4 Martingales

4.1 Definition and some examples

A martingale is a generalized version of a "fair game".

Definition 4.1. *A process* $(M_n)_{n\geq 0}$ *is a martingale if*

a) for every $n \geq 0$ *the expectation* $\mathsf{E}M_n$ *is finite, equivalently,* $\mathsf{E}|M_n| < \infty$*;*

b) for every $n \geq 0$ *and all* m_n, m_{n-1}, $\dots$, m_0 *we have*

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

For those familiar with the notion of conditioning on a random variable (see next section), (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* $(M_n)_{n\geq 0}$ *is a* supermartingale[24] *if the equality in* (4.1) *holds with* $\leq$, *ie.,* $\mathsf{E}\big(M_{n+1} \mid M_n, \dots, M_0\big) \leq M_n$*; and we say that* $(M_n)_{n\geq 0}$ *is a* submartingale, *if* (4.1) *holds with* $\geq$, *ie.,* $\mathsf{E}\big(M_{n+1} \mid M_n, \dots, M_0\big) \geq M_n$.

Example 4.3. *Let* $(\xi_n)_{n\geq 1}$ *be independent random variables*[25] *with*

$$\mathsf{E}\xi_n = 0$$

for all $n \geq 1$. *Then the process* $(M_n)_{n\geq 0}$ *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* $(\xi_n)_{n\geq 1}$ *and* $\mathsf{E}|M_0| < \infty$. *For example, we will often take* $M_0 = 0$ *or some other deterministic constant.*

Indeed, by the triangle inequality, $\mathsf{E}|M_n| \leq \mathsf{E}|M_0| + \sum_{j=1}^n \mathsf{E}|\xi_j| < \infty$ *for all* $n \geq 0$*; on the other hand, the independence property implies*

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

Remark 4.4. *Notice that if* $\mathsf{E}\xi_n \geq 0$ *for all* $n \geq 1$, *then* $(M_n)_{n\geq 0}$ *is a submartingale, whereas if* $\mathsf{E}\xi_n \leq 0$ *for all* $n \geq 1$, *then* $(M_n)_{n\geq 0}$ *is a supermartingale. More generally, if* $(\xi_n)_{n\geq 1}$ *are independent random variables with* $\mathsf{E}|\xi_n| < \infty$ *for all* $n \geq 1$, *then the process* $M_n = M_0 + (\xi_1 - \mathsf{E}\xi_1) + \dots + (\xi_n - \mathsf{E}\xi_n)$, $n \geq 0$, *is a martingale.*

Exercise 4.5. *Given a sequence* $(\xi_k)_{k\geq 1}$ *of independent Bernoulli variables with common distribution* $\mathsf{P}(\xi = 1) = p$ *and* $\mathsf{P}(\xi = -1) = q = 1 - p$, *define the generated random walk via* $X_n = \sum_{k=1}^n \xi_k$, $n \geq 0$. *Show that the process* $M_n = (q/p)^{X_n}$ *is a martingale.*

[24] If M_n traces your fortune, then "there is nothing super about a supermartingale".

[25] Notice that we do not assume that all ξ_n have the same distribution!

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 $M_n = \rho^{Z_n}$ is a martingale.*

We can also construct many examples of submartingales and supermartingales if we have a martingale and apply certain functions to it. This is a consequence of:

Lemma 4.8 (Conditional Jensen's inequality)**.** *Suppose that X is an integrable random variable on $(\Omega, \mathcal{F}, \mathsf{P})$ and $A \in \mathcal{F}$ with $\mathsf{P}(A) > 0$.*

Suppose that φ is a convex *function such that $\varphi(X)$ is also integrable. Then*

$$\mathsf{E}(\varphi(X)|A) \geq \varphi(\mathsf{E}(X|A)).$$

Proof. This follows from the non-conditional version of Jensen's inequality, and using the fact that $\mathsf{P}(\cdot|A)$ is a valid probability measure. □

Example 4.9. *Let $(X_n)_{n\geq 0}$ be a martingale. If, for some* convex *function $f(\,\cdot\,)$ we have $\mathsf{E}\big(|f(X_n)|\big) < \infty$ for all $n \geq 0$, then the process $f(X) := \big(f(X_n)\big)_{n\geq 0}$ is a submartingale.*

Similarly, if for some concave *$f(\,\cdot\,)$ we have $\mathsf{E}\big(|f(X_n)|\big) < \infty$ for all $n \geq 0$, then the process $f(X)$ is a supermartingale.*

Both these claims follow immediately from the Jensen inequalities for conditional expectations.

4.2 A few remarks on conditional expectation

Recall the following basic definition:

Example 4.10. *Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability triple and let X and Y be random variables taking values x_1, x_2, $\dots$, x_m and y_1, y_2, $\dots$, 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*. Of course, if X and Y are independent, we have $Z(\omega) \equiv \mathsf{E}X$.*

Definition 4.11. *If $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$ is a finite partition*[26] *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 4.10 above we see that Z is constant on the events $\{Y = y_j\}$; ie., Z is measurable w.r.t. the σ-field $\sigma(Y) \equiv \sigma(\{Y = y_j\}_j)$. Moreover, for every $G_j \equiv \{Y = y_j\}$, we have

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

where the last equality follows from the observation that the random variable $X1_{G_j}$ 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 4.12. *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 the unique random variable Z such that: Z is $\mathcal{G}$ measurable and for every set $G \in \mathcal{G}$ we have $\int_G Z\mathsf{dP} \equiv \mathsf{E}\big(Z1_G\big) = \mathsf{E}\big(X1_G\big) \equiv \int_G X\mathsf{dP}$.

Notice that when $\mathcal{G} = \sigma(D)$, the definition in the remark above coincides with Definition 4.11.

The following are the most important properties of conditional expectation:

Lemma 4.13. *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

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

[26] 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 $\mathcal{D}$.

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

g) If Z is independent of $\mathcal{G}$, then $\mathsf{E}\big[Z \mid \mathcal{G}\big] = \mathsf{E}[Z]$.

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

$$\mathcal{F}_k^X \stackrel{\mathsf{def}}{=} \sigma(X_1, X_2, \dots, X_k)\,; \tag{4.2}$$

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

$$\mathcal{F}_1^X \subseteq \mathcal{F}_2^X \subseteq \ldots \subseteq \mathcal{F}_k^X \subseteq \ldots. \tag{4.3}$$

The notion of filtration is very useful when working with martingales. Indeed, a process $(M_n)_{n\ge0}$ is a martingale if for all $n \ge 0$, $\mathsf{E}|M_n| < \infty$ and $\mathsf{E}\big(M_{n+1} \mid \mathcal{F}_n^M\big) = M_n$. We can also generalise the definition of a martingale:

Definition 4.14. *$(M_n)_{n\ge0}$ is a martingale with respect to a filtration $(\mathcal{F}_n)_{n\ge0}$ if for all $n \ge 0$, $\mathsf{E}(|M_n|) < \infty$ and*

$$\mathsf{E}(M_{n+1}|\mathcal{F}_n) = M_n.$$

The original definition of martingale is sometimes referred to as "being a martingale with respect to the natural filtration".

We say that M is a martingale with respect to a sequence $(X_n)_{n\ge0}$ if it is a martingale with respect to the filtration $(\mathcal{F}_n^X)_{n\ge0}$.

We also notice that by repeatedly applying the tower property in claim e) of Lemma 4.13 above to the sequence (4.3), 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 4.16. *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.*

4.3 Martingales and stopping times

Martingales are extremely useful in studying stopping times:

Example 4.17. *Let $(X_k)_{k\ge1}$ 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 $p \in (0,1)$, $p \ne 1/2$. For fixed $k \in \{0, 1, \dots, N\}$, define the random walk $(S_n)_{n\ge0}$ defined via $S_n = k + \sum_{j=1}^n X_j$, $n \ge 0$, and consider the process $(Y_n)_{n\ge0}$, where $Y_n = (q/p)^{S_n}$. Clearly, Y_n is an $(\mathcal{F}_n^X)_{n\ge0}$-martingale, so that $\mathsf{E}(Y_n) = \mathsf{E}(Y_0) = (q/p)^k$ for all $n \ge 0$, recall Lemma 4.15.*

Let T be the first time S_n hits 0 or N. If an analogue of the above equality, $\mathsf{E}(Y_T) = \mathsf{E}(Y_0) = (q/p)^k$ were true for (random) time T, we could find the exit probability $\mathsf{p}_k = \mathsf{P}(\, S$ hits 0 before $N \mid S_0 = k)$ from the expression

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

thus obtaining $\mathsf{p}_k = \big((q/p)^k - (q/p)^N\big)/\big(1 - (q/p)^N\big)$.

Remark 4.18. *The method used in the previous example relies on the assumption $\mathsf{E}(Y_T) = \mathsf{E}(Y_0)$ and on the formula for $\mathsf{E}(Y_T)$ for a certain random variable T. An important part of the theory of martingales is to study random variables T for which the above statements are true.*[27]

Example 4.19. *Let $S_n = \sum_{k=1}^n X_k$ be the simple symmetric random walk in $\{-K, \dots, K\}$, generated by a sequence of i.i.d. symmetric Bernoulli r.v. X_k, where $\mathsf{P}(X_1 = \pm 1) = 1/2$. Similarly to Example 4.17 one can study the hitting time T of the boundary, $T = \inf\big\{n \ge 0 : |S_n| = K\big\}$: namely, since*[28]

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

the same heuristics applied to the martingale $(Y_n)_{n\ge0}$, $Y_n \stackrel{\mathsf{def}}{=} \big(S_n\big)^2 - n$, leads to $0 = \mathsf{E}(Y_0) = \mathsf{E}(Y_T) = \mathsf{E}\big((S_T)^2\big) - \mathsf{E}(T)$; i.e., 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 4.20. *A variable T is a stopping time for a process $(X_n)_{n\ge0}$, if the occurrence/non-occurrence of the event $\{T = n\}$ =*"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, if we have $\{T \le n\} \in \mathcal{F}_n^X \equiv \sigma(X_0, \dots, X_n)$ for every $n \ge 0$.*

Example 4.21. *Let $(X_n)_{n\ge0}$ be a stochastic process with values in S, and let T be the hitting time of a set $A \subset S$, namely, $T = \min\big\{n \ge 0 : X_n \in A\big\}$. Then T is a stopping time for $(X_n)_{n\ge0}$.*

Indeed, for every fixed $n \ge 0$, 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 $\{T \le n\}$ both belong to $\mathcal{F}_n^X$.

By contrast, the last time that X visits A, $\tilde{T} = \max\{n \ge 0 : X_n \in A\}$ is not generally a stopping time. Because we generally cannot tell just by looking at $X_0, \dots, X_n$, whether the process will visit A after time n or not.

Exercise 4.22. *Let $k \in \mathbb{N}$ be fixed, and let S and T be stopping times for a process $(X_n)_{n\ge0}$. Show that the following are stopping times:*

(a) $T \equiv k$,

[27]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.

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

(b) $S \wedge T \equiv \min(S, T)$,

(c) $S \vee T \equiv \max(S, T)$.

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 $(X_n)_{n \geq 0}$. Is the variable $L = \max\{n \geq 0 : X_n \in A\}$ a stopping time for $(X_n)_{n \geq 0}$?

Exercise 4.23. *Let $(M_n)_{n \geq 0}$ be a submartingale w.r.t. $(X_n)_{n \geq 0}$, and let T be a stopping time for $(X_n)_{n \geq 0}$. Show that the process $(L_n^T)_{n \geq 0}$,*

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

is a submartingale w.r.t. $(X_n)_{n \geq 0}$. Deduce that if $(M_n)_{n \geq 0}$ is a martingale w.r.t. $(X_n)_{n \geq 0}$, then the stopped process $(L_n^T)_{n \geq 0}$ is also a martingale w.r.t. $(X_n)_{n \geq 0}$.

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 4.24 (Optional Stopping Theorem)**.** *Let $(M_n)_{n \geq 0}$ be a martingale w.r.t. $(X_n)_{n \geq 0}$, and let T be a stopping time for $(X_n)_{n \geq 0}$. Then the equality*

$$\mathsf{E}\big[M_T\big] = \mathsf{E}[M_0] \tag{4.4}$$

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

(OST-1) *the stopping time T is bounded, i.e., $\mathsf{P}(T \leq N) = 1$ for some $N < \infty$;*

(OST-2) $\mathsf{E}T < \infty$*, and the martingale $(M_n)_{n \geq 0}$ has bounded increments, i.e., $|M_{n+1} - M_n| \leq K$ for all n and some constant K;*

(OST-3) $\mathsf{P}(T < \infty) = 1$*, and the martingale $(M_n)_{n \geq 0}$ is bounded, i.e., $|M_n| \leq K$ for all n and some constant K.*

Remark 4.25. *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 4.26. *One can generalize* (OST-3) *by replacing the condition of boundedness, $|M_n| \leq K$, by that of uniform integrability for the martingale $(M_n)_{n \geq 0}$:* a sequence of random variables $(Y_n)_{n \geq 0}$ is *uniformly integrable* if

$$\lim_{K \to \infty} \sup_{n \geq 0} \mathsf{E}\big(|Y_n| \mathbb{1}_{\{|Y_n| > K\}}\big) = 0\,. \tag{4.5}$$

Example 4.27. *Let the SSRW $(S_n)_{n\geq 0}$ be generated as in Example 4.19. Put*

$$H \overset{\mathsf{def}}{=} \inf\{n \geq 0 : S_n = 1\}\,.$$

Since this RW is recurrent,[29] *we deduce that* $\mathsf{P}(H < \infty) = 1$*. However, the* (OST) *does not apply, as* $0 = \mathsf{E}(S_0) \neq \mathsf{E}(S_H) \equiv 1$*. It is an instructive* Exercise *to check which conditions in each of the above* (OST) *are violated.*

We now give the proof of Theorem 4.24.

Proof. Let us first assume that T satisfies (OST-1); that is, there is some $N < \infty$ with $0 \leq T \leq N$. Then the decomposition

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

holds. Now for each $0 \leq n \leq N-1$, (a) of Lemma 4.13 allows us to rewrite $\mathsf{E}\big((M_{n+1} - M_n)\mathsf{1}_{\{T>n\}}\big)$ as

$$\mathsf{E}\big(\mathsf{E}\big((M_{n+1} - M_n)\mathsf{1}_{\{T>n\}} \,|\, \mathcal{F}_n^X\big)\big) = \mathsf{E}\big(\mathsf{1}_{\{T>n\}}\mathsf{E}\big(M_{n+1} - M_n \,|\, \mathcal{F}_n^X\big)\big)$$

where we have used that $\{T > n\}$ is $\mathcal{F}_n^X$-measurable (by definition of a stopping time). Since M is a martingale, $\mathsf{E}\big(M_{n+1} - M_n \,|\, \mathcal{F}_n^X\big) = 0$ and hence $\mathsf{E}\big((M_{n+1} - M_n)\mathsf{1}_{\{T>n\}}\big) = 0$ for each n. Linearity of expectation then implies that $\mathsf{E}(M_T) = \mathsf{E}(M_0)$ as required.

Now let us assume that T satisfies (OST-2). For fixed n, $\min(T, n) = T \wedge n$ is clearly a bounded stopping time (it is less than n with probability one), so by (OST-1) we have that $\mathsf{E}M_{T\wedge n} = \mathsf{E}M_0$ for all $n \geq 0$. We now want to take a limit as $n \to \infty$. For this we write

$$|M_T - M_{T\wedge n}| = \big|\sum_{k>n} (M_k - M_n)\mathsf{1}_{\{T=k\}}\big| = \big|\sum_{k>n} \big(M_k - M_{k-1}\big)\mathsf{1}_{\{T\geq k\}}\big|$$

(in particular, noting that this gives $|M_T| \leq KT$ when $n = 0$ and so M_T is integrable). Taking expectations then gives that $\mathsf{E}(|M_T - M_{T\wedge n}|) \leq K \sum_{k>n} \mathsf{P}(T \geq k)$ which tends to 0 as $n \to \infty$. Hence $\mathsf{E}(M_T) = \lim_{n\to\infty} \mathsf{E}(M_{T\wedge n}) = \lim_{n\to\infty} \mathsf{E}(M_0) = \mathsf{E}(M_0)$.

Finally, we can assume (OST-3) and deduce the result in a similar way. The strategy is the same, but instead of writing $M_T - M_{T\wedge n}$ as a telescoping sum, we note that $|M_T - M_{T\wedge n}| \leq 2K\mathsf{1}_{\{T>n\}}$ so that

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

□

Remark 4.28. *For those who are familiar with the dominated convergence theorem. In* (OST-3) *it is sufficient to assume that* $\big|M_{n\wedge T}\big| \leq K$ *for all* $n \geq 0$ *and* $\mathsf{P}(T < \infty) = 1$*. Indeed, by the dominated convergence theorem we deduce that then* $\big|M_T\big| \leq K$*, so that the argument in the proof above applies.*

[29] alternatively, recall the result in Example 1.26.

With suitable martingales (OST) gives very powerful results.
The following instructive example is due to D. Williams:

Example 4.29. (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 ($n \ge 1$) bets £1 on the event $\{n$th letter will be $\mathsf{A}\}$. If he loses, he leaves. If he wins, he receives £26 all of which he bets on the event $\{n+1$ st letter will be $\mathsf{B}\}$. If he loses, he leaves. If he wins, he receives £26^2 all of which he bets on the event $\{n+2$ nd letter will be $\mathsf{R}\}$ 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 until the sequence ABRACADABRA *appears. At time N, gambler $N-10$ would have won £$26^{11}-1$, gambler $N-3$ 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*

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

Remark 4.30. *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.,* $\bigl\{$ABRACADABRA, ABRACADABR, $\dots$, ABRA, ABR, AB, A, $\varnothing\bigr\}$, *as there N is just the hitting time of the state* ABRACADABRA *from the initial condition $X_0=\varnothing$.*

Exercise 4.31. *Use an appropriate* (OST) *to carefully derive the probability p_k in Example 4.17.*

Exercise 4.32. *Use an appropriate* (OST) *to carefully derive the expectation $\mathsf{E}(T)$ in Example 4.19.*

Exercise 4.33. *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 $T=\inf\bigl\{n\ge0 : S_n\notin(a,b)\bigr\}\equiv\inf\bigl\{n\ge0 : S_n\in\{a,b\}\bigr\}$.*

(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 $(M_n)_{n\geq 0}$ defined via $M_n = (S_n)^2 - n$ is a martingale w.r.t. the process $(S_n)_{n\geq 0}$.

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

Exercise 4.34. *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'*,* 'HTH'*,* 'HHTTHH'*.*

Lemma 4.35. *Let $(Y_n)_{n\geq 0}$ be a supermartingale w.r.t. a sequence $(X_n)_{n\geq 0}$ and let $H_n \in \mathcal{F}^X_{n-1} = \sigma(X_0, \dots, X_{n-1})$ satisfy $0 \leq H_n \leq c_n$, where the constant c_n might depend on n. Then the process $W_n = W_0 + \sum_{m=1}^{n} H_m\big(Y_m - Y_{m-1}\big)$, $n \geq 0$, is a supermartingale w.r.t. $(X_n)_{n\geq 0}$.*

Proof. Following the proof of the optional stopping theorem, we observe that since $(Y_n)_{n\geq 0}$ is a supermartingale w.r.t. $(X_n)_{n\geq 0}$,

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

□

Example 4.36. *If $(Y_n)_{n\geq 0}$ 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 $(W_n)_{n\geq 0}$ is a supermartingale w.r.t. $(X_n)_{n\geq 0}$, we have $\mathsf{E}W_n \leq \mathsf{E}W_0$ for all $n \geq 0$.*

Remark 4.37. *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 $(W_{n\wedge T})_{n\geq 0}$ is a supermartingale, for all $n \geq 0$ we have $\mathsf{E}(W_{n\wedge T}) \leq \mathsf{E}(W_0)$, ie., on average, the doubling strategy does not produce money if one bets against a (super)martingale.*

Example 4.38. [Wald's equation] Let $(S_n)_{n\geq 0}$ be a random walk generated by a sequence $(X_n)_{n\geq 0}$ of i.i.d. steps with $\mathsf{E}|X| < \infty$ and $\mathsf{E}(X) = m$. If T is a stopping time for $(X_n)_{n\geq 0}$ with $\mathsf{E}(T) < \infty$, then the optional stopping theorem implies that S_T is integrable and

$$\mathsf{E}(S_T - S_0) = m\,\mathsf{E}(T).$$

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

$$\mathsf{E}(S_0) = \mathsf{E}\big(S_{n\wedge T} - m(n \wedge T)\big). \tag{4.6}$$

This rearranges to

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

for every n. Now, the RHS converges to $\mathsf{E}(T)$ as $n \to \infty$ since $|\mathsf{E}(T) - \mathsf{E}(n\wedge T)| \le \mathsf{E}(T\mathsf{1}_{\{T>n\}}) = \sum_{k>n} k\mathsf{P}(T = k)$, where the tail sums on the right go to zero as $n \to \infty$ by the assumption that T is integrable. Next, by writing $|S_T| = \sum_{k\ge 0} |X_k|\mathsf{1}_{\{T\ge k\}}$ as a telescoping sum, where $\mathsf{E}(|X_k|\mathsf{1}_{\{T\ge k\}}) = \mathsf{P}(T \ge k)\mathsf{E}(|X_1|)$ and $\sum_k \mathsf{P}(T \ge k) < \infty$, we see that $|S_T|$ is integrable. [30] Similarly, we can bound $\mathsf{E}(|S_T - S_{T\wedge n}|) \le \mathsf{E}(|X_1|)\sum_{k>n} \mathsf{P}(T \ge k)$ which tends to 0 as $n \to \infty$. This implies that $\mathsf{E}(S_{T\wedge n}) \to \mathsf{E}(S_T)$ as $n \to \infty$, and combining the above completes the argument.

4.5 Martingale convergence theorem

This subsection is optional and will not be examined.

The following example has a number of important applications.

Example 4.39 (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 $(\mathcal{F}^X_n)_{n\ge 0}$.*

Indeed, as $|M_n| \le 1$ we have $\mathsf{E}|M_n| \le 1$ for all $n \ge 0$, 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 4.40. *Show that $\mathsf{P}(M_n = \frac{k}{n+2}) = \frac{1}{n+1}$ for $1 \le k \le n + 1$, ie., M_n is uniformly distributed in $\big\{\frac{1}{n+2}, \frac{2}{n+2}, \dots, \frac{n+1}{n+2}\big\}$.*

Exercise 4.40 suggests that in the limit $n \to \infty$ the distribution of M_n becomes uniform in $(0, 1)$:

Exercise 4.41. *Show that $\lim\limits_{n\to\infty} \mathsf{P}\big(M_n < x\big) = x$ for every $x \in (0, 1)$.*

In view of Exercise 4.40, 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.

[30] Here we are really using the monotone convergence theorem: if $0 \le Z_n \le Z$ for every n and $Z_n \uparrow Z$ as $n \to \infty$, then $\mathsf{E}(Z_n) \to \mathsf{E}(Z)$ as $n \to \infty$.

Example 4.42. *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,*

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

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 $k \geq 0$,

$$T_k = \min\big\{n > S_k : M_n > b\big\}\,, \qquad S_{k+1} = \min\big\{n > T_k : M_n < a\big\}\,. \tag{4.7}$$

The argument above implies that

$$\mathsf{P}(S_k < \infty) \leq \prod_{j=1}^{k}\Big(\mathsf{P}(T_{j-1} < \infty \mid S_{j-1} < \infty)\mathsf{P}(S_j < \infty \mid T_{j-1} < \infty)\Big) \tag{4.8}$$

with the RHS bounded above by $\big(\frac{a}{b}\big)^k\big(\frac{1-b}{1-a}\big)^k \to 0$ 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,*[31] *trajectories of M_n eventually converge to a particular value.*[32]

Exercise 4.43. *Write $U_{(a,b)}$ for the total number of* upcrossings *of the strip (a,b) by the process $(M_n)_{n\geq 0}$. By using the approach of Example 4.42 and noticing that $\big\{U_{(a,b)} \geq m\big\} \subset \big\{S_{m-1} < \infty\big\}$ or otherwise, show that $\mathsf{E}U_{(a,b)} < \infty$.*

The argument in Example 4.42 also works in general. Let $(M_n)_{n\geq 0}$ be a martingale w.r.t. filtration $(\mathcal{F}_n^X)_{n\geq 0}$. 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 4.44. *Let the martingale $(M_n)_{n\geq 0}$ have uniformly bounded expectations, ie., for some constant K and all $n \geq 0$, $\mathsf{E}|M_n| < K < \infty$. 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 (4.7), put $H_n = 1$ if $S_m < n \leq T_m$ 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. $(M_n)_{n\geq 0}$, cf. Lemma 4.35. It is easy to check that $W_n \geq (b-a)\,U_n - |M_n - a|$ (draw the picture!), where $U_n = \max\{m \geq 0 : T_m \leq n\}$ is the number of upcrossings of the strip (a,b) up to time n. As a result

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

so that $\mathsf{E}U_n \leq (K + |a|)/(b-a)$ for all $n \geq 0$, and thus $\mathsf{E}U_{(a,b)} < \infty$. □

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

[32] which is random and depends on the trajectory

Theorem 4.45. *Let $(M_n)_{n\geq 0}$ be a martingale as in Lemma 4.44. 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 4.44, $\mathsf{E}U_{(a,b)} < \infty$ implying that $\mathsf{P}(U_{(a,b)} = \infty) = 0$. As the number of such pairs (a, b) is countable, the result follows. □

Exercise 4.46. *Let $(X_k)_{k\geq 1}$ 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 $(\mathcal{F}_n^X)_{n\geq 0}$ martingale. Deduce that $M_n \to M_\infty$ with probability one. Can you compute $\mathsf{E}(M_\infty)$?

4.6 Additional problems

Exercise 4.47. *Let $(\eta_n)_{n\geq 1}$ be independent positive random variables with $\mathsf{E}\eta_n = 1$ for all $n \geq 1$. 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 w.r.t. $(\eta_n)_{n\geq 1}$.*

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 4.48. *Let $(S_n)_{n\geq 0}$ be the random walk from Example 4.19. Find constants a, b, c such that the process $(S_n)^4 + an(S_n)^2 + bn^2 + cn$ is an $(X_n)_{n\geq 0}$-martingale. Use the heuristc in Example 4.19 to predict the value of the second moment $\mathsf{E}(T^2)$ of the exit time T.*

Exercise 4.49. *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* '123321'*.*

Exercise 4.50. *Suppose that the process in Example 4.39 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. $(\mathcal{F}_n^X)_{n\geq 0}$.*

Exercise 4.51. *Find the large-n limit of the distribution of the martingale $(M_n)_{n\geq 0}$ from Exercise 4.50.*

Exercise 4.52. *Let $(X_n)_{n\geq 0}$ 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 $p_{m,m-1} = q_m$, $p_{m,m} = r_m$, $p_{m,m+1} = p_m$ for $m > 0$, while $p_{m,k} = 0$ for all other pairs $(m, k) \in \mathcal{S}^2$. Let $X_0 = x$, and for $y \geq 0$ denote $T_y \stackrel{\mathsf{def}}{=} \min\{n \geq 0 : X_n = y\}$.*

(a) *Show that the process $\bigl(\varphi(X_n)\bigr)_{n\geq 0}$ 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 $0 \leq a < X_0 = x < b$ we have*

$$\mathsf{P}\bigl(T_b < T_a\bigr) = \bigl(\varphi(x) - \varphi(a)\bigr)/\bigl(\varphi(b) - \varphi(a)\bigr)\,.$$

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

(c) *Now suppose that $p_m \equiv p$, $q_m \equiv q = 1 - p$, and $r_m = 0$ for $m > 0$, whereas $p_0 = p$ and $r_0 = q$. Show that in this case the result in part b) above becomes*

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

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

Exercise 4.53. *Let $(\xi_k)_{k\geq 1}$ be i.i.d. random variables with $\mathsf{P}(\xi = 1) = p < \frac{1}{2}$, $\mathsf{P}(\xi = -1) = q = 1 - p$, 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 1.26.*

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

(a) *Let $m > 1$. By using Exercise 2.3 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 $m \leq 1$? Compute $\mathsf{E}(M_\infty)$.*

Hint: *Recall Exercise 4.56.*

Exercise 4.55. *Let $(X_n)_{n\geq 0}$ 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 $(S_n)_{n\geq 0}$ be the generated random walk with $S_0 = x > 0$, and let $T = \min\{n \geq 0 : S_n = 0\}$ be the hitting time of the origin. Example 1.26 suggests that $\mathsf{E}(T) = x/(q - p) < \infty$ for $q > p$.*

1. *Use (4.6) to deduce that* $(q-p)\mathsf{E}(n \wedge T) = \mathsf{E}\big(S_0 - S_{n\wedge T}\big) \le \mathsf{E}(S_0) = x$*; then take* $n \to \infty$ *to show that* $\mathsf{E}(T) < \infty$*;*

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

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

4. *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 4.56. *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 4.57. *Let* $(X_n)_{n\ge0}$ *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* $K \ge 0$ *be such that*

$$\mathsf{E}_x\varphi(X_1) \stackrel{\mathsf{def}}{=} \mathsf{E}\big[\varphi(X_1)|X_0 = x\big] \le \varphi(x)$$

for all $x \ge K$*.*

(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* $(X_n)_{n\ge0}$*.*
Hint *If* $H_K = \min\{n \ge 0 : 0 \le X_n \le K\}$*, show that* $\varphi(X_{n\wedge H_K})$ *is a supermartingale. Deduce that if* $T_M = \min\{n \ge 0 : X_n \ge M\}$*, then* $\varphi(x) \ge \varphi(M)\mathsf{P}(T_M < H_K)$*.*

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

Exercise 4.58. *Let* $(X_k)_{k\ge1}$ *be independent variables with* $\mathsf{P}(X = \pm1) = \frac{1}{2}$*. Show that the process*

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

is a martingale w.r.t. $(\mathcal{F}^X_n)_{n\ge0}$ *and that* $\mathsf{E}\big[(M_n)^2\big] < K < \infty$ *for some constant* K *and all* $n \ge 0$*. By using Exercise 4.56 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.

Exercise 4.59. [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*

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

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 4.17 but need to justify your computation!*

Exercise 4.60. *Let $(X_n)_{n\ge0}$ 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*[33]

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

Show that $(M_n)_{n\ge0}$ with $M_n = h(X_n,n)$ is a martingale w.r.t. $(X_n)_{n\ge0}$.

Exercise 4.61. *Let $(X_n)_{n\ge0}$ 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 $M_n = \lambda^{-n}\psi\bigl(X_n\bigr)$ is a martingale w.r.t. $(X_n)_{n\ge0}$.*

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