4 Martingales

4.1 Definition and some examples

A martingale is a generalized version of a fair game.

Definition 4.1. A process $(M_n)_{n\ge0}$ is a *martingale* if

- for every $n \ge 0$ the expectation $\mathsf{E}M_n$ is finite, equivalently, $\mathsf{E}|M_n| < \infty$;
- for every $n \ge 0$ and all m_n, m_{n-1}, ..., m_0 we have

$$\mathsf{E}\big(M_{n+1} \mid M_n = m_n, \dots, M_0 = m_0\big) = m_n\,. \tag{4.1}$$

Of course, (4.1) can just be written as $\mathsf{E}\big(M_{n+1} \mid M_n, \dots, M_0\big) = M_n$.

Definition 4.2. We say that $(M_n)_{n\ge0}$ is a *supermartingale*[32] if the equality in (4.1) holds with $\le$, ie., $\mathsf{E}\big(M_{n+1} \mid M_n, \dots, M_0\big) \le M_n$; and we say that $(M_n)_{n\ge0}$ is a *submartingale*, if (4.1) holds with $\ge$, ie., $\mathsf{E}\big(M_{n+1} \mid M_n, \dots, M_0\big) \ge M_n$.

Definition 4.3. A process $(M_n)_{n\ge0}$ is a *martingale* w.r.t. a sequence $(X_n)_{n\ge0}$ of random variables, if

- for every $n \ge 0$ the expectation $\mathsf{E}M_n$ is finite, equivalently, $\mathsf{E}|M_n| < \infty$;
- for every $n \ge 0$ and all x_n, x_{n-1}, ..., x_0 we have

$$\begin{aligned}&\mathsf{E}\big(M_{n+1} - M_n \mid X_n = x_n, \dots, X_0 = x_0\big)\\ &\qquad\equiv \mathsf{E}\big(M_{n+1} - M_n \mid X_n, \dots, X_0\big) = 0\,.\end{aligned} \tag{4.2}$$

Example 4.4. *Let $(\xi_n)_{n\ge1}$ be independent random variables*[33] *with $\mathsf{E}\xi_n = 0$ for all $n \ge 1$. Then the process $(M_n)_{n\ge0}$ defined via $M_n \stackrel{\mathsf{def}}{=} M_0 + \xi_1 + \dots + \xi_n$ is a martingale as long as the random variable M_0 is independent of $(\xi_n)_{n\ge1}$ and $\mathsf{E}|M_0| < \infty$.*

Solution. Indeed by the triangle inequality,

$$\mathsf{E}|M_n| \le \mathsf{E}|M_0| + \sum_{j=1}^{n} \mathsf{E}|\xi_j| < \infty \qquad \text{for all } n \ge 0,$$

whereas the independence property implies

$$\mathsf{E}\big(M_{n+1} - M_n \mid M_n, \dots, M_0\big) \equiv \mathsf{E}\big(\xi_{n+1} \mid M_n, \dots, M_0\big) = \mathsf{E}\xi_{n+1} = 0\,. \qquad \square$$

Remark 4.4.1. Notice that if $\mathsf{E}\xi_n \ge 0$ for all $n \ge 1$, then $(M_n)_{n\ge0}$ is a submartingale, whereas if $\mathsf{E}\xi_n \le 0$ for all $n \ge 1$, then $(M_n)_{n\ge0}$ is a supermartingale. More generally, if $(\xi_n)_{n\ge1}$ are independent random variables with $\mathsf{E}|\xi_n| < \infty$ for all $n \ge 1$, then the process $M_n = M_0 + (\xi_1 - \mathsf{E}\xi_1) + \dots + (\xi_n - \mathsf{E}\xi_n)$, $n \ge 0$, is a martingale.

[32] If M_n traces your fortune, then "there is nothing super about a supermartingale".
[33] Notice that we do not assume that all ξ_n have the same distribution!

Example 4.5. *If $(\xi_n)_{n\geq 1}$ are independent random variables with $\mathsf{E}\xi_n = 0$ and $\mathsf{E}(\xi_n)^2 < \infty$ for all $n \geq 1$, then the process $(T_n)_{n\geq 0}$ defined via $T_n = (M_n)^2$, where $M_n \stackrel{\mathsf{def}}{=} M_0 + \xi_1 + \cdots + \xi_n$, is a submartingale w.r.t. $(M_n)_{n\geq 0}$.*

Solution. By the Cauchy inequality $\bigl(\sum_{j=1}^n a_j\bigr)^2 \leq n\sum_{j=1}^n (a_j)^2$, we obviously have $\mathsf{E}T_n < \infty$ for all $n \geq 0$. Now, because $T_{n+1} - T_n = 2M_n\xi_{n+1} + (\xi_{n+1})^2$ and

$$\mathsf{E}\bigl(2M_n\xi_{n+1} \mid M_n, \dots, M_0\bigr) = 2M_n\mathsf{E}(\xi_{n+1}) = 0\,,$$
$$\mathsf{E}\bigl((\xi_{n+1})^2 \mid M_n, \dots, M_0\bigr) = \mathsf{E}\bigl((\xi_{n+1})^2\bigr) \geq 0\,,$$

we get $\mathsf{E}\bigl(T_{n+1} - T_n \mid M_n, \dots, M_0\bigr) \geq 0$, ie., the process $(T_n)_{n\geq 0}$ is a submartingale. □

Example 4.6. *Let $(Z_n)_{n\geq 0}$ be a branching process with $\mathsf{E}Z_1 = m < \infty$. We have $\mathsf{E}|Z_n| < \infty$ and $\mathsf{E}(Z_{n+1} \mid Z_n, \dots, Z_0) = mZ_n$ for all $n \geq 0$. In other words, the process $(Z_n)_{n\geq 0}$ is a martingale, a submartingale or a supermartingale depending on whether $m = 1$, $m > 1$ or $m < 1$.*
Notice also, that for every $m \in (0, \infty)$ the process $(Z_n/m^n)_{n\geq 0}$ is a martingale.

Exercise 4.7. Let ρ be the extinction probability for a branching process $(Z_n)_{n\geq 0}$; show that ρ^{Z_n} is a martingale w.r.t. $(Z_n)_{n\geq 0}$.

☞ **Example 4.8.** *Let $(X_n)_{n\geq 0}$ be a martingale. If, for some* convex *function $f(\,\cdot\,)$ we have $\mathsf{E}\bigl(|f(X_n)|\bigr) < \infty$ for all $n \geq 0$, then the process $\bigl(f(X_n)\bigr)_{n\geq 0}$ is a submartingale. Similarly, if for some* concave *$f(\,\cdot\,)$ we have $\mathsf{E}\bigl(|f(X_n)|\bigr) < \infty$ for all $n \geq 0$, then the process $\bigl(f(X_n)\bigr)_{n\geq 0}$ is a supermartingale.*

☞ **Exercise 4.9.** Let $(\eta_n)_{n\geq 0}$ be independent positive random variables with $\mathsf{E}\eta_n = 1$ for all $n \geq 0$. If a random variable $M_0 > 0$ is independent of $(\eta_n)_{n\geq 0}$ and $\mathsf{E}M_0 < \infty$, then the process $(M_n)_{n\geq 0}$ defined via $M_n = M_0\prod_{j=1}^n \eta_j$, is a martingale.

Remark 4.9.1. Interpreting $\eta_n - 1$ as the (fractional) change in the value of a stock during the nth time interval, the martingale $(M_n)_{n\geq 0}$ can be used to model stock prices. Two often used examples are:

Discrete Black-Sholes model: take $\eta_j = e^{\zeta_j}$, where ζ_j is Gaussian, $\zeta_j \sim \mathcal{N}(\mu, \sigma^2)$;

Binomial model:take $\eta_j = (1 + a)e^{-r}$ and $\eta_j = (1 + a)^{-1}e^{-r}$ with probabilities p and $1 - p$ respectively.

Exercise 4.10. Let $(X_n)_{n\geq 0}$ be a Markov chain with a (countable) state space $\mathcal{S}$ and the transition matrix $\mathbf{P}$. If $\boldsymbol{\psi}$ is a right eigenvector of $\mathbf{P}$ corresponding to the eigenvalue $\lambda > 0$, ie., $\mathbf{P}\boldsymbol{\psi} = \lambda\boldsymbol{\psi}$, show that the process $\lambda^{-n}\psi\bigl(X_n\bigr)$ is a martingale w.r.t. $(X_n)_{n\geq 0}$.

☞ **Exercise 4.11.** Let $(X_n)_{n\geq 0}$ be a Markov chain with a (countable) state space $\mathcal{S}$ and the transition matrix $\mathbf{P}$, and let $h(x, n)$ be a function of the state x and time n such that[34]

$$h(x, n) = \sum_y p_{xy}h(y, n + 1)\,.$$

Show that $(M_n)_{n\geq 0}$ with $M_n = h(X_n, n)$ is a martingale w.r.t. $(X_n)_{n\geq 0}$.

[34] This result is useful, eg., if $h(x, n) = x^2 - cn$ or $h(x, n) = \exp\{x - cn\}$ for a suitable c.

4.2 A few remarks on conditional expectation

Recall the following basic definition:

Example 4.12. *Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability triple and let X and Y be random variables taking values x_1, x_2, ..., x_m and y_1, y_2, ..., y_n respectively. On the event $\{Y = y_j\}$ one defines the conditional probability*

$$\mathsf{P}\big(X = x_i \mid Y = y_j\big) \stackrel{\mathsf{def}}{=} \frac{\mathsf{P}(X = x_i, Y = y_j)}{\mathsf{P}(Y = y_j)}$$

and the conditional expectation: $\mathsf{E}\big(X \mid Y = y_j\big) \equiv \sum_i x_i \mathsf{P}\big(X = x_i \mid Y = y_j\big)$. *Then the* conditional expectation $Z = \mathsf{E}(X \mid Y)$ *of X given Y is defined as follows:*

$$\text{if } Y(\omega) = y_j, \qquad \text{then} \qquad Z(\omega) \stackrel{\mathsf{def}}{=} z_j \equiv \mathsf{E}\big(X \mid Y = y_j\big)\,.$$

Notice that the value z_j is completely determined by y_j; in other words, Z is a function of Y, and as such, a random variable.

Definition 4.13. If $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$ is a finite partition[35] of Ω, the collection $\mathcal{G} = \sigma(\mathcal{D})$ of all 2^m possible subsets of Ω constructed from blocks D_i is called the *σ-field generated by $\mathcal{D}$*. A random variable Y is *measurable* with respect to $\mathcal{G}$ if it is constant on *every* block D_i of the initial partition $\mathcal{D}$.

In Example 4.12 above we see that Z is constant on the events $\{Y = y_j\}$ (ie., Z is measurable w.r.t. the σ-field $\mathcal{G} \stackrel{\mathsf{def}}{=} \sigma(Y) \equiv \sigma(\{Y = y_j\})$); moreover, for every $G_j \equiv \{Y = y_j\}$, we have

$$\begin{aligned}\mathsf{E}\big(Z\,\mathbb{1}_{G_j}\big) = z_j \mathsf{P}(Y = y_j) &= \sum_i x_i \mathsf{P}\big(X = x_i \mid Y = y_j\big)\mathsf{P}(Y = y_j)\\ &= \sum_i x_i \mathsf{P}\big(X = x_i, Y = y_j\big) = \mathsf{E}\big(X\mathbb{1}_{G_j}\big)\,,\end{aligned}$$

where the last equality follows from the observation that the random variable $X\mathbb{1}_{G_j}$ equals x_i on every event $G_j \cap \{X = x_i\} = \{Y = y_j\} \cap \{X = x_i\}$ and vanishes identically outside the set G_j.

Remark 4.13.1. The last two properties can be used to define conditional expectation in general: Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability triple, let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-field, and let X be a rv, $X : \Omega \to \mathbb{R}$. The *conditional expectation* of X w.r.t. $\mathcal{G}$ is a random variable Z such that: Z is $\mathcal{G}$ measurable and for every set $G \in \mathcal{G}$ we have $\int_G Z \mathsf{dP} \equiv \mathsf{E}\big(Z\mathbb{1}_G\big) = \mathsf{E}\big(X\mathbb{1}_G\big) \equiv \int_G X \mathsf{dP}$.

[35] In the general countable setting, if $\mathcal{D} = \{D_1, D_2, \dots\}$ forms a denumerable (ie., infinite countable) partition of Ω, then the generated σ-field $\mathcal{G} = \sigma(\mathcal{D})$ consists of all possible subsets of Ω which are obtained by taking countable unions of blocks of $\mathcal{D}$. Similarly, a variable Y is measurable w.r.t. $\mathcal{G}$, if for every real y the event $\{\omega : Y(\omega) \le y\}$ belongs to $\mathcal{G}$ (equivalently, can be expressed as a countable union of blocks of $\mathcal{D}$.

Notice that the last remark suggests the following interpretation of the conditional expectation w.r.t. a partition $\mathcal{D}$ of the probability space Ω: if X is any[36] random variable, then $\mathsf{E}(X \mid \sigma(\mathcal{D}))$ is a new random variable taking constant values on all blocks of $\mathcal{D}$, and such that on every block $D \in \mathcal{D}$ it has the same integral (volume below its profile) as the original variable X.

The following are the most important properties of conditional expectation:

✍ **Lemma 4.14.** *Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability triple, and let $\mathcal{G} \subset \mathcal{F}$ a σ-field. Then for all random variables X, X_1, X_2 and constants a_1, a_2, the following properties hold:*

a) If $Z = \mathsf{E}\big(X \mid \mathcal{G}\big)$ then $\mathsf{E}Z = \mathsf{E}X$;

b) If X is $\mathcal{G}$-measurable, then $\mathsf{E}\big(X \mid \mathcal{G}\big) = X$;

c) $\mathsf{E}\big(a_1X_1 + a_2X_2 \mid \mathcal{G}\big) = a_1\mathsf{E}\big(X_1 \mid \mathcal{G}\big) + a_2\mathsf{E}\big(X_2 \mid \mathcal{G}\big)$;

d) If $X \ge 0$, then $\mathsf{E}\big(X \mid \mathcal{G}\big) \ge 0$;

e) If $\mathcal{H}$ and $\mathcal{G}$ are two σ-fields in Ω such that $\mathcal{H} \subseteq \mathcal{G}$, then

$$\mathsf{E}\Big[\mathsf{E}\big(X \mid \mathcal{G}\big) \mid \mathcal{H}\Big] = \mathsf{E}\Big[\mathsf{E}\big(X \mid \mathcal{H}\big) \mid \mathcal{G}\Big] = \mathsf{E}\big(X \mid \mathcal{H}\big)\,;$$

f) If Z is $\mathcal{G}$-measurable, then $\mathsf{E}\big[ZX \mid \mathcal{G}\big] = Z\,\mathsf{E}(X \mid \mathcal{G})$.

If $(X_k)_{k\ge1}$ is a sequence of random variables, we can define the generated σ-fields $\mathcal{F}_1^X$, $\mathcal{F}_2^X$, ..., via

$$\mathcal{F}_k^X \stackrel{\mathsf{def}}{=} \sigma(X_1, X_2, \dots, X_k)\,; \tag{4.3}$$

here, the σ-field $\mathcal{F}_k^X$ stores all information about the process $(X_n)_{n\ge1}$ up to time k. Observe that these σ-fields form a filtration $\big(\mathcal{F}_n^X\big)_{n\ge1}$ in the sense that

$$\mathcal{F}_1^X \subseteq \mathcal{F}_2^X \subseteq \ldots \subseteq \mathcal{F}_k^X \subseteq \ldots\,. \tag{4.4}$$

☞ The notion of filtration is very useful when working with martingales. Indeed, a process $(M_n)_{n\ge0}$ is a martingale w.r.t. $(X_n)_{n\ge0}$ if for all $n \ge 0$ the variable M_n is $\mathcal{F}_n^X$-measurable, with $\mathsf{E}|M_n| < \infty$ and $\mathsf{E}\big(M_{n+1} \mid \mathcal{F}_n^X\big) = M_n$. Here, one says
☞ that M_n is $\mathcal{F}_n^X$-measurable if for every real y, the event $\{M_n \le y\}$ belongs to $\mathcal{F}_n^X$, recall footnote 35 on page 27.

We also notice that by repeatedly applying the tower property in claim e) of Lemma 4.14 above to the sequence (4.4), we get the following result:

✍ **Lemma 4.15.** *If $(M_n)_{n\ge0}$ is a submartingale w.r.t. $(X_n)_{n\ge0}$, then for all integer $n \ge k \ge 0$, we have $\mathsf{E}M_n \ge \mathsf{E}M_k$.*

☞ **Remark 4.15.1.** Changing signs, we get the inequality $\mathsf{E}M_n \le \mathsf{E}M_k$ for supermartingales, and therefore the equality $\mathsf{E}M_n = \mathsf{E}M_k$ for martingales.

[36] If X is not a discrete variable with finite number of values, one has to additionally assume that $\mathsf{E}(|X|) < \infty$.

4.3 Martingales and stopping times

Martingales are extremely useful in studying stopping times:

Example 4.16. *Let $(X_k)_{k\geq 1}$ be a sequence of i.i.d. Bernoulli random variables with common distribution $\mathsf{P}(X_1 = 1) = p$, $\mathsf{P}(X_1 = -1) = q = 1 - p$, where $p \in (0, 1)$, $p \neq 1/2$. For fixed $k \in \{0, 1, \dots, N\}$, define the random walk $(S_n)_{n\geq 0}$ defined via $S_n = k + \sum_{j=1}^{n} X_j$, $n \geq 0$, and consider the process $(Y_n)_{n\geq 0}$, where $Y_n \stackrel{\mathsf{def}}{=} \bigl(q/p\bigr)^{S_n}$. It is straightforward to check that Y_n is an $\mathcal{F}$-martingale (where $\mathcal{F}_n \equiv \mathcal{F}_n^X$). In particular, the martingale property implies that $\mathsf{E}(Y_n) = \mathsf{E}(Y_0) = (q/p)^k$ for all $n \geq 0$, recall Lemma 4.15.*

Let T be the absorption time for S_n, ie., the hitting time of $\{0, N\}$. If an analogue of the above equality holds at (random) time T, we could find the probability $\mathsf{p}_k = \mathsf{P}(\,S$ absorbed at $0\ |S_0 = k)$ from

$$\mathsf{E}\bigl(Y_T\bigr) = (q/p)^0 \mathsf{p}_k + (q/p)^N \bigl(1 - \mathsf{p}_k\bigr)\,,$$

and thus obtaining (as long as $p \neq q$) $\mathsf{p}_k = \bigl((q/p)^k - (q/p)^N\bigr)/\bigl(1 - (q/p)^N\bigr)$.

Remark 4.16.1. The method used in the previous example is very attractive, and relies on the assumption $\mathsf{E}\bigl(Y_T\bigr) = \mathsf{E}\bigl(Y_0\bigr)$ and on the formula for $\mathsf{E}\bigl(Y_T\bigr)$ for a certain random variable T. An important part of the theory of martingales is to determine conditions on such random variables T which ensure that the above statements are true.[37]

Example 4.17. *We now apply the idea from the previous example to the simple symmetric random walk on $\{-K, \dots, K\}$, ie., the sequence of partial sums $S_n = \sum_{k=1}^{n} X_k$, generated by a sequence of i.i.d. Bernoulli r.v. X_k with $\mathsf{P}(X_1 = \pm 1) = 1/2$. Indeed, if T is the stopping time*

$$T = \inf\bigl\{n \geq 0 : |S_n| = K\bigr\}\,,$$

then[38]

$$\mathsf{E}\bigl((S_T)^2\bigr) = K^2 \mathsf{P}(S_T = K) + K^2 \mathsf{P}(S_T = -K) = K^2\,,$$

and the same heuristics applied to the martingale $(Y_n)_{n\geq 0}$, $Y_n \stackrel{\mathsf{def}}{=} \bigl(S_n\bigr)^2 - n$, results in $0 = \mathsf{E}(Y_0) = \mathsf{E}(Y_T) = \mathsf{E}\bigl((S_T)^2\bigr) - \mathsf{E}(T)$; equivalently, it suggests that

$$\mathsf{E}(T) = K^2\,.$$

One of our aims is to discuss general results that justify the above heuristics. To this end, we need to carefully define what we mean by a "stopping time".

Definition 4.18. A variable T is a stopping time for a process $(X_n)_{n\geq 0}$, if the occurrence/non-occurrence of the event *"we stop at time n"* can be determined by looking at the values X_0, X_1, ..., X_n of the process up to time n. Equivalently, for every $n \geq 0$, we have $\{T \leq n\} \in \mathcal{F}_n^X \equiv \sigma(X_0, \dots, X_n)$.

[37] Notice that the gambler's ruin problem can be solved by using the methods of finite Markov chains, so we indeed know that the result above is correct.

[38] the equality is correct, because $\mathsf{P}(T < \infty) = 1$ here!

Example 4.19. *Let $(X_n)_{n\geq0}$ be a stochastic process with values in S, and let T be the hitting time of a set $A \subset S$, namely, $T = \min\{n \geq 0 : X_n \in A\}$. Then T is a stopping time for $(X_n)_{n\geq0}$.*

Solution. For every fixed $n \geq 0$, we have $\{T > n\} \equiv \{X_0 \notin A, X_1 \notin A, \dots, X_n \notin A\}$; therefore, the event $\{T > n\}$ and its complement $\{T \leq n\}$ both belong to $\mathcal{F}_n^X$. □

☞ **Exercise 4.20.** Let $k \in \mathbb{N}$ be fixed, and let S and T be stopping times for a process $(X_n)_{n\geq0}$. Show that the following are stopping times:

a) $T \equiv k$, b) $S \wedge T \equiv \min(S,T)$, c) $S \vee T \equiv \max(S,T)$.

d) Let A be a set of states, and let $T = T_A$ be the moment of the first visit to A, ie., $T = \min\{n \geq 0 : X_n \in A\}$. Consider $S = S_A = \min\{n > T_A : X_n \in A\}$, the moment of the second visit to A. Show that S_A is a stopping time for $(X_n)_{n\geq0}$.
e) Is the variable $L = \max\{n \geq 0 : X_n \in A\}$ a stopping time for $(X_n)_{n\geq0}$?

Exercise 4.21. Let $(M_n)_{n\geq0}$ be a submartingale w.r.t. $(X_n)_{n\geq0}$, and let T be a stopping time for $(X_n)_{n\geq0}$. Show carefully that the process $(L_n^T)_{n\geq0}$,

$$L_n^T \stackrel{\text{def}}{=} M_{n\wedge T} \equiv \begin{cases} M_n\,, & n \leq T\,, \\ M_T\,, & n > T\,, \end{cases}$$

is a submartingale w.r.t. $(X_n)_{n\geq0}$. Deduce that if $(M_n)_{n\geq0}$ is a martingale w.r.t. $(X_n)_{n\geq0}$, then the stopped process $(L_n^T)_{n\geq0}$ is also a martingale w.r.t. $(X_n)_{n\geq0}$.

4.4 Optional stopping theorem

The optional stopping (or sampling) theorem (OST) tells us that, under quite general assumptions, whenever X_n is a martingale, then $X_{T\wedge n}$ is a martingale for a stopping time T. Such results are very useful in deriving inequalities and probabilities of various events associated with such stochastic processes.

✍ **Theorem 4.22** (Optional Stopping Theorem). *Let $(M_n)_{n\geq0}$ be a martingale w.r.t. $(X_n)_{n\geq0}$, and let T be a stopping time for $(X_n)_{n\geq0}$. Then the equality*

$$\mathsf{E}\big[M_T\big] = \mathsf{E}[M_0] \tag{4.5}$$

holds whenever **one** *of the following conditions holds:*

(OST-1) *the stopping time T is bounded;*

(OST-2) *the martingale $(M_n)_{n\geq0}$ has bounded increments, $|M_{n+1} - M_n| \leq K$ for all n and some constant K, whereas $\mathsf{E}T < \infty$;*

(OST-3) *the martingale $(M_n)_{n\geq0}$ is bounded, $|M_n| \leq K$ for all n and some constant K, whereas T is a finite stopping time, $\mathsf{P}(T < \infty) = 1$.*

Remark 4.22.1. If M_n records gambler's fortune, by (OST-3), one cannot make money from a fair game, unless an unlimited amount of credit is available.

☞ **Remark 4.22.2.** One can generalize (OST-3) by replacing the condition of boundedness, $|M_n| \le K$, by that of *uniform integrability* for the martingale $(M_n)_{n\ge0}$: a sequence of random variables $(Y_n)_{n\ge0}$ is *uniformly integrable* if

$$\lim_{K\to\infty} \sup_{n\ge0} \mathsf{E}\big(|Y_n| \mathbb{1}_{\{|Y_n|>K\}}\big) = 0 . \tag{4.6}$$

Example 4.23. *Let the SSRW $(S_n)_{n\ge0}$ be generated as in Example 4.17. Put*

$$H \stackrel{\mathsf{def}}{=} \inf\{n \ge 0 : S_n = 1\} .$$

Since this RW is recurrent,[39] *we deduce that $\mathsf{P}(H < \infty) = 1$. However, the* ☞ (OST) *does not apply, as $0 = \mathsf{E}(S_0) \ne \mathsf{E}(S_H) \equiv 1$. It is an instructive* Exercise *to check which conditions in each of the above* (OST) *are violated.*

We now sketch the proof of Theorem 4.22.

☞ **Exercise 4.24.** Let T in (OST-1) satisfy $0 \le T \le N$. Decompose

$$M_T = \sum_{n=0}^{N} M_T \mathbb{1}_{\{T=n\}} = \sum_{n=0}^{N} M_n \mathbb{1}_{\{T=n\}} = M_0 + \sum_{n=0}^{N-1} \big(M_{n+1} - M_n\big) \mathbb{1}_{\{T>n\}}$$

and use the properties of conditional expectations in Lemma 4.14 to derive (4.5) under (OST-1).
[**Hint:** Check that $\mathsf{E}\big((M_{n+1} - M_n)\mathbb{1}_{\{T>n\}} \,|\, \mathcal{F}_n^X\big) = \mathbb{1}_{\{T>n\}} \mathsf{E}\big(M_{n+1} - M_n \,|\, \mathcal{F}_n^X\big) = 0$.]

☞ **Exercise 4.25.** Derive (OST-2) from (OST-1) in the following way: first, notice that $T \wedge n$ is a bounded stopping time, so that (OST-1) implies $\mathsf{E}M_{T\wedge n} = \mathsf{E}M_0$ for all $n \ge 0$. Next, using the decomposition from Exercise 4.24, rewrite

$$M_T - M_{T\wedge n} = \sum_{k>n} M_T \mathbb{1}_{\{T=k\}} = M_{n+1} \mathbb{1}_{\{T\ge n+1\}} + \sum_{k>n+1} \big(M_k - M_{k-1}\big) \mathbb{1}_{\{T\ge k\}}$$

and deduce (OST-2) from the estimate

$$\big|\mathsf{E}M_T - \mathsf{E}M_0\big| \le \big|\mathsf{E}M_0\big| \mathsf{P}(T \ge n+1) + K \sum_{k>n+1} \mathsf{P}(T \ge k) \to 0 \qquad \text{as } n \to \infty .$$

Remark 4.25.1. Notice that the previous argument can be generalized to the case when $\big|\mathsf{E}\big(M_k - M_{k-1} | \mathcal{F}_{k-1}^X\big)\big| \le K$ for all $k \ge 1$; indeed, one then has

$$\Big|\mathsf{E}\Big(\big(M_k - M_{k-1}\big)\mathbb{1}_{\{T\ge k\}}\Big)\Big| \le \mathsf{E}\big(\mathbb{1}_{\{T\ge k\}} \big|\mathsf{E}\big(M_k - M_{k-1} | \mathcal{F}_{k-1}^X\big)\big|\big) ,$$

where the last expression is bounded above by $K\mathsf{P}(T \ge k)$.

☞ **Exercise 4.26.** Derive (OST-3) from (OST-1) in the following way: first, notice that $T \wedge n$ is a bounded stopping time, so that (OST-1) implies $\mathsf{E}M_{T\wedge n} = \mathsf{E}M_0$ for all $n \ge 0$. Then deduce the result from

$$\big|\mathsf{E}M_T - \mathsf{E}M_0\big| = \big|\mathsf{E}M_T - \mathsf{E}M_{T\wedge n}\big| \le 2K\mathsf{P}(T > n) \to 0 \qquad \text{as } n \to \infty .$$

[39] alternatively, recall the result in Example 1.14.

Remark 4.26.1. In (OST-3) it is sufficient to assume that $\big|M_{n\wedge T}\big| \le K$ for all $n \ge 0$ and $\mathsf{P}(T < \infty) = 1$. Indeed, by (DOM) we deduce that then $\big|M_T\big| \le K$, so that the argument in Exercise 4.26 applies.

With suitable martingales (OST) gives very powerful results:

✍ **Example 4.27.** *In the setup of Example 4.17, put*

$$U_n \equiv \exp\bigl\{-\alpha S_n - \beta n\bigr\}$$

with some α, $\beta \ge 0$. Fix real α and find $\beta \ge 0$ so that $(U_n)_{n\ge 0}$ is a martingale. Use this martingale to study the stopping time $T = \inf\{n \ge 0 : |S_n| = K\}$.

Solution. By a straightforward computation,

$$\mathsf{E}\bigl(U_{n+1} \mid \mathcal{F}_n\bigr) = e^{-\beta}\,\frac{e^{\alpha} + e^{-\alpha}}{2}\,U_n\,,$$

so that U_n is a martingale iff (**check** the remaining conditions!) $e^{\beta} = \cosh(\alpha)$, equivalently, iff $\beta = \log\cosh(\alpha) \ge 0$. Next, observe[40] that $\mathsf{E}(T) < \infty$. Notice however that despite S_n has bounded increments ($|S_{n+1} - S_n| \le 1$), the increments of U_n are **not** uniformly bounded. We overcome this difficulty by considering the stopped[41] process $U_{n\wedge T}$. Since $n\wedge T$ is bounded, any of the (OST) above implies that $1 = \mathsf{E}(U_0) = \mathsf{E}(U_{n\wedge T})$; taking $n \to \infty$ we deduce[42]

$$\mathsf{E}(U_{n\wedge T}) \to \mathsf{E}(U_T) \equiv \Bigl(\frac{e^{\alpha K}}{2} + \frac{e^{-\alpha K}}{2}\Bigr)\,\mathsf{E}e^{-\beta T}\,, \qquad \text{as } n \to \infty\,,$$

so that the Laplace transform $\mathsf{E}e^{-\beta T}$ of T satisfis

$$\mathsf{E}e^{-\beta T} = \bigl(\cosh(\alpha K)\bigr)^{-1}\,, \qquad \text{where} \qquad \beta = \log\cosh(\alpha)\,.$$

Notice that the Laplace transform $\mathsf{E}e^{-\beta T}$ is essentially the moment generating function of the stopping time T, and as such completely describes its distribution. In particular, (right) differentiating $\mathsf{E}e^{-\beta T}$ w.r.t. β at $\beta = 0$ one can compute $\mathsf{E}T^m$ for all $m \ge 1$. □

The argument in footnote 40 can be easily generalized:

---

[40] To this end, notice that $2K$ consecutive jumps in the same direction force S_n to exit the segment $[-K, K]$ irrespectively of the initial state; therefore, $\mathsf{P}(T > 2K + m \,|\, T > m) \le \mathsf{P}(|S_{2K+m}| < K \,|\, |S_m| < K) \le 1 - 2^{-2K}$, so that by induction, $\mathsf{P}(T > 2Kn) \le r^n$ for $n \ge 0$, where $r = 1 - 2^{-2K} < 1$. We finally observe that

$$\mathsf{E}T \equiv \sum_{n\ge 0} \mathsf{P}(T > n) \le 2K \sum_{n\ge 0} \mathsf{P}(T > 2Kn) \le \frac{2K}{1-r} < \infty\,. \tag{4.7}$$

Alternatively, notice that $(S_{(2Km)\wedge T})_{m\ge 0}$ is a finite Markov chain in $\{-K, \dots, K\}$ with uniformly positive transition probabilities, and deduce the bound $\mathsf{E}(T) < \infty$ in a usual way.

[41] ie., we assume that the boundary $\{-K, K\}$ is absorbing and the process $(S_n)_{n\ge 0}$ (and thus $(U_n)_{n\ge 0}$) does not move after hitting it;

[42] think (DOM): for all $n \ge 0$ we have $|S_{n\wedge T}| \le K$ so that $|U_{n\wedge T}| \le e^{|\alpha|K - \beta(n\wedge T)} \le e^{|\alpha|K}$ with integrable upper bound, and since $\mathsf{P}(T < \infty) = 1$, we have $U_{n\wedge T} \to U_T$ as $n \to \infty$.

✍ **Lemma 4.28.** *Let $\big(\mathcal{F}_n^X\big)_{n\ge0}$ be the filtration*[43] *generated by a process $(X_n)_{n\ge0}$, and let T be a stopping time for $(X_n)_{n\ge0}$. Assume that there exist positive ε and K such that for all $n\ge0$ and all events $A\in\mathcal{F}_n^X$,*

$$\mathsf{P}\big(T\le n+K\,|\,A\big)\ge\varepsilon\,. \tag{4.8}$$

Then $\mathsf{E}(T)<\infty$.

Remark 4.28.1. In other words, 'whatever always stands a reasonable chance of happening, will almost surely happen — sooner rather than later'.

Proof. A straightforward induction based upon (4.8) implies $\mathsf{P}(T>mK)\le(1-\varepsilon)^m$ for all $m\ge0$. The result now follows as in (4.7). □

Exercise 4.29. Use an appropriate (OST) to carefully derive the probability p_k in Example 4.16.

Exercise 4.30. Use an appropriate (OST) to carefully derive the expectation $\mathsf{E}(T)$ in Example 4.17.
[**Hint:** Notice that the martingale $(Y_n)_{n\ge0}$ has unbounded jumps, so that none of the (OST) applies without truncation.]

Exercise 4.31. Consider the simple symmetric random walk $(S_n)_{n\ge0}$, generated by a sequence of i.i.d. Bernoulli r.v. X_k with $\mathsf{P}(X_1=\pm1)=1/2$, ie., $S_n=\sum_{k=1}^n X_k$. For integer $a<0<b$, let T be the stopping time $T=\inf\big\{n\ge0: S_n\notin(a,b)\big\}$.

a) Show that $(S_n)_{n\ge0}$ is a martingale and use an appropriate (OST) to find $\mathsf{P}(S_T=a)$ and $\mathsf{P}(S_T=b)$.

b) Show that $(M_n)_{n\ge0}$ defined via $M_n=(S_n)^2-n$ is a martingale w.r.t. the process $(S_n)_{n\ge0}$.

c) For fixed integer $K>0$, carefully apply an appropriate (OST) to M_n and prove that $\mathsf{E}(T\wedge K)=\mathsf{E}\big[(S_{T\wedge K})^2\big]$.

d) Deduce that $\mathsf{E}(T)=-ab$.

☞ **Example 4.32.** (ABRACADABRA)*A monkey types symbols at random, one per unit time, on a typewriter having 26 keys. How long on average will it take him to produce the sequence 'ABRACADABRA'?*

Solution. To compute the expected time, imagine a sequence of gamblers, each initially having £1, playing at a fair gambling casino. Gambler arriving just before time n ($n\ge1$) bets £1 on the event $\{n\text{th letter will be A}\}$. If he loses, he leaves. If he wins, he receives £26 all of which he bets on the event $\{n+1\,\text{st letter will be B}\}$. If he loses, he leaves. If he wins, he receives £26^2 all of which he bets on the event $\{n+2\,\text{nd letter will be R}\}$ and so on through the whole ABRACADABRA sequence.

Let X_n denote the total winnings of the casino after the nth day. Since all bets are fair the process $(X_n)_{n\ge0}$ is a martingale with mean zero. Let N denote the time

[43] ie., $\mathcal{F}_n^X\equiv\sigma\big(X_0,\dots,X_n\big)$, recall the definition in (4.4) above

until the sequence ABRACADABRA appears. At time N, gambler $N-11$ would have won £$26^{11}-1$, gambler $N-4$ would have won £26^4-1, gambler N would have won £$26-1$ and all other $N-3$ gamblers would have lost their initial fortune. Therefore,

$$X_N = N - 3 - (26^{11}-1) - (26^4-1) - (26-1) = N - 26^{11} - 26^4 - 26$$

and since (OST-2) can be applied (check this!), we deduce the $\mathsf{E}(X_N) = \mathsf{E}(X_0) = 0$, that is

$$\mathsf{E}(N) = 26 + 26^4 + 26^{11}\,. \qquad \square$$

Remark 4.32.1. Notice that the answer could also be obtained by considering a finite state Markov chain X_n on the state space of 12 strings representing the longest possible intersection of the tail of the typed text with the target word ABRACADABRA, ie., $\{\mathsf{ABRACADABRA}, \mathsf{ABRACADABR}, \dots, \mathsf{ABRA}, \mathsf{ABR}, \mathsf{AB}, \mathsf{A}, \varnothing\}$, as there N is just the hitting time of the state ABRACADABRA from the initial condition $X_0 = \varnothing$.

Exercise 4.33. A coin showing heads with probability p is tossed repeatedly. Let w be a fixed sequence of outcomes such as 'HTH', and let N denote the number of (independent) tosses until the word w is observed. Using an appropriate martingale, find the expectation $\mathsf{E}N$ for each of the following sequences: 'HH', 'HT', 'TH', 'TT', 'HTH', 'HHTTHH'.

Exercise 4.34. A standard symmetric dice is tossed repeatedly. Let N be the number of (independent) tosses until a fixed pattern is observed. Using an appropriate martingale, find $\mathsf{E}N$ for the sequences '123456' and '666666'.

☞ **Lemma 4.35.** *Let $(Y_n)_{n\ge0}$ be a supermartingale w.r.t. a sequence $(X_n)_{n\ge0}$ and let $H_n \in \mathcal{F}^X_{n-1} = \sigma(X_0,\dots,X_{n-1})$ satisfy $0 \le H_n \le c_n$, where the constant c_n might depend on n. Then the process $(W_n)_{n\ge0}$ with*

$$W_n = W_0 + \sum_{m=1}^{n} H_m\big(Y_m - Y_{m-1}\big)$$

is a supermartingale w.r.t. $(X_n)_{n\ge0}$.

Proof. Following the hint in Exercise 4.24, we observe that since $(Y_n)_{n\ge0}$ is a supermartingale w.r.t. $(X_n)_{n\ge0}$,

$$\mathsf{E}\big(H_m(Y_m - Y_{m-1})\big) = \mathsf{E}\big[H_m\,\mathsf{E}(Y_m - Y_{m-1}\,|\,\mathcal{F}_{m-1})\big] \le 0\,. \qquad \square$$

Example 4.36. *If $(Y_n)_{n\ge0}$ describes the stock price process, and H_m is the number of stocks held during the time $(m-1,m]$ (decided when the price Y_{m-1} is known), then W_n describes the fortune of an investor at time n. As $(W_n)_{n\ge0}$ is a supermartingale w.r.t. $(X_n)_{n\ge0}$, we have $\mathsf{E}W_n \le \mathsf{E}W_0$ for all $n \ge 0$.*

Remark 4.36.1. The famous "doubling martingale" corresponds to doubling the bet size until one wins, ie., to taking $H_m = 2^{m-1}\mathbb{1}_{\{T>m\}}$, where T is the first moment when the price goes up, ie., $T = \min\{m>0 : Y_m - Y_{m-1} = 1\}$. Since the stopped process $(W_{n\wedge T})_{n\ge0}$ is a supermartingale, for all $n \ge 0$ we have $\mathsf{E}(W_{n\wedge T}) \le \mathsf{E}(W_0)$, ie., on average, the doubling strategy does not produce money if one bets against a (super)martingale.

✍ **Example 4.37.** [Wald's equation] *Let $(S_n)_{n\ge0}$ be a random walk generated by a sequence $(X_n)_{n\ge0}$ of i.i.d. steps with $\mathsf{E}|X| < \infty$ and $\mathsf{E}(X) = m$. If T is a stopping time for $(X_n)_{n\ge0}$ with $\mathsf{E}(T) < \infty$, then $\mathsf{E}(S_T - S_0) = m\,\mathsf{E}(T)$.*

Solution. Notice that $S_n - n\mathsf{E}X = S_n - mn$ is a martingale and for every $n \ge 0$ the variable $T \wedge n$ is a bounded stopping time. By (OST-1), we have

$$\mathsf{E}(S_0) = \mathsf{E}\big(S_{n\wedge T} - m(n \wedge T)\big)\,; \tag{4.9}$$

now rearrange to $\mathsf{E}\big(S_{n\wedge T} - S_0\big) = m\mathsf{E}(n \wedge T)$ and let $n \to \infty$ to get the result.[44] □

Exercise 4.38. Let $(X_n)_{n\ge0}$ be a sequence of $i.i.d.$ Bernoulli random variables with $\mathsf{P}(X = 1) = p$ and $\mathsf{P}(X = -1) = 1 - p = q$. Let $(S_n)_{n\ge0}$ be the generated random walk with $S_0 = x > 0$, and let $T = \min\{n \ge 0 : S_n = 0\}$ be the hitting time of the origin. Example 1.14 suggests that $\mathsf{E}(T) = x/(q - p) < \infty$ for $q > p$.

a) Use (4.9) to deduce that

$$-m\mathsf{E}(n \wedge T) = \mathsf{E}\big(S_0 - S_{n\wedge T}\big) \le \mathsf{E}(S_0) = x\,;$$

and take $n \to \infty$ to show that $\mathsf{E}(T) < \infty$;

b) Use the Wald equation to deduce that indeed $\mathsf{E}(T) = x/(q - p)$. Can you give a heuristic explanation of this result?

c) Argue, without using the Wald equation, that $\mathsf{E}(T) = cx$ for some constant c.

d) Use the Wald equation and an argument by contradiction to show that if $p \ge q$, then $\mathsf{E}(T) = \infty$ for all $x > 0$.

Exercise 4.39. [Wright-Fischer model] Thinking of a population of N haploid individuals who have one copy of each of their chromosomes, consider a fixed population of N genes that can be of two types A or a. In the simplest version of this model the population at time $n+1$ is obtained by sampling with replacement from the population at time n. If we let X_n to be the number of A alleles at time n, then X_n is a Markov chain with transition probability

$$p_{ij} = \binom{N}{j}\Big(\frac{i}{N}\Big)^j\Big(1 - \frac{i}{N}\Big)^{N-j}.$$

Starting from i of the A alleles and $N - i$ of a alleles, what is the probability that the population fixates in the all A state?
[**Hint:** You can use the heuristics of Example 4.16 but need to justify your computation!]

[44] The the RHS converges by (MON), and, since $S_T = \sum_{n=0}^{\infty} S_n \mathbb{1}_{T=n} = \sum_{n=0}^{\infty} X_n \mathbb{1}_{T\ge n}$, the LHS being absolutely bounded above by an integrable function $\sum_{n=0}^{\infty} |X_n|\mathbb{1}_{T\ge n}$ (notice that $\{T \ge n\} \in \mathcal{F}_{n-1}$ and thus is independent of X_n) converges by (DOM); alternatively, apply (OST-2) as generalized in Remark 4.25.1.

Example 4.40. *Let* $(\xi_k)_{k\ge1}$ *be* i.i.d.r.v. *with* $\varphi(\rho) \stackrel{\mathsf{def}}{=} \mathsf{E}e^{\rho\xi} < \infty$ *for all real* ρ. *If* $(S_n)_{n\ge0}$ *is the generated random walk,* $S_n = \xi_1 + \cdots + \xi_n$, *then the process* $(M_n)_{n\ge0}$, *where* $M_n = e^{\rho S_n}/\varphi(\rho)^n$, *is a martingale w.r.t.* $(S_n)_{n\ge0}$.

In addition, let ξ *take integer values with* $\mathsf{P}(\xi \le 1) = 1$, $\mathsf{P}(\xi = 1) > 0$, *and negative expectation* $\mathsf{E}\xi < 0$. *Let, further,* T_a *be the hitting time of state* $a > 0$, *ie.,* $T_a = \min\{n \ge 0 : S_n = a\}$. *If* $\alpha > 0$ *is defined from the condition* $\varphi(\alpha) = 1$, *show that* $\mathsf{P}(T_a < \infty) = e^{-\alpha a}$ *for all* $a > 0$.

Solution. It is immediate to check that $(M_n)_{n\ge0}$ is a martingale w.r.t. $(S_n)_{n\ge0}$.
Let now $\alpha > 0$ be such that $\varphi(\alpha) = 1$ and consider $T = T_1$, ie., $a = 1$. Then (OST-3) applied to the martingale $M_n = e^{\alpha S_n}$ gives

$$1 = \mathsf{E}\big(e^{\alpha S_{n\wedge T_a}}\big) = e^{\alpha}\,\mathsf{P}(T \le n) + \mathsf{E}\big(e^{\alpha S_n}\mathbb{1}_{\{T>n\}}\big)$$

and since, by the strong law of large numbers $S_n/n \to \mathsf{E}\xi < 0$ as $n \to \infty$, the second term[45] goes to 0, we get $1 = e^{\alpha}\,\mathsf{P}(T < \infty)$, ie., $\mathsf{P}(T < \infty) = e^{-\alpha} < 1$. Now, by the strong Markov property, $\mathsf{P}(T_a < \infty) = \big(\mathsf{P}(T_1 < \infty)\big)^a \equiv e^{-\alpha a}$ for all $a > 0$.
It remains to show existence of such $\alpha > 0$. To this end we notice that $\varphi(x)$ is convex and satisfies the following properties: $\varphi(0) = 1$, $\varphi'(0) = \mathsf{E}\xi < 0$, and for every $\rho > 0$ we have $\varphi(\rho) \ge e^{\rho}\mathsf{P}(\xi = 1)$, so that $\varphi(\rho) \to \infty$ as $\rho \to \infty$. □

☞ **Exercise 4.41.** Let $(X_n)_{n\ge0}$ be a birth-and-death process in $\mathcal{S} = \{0, 1, \dots\}$, ie., a Markov chain in $\mathcal{S}$ with transition probabilities $p_{00} = r_0$, $p_{01} = p_0$, and

$$p_{n,n-1} = q_n\,, \qquad p_{n,n} = r_n\,, \qquad p_{n,n+1} = p_n\,, \qquad n > 0\,,$$

while $p_{n,m} = 0$ for all other pairs $(n, m) \in \mathcal{S}^2$. Let $X_0 = x$, and for every $y \ge 0$ denote $T_y \stackrel{\mathsf{def}}{=} \min\{n \ge 0 : X_n = y\}$.

a) Show that the process $\big(\varphi(X_n)\big)_{n\ge0}$ with $\varphi(z) \stackrel{\mathsf{def}}{=} \sum_{y=1}^{z}\prod_{x=1}^{y-1}\frac{q_x}{p_x}$ is a martingale.

b) Show that for all $0 \le a < X_0 = x < b$ we have[46]

$$\mathsf{P}\big(T_b < T_a\big) = \frac{\varphi(x) - \varphi(a)}{\varphi(b) - \varphi(a)}\,. \tag{4.10}$$

c) Deduce that state 0 is recurrent iff $\varphi(b) \to \infty$ as $b \to \infty$.

d) Now suppose that $p_n \equiv p$, $q_n \equiv q = 1 - p$, and $r_n = 0$ for $n > 0$, whereas $p_0 = p$ and $r_0 = q$. Show that in this case (4.10) becomes (cf. Example 4.16)

$$\mathsf{P}\big(T_b < T_a\big) = \frac{\big(q/p\big)^a - \big(q/p\big)^x}{\big(q/p\big)^a - \big(q/p\big)^b}\,.$$

e) Find $\mathsf{P}\big(T_b < T_a\big)$ if in the setup of part d) one has $p = q = 1/2$.

[45] notice that $\mathsf{E}\big(e^{\alpha S_n}\mathbb{1}_{\{T>n\}}\big) \le \mathsf{P}(T > n)$ here!
[46] Recall the argument in Example 4.16.

Exercise 4.42. Let $(\xi_k)_{k\geq 1}$ be i.i.d.r.v. with $\mathsf{P}(\xi = 1) = p$, $\mathsf{P}(\xi = -1) = q = 1-p$, and $\mathsf{E}\xi > 0$. Let $(S_n)_{n\geq 0}$ be the generated random walk, $S_n = x+\xi_1+\cdots+\xi_n$, and let $T_0 = \min\{n \geq 0 : S_n = 0\}$ be the hitting time of 0. Deduce that for all $x > 0$, $\mathsf{P}(T_0 < \infty) = (q/p)^x$. Compare this to the result of Example 1.14.

Exercise 4.43. Let $(X_n)_{n\geq 0}$ be an irreducible Markov chain in $\mathcal{S} = \{0, 1, \dots\}$ with bounded jumps, and let a function $\varphi : \mathcal{S} \to \mathbb{R}^+$ satisfy $\varphi(x) \to \infty$ as $x \to \infty$. Let $K \geq 0$ be such that $\mathsf{E}_x\varphi(X_1) \stackrel{\mathsf{def}}{=} \mathsf{E}\big[\varphi(X_1)|X_0 = x\big] \leq \varphi(x)$ for all $x \geq K$.

a) If the function $\varphi(x)$ is monotone, show that the set of states $\{0, 1, \dots, K\}$, and thus the whole space $\mathcal{S}$ is recurrent for $(X_n)_{n\geq 0}$.
[**Hint:** If $H_K = \min\{n \geq 0 : 0 \leq X_n \leq K\}$, show that $\varphi(X_{n\wedge H_K})$ is a supermartingale. Deduce that if $T_M = \min\{n \geq 0 : X_n \geq M\}$, then $\varphi(x) \geq \varphi(M)\mathsf{P}(T_M < H_K)$.]

b) Argue that the result holds for $\varphi(x) \geq 0$ not necessarily monotone, but only satisfying $\varphi(x) \to \infty$ as $x \to \infty$.
[**Hint:** With T_M as above, show that $\varphi^*_M \stackrel{\mathsf{def}}{=} \min\{\varphi(x) : x \geq M\} \to \infty$ as $M \to \infty$.]

4.5 Long time behaviour of martingales

4.5.1 Pólya's urn martingale

The following example has a number of important applications.

Example 4.44 (Pólya's urn). *An urns contains one green and one red ball. At every step a ball is selected at random, and then replaced together with another ball of the same colour. Let X_n be the number of green balls after nth draw, $X_0 = 1$. Then the fraction $M_n = X_n/(n+2)$ of green balls is a martingale w.r.t. the filtration $(\mathcal{F}^X_n)_{n\geq 0}$.*

Solution. As $|M_n| \leq 1$ we have $\mathsf{E}|M_n| \leq 1$ for all $n \geq 0$, and since

$$\mathsf{P}(X_{n+1} = k+1 \mid X_n = k) = \frac{k}{n+2}\,, \qquad \mathsf{P}(X_{n+1} = k \mid X_n = k) = 1 - \frac{k}{n+2}\,,$$

we get $\mathsf{E}\big(X_{n+1} \mid \mathcal{F}^X_n\big) = \frac{n+3}{n+2}\,X_n$, equivalently, $\mathsf{E}(M_{n+1} \mid \mathcal{F}^X_n) = M_n$. □

Exercise 4.45. Show that the distribution of M_n is uniform, $\mathsf{P}(M_n = \frac{k}{n+2}) = \frac{1}{n+1}$ for $n \geq 0$ and $k = 1, \dots, n+1$.

Exercise 4.46. Suppose that the process in Example 4.44 is modified as follows: for a fixed integer $c > 1$, every time a random ball is selected, it is replaced together with other c balls of the same colour. If, as before, X_n denotes the total number of green balls after n draws, show that the the fraction $M_n = \frac{X_n}{2+nc}$ of green balls forms a martingale w.r.t. $(\mathcal{F}^X_n)_{n\geq 0}$.

Exercise 4.45 suggests that in the limit $n \to \infty$ the distribution of M_n is asymptotically uniform:[47]

[47] see also the R-script available from the course webpage!

Exercise 4.47. Let $(M_n)_{n\geq 0}$ be the process from Example 4.44. Show that for every $x \in (0,1)$ we have $\lim_{n\to\infty} \mathsf{P}\bigl(M_n < x\bigr) = x$.

Exercise 4.48.$^{\circledast}$ Find the large-n limit of the distribution of the martingale $(M_n)_{n\geq 0}$ from Exercise 4.46.

In view of Exercise 4.45, a natural question is: does the proportion M_n of green balls fluctuate between 0 and 1 infinitely often or does it eventually settle down to a particular value? The following example shows that the latter is true. Our argument is based upon the following observation: if a real sequence y_n does not converge, for some *real* a, b with $-\infty < a < b < \infty$ the sequence y_n must go from the region below a to the region above b (and back) infinitely often.

Example 4.49. *For fixed $n \geq 0$ let $M_n < a \in (0,1)$, and let $N = \min\{k > n : M_n > b\}$ for some $b \in (a,1)$. Since $N_m = N \wedge m$ is a bounded stopping time, by* (OST-1) *we have $\mathsf{E}M_{N_m} = \mathsf{E}M_n < a$ if only $m > n$. On the other hand,*

$$\mathsf{E}M_{N_m} \geq \mathsf{E}\bigl(M_{N_m}\mathbb{1}_{N\leq m}\bigr) \equiv \mathsf{E}\bigl(M_N\mathbb{1}_{N\leq m}\bigr) > b\,\mathsf{P}(N \leq m)\,.$$

In other words, $\mathsf{P}(N \leq m) < \frac{a}{b}$ and consequently $\mathsf{P}(N < \infty) \leq \frac{a}{b} < 1$, ie., the fraction M_n of green balls ever gets above level b with probability at most $\frac{a}{b}$. Suppose that at certain moment $N \in (n,\infty)$ the fraction of green balls became bigger than b. Then a similar argument shows that with probability at most $(1-b)/(1-a)$ the value M_n becomes smaller than a at a later moment.

Put $S_0 = \min\{n \geq 0 : M_n < a\}$, and then, inductively, for $k \geq 0$,

$$T_k = \min\bigl\{n > S_k : M_n > b\bigr\}\,, \qquad S_{k+1} = \min\bigl\{n > T_k : M_n < a\bigr\}\,. \tag{4.11}$$

The argument above implies that

$$\mathsf{P}(S_k < \infty) \leq \prod_{j=1}^{k}\Bigl(\mathsf{P}(T_{j-1} < \infty \mid S_{j-1} < \infty)\mathsf{P}(S_j < \infty \mid T_{j-1} < \infty)\Bigr)$$

with the RHS bounded above by $\bigl(\frac{a}{b}\bigr)^k\bigl(\frac{1-b}{1-a}\bigr)^k \to 0$ as $k \to \infty$. As a result, the probability of infinitely many crossing (ie., $S_k < \infty$ for all $k \geq 0$) vanishes.

Clearly, the argument above applies to all strips $(a,b) \subset (0,1)$ with rational *endpoints. Thus, with probability one,*[48] *trajectories of M_n eventually converge to a particular value.*[49]

Remark 4.49.1. Let $U_{(a,b)}$ be the total number of upcrossings of the strip (a,b) for the process $(M_n)_{n\geq 0}$. We clearly have $\bigl\{U_{(a,b)} \geq m\bigr\} \equiv \bigl\{T_m < \infty\bigr\}$. Notice that the argument in Example 4.49 implies that $\mathsf{E}U_{(a,b)} < \infty$.

Exercise 4.50. By noticing that $\bigl\{U_{(a,b)} \geq m\bigr\} \subset \bigl\{S_{m-1} < \infty\bigr\}$ or otherwise, show that $\mathsf{E}U_{(a,b)} < \infty$.

[48] If M_n does not converge, it must cross at least one of countably many strips (a,b) with rational points infinitely many times.

[49] which is random and depends on the trajectory, see the R-script on the course webpage!

4.5.2 Martingale convergence theorem

The argument in Example 4.49 also works in general. Let $(M_n)_{n\ge0}$ be a martingale w.r.t. filtration $(\mathcal{F}_n^X)_{n\ge0}$. For real a, b with $-\infty < a < b < \infty$ let $U_{(a,b)}$ be the total number of upcrossings of the strip (a,b). The following result (or some of its variants) is often referred to as Doob's Upcrossing Lemma:

Lemma 4.51. *Let the martingale $(M_n)_{n\ge0}$ have uniformly bounded expectations, ie., for some constant K and all $n\ge0$, $\mathsf{E}|M_n| < K < \infty$. If $U_{(a,b)}$ is the number of upcrossings of a strip (a,b), then $\mathsf{E}U_{(a,b)} < \infty$.*

Proof. With stopping times as in (4.11), put $H_n = 1$ if $S_m < n \le T_m$ and put $H_n = 0$ otherwise. Then the process

$$W_n = \sum_{k=1}^{n} H_k (M_k - M_{k-1})$$

is a martingale w.r.t. $(M_n)_{n\ge0}$, cf. Lemma 4.35. It is easy to check that $W_n \ge (b-a)\,U_n - |M_n - a|$ (draw the picture!), where $U_n = \max\{m\ge0 : T_m \le n\}$ is the number of upcrossings of the strip (a,b) up to time n. As a result

$$0 = \mathsf{E}W_0 = \mathsf{E}W_n \ge (b-a)\,\mathsf{E}U_n - \mathsf{E}|M_n - a| \ge (b-a)\,\mathsf{E}U_n - (K+|a|)\,,$$

so that $\mathsf{E}U_n \le (K+|a|)/(b-a)$ for all $n\ge0$, and thus $\mathsf{E}U_{(a,b)} < \infty$. □

Theorem 4.52. *Let $(M_n)_{n\ge0}$ be a martingale as in Lemma 4.51. Then there exists a random variable M_∞ such that $M_n \to M_\infty$ with probability one.*

Proof. If M_n does not converge, for some rational a, b with $-\infty < a < b < \infty$ we must have $U_{(a,b)} = \infty$. However, by Lemma 4.51, $\mathsf{E}U_{(a,b)} < \infty$ implying that $\mathsf{P}(U_{(a,b)} = \infty) = 0$. As the number of such pairs (a,b) is countable, the result follows. □

In applications of the convergence theorem, the following fact is often useful:

Exercise 4.53. Let a variable Y satisfy $\mathsf{E}(Y^2) < \infty$. Show that $\mathsf{E}|Y| < \infty$.
[**Hint:** Notice that $\mathsf{Var}(|Y|) \ge 0$.]

Exercise 4.54. Let $(Z_n)_{n\ge0}$ be a homogeneous branching process with $Z_0 = 1$, $m = \mathsf{E}Z_1 > 0$ and $\sigma^2 = \mathsf{Var}(Z_1) < \infty$. Show that $M_n = Z_n/m^n$ is a martingale.

a) Let $m > 1$. By using Exercise 2.2 or otherwise show that $\mathsf{E}(M_n)$ is uniformly bounded. Deduce that $M_n \to M_\infty$ almost surely. What can you say about $\mathsf{E}M_\infty$?

b) What happens if $m \le 1$? Compute $\mathsf{E}(M_\infty)$. [**Hint:** Recall Exercise 4.53.]

Exercise 4.55. Let $(X_k)_{k\ge1}$ be independent variables with

$$\mathsf{P}\Big(X = \frac{3}{2}\Big) = \mathsf{P}\Big(X = \frac{1}{2}\Big) = \frac{1}{2}\,.$$

Put $M_n = X_1 \cdot \ldots \cdot X_n$ with $M_0 = 1$. Show that M_n is an $(\mathcal{F}_n^X)_{n\ge0}$ martingale. Deduce that $M_n \to M_\infty$ with probability one. Can you compute $\mathsf{E}(M_\infty)$?
[**Hint:** Consider $L_n \stackrel{\mathsf{def}}{=} \ln(M_n)$.]

Exercise 4.56. Let $(X_k)_{k\geq 1}$ be independent variables with $\mathsf{P}(X = \pm 1) = \frac{1}{2}$. Show that the process

$$M_n = \sum_{k=1}^{n} \frac{1}{k} X_k$$

is a martingale w.r.t. $(\mathcal{F}_n^X)_{n\geq 0}$ and that $\mathsf{E}\big[(M_n)^2\big] < K < \infty$ for some constant K and all $n \geq 0$. By using Exercise 4.53 or otherwise, deduce that with probability one, $M_n \to M_\infty$ for some random variable M_∞.
In other words, the *random sign harmonic series converges with probability one*.