] = \mathbb{E}\left[X_0\right]$ (resp. $\le, \ge$).*

Proof. Apply Theorem 2.3.8 to $(X_{n\wedge\tau})_{n\in\mathbb{N}}$, which is a martingale according to Proposition 2.3.9, and use the Dominated Convergence Theorem. ☐

2.4 Almost sure convergence for (super)martingales

Definition 2.4.1 (Upcrossings)**.** *Let $(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Consider two real numbers $a < b$. Define two sequences $(S_k(x))_{k\in\mathbb{N}^*}$ and $(T_k(x))_{k\in\mathbb{N}^*}$ by induction:*

$$S_1(x) = \inf\left\{n \ge 0,\ x_n < a\right\} \in \overline{\mathbb{N}} \quad \textit{and} \quad T_1(x) = \inf\left\{n \ge S_1(x),\ x_n \ge b\right\} \in \overline{\mathbb{N}},$$

and for all $k \in \mathbb{N}^*$:

$$S_{k+1}(x) = \inf\left\{n \ge T_k(x),\ x_n < a\right\} \quad \textit{and} \quad T_{k+1}(x) = \inf\left\{n \ge S_{k+1}(x),\ x_n \ge b\right\}.$$

The number of upcrossings *before time n is defined as $\mathcal{N}_n(x,a,b) = \sup\left\{k \ge 1,\ T_k(x) \le n\right\}$; the total number of upcrossings is therefore $\mathcal{N}_\infty(x,a,b) = \sup_{n\in\mathbb{N}} \mathcal{N}_n(x,a,b)$.*

Proposition 2.4.2. *A sequence $x \in \mathbb{R}^{\mathbb{N}}$ converges in $\overline{\mathbb{R}}$* iff

$$\forall (a,b) \in \mathbb{Q}^2,\ a < b \Longrightarrow \mathcal{N}_\infty(x,a,b) < +\infty.$$

Proof. Use the fact that x converges in $\overline{\mathbb{R}}$ iff $\liminf_{n\to+\infty} x_n = \limsup_{n\to+\infty} x_n$. $\square$

Lemma 2.4.3 (Doob's Upcrossings Lemma)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-supermartingale. Then, for every real numbers $a < b$:*

$$\mathbb{E}\left[\mathcal{N}_n(X,a,b)\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_n - a)^-\right].$$

Proof. Fix two real numbers $a < b$. Note that, for $k \in \mathbb{N}^*$, $S_k(X)$ and $T_k(X)$ are $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times. For $n \in \mathbb{N}^*$, define:

$$C_n = \sum_{k\in\mathbb{N}^*} \mathbb{1}_{(S_k(X)<n\leq T_k(X))}.$$

Hence $(C_n)^*_{n\in\mathbb{N}}$ is a bounded $(\mathcal{F}_n)_{n\in\mathbb{N}}$-previsible process. According to Proposition 2.2.2, $C \cdot X$ is a supermartingale. But:

$$\begin{aligned}
\forall n \in \mathbb{N},\ (C\cdot X)_n &= \sum_{k=1}^{\mathcal{N}_n(X,a,b)} \left(\underbrace{X_{T_k(X)}}_{\geq b} - \underbrace{X_{S_k(X)}}_{<a} \right) + \left(X_n - \underbrace{X_{S_{\mathcal{N}_n(X,a,b)+1}(X)}}_{<a} \right) \mathbb{1}_{\left(S_{\mathcal{N}_n(X,a,b)+1}(X)\leq n\right)} \\
&\geq (b-a)\mathcal{N}_n(X,a,b) + (X_n - a)\, \mathbb{1}_{\left(S_{\mathcal{N}_n(X,a,b)+1}(X)\leq n\right)} \\
&\geq (b-a)\mathcal{N}_n(X,a,b) + (X_n - a)\, \mathbb{1}_{(X_n<a)} \\
&= (b-a)\mathcal{N}_n(X,a,b) - (X_n - a)^- .
\end{aligned}$$

As $(C\cdot X)$ is a supermartingale, we have $\forall n \in \mathbb{N}$, $\mathbb{E}\left[(C\cdot X)_n\right] \leq \mathbb{E}\left[(C\cdot X)_0\right] = 0$, which gives the result. $\square$

Theorem 2.4.4. *Let $(X_n)_{n\in\mathbb{N}}$ be a supermartingale that is bounded in L^1 (i.e.* $\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < +\infty$, *or equivalently* $\sup_{n\in\mathbb{N}} \mathbb{E}\left[X_n^-\right] < +\infty$*). Then there exists $X_\infty \in L^1(\Omega,\mathcal{F},\mathbb{P})$ s.t. $X_n \xrightarrow[n\to+\infty]{} X_\infty$ a.s.*

Proof. Let $a < b$ be two rational numbers. According to Lemma 2.4.3, we have:

$$\mathbb{E}\left[\mathcal{N}_n(X,a,b)\right] \leq \frac{1}{b-a}\left(\mathbb{E}\left[|X_n|\right] + |a|\right) \leq \frac{M+|a|}{b-a},$$

where $M = \sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < +\infty$. By monotone convergence, $\mathbb{E}\left[\mathcal{N}_\infty(X,a,b)\right] \leq \frac{M+|a|}{b-a} < +\infty$, so $\mathcal{N}_\infty(X,a,b) < +\infty$ a.s. This shows that:

$$\mathbb{P}\left(\bigcup_{\substack{(a,b)\in\mathbb{Q}^2 \\ a<b}} \left(\mathcal{N}_\infty(X,a,b) = +\infty\right) \right) \leq \sum_{\substack{(a,b)\in\mathbb{Q}^2 \\ a<b}} \mathbb{P}\left(\mathcal{N}_\infty(X,a,b) = +\infty\right) = 0.$$

Using Proposition 2.4.2, $(X_n)_{n\in\mathbb{N}}$ converges a.s. in $\overline{\mathbb{R}}$ to some limit X_∞. But:

$$\mathbb{E}\left[|X_\infty|\right] = \mathbb{E}\left[\liminf_{n\to+\infty} |X_n|\right] \leq \liminf_{n\to+\infty} \mathbb{E}\left[|X_n|\right] \leq \sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < +\infty.$$

Therefore, X_∞ is a.s. finite and $X_\infty \in L^1(\Omega,\mathcal{F},\mathbb{P})$. $\square$

Corollary 2.4.5. *If $(X_n)_{n\in\mathbb{N}}$ is a nonnegative supermartingale, then there exists $X_\infty \in L^1(\Omega,\mathcal{F},\mathbb{P})$ s.t. $X_n \xrightarrow[n\to+\infty]{} X_\infty$ a.s.*

Proof. If $(X_n)_{n\in\mathbb{N}}$ is a nonnegative supermartingale, then $\forall n \in \mathbb{N}$, $\mathbb{E}\left[|X_n|\right] = \mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X_0\right]$, so Theorem 2.4.4 applies. $\square$

Corollary 2.4.6. *Let $(X_n)_{n\in\mathbb{N}}$ be a submartingale that is bounded in L^1 (i.e.* $\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < +\infty$, *or equivalently* $\sup_{n\in\mathbb{N}} \mathbb{E}\left[X_n^+\right] < +\infty$*). Then there exists $X_\infty \in L^1(\Omega,\mathcal{F},\mathbb{P})$ s.t. $X_n \xrightarrow[n\to+\infty]{} X_\infty$ a.s.*

Proof. Apply Theorem 2.4.4 to $(-X_n)_{n\in\mathbb{N}}$. $\square$

Corollary 2.4.7. *Let $(X_n)_{n\in\mathbb{N}}$ be a martingale with bounded increments: $\sup_{n\in\mathbb{N}} \|X_{n+1}-X_n\|_\infty < +\infty$. Set:*

$$A = \Big((X_n)_{n\in\mathbb{N}} \text{ converges in } \mathbb{R}\Big) \quad \text{and} \quad B = \left(\liminf_{n\to+\infty} X_n = -\infty \text{ and } \limsup_{n\to+\infty} X_n = +\infty\right).$$

Then $\mathbb{P}(A\cup B)=1$.

Proof. Introduce $\tau_k = \inf\{n\in\mathbb{N},\ X_n \geq k\}$ and $\tau_{-k} = \inf\{n\in\mathbb{N},\ X_n < -k\}$ for $k\in\mathbb{N}$; τ_k and τ_{-k} are stopping times. According to Proposition 2.3.9, X^{τ_k} is a martingale that is bounded by $k+\sup_{n\in\mathbb{N}}\|X_{n+1}-X_n\|_\infty$, so it converges a.s. according to Theorem 2.4.4. Do the same thing for $X^{\tau_{-k}}$ and obtain the result. $\square$

2.5 Doob's L^p-inequality and convergence in L^p

Lemma 2.5.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a submartingale. Set $\overline{X}_n = \max_{0\leq k\leq n} X_k$ for $n\in\mathbb{N}$. Then:*

$$\forall x>0,\ \mathbb{P}\left(\overline{X}_n > x\right) \leq \frac{1}{x}\mathbb{E}\left[X_n \mathbb{1}_{(\overline{X}_n > x)}\right].$$

Proof. Let $\tau = \inf\{n\in\mathbb{N},\ X_n > x\}$. τ is a stopping time and $\left(\overline{X}_n > x\right) = (\tau\leq n)$. According to Proposition 2.3.9, X^τ is a submartingale, so $\mathbb{E}[X_{n\wedge\tau}] \leq \mathbb{E}[X_n]$ for all $n\in\mathbb{N}$. But:

$$\mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}\left[X_n\mathbb{1}_{(\tau>n)} + X_\tau \mathbb{1}_{(\tau\leq n)}\right] \geq \mathbb{E}[X_n] - \mathbb{E}\left[X_n \mathbb{1}_{(\tau\leq n)}\right] + x\mathbb{P}(\tau\leq n).$$

This gives the desired result. $\square$

Theorem 2.5.2 (Doob's L^p-inequality)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a martingale. For $n\in\mathbb{N}$, set $X_n^* = \max_{0\leq k\leq n}|X_k|$. Then:*

$$\forall p>1,\ \forall n\in\mathbb{N},\ \|X_n^*\|_p \leq \frac{p}{p-1}\|X_n\|_p.$$

Proof. Apply Lemma 2.5.1 to the submartingale $(Y_n)_{n\in\mathbb{N}} = (|X_n|)_{n\in\mathbb{N}}$. Note that $\overline{Y}_n = X_n^*$ for all $n\in\mathbb{N}$. Hence, for $n\in\mathbb{N}$:

$$\begin{aligned}
\mathbb{E}[(X_n^*)^p] &= \int_0^\infty px^{p-1}\mathbb{P}(X_n^* > x)\,\mathrm{d}x \\
&\leq \int_0^\infty px^{p-1}\frac{1}{x}\mathbb{E}\left[|X_n|\,\mathbb{1}_{(X_n^*>x)}\right]\mathrm{d}x \\
&= p\mathbb{E}\left[|X_n|\,\frac{1}{p-1}(X_n^*)^{p-1}\right] \\
&\leq \frac{p}{p-1}\mathbb{E}[|X_n|^p]^{1/p}\,\mathbb{E}\left[(X_n^*)^{q(p-1)}\right]^{1/q},
\end{aligned}$$

where $\frac{1}{p}+\frac{1}{q}=1$. Since $(p-1)q=p$, we obtain $\|X_n^*\|_p = \mathbb{E}[(X_n^*)^p]^{1-1/q} \leq \frac{p}{p-1}\mathbb{E}[|X_n|^p]^{1/p} = \frac{p}{p-1}\|X_n\|_p$ $\square$

Corollary 2.5.3. *Fix $p>1$. Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale that is bounded in L^p (i.e. $\sup_{n\in\mathbb{N}}\|X_n\|_p < +\infty$). Then $(X_n)_{n\in\mathbb{N}}$ converges a.s. and in L^p to a limit X_∞. Moreover, for every bounded stopping time τ, we have $X_\tau = \mathbb{E}[X_\infty \mid \mathcal{F}_\tau]$.*

Proof. Our assumptions imply that $(X_n)_{n\in\mathbb{N}}$ is bounded in L^1, so according to Theorem 2.4.4, it converges a.s. to some limit $X_\infty\in L^1(\Omega,\mathcal{F},\mathbb{P})$. Since $(X_n)_{n\in\mathbb{N}}$ is bounded in L^p, Fatou's Lemma shows that $X_\infty \in L^p(\Omega,\mathcal{F},\mathbb{P})$. Now, Doob's L^p-inequality (Theorem 2.5.2) gives:

$$\mathbb{E}[(X_n^*)^p] \leq \left(\frac{p}{p-1}\right)^p \sup_{n\in\mathbb{N}}\mathbb{E}[|X_n|^p] < +\infty.$$

Moreover, $X_n^* \xrightarrow[n\to+\infty]{} X_\infty^* = \sup_{k\in\mathbb{N}} |X_k|$; by monotone convergence, we get $\mathbb{E}\left[(X_\infty^*)^p\right] < +\infty$. But $|X_n - X_\infty|^p \leq 2^p \left(|X_\infty|^p + |X_n|^p\right) \leq 2^{p+1} \left(X_\infty^*\right)^p \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. By dominated convergence, we obtain $X_n \xrightarrow[n\to+\infty]{} X_\infty$ in L^p. Moreover, as $\mathbb{E}[\cdot \mid \mathcal{G}]$ sends L^p to L^p, we obtain that $\mathbb{E}[X_\infty \mid \mathcal{F}_n] = X_n$, and similarly for stopping times τ. □

Corollary 2.5.4. *If $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale that is bounded in L^p, then $(X_n)_{n\in\mathbb{N}}$ is* closed *in L^p, i.e. there exists $Z \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ s.t. $\forall n \in \mathbb{N}$, $X_n = \mathbb{E}[Z \mid \mathcal{F}_n]$.*

2.6 Martingales in L^2

Proposition 2.6.1 (Doob's decomposition)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-submartingale. Then there exists a unique $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale $(M_n)_{n\in\mathbb{N}}$ and a nondecreasing $(\mathcal{F}_n)_{n\in\mathbb{N}}$-previsible process $(A_n)_{n\in\mathbb{N}}$ with $A_0 = 0$ s.t.*

$$\forall n \in \mathbb{N},\ X_n = M_n + A_n.$$

Moreover, for all $n \in \mathbb{N}$, we have $A_{n+1} - A_n = \mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$.

Definition 2.6.2 (Angle bracket)**.** *If $(X_n)_{n\in\mathbb{N}} \in L^2(\Omega, \mathcal{F}, \mathbb{P})^{\mathbb{N}}$ is a martingale, then $(X_n^2)_{n\in\mathbb{N}}$ is a submartingale, and we define the* angle bracket *$(\langle X\rangle_n)_{n\in\mathbb{N}}$ as the nondecreasing previsible process arising in Doob's decomposition of $(X_n^2)_{n\in\mathbb{N}}$ (c.f. Proposition 2.6.1). Moreover, we set $\langle X\rangle_\infty = \lim_{n\to+\infty} \langle X\rangle_n \in [0, +\infty]$.*

Proposition 2.6.3. *Let $(X_n)_{n\in\mathbb{N}} \in L^2(\Omega, \mathcal{F}, \mathbb{P})^{\mathbb{N}}$ be a martingale. Then:*

(i) $(\langle X\rangle_\infty < +\infty) \overset{a.s.}{\Longrightarrow} \Big((X_n)_{n\in\mathbb{N}}$ *converges*$\Big)$.

(ii) *If in addition* $\sup_{n\in\mathbb{N}} \|X_{n+1} - X_n\|_\infty < +\infty$, *then* $(\langle X\rangle_\infty < +\infty) \overset{a.s.}{\Longleftrightarrow} \Big((X_n)_{n\in\mathbb{N}}$ *converges*$\Big)$.

Proof. (i) Fix $K > 0$ and define:

$$T_K = \inf\left\{n \in \mathbb{N},\ \langle X\rangle_{n+1} > K\right\}.$$

Then T_K is a stopping time and the stopped process $\left(\langle X\rangle_{n\wedge T_k}\right)_{n\in\mathbb{N}}$ is previsible. Therefore, Doob's decomposition for the submartingale $\left(X_{n\wedge T_K}^2\right)_{n\in\mathbb{N}}$ is given by:

$$\forall n \in \mathbb{N},\ X_{n\wedge T_K}^2 = M_{n\wedge T_K} + \langle X\rangle_{n\wedge T_K}.$$

Hence, $(X_{n\wedge T_K})_{n\in\mathbb{N}}$ is bounded in L^2 by $\mathbb{E}[M_0] + K$. According to Corollary 2.5.3, $(X_{n\wedge T_K})_{n\in\mathbb{N}}$ converges a.s. Thus, on the set $(\langle X\rangle_\infty \leq K) = (T_K = +\infty)$, one has that $(X_n)_{n\in\mathbb{N}}$ converges a.s. Taking the union over all $K \in \mathbb{N}$ gives the desired result. (ii) Since $\forall n \in \mathbb{N}$, $X_n^2 = M_n + \langle X\rangle_n$, we have $\left(M_n \xrightarrow[n\to+\infty]{} -\infty\right)$ on the event $(\langle X\rangle_\infty = +\infty) \cap \Big((X_n)_{n\in\mathbb{N}}$ converges$\Big)$. As in Corollary 2.4.7, we prove that:

$$\mathbb{P}\Big((\langle X\rangle_\infty = +\infty) \cap \Big((X_n)_{n\in\mathbb{N}} \text{ converges}\Big)\Big) = 0.$$

□

Corollary 2.6.4 (Conditioned Borel-Cantelli Lemma)**.** *Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration and consider a sequence $(A_n)_{n\in\mathbb{N}^*}$ of events s.t. A_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}^*$. Then:*

$$\left(\sum_{k\in\mathbb{N}^*} \mathbb{P}(A_k \mid \mathcal{F}_{k-1}) < +\infty\right) \overset{a.s.}{\Longrightarrow} \left(\sum_{k\in\mathbb{N}^*} \mathbb{1}_{A_k} < +\infty\right).$$

Proof. Set $Y_n = \sum_{k=1}^{n} \mathbb{1}_{A_k}$, $Z_n = \sum_{k=1}^{n} \mathbb{P}(A_k \mid \mathcal{F}_{k-1})$ and $X_n = Y_n - Z_n$. Then $(X_n)_{n\in\mathbb{N}}$ is a martingale and Doob's decomposition for the submartingale $(Y_n)_{n\in\mathbb{N}}$ is given by $\forall n \in \mathbb{N}$, $Y_n = X_n + Z_n$. Moreover, for $n \in \mathbb{N}$, we have:

$$\begin{aligned}\langle X\rangle_{n+1} - \langle X\rangle_n &= \mathbb{E}\left[X_{n+1}^2 - X_n^2 \mid \mathcal{F}_n\right] = \mathbb{E}\left[(X_{n+1} - X_n)^2 \mid \mathcal{F}_n\right] \\ &= \mathbb{P}(A_{n+1} \mid \mathcal{F}_n)\left(1 - \mathbb{P}(A_{n+1} \mid \mathcal{F}_n)\right) \leq \mathbb{P}(A_{n+1} \mid \mathcal{F}_n).\end{aligned}$$

Therefore, for $n \in \mathbb{N}$, $\langle X\rangle_n \leq Z_n$. Thus, if $Z_\infty = \sum_{k\in\mathbb{N}^*} \mathbb{P}(A_k \mid \mathcal{F}_{k-1}) < +\infty$, then $\langle X\rangle_\infty < +\infty$, so $(X_n)_{n\in\mathbb{N}}$ converges according to Proposition 2.6.3, and so does $(Y_n)_{n\in\mathbb{N}}$. □

2.7 Uniform integrability

Proposition 2.7.1. *Let X be an integrable random variable. Then:*

(i) $\mathbb{E}\left[|X| \mathbb{1}_{(|X|>a)}\right] \xrightarrow[a\to+\infty]{} 0$.

(ii) $\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall A \in \mathcal{F},\ \mathbb{P}(A) \leq \delta \Longrightarrow \mathbb{E}[|X| \mathbb{1}_A] \leq \varepsilon$.

Definition 2.7.2 (Uniform integrability)**.** *A family $(X_i)_{i\in I}$ is said to be* uniformly integrable *if:*

$$\sup_{i\in I} \mathbb{E}\left[|X_i| \mathbb{1}_{(|X_i|>a)}\right] \xrightarrow[a\to+\infty]{} 0.$$

Example 2.7.3. *If X is an integrable random variable, then the family $\{X\}$ is uniformly integrable.*

Proposition 2.7.4. *A uniformly integrable family of random variables is bounded in L^1, but the converse is false.*

Proposition 2.7.5. *If the family $(X_i)_{i\in I}$ of random variables is dominated in the sense that there exists an integrable random variable Y s.t. $\forall i \in I$, $|X_i| \leq Y$, then $(X_i)_{i\in I}$ is uniformly integrable.*

Proposition 2.7.6. *Let $(X_i)_{i\in I}$ be a family of random variables. The following assertions are equivalent:*

(i) *$(X_i)_{i\in I}$ is uniformly integrable.*

(ii) *$(X_i)_{i\in I}$ is bounded in L^1 and:*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall A \in \mathcal{F},\ \mathbb{P}(A) \leq \delta \Longrightarrow \sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_A] \leq \varepsilon.$$

(iii) *There exists a nondecreasing function $G : \mathbb{R}_+ \to \mathbb{R}_+$ s.t.*

$$\frac{G(x)}{x} \xrightarrow[x\to+\infty]{} +\infty \quad \text{and} \quad \sup_{i\in I} \mathbb{E}\left[G(|X_i|)\right] < +\infty.$$

Corollary 2.7.7. *A family of random variables that is bounded in L^p for some $p > 1$ is uniformly integrable.*

Lemma 2.7.8. *If $(X_i)_{i\in I}$ and $(Y_j)_{j\in J}$ are uniformly integrable families of random variables, then $(X_i + Y_j)_{(i,j)\in I\times J}$ is uniformly integrable.*

Theorem 2.7.9. *Let $(X_n)_{n\in\mathbb{N}}$ and X be random variables. The following assertions are equivalent:*

(i) $X_n \xrightarrow[n\to+\infty]{L^1} X$.

(ii) $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ and *$(X_n)_{n\in\mathbb{N}}$ is uniformly integrable.*

Proof. (ii) ⇒ (i) Let us show that $\mathbb{E}\left[|X_n - X|\right] \xrightarrow[n\to+\infty]{} 0$, assuming that $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ and $(X_n)_{n\in\mathbb{N}}$ is uniformly integrable. Fix $\varepsilon > 0$. Note that:

$$\mathbb{E}\left[|X_n - X|\right] = \mathbb{E}\left[|X_n - X| \mathbb{1}_{(|X_n - X| \leq \varepsilon)}\right] + \mathbb{E}\left[|X_n - X| \mathbb{1}_{(|X_n - X| > \varepsilon)}\right] \leq \varepsilon + \mathbb{E}\left[|X_n - X| \mathbb{1}_{(|X_n - X| > \varepsilon)}\right].$$

By Lemma 2.7.8, $(X_n - X)_{n\in\mathbb{N}}$ is uniformly integrable, therefore there exists $\delta > 0$ s.t. $\mathbb{P}(A) \leq \delta \implies \forall n \in \mathbb{N},\ \mathbb{E}\left[|X_n - X| \mathbb{1}_A\right] \leq \varepsilon$. Now, since $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$, choose $N \in \mathbb{N}$ s.t. $\forall n \geq N,\ \mathbb{P}\left(|X_n - X| > \varepsilon\right) \leq \delta$. This yields $\forall n \geq N,\ \mathbb{E}\left[|X_n - X|\right] \leq 2\varepsilon$. (i) ⇒ (ii) If $X_n \xrightarrow[n\to+\infty]{L^1} X$, then $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ because of Markov's inequality. Now let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ s.t. $\forall n \geq N,\ \mathbb{E}\left[|X_n - X|\right] \leq \frac{\varepsilon}{2}$, then choose $\delta > 0$ s.t. for any event A with $\mathbb{P}(A) \leq \delta$, we have $\mathbb{E}\left[|X| \mathbb{1}_A\right] \leq \frac{\varepsilon}{2}$ and $\max_{0\leq n<N} \mathbb{E}\left[|X_n - X| \mathbb{1}_A\right] \leq \frac{\varepsilon}{2}$. Hence, if A is an event s.t. $\mathbb{P}(A) \leq \delta$, we have:

$$\mathbb{E}\left[|X_n| \mathbb{1}_A\right] \leq \mathbb{E}\left[|X| \mathbb{1}_A\right] + \mathbb{E}\left[|X_n - X| \mathbb{1}_A\right] \leq \varepsilon.$$

□

2.8 Martingales in L^1

Theorem 2.8.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale. Then $(X_n)_{n\in\mathbb{N}}$ converges in L^1* iff *$(X_n)_{n\in\mathbb{N}}$ is uniformly integrable. In this case, $(X_n)_{n\in\mathbb{N}}$ is* closed *in L^1 by its limit X_∞:*

$$\forall n \in \mathbb{N},\ X_n = \mathbb{E}\left[X_\infty \mid \mathcal{F}_n\right].$$

Proof. Use Proposition 2.7.4, Theorem 2.4.4 and Theorem 2.7.9 (as well as the fact that if a sequence of random variables converges almost surely, then it converges in probability). □

Corollary 2.8.2 (Doob's Stopping Theorem, uniformly integrable version)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a uniformly integrable martingale and consider two stopping times σ, τ s.t. $\sigma \leq \tau$. Then:*

$$\mathbb{E}\left[X_\tau \mid \mathcal{F}_\sigma\right] = X_\sigma.$$

In particular, if τ is a stopping time, then:

$$\mathbb{E}\left[X_\tau\right] = \mathbb{E}\left[X_0\right].$$

Proof. Note that, according to Theorem 2.8.1 and Theorem 2.4.4, $(X_n)_{n\in\mathbb{N}}$ converges almost surely and in L^1 towards a random variable $X_\infty \in L^1$. Now let τ be a stopping time. We shall show that $\mathbb{E}\left[X_\infty \mid \mathcal{F}_\tau\right] = X_\tau$. For $A \in \mathcal{F}_\tau$, we have:

$$\begin{aligned}
\mathbb{E}\left[X_\infty \mathbb{1}_A\right] &= \sum_{n\in\mathbb{N}} \mathbb{E}\left[X_\infty \mathbb{1}_{A\cap(\tau=n)}\right] + \mathbb{E}\left[X_\infty \mathbb{1}_{A\cap(\tau=\infty)}\right] \\
&= \sum_{n\in\mathbb{N}} \mathbb{E}\left[\mathbb{E}\left[X_\infty \mid \mathcal{F}_n\right] \mathbb{1}_{A\cap(\tau=n)}\right] + \mathbb{E}\left[X_\tau \mathbb{1}_{A\cap(\tau=\infty)}\right] \\
&= \sum_{n\in\mathbb{N}} \mathbb{E}\left[X_n \mathbb{1}_{A\cap(\tau=n)}\right] + \mathbb{E}\left[X_\tau \mathbb{1}_{A\cap(\tau=\infty)}\right] = \mathbb{E}\left[X_\tau \mathbb{1}_A\right].
\end{aligned}$$

Hence $\mathbb{E}\left[X_\infty \mid \mathcal{F}_\tau\right] = X_\tau$. We easily obtain the desired result using the fact that $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$. □

3 Applications of martingales

3.1 Lévy's Convergence Theorem

Lemma 3.1.1. *If $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then the family $\left\{\mathbb{E}\left[Z \mid \mathcal{G}\right],\ \mathcal{G}\ \text{sub-}\sigma\text{-algebra of}\ \mathcal{F}\right\}$ is uniformly integrable.*

Proof. For $a > 0$, we have:

$$\mathbb{E}\left[|\mathbb{E}[Z \mid \mathcal{G}]| \mathbb{1}_{(|\mathbb{E}[Z|\mathcal{G}]|>a)}\right] \leq \mathbb{E}\left[\mathbb{E}[|Z| \mid \mathcal{G}] \mathbb{1}_{(\mathbb{E}[|Z||\mathcal{G}]>a)}\right] = \mathbb{E}\left[|Z| \mathbb{1}_{(\mathbb{E}[|Z||\mathcal{G}]>a)}\right].$$

Now by Markov's Inequality, $\mathbb{P}(\mathbb{E}[|Z| \mid \mathcal{G}] > a) \leq \frac{1}{a}\mathbb{E}[|Z|] \xrightarrow[a\to+\infty]{} 0$. We conclude by uniform integrability of $\{Z\}$. □

Theorem 3.1.2 (Lévy's Convergence Theorem)**.** *Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration and $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then $(\mathbb{E}[Z \mid \mathcal{F}_n])_{n\in\mathbb{N}}$ converges a.s. and in L^1 to $\mathbb{E}[Z \mid \mathcal{F}_\infty]$.*

Proof. Define $X_n = \mathbb{E}[Z \mid \mathcal{F}_n]$ for $n \in \mathbb{N}$. By Lemma 3.1.1, the martingale $(X_n)_{n\in\mathbb{N}}$ is uniformly integrable. By Theorem 2.8.1, $(X_n)_{n\in\mathbb{N}}$ converges a.s. and in L^1 to some limit $X_\infty \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. We easily check that $\forall A \in \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$, $\mathbb{E}[X_\infty \mathbb{1}_A] = \mathbb{E}[Z\mathbb{1}_A]$. Since $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ is stable by finite intersections and generates $\mathcal{F}_\infty$, this is actually true for all $A \in \mathcal{F}_\infty$, which shows that $X_\infty = \mathbb{E}[Z \mid \mathcal{F}_\infty]$. □

Corollary 3.1.3. *Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration.*

(i) *If $\mathcal{M}_{UI}$ denotes the set of uniformly integrable $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingales, then the map:*

$$Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \longmapsto (\mathbb{E}[Z \mid \mathcal{F}_n])_{n\in\mathbb{N}} \in \mathcal{M}_{UI}$$

is a bijection.

(ii) *Let $p \in]1, +\infty[$. If $\mathcal{M}_{L^p}$ denotes the set of $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingales that are bounded in L^p, then the map:*

$$Z \in L^p(\Omega, \mathcal{F}, \mathbb{P}) \longmapsto (\mathbb{E}[Z \mid \mathcal{F}_n])_{n\in\mathbb{N}} \in \mathcal{M}_{L^p}$$

is a bijection.

In particular, a martingale is closed iff *it is uniformly integrable.*

Corollary 3.1.4 (Kolmogorov's Zero-One Law)**.** *Let $(\mathcal{H}_n)_{n\in\mathbb{N}}$ be a family of independent sub-σ-algebras of $\mathcal{F}$. Let $\mathcal{G}_n = \bigvee_{k\geq n} \mathcal{H}_k$ for $n \in \mathbb{N}$ and $\mathcal{G}_\infty = \bigcap_{n\in\mathbb{N}} \mathcal{G}_n$. Then $\mathcal{G}_\infty$ is trivial:*

$$\forall A \in \mathcal{G}_\infty,\ \mathbb{P}(A) \in \{0, 1\}.$$

Proof. Let $\mathcal{F}_n = \bigvee_{0\leq k\leq n} \mathcal{H}_k$ for $n \in \mathbb{N}$. Then $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration and $\mathcal{F}_\infty = \mathcal{G}_0 \supseteq \mathcal{G}_\infty$. Hence, if $A \in \mathcal{G}_\infty$, we have $\mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_n] \xrightarrow[n\to+\infty]{a.s.} \mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_\infty] = \mathbb{1}_A$ by Lévy's Convergence Theorem (Theorem 3.1.2). But note that $A \in \mathcal{G}_{n+1}$, so A is independent from $\mathcal{F}_n$; therefore $\mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_n] = \mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$. Therefore, $\mathbb{P}(A) = \mathbb{1}_A$ a.s. □

Theorem 3.1.5 (Hewitt–Savage Zero-One Law)**.** *Let $(\xi_n)_{n\in\mathbb{N}}$ be independent random variables. Consider a map $F : \mathbb{R}^\mathbb{N} \to \mathbb{R}$ s.t.*

$$\forall \sigma \in \mathfrak{S}_n,\ F\left((\xi_n)_{n\in\mathbb{N}}\right) = F\left(\left(\xi_{\sigma(n)}\right)_{n\in\mathbb{N}}\right).$$

Then $F\left((\xi_n)_{n\in\mathbb{N}}\right)$ is a.s. constant.

3.2 Backwards martingales

Definition 3.2.1 (Backwards martingale)**.** *Consider a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}_-}$ (i.e. s.t. $\mathcal{F}_{n-1} \subseteq \mathcal{F}_n$ for all $n \in \mathbb{Z}_-$; we then write $\mathcal{F}_{-\infty} = \bigcap_{n\in\mathbb{Z}_-} \mathcal{F}_n$) and a sequence $(X_n)_{n\in\mathbb{Z}_-}$ of integrable random variables. We say that $(X_n)_{n\in\mathbb{Z}_-}$ is a $(\mathcal{F}_n)_{n\in\mathbb{Z}_-}$-*backwards martingale *if $\forall n \in \mathbb{Z}_-$, $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = X_{n-1}$.*

Theorem 3.2.2. *Let $(X_n)_{n\in\mathbb{Z}_-}$ be a $(\mathcal{F}_n)_{n\in\mathbb{Z}_-}$-backwards martingale. Then $(X_n)_{n\in\mathbb{Z}_-}$ converges a.s. and in L^1 to $X_{-\infty} = \mathbb{E}[X_0 \mid \mathcal{F}_{-\infty}]$.*

Proof. Note that $(X_n)_{n\in\mathbb{Z}_-} = (\mathbb{E}[X_0 \mid \mathcal{F}_n])_{n\in\mathbb{Z}_-}$ is uniformly integrable by Lemma 3.1.1. Adapt the proof of Theorem 2.4.4 to prove that $(X_n)_{n\in\mathbb{Z}_-}$ converges a.s. The convergence is also L^1 since $(X_n)_{n\in\mathbb{Z}_-}$ is uniformly integrable. To show that $X_{-\infty} = \mathbb{E}[X_0 \mid \mathcal{F}_{-\infty}]$, adapt the proof of Lévy's Convergence Theorem (Theorem 3.1.2). $\square$

Corollary 3.2.3 (Strong Law of Large Numbers)**.** *Consider i.i.d. random variables $(\xi_n)_{n\in\mathbb{N}^*}$ in $L^1(\Omega,\mathcal{F},\mathbb{P})$. Set $S_n = \sum_{k=1}^n \xi_k$ for $n\in\mathbb{N}$. Then $\left(\frac{1}{n}S_n\right)_{n\in\mathbb{N}}$ converges a.s. and in L^1 to $\mathbb{E}[\xi_1]$.*

Proof. Let $\mathcal{G}_n = \sigma(S_n, S_{n+1},\dots) = \sigma(S_n, \xi_{n+1}, \xi_{n+2},\dots)$ for $n\in\mathbb{N}$. Hence, $(\mathcal{G}_{-n})_{n\in\mathbb{Z}_-}$ is a backwards filtration. Let us show that $\left(\frac{1}{-n}S_{-n}\right)_{n\in\mathbb{Z}_-}$ is a backwards martingale. It is clearly integrable and adapted. Moreover, for $n\in\mathbb{N}$:

$$\mathbb{E}\left[\frac{1}{n}S_n \mid \mathcal{G}_{n+1}\right] = \frac{1}{n}\mathbb{E}[S_{n+1} - \xi_{n+1} \mid \mathcal{G}_{n+1}] = \frac{1}{n}\left(S_{n+1} - \mathbb{E}[\xi_{n+1} \mid S_{n+1}]\right),$$

using the fact that S_{n+1} is $\mathcal{G}_{n+1}$-measure and Proposition 1.4.4. Now, by symmetry, we have $\mathbb{E}[\xi_{n+1} \mid S_{n+1}] = \mathbb{E}[\xi_k \mid S_{n+1}]$ for all $k\in\{1,\dots,n+1\}$, so $S_{n+1} = \sum_{k=1}^{n+1}\mathbb{E}[\xi_k \mid S_{n+1}] = (n+1)\mathbb{E}[\xi_{n+1} \mid S_{n+1}]$, which gives:

$$\mathbb{E}\left[\frac{1}{n}S_n \mid \mathcal{G}_{n+1}\right] = \frac{1}{n+1}S_{n+1}.$$

So $\left(\frac{1}{-n}S_{-n}\right)_{n\in\mathbb{Z}_-}$ is indeed a backwards martingale. By Theorem 3.2.2, $\left(\frac{1}{n}S_n\right)_{n\in\mathbb{N}}$ converges a.s. and in L^1 to $Y_\infty \in L^1(\Omega,\mathcal{F},\mathbb{P})$. If $k\in\mathbb{N}$ is fixed, we have $Y_\infty = \lim_{n\to+\infty}\frac{1}{n}(\xi_{k+1}+\dots+\xi_n)$, so Y_∞ is measurable w.r.t. the asymptotic σ-algebra $\bigcap_{n\in\mathbb{N}}\bigvee_{k\geq n}\sigma(\xi_k)$. By Kolmogorov's Zero-One Law (Corollary 3.1.4), Y_∞ is a.s. constant, so $Y_\infty = \mathbb{E}[Y_\infty] = \lim_{n\to+\infty}\mathbb{E}\left[\frac{1}{n}S_n\right] = \mathbb{E}[\xi_1]$. $\square$

3.3 Radon-Nikodym Theorem

Lemma 3.3.1. *Let μ and ν be two finite measures on a measurable space $(\Omega,\mathcal{F})$. Assume that $\mu \ll \nu$, i.e. $\forall A\in\mathcal{F}$, $\nu(A) = 0 \Longrightarrow \mu(A) = 0$. Then:*

$$\forall\varepsilon>0,\ \exists\delta>0,\ \forall A\in\mathcal{F},\ \nu(A)\leq\delta \Longrightarrow \mu(A)\leq\varepsilon.$$

Proof. Assume for contradiction the existence of $\varepsilon_0>0$ s.t. for all $n\in\mathbb{N}^*$, there exists $A_n\in\mathcal{F}$ s.t. $\nu(A_n)\leq\frac{1}{2^n}$ and $\mu(A_n)>\varepsilon_0$. Then $\nu\left(\limsup_{n\to+\infty}A_n\right) = 0$ but $\mu\left(\limsup_{n\to+\infty}A_n\right)\geq\varepsilon_0$. This is a contradiction. $\square$

Theorem 3.3.2 (Radon-Nikodym Theorem)**.** *Consider a measurable space $(\Omega,\mathcal{F})$ that is separable, i.e. s.t. there exists $(F_n)_{n\in\mathbb{N}}\in\mathcal{P}(\Omega)^{\mathbb{N}}$ s.t. $\mathcal{F} = \sigma(\{F_n\},\ n\in\mathbb{N})$. Let $\mathbb{P}$ and $\mathbb{Q}$ be finite measures on $(\Omega,\mathcal{F})$ with $\mathbb{P}$ a probability measure. If $\mathbb{Q}\ll\mathbb{P}$, then there exists a unique random variable X that is integrable and s.t. $\forall A\in\mathcal{F}$, $\mathbb{Q}(A) = \mathbb{E}[X\mathbb{1}_A]$.*

Proof. Let $\mathcal{F}_n = \sigma(\{F_0\},\dots,\{F_n\})$ for $n\in\mathbb{N}$. We have $\mathcal{F}_n = \sigma(\{A_\varepsilon\},\ \varepsilon\in\{-1,1\}^{n+1})$, where $A_\varepsilon = \bigcap_{i=0}^n F_i^{\varepsilon_i}$, with the notation $F^1 = F$ and $F^{-1} = \Omega\backslash F$. We now define a $\mathcal{F}_n$-measurable random variable X_n as follows:

$$X_n = \sum_{\varepsilon\in\{-1,1\}^{n+1}}\frac{\mathbb{Q}(A_\varepsilon)}{\mathbb{P}(A_\varepsilon)}\mathbb{1}_{A_\varepsilon},$$

with the convention $\frac{\mathbb{Q}(A_\varepsilon)}{\mathbb{P}(A_\varepsilon)} = 0$ if $\mathbb{P}(A_\varepsilon) = 0$. Hence, we have $\forall A\in\mathcal{F}_n$, $\mathbb{Q}(A) = \mathbb{E}[X_n\mathbb{1}_A]$ (i.e. X_n is the Radon-Nikodym derivative of $\mathbb{Q}_{|\mathcal{F}_n}$ w.r.t. $\mathbb{P}_{|\mathcal{F}_n}$). Now, $(X_n)_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)_{n\in\mathbb{N}}$ martingale. With Lemma 3.3.1, we show that $(X_n)_{n\in\mathbb{N}}$ is uniformly integrable. By Theorem 2.8.1, $(X_n)_{n\in\mathbb{N}}$ converges a.s. and in L^1 to a limit X. Therefore, $\forall A\in\bigcup_{n\in\mathbb{N}}\mathcal{F}_n$, $\mathbb{Q}(A) = \mathbb{E}[X\mathbb{1}_A]$. Since $\bigcup_{n\in\mathbb{N}}\mathcal{F}_n$ is stable by finite intersections and generates $\mathcal{F}$, we obtain the result for all $A\in\mathcal{F}$. $\square$

4 Markov chains

4.1 Definitions and first properties

Definition 4.1.1 (Markov transition function)**.** *Let S be a countable set. The elements of S will be viewed as* states*. A* Markov transition function *(or transition matrix, or transition kernel) on S is a map $Q : S^2 \to \mathbb{R}_+$ s.t.*

$$\forall x \in S, \sum_{y \in S} Q(x,y) = 1.$$

Hence, for all $x \in S$, $Q(x,\cdot)$ defines a probability distribution on S.

Notation 4.1.2. *Let S be a countable set.*

(i) *If Q, Q' are two transition functions on S (seen as* matrices*), we define a transition function $QQ' : (x,y) \in S^2 \longmapsto \sum_{z \in S} Q(x,z)Q(z,y)$.*

(ii) *If Q is a transition function and $n \in \mathbb{N}$, we define $Q^n = Q \cdots Q$.*

(iii) *If Q is a transition function and $f \in \mathbb{R}^S$ is a bounded function (seen as a* column vector*), we define a bounded function $Qf : x \in S \longmapsto \sum_{y \in S} Q(x,y)f(y)$.*

(iv) *If Q is a transition function and $\mu \in \mathbb{R}^S$ is a bounded function (seen as a* row vector*), we define a bounded function $\mu Q : y \in S \longmapsto \sum_{x \in S} \mu(x)Q(x,y)$.*

(v) *If Q is a transition function, f is a column vector and μ is a row vector, we define $\mu Q f = \sum_{x,y \in S} \mu(x)Q(x,y)f(y)$.*

Row vectors should be seen as measures on S, while column vector should be seen as functions on S; they play different roles.

Definition 4.1.3 (Markov chains)**.** *A* Markov chain *with transition function Q (or a Q-Markov chain) is a random process $(X_n)_{n \in \mathbb{N}}$ with values in S and s.t., for all $n \in \mathbb{N}$ and $y, x_0, \ldots, x_n \in S$, we have:*

$$\mathbb{P}\left(X_{n+1} = y \mid X_0 = x_0, \ldots, X_n = x_n\right) = Q\left(x_n, y\right),$$

as soon as this probability is well-defined.

Proposition 4.1.4. *If $(X_n)_{n \in \mathbb{N}}$ is a Q-Markov chain, then for all $n \in \mathbb{N}$ and $x_0, \ldots, x_n \in S$:*

$$\mathbb{P}\left(X_0 = x_0, \ldots, X_n = x_n\right) = \mathbb{P}\left(X_0 = x_0\right) Q\left(x_0, x_1\right) Q\left(x_1, x_2\right) \cdots Q\left(x_{n-1}, x_n\right).$$

Proposition 4.1.5. *The random process $(X_n)_{n \in \mathbb{N}}$ is a Q-Markov chain* iff *for all $y \in S$, we have:*

$$\mathbb{P}\left(X_{n+1} = y \mid X_0, \ldots, X_n\right) = Q\left(X_n, y\right).$$

In other words, the conditional law of X_{n+1} given $X_0, \ldots, X_n$ is $Q(X_n, \cdot)$.

Corollary 4.1.6. *Let $(X_n)_{n \in \mathbb{N}}$ be a Q-Markov chain. Let $f : S \to \mathbb{R}_+$ be a bounded function.*

(i) *For all $n \in \mathbb{N}$, $\mathbb{E}\left[f\left(X_{n+1}\right) \mid X_0, \ldots, X_n\right] = Qf\left(X_n\right)$.*

(ii) *For all $x \in S$, $\mathbb{E}\left[f\left(X_1\right) \mid X_0 = x\right] = Qf(x)$.*

(iii) *If μ is the probability law of X_0, then $\mathbb{E}\left[f\left(X_1\right)\right] = \mu Q f$.*

Proposition 4.1.7. *Let $(X_n)_{n \in \mathbb{N}}$ be a Q-Markov chain.*

(i) *For $x_0, x_n \in S$, $\mathbb{P}\left(X_n = x_n \mid X_0 = x_0\right) = Q^n\left(x_0, x_n\right)$.*

(ii) *For $y_1, \dots, y_k, x_0, \dots, x_n \in S$:*

$$\mathbb{P}\left(X_{n+1} = y_1, \dots, X_{n+k} = y_k \mid X_0 = x_0, \dots, X_n = x_n\right) = \mathbb{P}\left(X_1 = y_1, \dots, X_k = y_k \mid X_0 = x_n\right).$$

Therefore, $(X_{k+np})_{n\in\mathbb{N}}$ is a Q^p-Markov chain for all $p, k \in \mathbb{N}^*$.

Example 4.1.8.

(i) Independent and identically distributed sequences. *Let μ be a probability distribution on a countable set S. Set $Q : (x, y) \in S^2 \longmapsto \mu(y)$. Then an i.i.d. sequence of random variables of law μ is a Q-Markov chain.*

(ii) Random walks in abelian groups. *Let G be a countable abelian group equipped with a probability law μ. Let $(\xi_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random variables of law μ. Then the sequence $\left(\sum_{k=1}^n \xi_k\right)_{n\in\mathbb{N}}$ is a Markov chain associated to the transition function $Q(x, y) = \mu(y - x)$.*

(iii) Branching processes. *Let μ be a probability distribution on $\mathbb{N}$, let $(\xi_{n,i})_{(n,i)\in\mathbb{N}^2}$ be i.i.d. random variables of law μ. Let X_0 be a random variable independent of $(\xi_{n,i})_{(n,i)\in\mathbb{N}^2}$, and define $(X_n)_{n\in\mathbb{N}}$ by induction by:*

$$X_{n+1} = \sum_{i=0}^{X_n - 1} \xi_{n,i}.$$

Then $(X_n)_{n\in\mathbb{N}}$ is a Markov chain associated to the transition function $Q(x, y) = \mu^{*x}(y) = \sum_{n_1+\cdots+n_x=y} \mu(n_1)\cdots\mu(n_x)$.

4.2 Existence of Markov chains, the canonical process

Theorem 4.2.1. *For every probability distribution μ and for every transition function Q on a countable set S, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables that is a Q-Markov chain with initial law μ.*

Proof. Take a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define a sequence $(U_n)_{n\in\mathbb{N}}$ of i.i.d. random variables with law $\mathcal{U}([0, 1])$. Now, arrange the elements of S into a list $(s_i)_{i\in\mathbb{N}}$ and define:

$$X_0 = \sum_{i\in\mathbb{N}^*} s_i \mathbb{1}_{\left\{\sum_{j<i} \mu(s_j) \le U_0 < \sum_{j\le i} \mu(s_j)\right\}}.$$

Thus, X_0 has law μ. Inductively, once $X_0, \dots, X_n$ have been constructed s.t. X_k is $\sigma(U_0, \dots, U_k)$-measurable and $(X_0, \dots, X_n)$ is a Q-Markov chain, set:

$$X_{n+1} = \sum_{i\in\mathbb{N}^*} s_i \mathbb{1}_{\left\{\sum_{j<i} Q(X_n, s_j) \le U_{n+1} < \sum_{j\le i} Q(X_n, s_j)\right\}}.$$

Thus, for $j \in \mathbb{N}^*$, $\mathbb{P}\left(X_{n+1} = s_j \mid \mathcal{F}_n^U\right) = Q(X_n, x_j)$. Hence, by induction, we construct a Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ with initial law μ. □

Proposition 4.2.2. *Let $(Y_n)_{n\in\mathbb{N}}$ be a Q-Markov chain defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider the measurable space $\left(S^{\mathbb{N}}, \mathcal{P}(S)^{\otimes\mathbb{N}}\right)$ and the (measurable) map $\varphi : \omega \in \Omega \longmapsto (Y_n(\omega))_{n\in\mathbb{N}} \in S^{\mathbb{N}}$. If $(Y_n)_{n\in\mathbb{N}}$ is s.t. $\mathbb{P}(Y_0 = x) = 1$ for some $x \in S$ (i.e. $(Y_n)_{n\in\mathbb{N}}$ has initial distribution δ_x), then we write $\mathbb{P}_x = \varphi_*\mathbb{P}$; it is a probability distribution on $\left(S^{\mathbb{N}}, \mathcal{P}(S)^{\otimes\mathbb{N}}\right)$. Now, the family of laws $(\mathbb{P}_x)_{x\in S}$ does not depend on the choice of $(\Omega, \mathcal{F}, \mathbb{P})$ and the sequence $(X_n)_{n\in\mathbb{N}}$ of random variables on $\left(S^{\mathbb{N}}, \mathcal{P}(S)^{\otimes\mathbb{N}}, \mathbb{P}_x\right)$ defined as the projections on each coordinate is called the* canonical Q-Markov chain.

4.3 The simple and strong Markov properties

Notation 4.3.1. *For any set S and for any $k \in \mathbb{N}$, we define the* shift operator $\theta_k : x \in S^{\mathbb{N}} \longmapsto (x_{n+k})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$.

Theorem 4.3.2 (Simple Markov Property)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain (write $\mathbb{P}_x$ for the law of the canonical Markov chain starting at x, and $\mathbb{E}_x$ for the corresponding expectation). Let $\mathcal{F} = \bigvee_{n\in\mathbb{N}} \mathcal{F}_n^X$. Let G be a nonnegative $\mathcal{F}$-measurable function. Then, for every $n \in \mathbb{N}$ and for every nonnegative $\mathcal{F}_n^X$-measurable function F, we have:*

$$\mathbb{E}_x \left[F \cdot (G \circ \theta_n)\right] = \mathbb{E}_x \left[F \cdot \mathbb{E}_{X_n} [G]\right].$$

In other words, for every $n \in \mathbb{N}$:

$$\mathbb{E}_x \left[G \circ \theta_n \mid \mathcal{F}_n^X\right] = \mathbb{E}_{X_n} [G].$$

Note that $\mathbb{E}_{X_n}[G]$ is the random variable $\omega \mapsto \mathbb{E}_{X_n(\omega)}[G]$.

Proof. Show that the statement is true for $F = \mathbb{1}_{\{X_0=x_0,\dots,X_n=x_n\}}$ and $G = \mathbb{1}_{\{X_0=x_n,\dots,X_k=x_{n+k}\}}$ with $x_0, \dots, x_{n+k} \in S$. Use the Monotone Class Theorem to generalise to $F = \mathbb{1}_A$, $G = \mathbb{1}_B$ with $A \in \mathcal{F}_n$ and $B \in \mathcal{F}$, and then use an approximation argument to obtain the result. □

Theorem 4.3.3 (Strong Markov Property)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $\mathcal{F} = \bigvee_{n\in\mathbb{N}} \mathcal{F}_n^X$. Let G be a nonnegative $\mathcal{F}$-measurable function. Then, for every $\left(\mathcal{F}_n^X\right)_{n\in\mathbb{N}}$-stopping time τ and for every nonnegative $\mathcal{F}_\tau$-measurable function F, we have:*

$$\mathbb{E}_x \left[F\mathbb{1}_{\{\tau<+\infty\}} \cdot (G \circ \theta_\tau)\right] = \mathbb{E}_x \left[F\mathbb{1}_{\{\tau<+\infty\}} \cdot \mathbb{E}_{X_\tau} [G]\right].$$

In other words, for every stopping time τ:

$$\mathbb{1}_{\{\tau<+\infty\}}\mathbb{E}_x \left[G \circ \theta_\tau \mid \mathcal{F}_\tau\right] = \mathbb{1}_{\{\tau<+\infty\}}\mathbb{E}_{X_\tau} [G].$$

Proof. Let F, G be as above. Then:

$$\begin{aligned}\mathbb{E}_x \left[F\mathbb{1}_{\{\tau<+\infty\}} \cdot (G \circ \theta_\tau)\right] &= \sum_{t\in\mathbb{N}} \mathbb{E}_x \left[F\mathbb{1}_{\{\tau=t\}} \cdot (G \circ \theta_t)\right] = \sum_{t\in\mathbb{N}} \mathbb{E}_x \left[F\mathbb{1}_{\{\tau=t\}} \cdot \mathbb{E}_{X_t} [G]\right] \\ &= \mathbb{E}_x \left[F\mathbb{1}_{\{\tau<+\infty\}} \cdot \mathbb{E}_{X_\tau} [G]\right].\end{aligned}$$

□

Remark 4.3.4. *Consider the canonical Q-Markov chain $(X_n)_{n\in\mathbb{N}}$. If μ is any distribution on S, write:*

$$\mathbb{P}_\mu = \sum_{x\in S} \mu(x)\mathbb{P}_x;$$

thus, under $\left(S^{\mathbb{N}}, \mathcal{P}(S)^{\otimes\mathbb{N}}, \mathbb{P}_\mu\right)$, $(X_n)_{n\in\mathbb{N}}$ is a Q-Markov chain with initial law μ. Hence, the simple and strong Markov properties remain valid if one replaces $\mathbb{E}_x$ by $\mathbb{E}_\mu$.

4.4 Classification of states

Proposition 4.4.1. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. For $x \in S$, define $T_x^+ = \inf\{n \in \mathbb{N}^*,\ X_n = x\}$ and $N_x = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{X_n=x\}}$. Then we are in one of the two following cases:*

(i) *Either $\mathbb{P}_x(T_x^+ < +\infty) = 1$ and $N_x = +\infty$ $\mathbb{P}_x$-a.s. We then say that x is a* recurrent state.

(ii) *Or $\mathbb{P}_x(T_x^+ < +\infty) < 1$ and $N_x < +\infty$ $\mathbb{P}_x$-a.s. Moreover:*

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

We then say that x is a transient state.

Proof. Note that $\{N_x \geq k+1\} = \left\{T_x^+ < +\infty,\ N_x \circ \theta_{T_x^+} \geq k\right\}$. Therefore, by the Strong Markov Property:

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

Hence:

$$\mathbb{P}_x\left(N_x \geq k\right) = \mathbb{P}_x\left(T_x^+ < +\infty\right)^k.$$

So, under $\mathbb{P}_x$, we have shown that N_x follows a geometric distribution with parameter $\mathbb{P}_x\left(T_x^+ = \infty\right)$; we easily deduce the result. $\square$

Definition 4.4.2 (Green function of a Markov chain)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. We define the* Green function *of $(X_n)_{n\in\mathbb{N}}$ by:*

$$G : (x,y) \in S \longmapsto \mathbb{E}_x\left[\sum_{n\in\mathbb{N}} \mathbb{1}_{\{X_n=y\}}\right].$$

We have shown that $x \in S$ is recurrent iff *$G(x,x) = +\infty$.*

Remark 4.4.3. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Then, for all $x, y \in S$:*

$$G(x,y) = \sum_{n\in\mathbb{N}} Q^n(x,y).$$

In particular, $G(x,y) > 0 \iff \exists n \in \mathbb{N},\ Q^n(x,y) > 0$.

Proposition 4.4.4 ("Recurrent states are contagious")**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $x \neq y$ be two states s.t. x is recurrent and $G(x,y) > 0$. Write $T_x = \inf\{n \in \mathbb{N},\ X_n = x\}$ and similarly for T_y. Then:*

(i) $\mathbb{P}_x\left(T_y < +\infty\right) = \mathbb{P}_y\left(T_x < +\infty\right) = 1$.

(ii) *y is a recurrent state.*

(iii) $G(y,x) > 0$.

Proof. Since x is recurrent, the Strong Markov Property implies that:

$$\begin{aligned}0 = \mathbb{P}_x\left(T_x^+ = +\infty\right) &\geq \mathbb{P}_x\left(T_x^+ = +\infty,\ T_y = +\infty\right) = \mathbb{P}_x\left(T_x^+ \circ \theta_{T_y} = +\infty,\ T_y = +\infty\right)\\ &= \mathbb{E}_x\left[\mathbb{1}_{\{T_y<+\infty\}}\mathbb{E}_{X_{T_y}}\left[\mathbb{1}_{\{T_x=+\infty\}}\right]\right] = \mathbb{P}_x\left(T_y < +\infty\right)\mathbb{P}_y\left(T_x = +\infty\right).\end{aligned}$$

Since $\mathbb{P}_x\left(T_y < +\infty\right) > 0$ (because $G(x,y) > 0$), we obtain $\mathbb{P}_y\left(T_x = +\infty\right) = 0$. In particular, $G(y,x) > 0$. Since $G(x,y) > 0$ and $G(y,x) > 0$, there exist $n_1, n_2 \in \mathbb{N}$ s.t. $Q^{n_1}(x,y) > 0$ and $Q^{n_2}(y,x) > 0$. Now:

$$G(y,y) \geq Q^{n_2}(y,x)G(x,x)Q^{n_1}(x,y) = +\infty,$$

so y is recurrent. Moreover, by symmetry, $\mathbb{P}_x\left(T_y < +\infty\right) = 1$. $\square$

Definition 4.4.5 (Irreducible chain)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. We say that the Markov chain is* irreducible *if $\forall x, y \in S$, $G(x,y) > 0$. In this case, either all states are recurrent (and we say that the chain is* recurrent*) or all states are transient (and we say that the chain is* transient*).*

Definition 4.4.6 (Recurrence classes)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $R \subseteq S$ be the set of recurrent states. Define an equivalence relation $\sim$ on R by:*

$$x \sim y \iff G(x,y) > 0.$$

The relation $\sim$ is indeed an equivalence relation by Proposition 4.4.4. Its equivalence classes are called the recurrence classes *of the Markov chain.*

Theorem 4.4.7. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. For $x \in R$, denote by R_x the recurrence class of x. For any $y \in S$, write $N_y = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{X_n=y\}}$.*

(i) *If $x \in R$ is a recurrent state, then $\mathbb{P}_x$-a.s.:*

$$\forall y \in R_x,\ N_y = +\infty \qquad \text{and} \qquad \forall y \in S\backslash R_x,\ N_y = 0.$$

(ii) *If $x \in S\backslash R$ is a transient state, we have $N_y < +\infty$ $\mathbb{P}_x$-a.s. for all $y \in S\backslash R$. Moreover, if we let $T = \inf\{n \in \mathbb{N},\ X_n \in R\}$, we have $\mathbb{P}_x$-a.s.:*

- *Either $T = +\infty$ and $\forall y \in R,\ N_y = 0$.*
- *Or $T < +\infty$ and $\forall y \in R_{X_T},\ N_y = +\infty$ and $\forall y \in R\backslash R_{X_T},\ N_y = 0$.*

Example 4.4.8.

(i) *Let $(\xi_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random variables with law μ on S. If ξ_0 is any random variable on S that is independent from $(\xi_n)_{n\in\mathbb{N}^*}$, then $(\xi_n)_{n\in\mathbb{N}}$ is a Markov chain. The set of recurrent states is the support of μ, and all recurrent states are in the same recurrence class.*

(ii) *Let μ be a law on $\mathbb{Z}$ and let $(\xi_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random variables with law μ. Let $S_n = \sum_{k=1}^n \xi_k$ for $n \in \mathbb{N}$. Then $(S_n)_{n\in\mathbb{N}}$ is a Markov chain on $\mathbb{Z}$. If we assume that ξ_1 is L^1, then:*

- *If $\mathbb{E}[\xi_1] \neq 0$, then all states are transient.*
- *If $\mathbb{E}[\xi_1] = 0$, then all states are recurrent, and the chain is irreducible* iff *the support of μ generates $\mathbb{Z}$ (as an additive group).*

(iii) *Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on a finite set S. Then the chain is irreducible* iff *the graph $(S, \{(x,y) \in S^2,\ Q(x,y) > 0\})$ is strongly connected. Moreover, if the chain is irreducible, then it is recurrent.*

(iv) *Let $(X_n)_{n\in\mathbb{N}}$ be the branching process with law μ, as in Example 4.1.8. Assume that $\mu(1) < 1$. Note that $(X_n)_{n\in\mathbb{N}}$ is a Markov chain, where 0 is the only recurrent state (and 0 is even absorbing).*

4.5 Invariant measures for Markov chains

Definition 4.5.1 (Invariant measure)**.** *Let μ be a nonnegative measure on S. We say that μ is* invariant *for the transition function Q if:*

$$\mu Q = \mu,$$

i.e. $\forall y \in S,\ \mu(y) = \sum_{x\in S} \mu(x)Q(x,y)$. In this case, we have $\forall n \in \mathbb{N},\ \mu Q^n = \mu$.

Remark 4.5.2. *Let μ be an invariant measure for Q. If $(X_n)_{n\in\mathbb{N}}$ is a Q-Markov chain with initial "law" μ (which does not always make sense because μ is not necessarily a probability distribution), then X_n also has "law" μ for all $n \in \mathbb{N}$.*

Definition 4.5.3 (Reversible measure)**.** *Let μ be a nonnegative measure on S. We say that μ is* reversible *for the transition function Q if:*

$$\forall x, y \in S,\ \mu(x)Q(x,y) = \mu(y)Q(y,x).$$

A reversible measure is always invariant.

Theorem 4.5.4. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. If x is a recurrent state, set:*

$$\mu_x : y \in S \longmapsto \mathbb{E}_x\left[\sum_{n=0}^{T_x^+-1} \mathbb{1}_{\{X_n=y\}}\right],$$

where $T_x^+ = \inf\{n \in \mathbb{N}^*,\ X_n = x\}$. Then μ_x is an invariant measure, $\mu_x(x) = 1$ and $\mu_x(y) > 0$ iff y is in the recurrence class of x.

Proof. Note that:

$$\begin{aligned}
\mu_x(y) &= \sum_{n\in\mathbb{N}^*} \mathbb{P}_x\left(X_n = y,\ n \le T_x^+\right) = \sum_{n\in\mathbb{N}^*}\sum_{z\in S} \mathbb{P}_x\left(X_{n-1} = z,\ X_n = y,\ n \le T_x^+\right) \\
&= \sum_{n\in\mathbb{N}^*}\sum_{z\in S} \mathbb{E}_x\left[\mathbb{1}_{\left\{X_{n-1}=z,\, n\le T_x^+\right\}}\mathbb{P}_{X_{n-1}}\left(X_1 = y\right)\right] \\
&= \sum_{n\in\mathbb{N}^*}\sum_{z\in S} \mathbb{P}_x\left(X_{n-1} = z,\ n \le T_x^+\right) Q(z,y) = \sum_{z\in S} Q(z,y)\mu_x(z).
\end{aligned}$$

□

Remark 4.5.5. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. If $x_1, \ldots, x_r$ are recurrent states then for all $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$, $\sum_{i=1}^r \alpha_i \mu_{x_i}$ is an invariant measure.*

Lemma 4.5.6. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and recurrent. If ν is an invariant measure, then:*

$$\forall x, y \in S,\ \forall p \in \mathbb{N},\ \nu(y) \ge \nu(x)\mathbb{E}_x\left[\sum_{k=0}^{\left(T_x^+-1\right)\wedge p} \mathbb{1}_{\{X_n=y\}}\right].$$

Proof. By induction on p, for x fixed and y arbitrary. The result is clear for $p = 0$. If it is true for p, then write $\nu(y) = \sum_{z\in S}\nu(z)Q(z,y)$ and apply the induction hypothesis to obtain a lower bound for $\nu(z)$. □

Proposition 4.5.7. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and recurrent. Then, for all $x \in S$, any invariant measure ν is equal to $\nu(x)\mu_x$.*

Proof. Fix $x \in S$ and let ν be an invariant measure. Using Lemma 4.5.6, letting $p \to +\infty$ and using the Monotone Convergence Theorem, we obtain:

$$\forall y \in S,\ \nu(y) \ge \nu(x)\mu_x(y).$$

But since ν and μ_x are both invariant, we have:

$$\forall n \in \mathbb{N},\ \nu(x) = \sum_{y\in S}\nu(y)Q^n(y,x) \ge \sum_{y\in S}\nu(x)\mu_x(y)Q^n(y,x) = \nu(x)\mu_x(x) = \nu(x).$$

Thus, equality must hold throughout, which gives $\forall y \in S,\ \forall n \in \mathbb{N},\ \left(\nu(y) - \nu(x)\mu_x(y)\right)Q^n(x,y) = 0$. Fixing $y \in S$ and summing over n gives $\left(\nu(y) - \nu(x)\mu_x(y)\right)G(x,y) = 0$, so $\nu(y) = \nu(x)\mu_x(y)$ because $G(x,y) > 0$ since the chain is irreducible. □

4.6 Invariant measures of finite mass

Remark 4.6.1. *For an irreducible and recurrent Markov chain, either all nonzero invariant measures have finite mass (in which case there exists an invariant probability distribution) or none of them does.*

Proposition 4.6.2. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and admits a nonzero invariant measure μ of finite mass. Then the Markov chain is recurrent.*

Proof. Let $x \in S$ s.t. $\mu(x) > 0$. We have $\forall n \in \mathbb{N}$, $\mu(x) = \sum_{y \in S} \mu(y) Q^n(y,x)$, therefore:

$$+\infty = \sum_{y \in S} \mu(y) G(y,x) \leq \sum_{y \in S} \mu(y) \frac{G(y,x)}{\mathbb{P}_y\left(T_x < +\infty\right)} = \sum_{y \in S} \mu(y) G(x,x) = \mu(S) G(x,x),$$

so $G(x,x) = +\infty$ and x is recurrent. ☐

Theorem 4.6.3. *Assume that the Q-Markov chain $(X_n)_{n \in \mathbb{N}}$ is irreducible and recurrent. Then we are in one of the following two situations:*

(i) *Either there exists an invariant probability distribution π (also called the* stationary probability distribution*), in which case $\mathbb{E}_x\left[T_x^+\right] < +\infty$ and $\pi(x) = \frac{1}{\mathbb{E}_x\left[T_x^+\right]}$ for all $x \in S$. We then say that the chain is* positive recurrent*.*

(ii) *Or all nonzero invariant measures have infinite mass, in which case $\mathbb{E}_x\left[T_x^+\right] = +\infty$ for all $x \in S$. We then say that the chain is* null recurrent*.*

Proof. The dichotomy is clear given Remark 4.6.1 and the fact that $\mu_x(S) = \mathbb{E}_x\left[T_x^+\right]$ for all $x \in S$. And in the case where the chain is positive recurrent, we have $\pi = \frac{\mu_x}{\mu_x(S)} = \frac{\mu_x}{\mathbb{E}_x\left[T_x^+\right]}$ for all $x \in S$. ☐

Corollary 4.6.4. *An irreductible Markov chain with finite state space is positive recurrent.*

Example 4.6.5. *Consider a nonoriented simple graph $G = (V,E)$. Define a transition function Q by $Q(x,y) = \frac{1}{\deg x}$ if $\{x,y\} \in E$, and $Q(x,y) = 0$ otherwise. Consider the Q-Markov chain.*

(i) *The chain is irreducible* iff *G is connected.*

(ii) *An invariant measure is given by $\mu(x) = \deg x$ for all $x \in V$.*

(iii) *Assume that G is connected and finite. Then the chain is positive recurrent, with invariant probability measure $\pi(x) = \frac{\deg x}{2|E|}$. In particular:*

$$\mathbb{E}_x\left[T_x^+\right] = \frac{2\,|E|}{\deg x}.$$

(iv) *Assume that G is connected and infinite. Then the chain is either transient or null recurrent.*

Example 4.6.6. *The simple random walk on $\mathbb{Z}$ is irreducible and null recurrent, and $\mu(n) = 1$ defines an invariant measure.*

4.7 Asymptotic behaviour of recurrent chains – an ergodic theorem

Theorem 4.7.1. *Assume that the Q-Markov chain $(X_n)_{n \in \mathbb{N}}$ is irreducible and recurrent. Let μ be a nonzero invariant measure. Let $f, g : S \to \mathbb{R}_+$ be s.t. the integrals $\int_S f \,\mathrm{d}\mu$ and $\int_S g \,\mathrm{d}\mu$ are not both infinite. Then, for all $x \in S$, the following holds $\mathbb{P}_x$-a.s.:*

$$\frac{\sum_{k=0}^n f\left(X_k\right)}{\sum_{k=0}^n g\left(X_k\right)} \xrightarrow[n \to +\infty]{} \frac{\int_S f \,\mathrm{d}\mu}{\int_S g \,\mathrm{d}\mu}.$$

Proof. Fix $x \in S$. We may assume that both integrals are finite (otherwise, approximate the function with infinite integral by a monotone sequence of functions with finite integrals). Define a sequence $\left(T^{(k)}\right)_{k \in \mathbb{N}}$ of random variable by $T^{(0)} = 0$, and:

$$\forall k \in \mathbb{N},\ T^{(k+1)} = T^{(k)} + T_x^+ \circ \theta_{T^{(k)}}.$$

Hence $0 = T^{(0)} < T^{(1)} < \cdots < T^{(k)} < \cdots$ are the consecutive visit times at x; they are all stopping times. Define another sequence $(Z_k)_{k\in\mathbb{N}^*}$ of random variables by:

$$\forall k \in \mathbb{N}^*,\ Z_k = \sum_{i=T^{(k-1)}+1}^{T^{(k)}} f\left(X_i\right).$$

If $h_1, \ldots, h_m$ are functions $S \to \mathbb{R}_+$, one shows by induction, using the Strong Markov Property (Theorem 4.3.3), that:

$$\mathbb{E}_x\left[h_1\left(Z_1\right)\cdots h_m\left(Z_m\right)\right] = \mathbb{E}_x\left[h_1\left(Z_1\right)\right]\cdots \mathbb{E}_x\left[h_m\left(Z_m\right)\right].$$

Therefore, $(Z_k)_{k\in\mathbb{N}^*}$ is a sequence of independent random variables. They are also identically distributed, and we have $\mathbb{E}_x\left[Z_1\right] = \frac{1}{\mu(x)}\int_S f\ \mathrm{d}\mu$. Using the Strong Law of Large Numbers (Corollary 3.2.3), we obtain, $\mathbb{P}_x$-a.s.:

$$\frac{1}{k}\sum_{i=1}^{k} Z_i \xrightarrow[k\to+\infty]{} \frac{1}{\mu(x)}\int_S f\ \mathrm{d}\mu.$$

Now, for $n \in \mathbb{N}^*$, there exists a unique random variable $k(n)$ s.t. $T^{(k(n))} < n \leq T^{(k(n)+1)}$. We have $k(n) \xrightarrow[n\to+\infty]{} +\infty$ $\mathbb{P}_x$-a.s. And, since $f \geq 0$:

$$\underbrace{\frac{1}{k(n)}\sum_{i=1}^{k(n)} Z_i}_{\xrightarrow[n\to+\infty]{} \frac{1}{\mu(x)}\int_S f\ \mathrm{d}\mu} \leq \frac{1}{k(n)}\sum_{j=1}^{n} f\left(X_j\right) \leq \underbrace{\frac{1}{k(n)}\sum_{i=1}^{k(n)+1} Z_i}_{\xrightarrow[n\to+\infty]{} \frac{1}{\mu(x)}\int_S f\ \mathrm{d}\mu}.$$

Therefore $\frac{1}{k(n)}\sum_{j=1}^{n} f\left(X_j\right) \xrightarrow[n\to+\infty]{} \frac{1}{\mu(x)}\int_S f\ \mathrm{d}\mu$. We obtain the result by dividing by the same sum for g instead of f. □

Remark 4.7.2. *Theorem 4.7.1 remains valid if we replace $\mathbb{P}_x$ by $\mathbb{P}_\gamma$ where γ is any probability distribution on S (because $\mathbb{P}_\gamma = \sum_{x\in S}\gamma(x)\mathbb{P}_x$).*

Corollary 4.7.3. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and recurrent.*

(i) *If $(X_n)_{n\in\mathbb{N}}$ is positive recurrent, then, for all $x, y \in S$, the following holds $\mathbb{P}_x$-a.s.:*

$$\frac{1}{n}\sum_{k=0}^{n} \mathbb{1}_{\{X_k=y\}} \xrightarrow[n\to+\infty]{} \pi(y),$$

where π is the stationary probability distribution.

(ii) *If $(X_n)_{n\in\mathbb{N}}$ is null recurrent, then, for all $x, y \in S$, the following holds $\mathbb{P}_x$-a.s.:*

$$\frac{1}{n}\sum_{k=0}^{n} \mathbb{1}_{\{X_k=y\}} \xrightarrow[n\to+\infty]{} 0.$$

Corollary 4.7.4. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and positive recurrent. Let $f : S \to \mathbb{R}$ be an integrable function w.r.t. the stationary probability distribution π. Then, for all $x \in S$, the following holds $\mathbb{P}_x$-a.s.:*

$$\frac{1}{n}\sum_{i=0}^{n} f\left(X_i\right) \xrightarrow[n\to+\infty]{} \int_S f\ \mathrm{d}\pi.$$

4.8 Asymptotic behaviour of Markov chains – convergence in probability

Definition 4.8.1 (Period of a state). *The* period *of a state $x \in S$ of the Q-Markov chain is defined by:*

$$d_x = \gcd \{n \in \mathbb{N},\ Q^n(x,x) > 0\}.$$

Lemma 4.8.2. *Let $x \in S$ be a state of the Q-Markov chain. Let $A_x = \{n \in \mathbb{N},\ Q^n(x,x) > 0\}$. Then the subgroup of $\mathbb{Z}$ generated by A_x is:*

$$\langle A_x \rangle = A_x - A_x = d_x \mathbb{Z}.$$

Proof. Show that A_x is stable under addition. □

Proposition 4.8.3. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible. Then all states have the same period, and this period is called the period of the chain and denoted by d.*

Proof. Let $x, y \in S$. Since the Markov chain is irreducible, there exist $n_1, n_2 \in \mathbb{N}^*$ s.t. $Q^{n_1}(y,x) > 0$ and $Q^{n_2}(x,y) > 0$. Thus, $Q^{n_1+n_2}(y,y) \geq Q^{n_1}(y,x)Q^{n_2}(x,y) > 0$, so $n_1 + n_2 \in A_y$. Now, for $n \in A_x$, we have $n_1 + n + n_2 \in A_y$ for the same reason, and therefore:

$$n = (n_1 + n + n_2) - (n_1 + n_2) \in \langle A_y \rangle.$$

This proves that $A_x \subseteq \langle A_y \rangle$, so $\langle A_x \rangle \subseteq \langle A_y \rangle$. By symmetry, $\langle A_x \rangle = \langle A_y \rangle$, and by Lemma 4.8.2, $d_x = d_y$. □

Definition 4.8.4 (Aperiodic chain). *An irreducible Markov chain is said to be* aperiodic *if it has period 1.*

Remark 4.8.5. *In an irreducible Markov chain, if there exists $x_0 \in S$ s.t. $Q(x_0, x_0) > 0$, then the chain is aperiodic.*

Lemma 4.8.6. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and aperiodic. Then:*

$$\forall x, y \in S,\ \exists n_0 \in \mathbb{N},\ \forall n \geq n_0,\ Q^n(x,y) > 0.$$

Proof. Note that we only need to prove the result for $x = y$ (indeed, if $x \neq y$, there exists $m_0 \in \mathbb{N}$ s.t. $Q^{m_0}(x,y) > 0$ and thus, if $Q^n(x,x) > 0$ then $Q^{n+m_0}(x,y) \geq Q^n(x,x) + Q^{m_0}(x,y) > 0$). Since $d_x = 1$, we have $\langle A_x \rangle = \mathbb{Z}$ by Lemma 4.8.2, so there exists $m \in A_x$ s.t. $m + 1 \in A_x$. If $m \in \{0, 1\}$, then $1 \in A_x$ and we are done. So assume that $m \geq 2$. Note that:

$$\forall k \in \{0, \dots, m\},\ m^2 + k = (m-k)\,m + k\,(m+1) \in A_x.$$

We have found $(m+1)$ consecutive integers in A_x. Thus, if $n \geq m^2$, write $n = km + i$, with $m^2 \leq i < m^2 + m$ and $k \in \mathbb{N}$. Thus, $m \in A_x$ and $i \in A_x$, so $n \in A_x$. This proves the result. □

Theorem 4.8.7. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible, aperiodic and positive recurrent. Then, for all $x \in S$:*

$$\sum_{y\in S} |\mathbb{P}_x(X_n = y) - \pi(y)| \xrightarrow[n\to+\infty]{} 0,$$

where π is the stationary probability distribution.

Proof. Define a transition function $\hat{Q}$ on $S \times S$ by:

$$\hat{Q}((x,y),(x',y')) = Q(x,x')\,Q(y,y').$$

Now, consider the canonical process $(X_n, Y_n)_{n\in\mathbb{N}}$ for the $\hat{Q}$-Markov chain under the law $\hat{\mathbb{P}}_{\delta_x\otimes\pi}$. Then $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are both Q-Markov chains, with respective initial distributions δ_x and π. Using the aperiodicity of Q and Lemma 4.8.6, we show that $(X_n, Y_n)_{n\in\mathbb{N}}$ is irreducible. Moreover, if π is the stationary probability distribution for Q, then $\pi\otimes\pi$ is an invariant probability distribution for $\hat{Q}$; by Proposition 4.6.2, $(X_n, Y_n)_{n\in\mathbb{N}}$ is positive recurrent. Let:

$$\tau = \inf\{n\in\mathbb{N},\ X_n = Y_n\}.$$

Note that $\tau \le T_{(0,0)} = \inf\{n\in\mathbb{N},\ (X_n, Y_n) = (0,0)\}$. But since the Markov chain is recurrent, $T_{(0,0)}$ is $\hat{\mathbb{P}}_{\delta_x\otimes\pi}$-a.s. finite, and so is τ. Now, for $n\in\mathbb{N}$:

$$\begin{aligned}\left|\mathbb{P}_x(X_n = y) - \pi(y)\right| &= \left|\hat{\mathbb{P}}_{\delta_x\otimes\pi}(X_n = y) - \hat{\mathbb{P}}_{\delta_x\otimes\pi}(Y_n = y)\right| = \left|\hat{\mathbb{E}}_{\delta_x\otimes\pi}\left[\mathbb{1}_{\{X_n=y\}} - \mathbb{1}_{\{Y_n=y\}}\right]\right| \\ &\le \left|\hat{\mathbb{E}}_{\delta_x\otimes\pi}\left[\left(\mathbb{1}_{\{X_n=y\}} - \mathbb{1}_{\{Y_n=y\}}\right)\mathbb{1}_{\{\tau\le n\}}\right]\right| + \hat{\mathbb{E}}_{\delta_x\otimes\pi}\left[\left(\mathbb{1}_{\{X_n=y\}} + \mathbb{1}_{\{Y_n=y\}}\right)\mathbb{1}_{\{\tau> n\}}\right].\end{aligned}$$

And:

$$\begin{aligned}\hat{\mathbb{P}}_{\delta_x\otimes\pi}(X_n = y, \tau\le n) &= \sum_{k=0}^{n}\sum_{z\in S}\hat{\mathbb{P}}_{\delta_x\otimes\pi}(X_n = y, X_k = z, \tau = k) \\ &= \sum_{k=0}^{n}\sum_{z\in S}\hat{\mathbb{P}}_{\delta_x\otimes\pi}(X_k = z, \tau = k)\,Q^{n-k}(z,y) \\ &= \sum_{k=0}^{n}\sum_{z\in S}\hat{\mathbb{P}}_{\delta_x\otimes\pi}(Y_k = z, \tau = k)\,Q^{n-k}(z,y) \\ &= \hat{\mathbb{P}}_{\delta_x\otimes\pi}(Y_n = y, \tau\le n).\end{aligned}$$

Thus $\hat{\mathbb{E}}_{\delta_x\otimes\pi}\left[\left(\mathbb{1}_{\{X_n=y\}} - \mathbb{1}_{\{Y_n=y\}}\right)\mathbb{1}_{\{\tau\le n\}}\right] = \hat{\mathbb{P}}_{\delta_x\otimes\pi}(X_n = y, \tau\le n) - \hat{\mathbb{P}}_{\delta_x\otimes\pi}(Y_n = y, \tau\le n) = 0$. From this, we obtain:

$$\sum_{y\in S}\left|\mathbb{P}_x(X_n = y) - \pi(y)\right| \le \sum_{y\in S}\hat{\mathbb{E}}_{\delta_x\otimes\pi}\left[\left(\mathbb{1}_{\{X_n=y\}} + \mathbb{1}_{\{Y_n=y\}}\right)\mathbb{1}_{\{\tau> n\}}\right] = 2\hat{\mathbb{P}}_{\delta_x\otimes\pi}(\tau > n) \xrightarrow[n\to+\infty]{} 0.$$

□

Remark 4.8.8. *If S is finite, then Theorem 4.8.7 is a consequence of the Perron-Frobenius Theorem: let M be a $N\times N$ matrix with nonnegative entries. Assume that there exists $n_0\in\mathbb{N}$ s.t. M^{n_0} has strictly positive entries. Then:*

(i) *The spectral radius of M is a real eigenvalue of M, which we denote by λ_*.*

(ii) *The eigenvalues of M other than λ_* have strictly smaller modules than λ_*.*

(iii) *The λ_*-eigenspace is of the form $\mathrm{Vect}(x_*)$, where x_* is a vector in $\mathbb{R}^N$ with strictly positive entries. Moreover, there is only one eigenvalue of M with this property.*

Remark 4.8.9. *If μ and ν are two measures on S, the* distance of total variation *between μ and ν is defined by:*

$$\|\mu-\nu\|_{\mathrm{TV}} = \sup_{A\subseteq S}|\mu(A)-\nu(A)| = \frac{1}{2}\sum_{x\in S}|\mu(x)-\nu(x)|.$$

Therefore, Theorem 4.8.7 can be restated as:

$$\forall x\in S,\ \|\mathbb{P}_x(X_n\in\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \xrightarrow[n\to+\infty]{} 0.$$

4.9 Harmonic functions

Remark 4.9.1. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $f : S \to \mathbb{R}$ be a function. In general, $\left(f(X_n) + \sum_{i=0}^{n-1} (I-Q) f(X_i)\right)_{n\in\mathbb{N}}$ is a martingale under $\mathbb{P}_x$ for all $x \in S$ if $\forall x \in S, Q|f|(x) < +\infty$.*

Definition 4.9.2 (Harmonic functions)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $f : S \to \mathbb{R}$ be a function.*

(i) *We say that f is* harmonic *at $x \in S$ if $Qf(x) = f(x)$. We say that f is harmonic if it is harmonic at every state. In this case, $(f(X_n))_{n\in\mathbb{N}}$ is a martingale under $\mathbb{P}_x$ for all $x \in S$.*

(ii) *We say that f is* superharmonic *(resp.* subharmonic*) at $x \in S$ if $Qf(x) \leq f(x)$ (resp. $Qf(x) \geq f(x)$). We say that f is superharmonic (resp. subharmonic) if it is superharmonic (resp. subharmonic) at every state. In this case, $(f(X_n))_{n\in\mathbb{N}}$ is a supermartingale (resp. submartingale) under $\mathbb{P}_x$ for all $x \in S$.*

Proposition 4.9.3. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $h : S \to \mathbb{R}$ be a function that is harmonic (resp. superharmonic, subharmonic) on some subset $A \subseteq S$. Then $\left(h\left(X_{n\wedge\tau_{S\backslash A}}\right)\right)_{n\in\mathbb{N}}$ is a martingale (resp. supermartingale, submartingale) under $\mathbb{P}_x$ for all $x \in S$, where $\tau_{S\backslash A} =$* $\inf\{n \in \mathbb{N},\ X_n \in S\backslash A\}$.

Proof. Note that:

$$\begin{aligned}
\mathbb{E}_x\left[h\left(X_{(n+1)\wedge\tau_{S\backslash A}}\right) - h\left(X_{n\wedge\tau_{S\backslash A}}\right) \mid \mathcal{F}_n^X\right] &= \mathbb{E}_x\left[\left(h\left(X_{(n+1)\wedge\tau_{S\backslash A}}\right) - h\left(X_{n\wedge\tau_{S\backslash A}}\right)\right)\mathbb{1}_{\{n<\tau_{S\backslash A}\}} \mid \mathcal{F}_n^X\right] \\
&= \mathbb{1}_{\{n<\tau_{S\backslash A}\}}\mathbb{E}_x\left[h(X_{n+1}) - h(X_n) \mid \mathcal{F}_n^X\right] \\
&= \mathbb{1}_{\{n<\tau_{S\backslash A}\}}\left(\mathbb{E}_x\left[h(X_{n+1}) \mid \mathcal{F}_n^X\right] - h(X_n)\right) \\
&= \mathbb{1}_{\{n<\tau_{S\backslash A}\}}\left(\mathbb{E}_{X_n}\left[h(X_1)\right] - h(X_n)\right) \\
&= \mathbb{1}_{\{n<\tau_{S\backslash A}\}}\left(Qh(X_n) - h(X_n)\right).
\end{aligned}$$

□

Remark 4.9.4. *Finding a harmonic function on some subset $A \subseteq S$ amounts to solving a problem of the form:*

$$\begin{cases} Qf - f = -\varphi & \text{on } A \\ f = g & \text{on } S\backslash A \end{cases}, \tag{P}$$

for some specified functions $\varphi : A \to \mathbb{R}$ and $g : S\backslash A \to \mathbb{R}$. This is the discrete analogue of the Poisson problem*:*

$$\begin{cases} \Delta f = -\varphi & \text{on } \mathcal{D} \\ f = g & \text{on } \partial\mathcal{D} \end{cases}.$$

We say that $(Q - I)$ is the discrete Laplacian*.*

Proposition 4.9.5. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Let $A \subseteq S$ be a finite subset. Consider two functions $\varphi : A \to \mathbb{R}$ and $g : S\backslash A \to \mathbb{R}$, with g bounded. Then the discrete Poisson problem (P) has a unique solution given by:*

$$f(x) = \mathbb{E}_x\left[g\left(X_{\tau_{S\backslash A}}\right) + \sum_{i=0}^{\tau_{S\backslash A}-1} \varphi(X_i)\right].$$

Proof. *Uniqueness.* If f is a solution of (P), set $M_n = f\left(X_{n\wedge\tau_{S\setminus A}}\right) + \sum_{i=0}^{(n-1)\wedge\left(\tau_{S\setminus A}-1\right)} (I-Q)\, f(X_i)$. Then $(M_n)_{n\in\mathbb{N}}$ is a martingale, so:

$$f(x) = \mathbb{E}_x[M_0] = \mathbb{E}_x[M_n] = \mathbb{E}_x\left[f\left(X_{n\wedge\tau_{S\setminus A}}\right) + \sum_{i=0}^{(n-1)\wedge\left(\tau_{S\setminus A}-1\right)} \varphi(X_i)\right].$$

Since A is finite, φ is bounded and $\mathbb{E}_x\left[\tau_{S\setminus A}\right] < +\infty$. Moreover, g is bounded, so we can use the Dominated Convergence Theorem and make $n \to +\infty$ in the above equation, which gives the result.
Existence. Define f as above. Then $f = g$ on $S\setminus A$. Now, if $x \in A$, we have:

$$f(x) = \mathbb{E}_x\left[g\left(X_{\tau_{S\setminus A}}\right)\circ\theta_1 + \left(\sum_{i=1}^{\tau_{S\setminus A}-1} \varphi(X_i)\right)\circ\theta_1\right] + \varphi(x) = \varphi(x) + \mathbb{E}_x\left[f(X_1)\right] = \varphi(x) + Qf(x).$$

$\square$

Remark 4.9.6. *In Proposition 4.9.5, the existence of the solution remains valid with the same proof if A is infinite.*

Proposition 4.9.7. *Assume that the Q-Markov chain $(X_n)_{n\in\mathbb{N}}$ is irreducible and recurrent. Then every bounded or nonnegative harmonic function $h : S \to \mathbb{R}$ is constant.*

Proof. Note that $(h(X_n))_{n\in\mathbb{N}}$ is a bounded or nonnegative martingale, so it converges a.s. to a random variable Z (by Theorem 2.4.4). Now, for all $x \in S$, $h(X_n) = h(x)$ for infinitely many values of n; therefore $Z = h(x)$ a.s. Since this is true for all x, h is constant. $\square$

Definition 4.9.8 (Invariant and tail σ-algebras)**.** *If $(X_n)_{n\in\mathbb{N}}$ is a sequence of random variables, we define:*

(i) *The* invariant σ-algebra $\mathcal{J} = \{A \in \sigma(X_n,\, n\in\mathbb{N}),\ \theta_1(A) = A\}$,

(ii) *The* tail σ-algebra $\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(X_n, X_{n+1}, \dots)$.

We have $\mathcal{J} \subseteq \mathcal{T}$.

Theorem 4.9.9. *Let $(X_n)_{n\in\mathbb{N}}$ be the canonical Q-Markov chain. Then the set $\mathscr{R}$ of bounded $\mathcal{J}$-measurable random variables is in bijection with the set $\mathscr{H}$ of bounded harmonic functions, via:*

$$Z \in \mathscr{R} \longmapsto (x \mapsto \mathbb{E}_x[Z]) \in \mathscr{H} \qquad \textit{and} \qquad h \in \mathscr{H} \longmapsto \lim_{n\to+\infty} h(X_n) \in \mathscr{R}.$$

Example 4.9.10. *Using Theorem 3.1.5, we can show that bounded harmonic functions on $\mathbb{Z}^d$ are constant. Therefore, by Theorem 4.9.9, the invariant σ-algebra of the simple random walk on $\mathbb{Z}^d$ is trivial. In particular, it cannot happen that the simple random walk in $\mathbb{Z}^d$ remains in some cone after a certain amount of time.*

4.10 The Poisson process

Definition 4.10.1 (Poisson process)**.** *Let $(\xi_n)_{n\in\mathbb{N}^*}$ be i.i.d. random variables with exponential law of parameter θ, for $\theta \in \mathbb{R}_+^*$. We set $T_n = \sum_{i=1}^n \xi_i$ for $n \in \mathbb{N}$ and:*

$$N_t = \sum_{n\in\mathbb{N}^*} \mathbb{1}_{\{T_n \le t\}}.$$

Then the collection $(N_t)_{t\in\mathbb{R}_+}$ is called the Poisson process *with intensity θ.*

Theorem 4.10.2. *Consider the Poisson process $(N_t)_{t\in\mathbb{R}_+}$ with intensity $\theta \in \mathbb{R}_+^*$. Then:*

(i) *For $t \in \mathbb{R}_+^*$, N_t has a Poisson law of parameter θt.*

(ii) *For $s, t \in \mathbb{R}_+^*$, the variables $(N_{t+s} - N_t)$ and N_t have the same law and are independent.*

References

[1] R. Durrett. *Probability: Theory and Examples.*

[2] O. Kallenberg. *Foundations of Modern Probability.*

[3] J. Neveu. *Martingales à temps discret.*

[4] D. Williams. *Probability with Martingales.*