Student Lecture Notes
taken by Mr. Duncan A. Bleich

Mathematics Section

Martingales and Brownian Motion

MATH-330 - Spring 2025

Last Update:
Wednesday 28 May, 2025

Teacher:
Robert DALANG

General comments

These notes are here to help follow the course; they do not replace it. The material presented in class is the material eligible for the exam. The notes will evolve as the course progresses, and updates will be posted on Moodle. If you have any comments/find any errors or typos/have suggestions for improving the notes, please send me an email at: duncan.bleich@epfl.ch.

Contents

1 Preliminaries

Before diving into the main topics of this course, we will first review some basic notions and results of probability theory; however, proofs of these concepts are not provided. In these notes, we consistently use both $A \subset B$ and $A \subseteq B$ to indicate that A is a subset of B, where A could be equal to B. We would use instead $A \subsetneq B$ to emphasize that A is a proper or strict subset of B, meaning $A \subset B$ but $A \neq B$. This section will be somewhat basic, as it primarily presents definitions and results. However, the subsequent sections will not be as dry, hopefully offering more intuitive explanations.

1.1 Probability spaces

Definition 1.1. *Let Ω be a nonempty set. A **σ-field** $\mathcal{F}$ on Ω is a collection of subsets of Ω, $\mathcal{F} \subset \mathcal{P}(\Omega)$, that satisfy*

(i) $\emptyset \in \mathcal{F}$,

(ii) if $A \in \mathcal{F}$, *then* $A^c \in \mathcal{F}$, *and*

(iii) if $A_i \in \mathcal{F}$ *is a countable sequence of sets, then* $\cup_i A_i \in \mathcal{F}$.

Definition 1.2. *A **measurable space**, is a pair $(\Omega, \mathcal{F})$, with Ω a nonempty set and $\mathcal{F}$ a σ-field on Ω.*

Theorem 1.3. ***Elementary properties of σ-fields.*** *Let Ω be a nonempty set and $\mathcal{F}$ a σ-field on Ω. We have:*

1. $\Omega \in \mathcal{F}$.
2. *If* $A, B \in \mathcal{F}$, *then* $A \cap B \in \mathcal{F}$.
3. *If* $A, B \in \mathcal{F}$, *then* $A \setminus B \in \mathcal{F}$.
4. *If* $A_1, A_2, \ldots \in \mathcal{F}$, *then* $\bigcap_{i \geq 1} A_i \in \mathcal{F}$.
5. *If* $A_i \uparrow A$ *(i.e.,* $A_i \subset A_{i+1}$ *and* $A := \lim_{i \to \infty} A_i = \cup_{i \geq 1} A_i$*), then* $A \in \mathcal{F}$.
6. *If* $A_i \downarrow A$ *(i.e.,* $A_i \supset A_{i+1}$ *and* $A := \lim_{i \to \infty} A_i = \cap_{i \geq 1} A_i$*), then* $A \in \mathcal{F}$.

Definition 1.4. *Let Ω be a nonempty set and $\mathcal{A} \subset \mathcal{P}(\Omega)$ be a collection of subsets of Ω. The **σ-field generated by** $\mathcal{A}$, defined by*

$$\sigma(\mathcal{A}) := \bigcap_{A \subset \mathcal{F}} \mathcal{F}$$

with $\mathcal{F}$ a σ-field on Ω, is the smallest σ-field containing $\mathcal{A}$. Note that to define $\sigma(\mathcal{A})$, we used the fact that if $\mathcal{F}_i, i \in I$ are σ-fields, then $\cap_{i \in I} \mathcal{F}_i$ is one too. This follows easily from Definition 1.1.

Definition 1.5. *Let $(\Omega, \mathcal{F})$ be a measurable space. A **measure** is a nonnegative countably additive set function; that is, a function $\mu : \mathcal{F} \to \mathbb{R}$ with*

(i) $\mu(A) \geq \mu(\emptyset) = 0$ *for all* $A \in \mathcal{F}$, *and*

(ii) if $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, *is a countable sequence of disjoint sets, then*

$$\mu\Big(\bigcup_{i \in \mathbb{N}} A_i\Big) = \sum_{i \in \mathbb{N}} \mu(A_i).$$

*If $\mu(\Omega) = 1$, we say that μ is a **probability measure**. We will usually denote probability measures by $\mathbb{P}$. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space**.*

Theorem 1.6. *Let μ be a measure on $(\Omega, \mathcal{F})$ and assume that the sets we mention are in $\mathcal{F}$. We have the following properties:*

(i) ***monotonicity:*** *If $A \subset B$, then $\mu(A) \leq \mu(B)$.*

(ii) ***subadditivity:*** *If $A \subset \bigcup_{i\geq 1} A_i$, then $\mu(A) \leq \sum_{i\geq 1} \mu(A_i)$.*

(iii) ***continuity from below:*** *If $A_i \uparrow A$, then $\mu(A_i) \uparrow \mu(A)$.*

(iv) ***continuity from above:*** *If $A_i \downarrow A$ and $\mu(A_1) < \infty$, then $\mu(A_i) \downarrow \mu(A)$.*

Remark 1.7. *Note that in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the condition $\mu(A_1) = \mathbb{P}(A_1) < \infty$ is automatically fulfilled, since $\mathbb{P}(A_1) \leq \mathbb{P}(\Omega) = 1$ by monotonicity.*

1.2 Random variables

Definition 1.8. *A real-valued function $X : \Omega \to \mathbb{R}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a **random variable** if for every $B \in \mathcal{B}$ we have $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$, where $\mathcal{B}$ denotes the Borel sets. In case we want to stress the σ-field we are referring to, we will say that X is **$\mathcal{F}$-measurable** or write $X \in \mathcal{F}$.*

Definition 1.9. *A random variable X is **discrete** if there is a finite or countable set of distinct values $\{x_1, x_2, \ldots\} \subset \mathbb{R}$ such that $p_i := \mathbb{P}(X = x_i) > 0$ for $i \geq 1$ and $\sum_{i\geq 1} p_i = 1$. If $\mathbb{P}(X = x) = 0$, for all $x \in \mathbb{R}$, the random variable X is called **continuous**. In case X has a probability density function $f_X(x)$, that is*

$$\mathbb{P}(X \in B) = \int_B f_X(x)\, dx$$

for all $B \in \mathcal{B}$, then it is necessarily continuous; however there exist continuous random variables which do not have probability densities.

Definition 1.10. *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable X, we have that X induces a probability measure on $(\mathbb{R}, \mathcal{B})$ called the law or distribution of X. This probability measure, usually denoted by μ_X or $\mathbb{P}_X$, is defined as follows:*

$$\mu_X(A) = vp_X(A) = \mathbb{P}(X^{-1}(A)).$$

We can easily check that $(\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$ is a probability space.

Definition 1.11. *Typically, we describe the distribution of a random variable X by giving its **distribution function**, $F_X(x) = \mathbb{P}(X \leq x)$. We will often write $\{X \leq x\}$ instead of $\{\omega \in \Omega : X(\omega) \leq x\}$ for clarity.*

Theorem 1.12. *Any distribution function F has the following properties:*

(i) F is nondecreasing.

(ii) $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

(iii) F *is right continuous, i.e.,* $\lim_{y \downarrow x} F(y) = F(x)$.

(iv) If $F(x-) = lim_{y \uparrow x} F(y)$, *then* $F(x-) = P(X < x)$.

(v) $\mathbb{P}(X = x) = F(x) - F(x-)$.

Theorem 1.13. *If* $X_1, X_2, \ldots$ *are random variables, then so are*

$$\sup_n X_n \quad \inf_n X_n \quad \limsup_n X_n \quad \liminf_n X_n.$$

1.3 Independence, conditional probability and independence

Definition 1.14. *We say that the* σ*-fields* $\mathcal{F}_1, \ldots, \mathcal{F}_n$, $n \in \mathbb{N}$, *are* ***independent*** *if whenever* $A_i \in \mathcal{F}_i$ *for* $i \in \{1, \ldots n\}$, *we have*

$$\mathbb{P}\big(\bigcap_{i=1}^n A_i\big) = \prod_{i=1}^n \mathbb{P}(A_i).$$

Definition 1.15. *We say that the random variables* $X_1, \ldots, X_n$, $n \in \mathbb{N}$, *are* ***independent*** *if whenever* $B_i \in \mathcal{B}$ *for* $i \in \{1, \ldots n\}$, *we have*

$$\mathbb{P}\big(\bigcap_{i=1}^n \{X_i \in B_i\}\big) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i).$$

Definition 1.16. *We say that the sets* $A_1, \ldots, A_n$, $n \in \mathbb{N}$, *are* ***independent*** *if whenever* $I \subset \{1, \ldots, n\}$, *we have*

$$\mathbb{P}\big(\bigcap_{i \in I} A_i\big) = \prod_{i \in I} \mathbb{P}(A_i).$$

Remark 1.17. *An infinite collection of objects (*σ*-fields, random variables, or sets) is said to be independent if every finite subcollection is.*

Theorem 1.18. *If the random variables* X *and* Y *are independent, then the* σ*-fields* $\sigma(X)$ *and* $\sigma(Y)$ *are also independent. Moreover, if* $X \in \mathcal{F}$ *and* $Y \in \mathcal{G}$, *with* $\mathcal{F}$ *and* $\mathcal{G}$ *independent, then* X *and* Y *are independent.*

Definition 1.19. *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and an event* $B \in \mathcal{F}$ *of nonzero probability, the* ***conditional probability*** *of event* $A \in \mathcal{F}$ *given* B *is defined by*

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Lemma 1.20. ***Law of Total Probabilities.*** *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *a partition of* Ω *into events* $A_1, \ldots A_N$ *and any event* B, *we have*

$$\mathbb{P}(B) = \sum_{i=1}^N \mathbb{P}(A_i)\mathbb{P}(A \,|\, A_i).$$

In case $\mathbb{P}(A_i) = 0$, *we can give* $\mathbb{P}(B|A_i)$ *any value in* $[0, 1]$ *and the above formula still holds. It will also hold for countable partitions.*

Remark 1.21. *As a reminder, we say that the events $A_1, \ldots, A_N \in \mathcal{F}$ form a partition of Ω if they are disjoint and such that $\Omega = \cup_{i=1}^N A_i$. The case of countable partitions is defined analogously.*

Lemma 1.22. *Given events $A_1, \ldots, A_N$, a simple induction argument gives*

$$\mathbb{P}\big(\bigcap_{i=1}^N A_i\big) = \mathbb{P}(A_1)\mathbb{P}(A_2 \,|\, A_1)\mathbb{P}(A_3 \,|\, A_1 \cap A_2) \ldots \mathbb{P}(A_N \,|\, A_1 \cap A_2 \cap \ldots \cap A_{N-1}).$$

Definition 1.23. *Two events A and B are said to be* ***conditionally independent*** *given a third event C if*

$$\mathbb{P}(A \cap B \,|\, C) = \mathbb{P}(A \,|\, C)\mathbb{P}(B \,|\, C).$$

Definition 1.24. *Two random variables X and Y are said to be* ***conditionally independent*** *given a third random variable Z if*

$$\mathbb{P}(X \in A, Y \in B \,|\, Z) = \mathbb{P}(X \in A \,|\, Z)\mathbb{P}(Y \in B \,|\, Z)$$

for all $A, B \in \mathcal{B}$.

1.4 Expectation

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $X \geq 0$, we define its **expected value** to be $\mathbb{E}[X] = \int_\Omega X d\mathbb{P}$, which will always make sense but may be infinite. To get the general case, define the **positive part** of x as $x^+ = \max(0, x)$ and the **negative part** of x as $x^- = \min(0, -x)$, respectively. We say that $\mathbb{E}[X]$ exists and set $\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$, whenever the subtraction makes sense, i.e., when $\mathbb{E}[X^+] < \infty$ or $\mathbb{E}[X^-] < \infty$.

Theorem 1.25. *Suppose $X, Y \geq 0$ or $E|X|, \mathbb{E}|Y| < \infty$, we have*

1. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
2. $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$, *for any real numbers a, b.*
3. *If $X \geq Y$, then $\mathbb{E}[X] \geq \mathbb{E}[Y]$.*

Theorem 1.26. ***Jensen's inequality.*** *Suppose φ is convex, that is,*

$$\lambda\varphi(x) + (1 - \lambda)\varphi(y) \geq \varphi(\lambda x + (1 - \lambda)y)$$

for all $\lambda \in [0, 1]$ and $x, y \in \mathbb{R}$. Then

$$\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X])$$

provided both expectations exist, i.e., $\mathbb{E}|X|$, $\mathbb{E}|\varphi(X)| < \infty$.

Theorem 1.27. ***Hölder's inequality.*** *If $p, q \in [1, \infty]$ with $1/p + 1/q = 1$, then*

$$\mathbb{E}|XY| \leq ||X||_p ||Y||_q$$

where $||X||_r = (\mathbb{E}|X|^r)^{1/r}$ for $r \in [1, \infty)$ and $||X||_\infty = \inf\{M : \mathbb{P}(|X| > M) = 0\}$.

Remark 1.28. *Before we state the next result, note that we will often write*

$$\mathbb{E}[X; A] = \int_A X d\mathbb{P}$$

when integrating over a subset $A \subset \Omega$.

Theorem 1.29. ***Markov's inequality.*** *Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ has $\varphi \geq 0$, let $A \in \mathcal{B}$ and $i_A = \inf\{\varphi(y) : y \in A\}$. We have*

$$i_A \mathbb{P}(X \in A) \leq \mathbb{E}[\varphi(X); X \in A] \leq \mathbb{E}[\varphi(X)].$$

Should you wish to gain a deeper understanding of the notions previously mentioned, I strongly recommend reading the book *Introduction à la théorie des probabilités* [1] by R. Dalang and D. Conus. In case French is not your forte, the first chapter of the book *Probability : Theory and Examples* [2] by R.Durrett is a good alternative. Both are available in digital formats for free online. We now give two special cases before moving on to the last part of the section.

Statement 1.30. *If X is a discrete random variable taking values in $\{x_1, x_2, \dots\} \subset \mathbb{R}$, then*

$$\mathbb{E}[X] = \sum_i x_i \mathbb{P}(X = x_i)$$

if X is integrable, i.e., if $\sum_i |x_i| \mathbb{P}(X = x_i) < \infty$. Moreover, for $h : \mathbb{R} \to \mathbb{R}$, we have that $\mathbb{E}[h(X)] = \sum_i h(x_i) \mathbb{P}(X = x_i)$ if $\sum_i |h(x_i)| \mathbb{P}(X = x_i) < \infty$.

Statement 1.31. *If X has a probability density function $f_X(x)$, then*

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

if $\mathbb{E}|X| = \int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$. Note that $f_X \geq 0$. Similarly, for $h : \mathbb{R} \to \mathbb{R}$, we have $\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx$ if $\mathbb{E}|h(X)| = \int_{-\infty}^{\infty} |h(x)| f_x(x) dx < \infty$.

1.5 Conditional expectation

Definition 1.32. ***Conditional expectation.*** *Let (X, Y) be a random vector in $\mathbb{R} \times \mathbb{R}^m$, for some $m \geq 1$, such that $\mathbb{E}|X| < \infty$. Then $\mathbb{E}[X \mid Y]$ is the integrable random variable of the form $\varphi(Y)$, where $\varphi : \mathbb{R}^m \to \mathbb{R}$ is such that for all bounded functions $h : \mathbb{R}^m \to \mathbb{R}$, we have*

$$\mathbb{E}\big(\mathbb{E}[X \mid Y] \cdot h(Y)\big) = \mathbb{E}[X \cdot h(Y)].$$

If $\mathbb{E}[X^2] < \infty$, then $\mathbb{E}[X \mid Y] = \varphi(Y)$ is such that

$$\mathbb{E}[(X - \varphi(Y))^2] \leq \mathbb{E}[(X - \psi(Y))^2]$$

for all $\psi : \mathbb{R}^m \to \mathbb{R}$.

Remark 1.33. *If we take the function $h \equiv 1$ in the above definition of $\mathbb{E}[X|Y]$, we obtain*

$$\mathbb{E}\big(\mathbb{E}[X|Y]\big) = \mathbb{E}[X].$$

In words, the expectation of the conditional expectation of X given Y is the "unconditional" expectation of X. Moreover, conditional expectation is unique almost surely, meaning that any two versions of the same conditional expectation differ only on a set of probability zero.

We now give some formulas, for $\mathbb{E}[X|Y]$, in both the discrete and continuous case.

Discrete case: Let (X, Y) be a discrete random vector with $\mathbb{E}|X| < \infty$. We proceed in two steps:

(a) Set $\mathbb{E}[X \,|\, Y = y] := \sum_x x\mathbb{P}(X = x \,|\, Y = y) = \mathbb{P}(Y = y)^{-1} \cdot \sum_x x\mathbb{P}(X = x, Y = y)$, if $\mathbb{P}(Y = y) > 0$.

(b) Define the function $\varphi : \mathbb{R}^m \to \mathbb{R}$ by $\varphi(y) = \mathbb{E}[X \,|\, Y = y]$, for y such that $\mathbb{P}(Y = y) > 0$, then $\mathbb{E}[X|Y] = \varphi(Y)$

Continuous case: Similarly, consider a continuous random vector (X, Y) taking values in $\mathbb{R} \times \mathbb{R}^m$, with $\mathbb{E}|X| < \infty$. Additionally, assume that the random vector (XY) has a joint probability density function given by $f_{X,Y} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, meaning

$$\mathbb{P}(X \in A, Y \in B) = \int_A dx \int_B dy \; f_{X,Y}(x, y),$$

for all $A \subset \mathbb{R}, B \subset \mathbb{R}^m$. Before going further, note that we will sometimes write

$$\int_{A_1} dx_1 \int_{A_2} dx_2 \ldots \int_{A_n} dx_n \, f(x_1, \ldots, x_n)$$

instead of the usual

$$\int_{A_1} \cdots \int_{A_n} f(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

for clarity. This notation may initially appear superfluous; however, it is essential for ensuring accurate integration over the appropriate sets, thereby avoiding potential confusion. Now recall that the marginal density of Y is given by

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) dx.$$

As done in the discrete case, we proceed in two steps:

(a) Set

$$\mathbb{E}[X \,|\, Y = y] = \int_{\mathbb{R}} x \, \frac{f_{X,Y}(x, y)}{f_Y(y)} \, dx, \qquad \text{for } f_y(y) > 0.$$

Where, $f_{X,Y}(x, y) \cdot f_Y(y)^{-1}$ is the conditional density of X given $Y = y$.

(b) Define the function $\varphi(y) = \mathbb{E}[X \,|\, Y = y]$ for $y \in \mathbb{R}^m$ such that $f_Y(y) > 0$, then $\mathbb{E}[X|Y] = \varphi(Y)$.

If we have a function $h : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, we can generalize the previous case by setting

$$\mathbb{E}[h(X, Y) \,|\, Y = y] = \int_{\mathbb{R}} h(x, y) \frac{f_{X,Y}(x, y)}{f_Y(y)} \, dx,$$

for $f_Y(y) > 0$. If we set $\psi(y) = \mathbb{E}[h(X, Y) \,|\, Y = y]$, then $\mathbb{E}[h(X, Y) \,|\, Y] = \psi(Y)$.

Proposition 1.34. ***Properties of conditional expectation.*** *Let (X, Y, Z) be a discrete (or continuous) random vector defined on a probability space, and let $f : \mathbb{R} \to \mathbb{R}$ be a bounded function.*

(A) ***Linearity:*** *For all* $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}[\alpha Y + \beta Z \mid X] = \alpha\, \mathbb{E}[Y|X] + \beta\, \mathbb{E}[Z|X]$.

(B) ***Monotonicity:*** *If* $Y \leq Z$, *then* $\mathbb{E}[Y|X] \leq \mathbb{E}[Z|X]$.

(C) ***Iteration:*** $\mathbb{E}\big(\mathbb{E}[Z \mid X, Y] \mid X\big) = \mathbb{E}[Z|X]$.

(D) $\mathbb{E}[Y f(X)|X] = \mathbb{E}[Y|X] f(X)$

(E) *If* X *and* Y *are independent, then* $\mathbb{E}[Y|X] = \mathbb{E}[Y]$.

(F) *If* $Y = f(X)$, *then* $\mathbb{E}[Y|X] = Y$.

(G) ***Jensen's inequality:*** *If* $g : \mathbb{R} \to \mathbb{R}$ *is a convex function and* $\mathbb{E}|g(Y)| < \infty$, *then* $g(\mathbb{E}[Y|X]) \leq \mathbb{E}[g(Y)|X]$.

Initially, we will denote the use of properties $(A), (B), ..., (G)$ *explicitly, but as the course progresses, we will phase out this systematic notation..*

Statement 1.35. *Using the notation* $\varphi(y) = \mathbb{E}[X \mid Y = y]$ *and referring to Remark 1.33, we can formulate the* ***law of total probability*** *in the* ***discrete*** *case:*

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\varphi(Y)] \\
&= \sum_y \varphi(y) \mathbb{P}(Y = y) \\
&= \sum_y \mathbb{E}[X \mid Y = y] \mathbb{P}(Y = y).
\end{aligned}
$$

The ***continuous*** *case is derived in a similar manner:*

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\varphi(Y)] \\
&= \int_{\mathbb{R}^m} \varphi(y)\, f_Y(y) dy \\
&= \int_{\mathbb{R}^m} \mathbb{E}[X \mid Y = y] f_Y(y) dy.
\end{aligned}
$$

Remark 1.36. *Conditional probabilities are a special case of conditional expectation. For a given event* G, *we have*

$$\mathbb{P}(G|Y) = \mathbb{E}[\mathbb{1}_G \mid Y]$$

with

$$\mathbb{1}_G(\omega) = \begin{cases} 1 & \text{if } \omega \in G \\ 0 & \text{if } \omega \notin G \end{cases}.$$

Recall that $\mathbb{P}(G) = \mathbb{E}[\mathbb{1}_G]$, *as* $\mathbb{E}[\mathbb{1}_G] = 1 \cdot \mathbb{P}(\mathbb{1}_G = 1) + 0 \cdot \mathbb{P}(\mathbb{1}_G = 0) = \mathbb{P}(G)$.

2 Martingales

2.1 Definitions and examples

Before we delve into the definitions, let us first consider a motivating scenario: Imagine a player engaged in a sequence of fair games of chance. At each stage of the game, the expected reward, given all previously observed outcomes, is zero. Let S_n denote the cumulative gain (or loss) after n games. A natural question to ask is whether the player can ensure a positive expected reward upon exiting the game. Specifically, is there a random time τ, not predetermined, such that $\mathbb{E}[S_\tau] > 0$? Here, we can think of the random time τ as a strategic decision point—the player plans to exit the game when certain known conditions are met, although it is not known when these conditions will occur.

Martingales provide a powerful framework for analyzing this scenario. They allow us to understand the dynamics of gains and losses under fair game assumptions and can help us assess strategies that involve exiting at a random time τ.

We will discover that for a 'reasonable' $\tau, \mathbb{E}[S_\tau] = 0$. Should $\mathbb{E}[S_\tau] > 0$, then τ would represent a strategy that, under typical circumstances, most people would choose to avoid.

Definition 2.1. *Let $(X_n)_{n\geq 1}$ be a sequence of random variables. A sequence $(S_n)_{n\geq 1}$ of random variables is a **martingale** relative to $(X_n)_{n\geq 1}$ if for all $n \geq 1$*

(a) $\mathbb{E}[\,|S_n|\,] < \infty$,

(b) $\mathbb{E}[S_{n+1} \,|\, X_1, \ldots, X_n] = S_n$.

Remark 2.2. *If $(S_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$, then S_n is a function of $X_1, \ldots, X_n$. Indeed, by point (b) of the definition above, $S_n = \psi_n(X_1, \ldots, X_n)$ where $\psi_n(x_1, \ldots, x_n) = \mathbb{E}[S_{n+1} \,|\, X_1 = x_1, \ldots, X_n = x_n]$.*

Remark 2.3. *Often, $S_n = X_n$.*

It is useful to think of $(X_1, \ldots, X_n)$ as the accumulated information or history up to stage n. In a gambling context, this historical record may encompass more than just the sequence of past fortunes; it could, for example, include the outcomes of plays in which the player did not bet.

Returning to our motivating example, let S_n represent the player's wealth at time n. While we do not assign a specific meaning to $(X_1, \ldots, X_n)$, it encompasses everything observed up to time n. Condition (b) of Definition 2.1 reflects one interpretation of a fair game, stipulating that the player's expected wealth in the next play, given all previous observations, should equal his current wealth: $\mathbb{E}[S_{n+1}|X_1, \ldots, X_n] = S_n$. Let us now consider a more concrete situation.

The "double or nothing" strategy: The rules of the game are straightforward: With a probability of $1/2$, the player doubles his bet; conversely, with the same probability, he loses his initial bet. Therefore, if the player bets x francs, the potential profit or loss can be expressed as follows:

$$\text{Profit} = \begin{cases} x & \text{, with proability } 1/2 \\ -x & \text{, with probability } 1/2 \end{cases}.$$

Strategy: Bet 1 franc, then $2, 4, \ldots, 2^n$ until I win a bet, then leave the game. The total gain, given I lose at games 1 to n and finally win at game $n+1$, is:

$$-1 - 2 - 2^2 - \ldots - 2^{n-1} + 2^n = 1.$$

The strategy effectively 'guarantees' a profit of 1. At first glance, this might seem like a viable approach. While the guaranteed profits are modest—arguably negligible—they are still profits. This raises the question: Why aren't people flocking to casinos to capitalize on this strategy? The following discussion will address this question.

Let T be the number of games played up to the first win. It is easy to see that

$$\mathbb{P}(T=1)=\frac{1}{2},\, \mathbb{P}(T=2)=\frac{1}{4},\dots,\, \mathbb{P}(T=n)=\frac{1}{2^n} \tag{1}$$

and therefore

$$\mathbb{P}(T<\infty)=\mathbb{P}\big(\bigcup_{n=1}^{\infty}\{T=n\}\big)=\sum_{n=1}^{\infty}\mathbb{P}(T=n)=\sum_{n=1}^{\infty}\frac{1}{2^n}=1$$

where the first equality arises from the fact that the sets $\{T=n\}$ are disjoint. If we now look at $L :=$ *the amount lost up to time* $T-1$, we find that

$$\begin{aligned}
\mathbb{E}[L] &= \sum_{n=1}^{\infty}\mathbb{E}[L\,|\,T=n]\cdot\mathbb{P}(T=n)\\
&= \sum_{n=2}^{\infty}\mathbb{E}[L\,|\,T=n]\cdot\mathbb{P}(T=n)\\
&= \sum_{n=2}(1+2+\ldots+2^{n-2})\cdot 2^{-n}\\
&= \sum_{n=2}\frac{2^{n-1}-1}{2^n}=+\infty \quad ,\text{since } \frac{2^{n-1}-1}{2^n}\overset{n\to\infty}{\longrightarrow}\frac{1}{2}.
\end{aligned}$$

From this, we can deduce that this strategy will usually lead to bankruptcy. To better understand martingales, let's examine the following four examples.

Example 2.4. *Let $(X_n)_{n\geq 1}$ be independent and identically distributed (i.i.d.) random variables with $\mathbb{E}[X_1]=0$ (includes $\mathbb{E}[\,|X_1|\,]<\infty$). Define $S_n=X_1+\ldots+X_n$, for $n\geq 1$. Then $(S_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$.*

Proof. We prove both (a) and (b) of Definition 2.1.

(a)

$$\mathbb{E}[\,|S_n|\,]=\mathbb{E}[\,|X_1+\ldots+X_n|\,]\leq\mathbb{E}[\,|X_1|+\ldots+|X_n|\,]\overset{(i.i.d.)}{=}n\mathbb{E}[\,|X_1|\,]<\infty$$

(b)

$$\begin{aligned}
\mathbb{E}[S_{n+1}\,|\,X_1,\dots,X_n] &= \mathbb{E}[X_{n+1}+S_n\,|\,X_1,\dots,X_n]\\
&\overset{(A)}{=} \mathbb{E}[X_{n+1}\,|\,X_1,\dots,X_n]+\mathbb{E}[S_n\,|\,X_1,\dots,X_n]\\
&\overset{(E),(F)}{=} \mathbb{E}[X_{n+1}]+S_n=S_n
\end{aligned}$$

□

Example 2.5. *Let $(X_n)_{n\geq 1}$ be i.i.d. with $\mathbb{E}[X_1]=0$ and $\sigma^2=\mathbb{E}[X_1^2]<\infty$. Define $Z_n=\big(\sum_{i=1}^{n}X_i\big)^2-n\sigma^2=S_n^2-n\sigma^2$. Then $(Z_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$.*

Proof. (a) We have

$$\mathbb{E}[\,|Z_n|\,] \leq \mathbb{E}[S_n^2] + n\sigma^2 \overset{(*)}{=} \mathbb{E}[\sum_{i=1}^{n} X_1^2] + n\sigma^2$$
$$\overset{(*)}{=} n\sigma^2 + n\sigma^2 = 2n\sigma^2 < \infty.$$

The equalities marked by $(*)$ hold true because the variables X_n are i.i.d. with a mean of 0.

(b)

$$\begin{aligned}
\mathbb{E}[Z_{n+1}\,|\,X_1,\dots,X_n] &= \mathbb{E}[\big(X_{n+1}+S_n\big)^2 - (n+1)\sigma^2\,|\,X_1,\dots,X_n] \\
&= \mathbb{E}[X_{n+1}^2 + 2X_{n+1}S_n + S_n^2 - (n+1)\sigma^2 n\,|\,X_1,\dots,X_n] \\
&\overset{(A)}{=} \mathbb{E}[X_{n+1}^2\,|\,X_1,\dots,X_n] + 2\mathbb{E}[X_{n+1}S_n\,|\,X_1,\dots,X_n] \\
&\quad + \mathbb{E}[S_n^2\,|\,X_1,\dots,X_n] - \mathbb{E}[(n+1)\sigma^2\,|\,X_1,\dots,X-n] \\
&\overset{(E),(D),(F)}{=} \mathbb{E}[X_{n+1}^2] + 2S_n\mathbb{E}[X_{n+1}\,|\,X_1,\dots,X_n] + S_n^2 - (n+1)\sigma^2 \\
&\overset{(E)}{=} \sigma^2 + 2S_n\mathbb{E}[X_{n+1}] + S_n^2 - (n+1)\sigma^2 \\
&= Z_n
\end{aligned}$$

□

Example 2.6. ***Likelihood ratio:*** *Let $(X_n)_{n\geq 1}$ be i.i.d. random variables and f_0, f_1 be two bounded positive density functions ($f_i > 0, \int_{\mathbb{R}} f_i(x)dx = 1$). Define*

$$R_n = \frac{f_1(X_1)\dots f_1(X_n)}{f_0(X_1)\dots f_0(X_n)} = \frac{\prod_{i=1}^n f_1(X_i)}{\prod_{i=1}^n f_0(X_i)}$$

Statement 2.7. *If the common density of the X_n is f_0, then $(R_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$.*

Proof. (a) We write $\mathbb{E}_0$ to indicate that we take the expectation with respect to the common density function f_0.

$$\begin{aligned}
\mathbb{E}_0[\,|R_n|\,] = \mathbb{E}_0[R_n] &= \mathbb{E}_0\left[\frac{\prod_{i=1}^n f_1(X_i)}{\prod_{i=1}^n f_0(X_i)}\right] \\
&= \int_{\mathbb{R}^n} \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=1}^n f_0(x_i)} \cdot \prod_{i=1}^n f_0(x_i)\,dx_1\dots dx_n \\
&= \int_{\mathbb{R}^n} f_1(x_1)\dots f_1(x_n)dx_1\dots dx_n \\
&= 1 < \infty
\end{aligned}$$

(b)

$$\begin{aligned}
\mathbb{E}_0[R_{n+1} \mid X_1, \ldots, X_n] &= \mathbb{E}_0\left[R_n \cdot \frac{f_1(X_{n+1})}{f_0(X_{n+1})} \mid X_1, \ldots, X_n\right] \\
&\overset{(D)}{=} R_n \cdot \mathbb{E}_0\left[\frac{f_1(X_{n+1})}{f_0(X_{n+1})} \mid X_1, \ldots, X_n\right] \\
&\overset{(E)}{=} R_n \cdot \mathbb{E}_0\left[\frac{f_1(X_{n+1})}{f_0(X_{n+1})}\right] \\
&\overset{(i.i.d.)}{=} R_n \cdot \mathbb{E}_0\left[\frac{f_1(X_1)}{f_0(X_1)}\right] \\
&\overset{(a)}{=} R_n \cdot 1 = R_n
\end{aligned}$$

□

Example 2.8. ***Doob martingale:*** *Consider arbitrary random variables $X_1, X_2, \ldots$ and an integrable random variable X, i.e., $\mathbb{E}[|X|] < \infty$. Define $M_n = \mathbb{E}[X \mid X_1, \ldots, X_n]$ for each n. Then, the sequence $(M_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$.*

Proof. (a) Since the function $x \mapsto |x|$ is convex, we can use Jensen's inequality:

$$\begin{aligned}
\mathbb{E}[\,|M_n|\,] &= \mathbb{E}\big(\,|\mathbb{E}[X \mid X_1, \ldots, X_n]|\,\big) \\
&\overset{(G)}{\leq} \mathbb{E}\big(\,\mathbb{E}[\,|X| \mid X_1, \ldots, X_n]\,\big) \\
&= \mathbb{E}[\,|X|\,] < \infty.
\end{aligned}$$

We noted in Remark 1.33 that the last equality is valid.

(b)

$$\begin{aligned}
\mathbb{E}[M_{n+1} \mid X_1, \ldots, X_n] &= \mathbb{E}\big(\,\mathbb{E}[X \mid X_1, \ldots, X_{n+1}] \mid X_1, \ldots, X_n\big) \\
&\overset{(C)}{=} \mathbb{E}[X \mid X_1, \ldots, X_n] \\
&= M_n
\end{aligned}$$

.

□

Lemma 2.9. *If $(S_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$, then*

(a) $\mathbb{E}[S_{n+m} \mid X_1, \ldots, X_n] = S_n$*, for all* $n, m \geq 1$*,*

(b) $\mathbb{E}[S_n] = \mathbb{E}[S_1]$*, for all* $n \geq 1$*.*

Proof. (a)

$$\begin{aligned}
\mathbb{E}[S_{n+m} \mid X_1, \ldots, X_n] &\overset{(C)}{=} \mathbb{E}\big(\,\mathbb{E}[S_{n+m} \mid X_1, \ldots, X_n, \ldots X_{n+m-1}] \mid X_1, \ldots, X_n\big) \\
&\overset{(*)}{=} \mathbb{E}[S_{n+m-1} \mid X_1, \ldots, X_n] \\
&= \ldots \\
&= \mathbb{E}[S_{n+1} \mid X_1, \ldots, X_n] \\
&\overset{(*)}{=} S_n.
\end{aligned}$$

The equalities denoted by $(*)$ hold by Definition 2.1.

(b)

$$\mathbb{E}[S_n] = \mathbb{E}\big(\mathbb{E}[S_n \,|\, X_1]\big) \overset{(a)}{=} \mathbb{E}[S_1]$$

□

2.2 Supermartingales and submartingales

Definition 2.10. *We say that $(S_n)_{n\geq 1}$ is a **supermartingale** relative to $(X_n)_{n\geq 1}$ if*

(a) $\mathbb{E}[\,|S_n|\,] < \infty$*,*

(b) $\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n] \leq S_n$ *and*

(c) S_n *is a function of* $X_1, \dots, X_n$*.*

Definition 2.11. *We say that $(S_n)_{n\geq 1}$ is a **submartingale** relative to $(X_n)_{n\geq 1}$ if*

(a') $\mathbb{E}[\,|S_n|\,] < \infty$*,*

(b') $\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n] \geq S_n$ *and*

(c') S_n *is a function of* $X_1, \dots, X_n$*.*

Remark 2.12. *(1) For a martingale, $S_n = \mathbb{E}[S_{n+1} \,|\, X_+, \dots, X_n]$, so S_n is a function of $X_1, \dots, X_n$.*

(2) If $(S_n)_{n\geq 1}$ is a supermartingale, then $(-S_n)_{n\geq 1}$ is a submartingale, and vice-versa.

(3) $(S_n)_{n\geq 1}$ is a martingale if and only if $(S_n)_{n\geq 1}$ is both a submartingale and a supermartingale.

Lemma 2.13. *(a) Let $(S_n)_{n\geq 1}$ be a martingale relative to $(X_n)_{n\geq 1}$. If $\varphi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}[\,|\varphi(S_n)|\,] < \infty$ for all n, then $(\varphi(S_n))_{n\geq 1}$ is a submartingale.*

(b) Let $(S_n)_{n\geq 1}$ be a submartingale and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex and nondecreasing function such that $\mathbb{E}[\,|\varphi(S_n)|\,] < \infty$ for all n. Then $(\varphi(S_n))_{n\geq 1}$ is a submartingale.

Proof. (a) Conditions (a') and (c') of Definition 2.11 are met by our assumptions. We only need to check (b'):

$$\mathbb{E}[\varphi(S_{n+1}) \,|\, X_1, \dots, X_n] \overset{(G)}{\geq} \varphi(\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n]) = \varphi(S_n)$$

(b) Again, conditions (a') and (c') are met by our assumptions. For (b'), we have

$$\mathbb{E}[\varphi(S_{n+1}) \,|\, X_1, \dots X_n] \overset{(G)}{\geq} \varphi(\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n]) \geq \varphi(S_n).$$

The last inequality follows from $\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n] \geq S_n$, since S_n is a submartingale and φ is nondecreasing.

□

Remark 2.14. *If $(S_n)_{n\geq 1}$ is a supermartingale (resp. submartingale), then for $m, n \geq 1$:*

$$\mathbb{E}[S_{n+m} \,|\, X_1, \dots, X_n] \leq S_n \quad \textit{(resp. } \geq S_n\textit{)}$$

and

$$\mathbb{E}[S_n] \leq \mathbb{E}[S_1] \quad \textit{(resp. } \geq \mathbb{E}[S_1]\textit{).}$$

2.3 Stopping times

Motivation: The concept of stopping times is motivated by scenarios such as a player choosing to exit a sequence of games at a random time that is determined by the outcomes of previous games, yet independent of future outcomes.

Definition 2.15. *Let $(X_n)_{n\geq 1}$ be an observation sequence (real-valued). A* ***stopping time*** *T, relative to $(X_n)_{n\geq 1}$, is a random variable taking values in $\mathbb{N} \cup \{\infty\}$ such that for all $k \in \mathbb{N}$, there is $B_k \subset \mathbb{R}^k$ with*

$$\{T = k\} = \{(X_1, \ldots, X_k) \in B_k\}.$$

Where $\{T = k\}$ can be interpreted as the decision to stop at time k and B_k as a property of the observations $X_1, \ldots, X_k$.

Remark 2.16. *From this point onward, we will write $A \in \sigma(X_1, \ldots, X_k)$ to say that the event A is determined by $(X_1, \ldots, X_k)$, where $\sigma(X_1, \ldots, X_k)$ is the σ-field generated by $X_1, \ldots, X_k$. This σ-field represents the information available from observing the process up to time k. Moreover, while we did not specify what type of sets B_k we are referring to in the definition, some might find it useful to know that they belong to the Borel sets over $\mathbb{R}^k$, i.e., $B_k \in \mathcal{B}^k$.*

Example 2.17. *Let $(X_n)_{n\geq 1} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. At stage n, we observe X_n. Let T be the first time that $X_n \geq 2$. Then T is a stopping time.*

Proof. For $k = 1$, we have

$$\{T = 1\} = \{X_1 \geq 2\} = \{X_1 \in B_1\} \in \sigma(X_1),$$

where $B_1 = [2, \infty)$. If $k \geq 2$,

$$\{T = k\} = \{X_1 < 2, X_2 < 2, \ldots, X_k \geq 2\} = \{(X_1, \ldots, X_k) \in B_k\} \in \sigma(X_1, \ldots, X_k),$$

with $B_k = \overbrace{(-\infty, 2) \times \ldots \times (-\infty, 2)}^{k-1 \text{ factors}} \times [2, \infty)$. □

Proposition 2.18. ***Some properties of stopping times.***

(a) If T is a stopping time, then $\{T \leq k\}, \{T > k\}, \{T < k\}$ and $\{T \geq k\}$ are events in $\sigma(X_1, \ldots, X_k)$.

(b) If $\{T \leq k\} \in \sigma(X_1, \ldots, X_k)$ for all $k \geq 1$, then T is a stopping time.

(c) If S and T are stopping times, then $S \wedge T := \min\{S, T\}$ and $S \vee T := \max\{S, T\}$ are also stopping times.

(d) If T is deterministic, i.e., if there exists $k_0 \in \mathbb{N}$ such that $\mathbb{P}(T = k_0) = 1$, then T is also a stopping time.

(e) If T is a stopping time, then for $k \geq 1$, $\mathbb{1}\{T = k\}$ is a function of $(X_1, \ldots, X_k)$

Remark 2.19. *We will often write $\mathbb{1}\{A\}$ instead of $\mathbb{1}_A$.*

Proof. (a) First, observe that

$$\{T \leq k\} = \bigcup_{i=1}^{k} \{T = i\} \in \sigma(X_1, \ldots, X_k)$$

since $\{T = i\} \in \sigma(X_1, \ldots, X_i) \subset \sigma(X_1, \ldots, X_i, \ldots, X_k)$ for $i \leq k$. We then immediately find that

$$\{T < k\} = \{T \leq k-1\} \in \sigma(X_1, \ldots, X_{k-1}) \subset \sigma(X_1, \ldots, X_k).$$

From this, using the properties of σ-fields gives

$$\{T > k\} = \{T \leq k\}^c \in \sigma(X_1, \ldots, X_k)$$

and

$$\{T \geq k\} = \{T < k\}^c \in \sigma(X_1, \ldots, X_k).$$

(b) For all $k \geq 1$,

$$\{T = k\} = \{T \leq k\} \cap \{T \leq k-1\}^c \in \sigma(X_1, \ldots, X_k).$$

(c) Follows from (b) and the following observations

$$\{S \wedge T \leq k\} = \{S \leq k\} \cup \{T \leq k\},$$
$$\{S \vee T \leq k\} = \{S \leq k\} \cap \{T \leq k\}.$$

(d) For $k \neq k_0$,

$$\{T = k\} = \{(X_1, \ldots, X_k) \in \emptyset\} \in \sigma(X_1, \ldots, X_k).$$

For $k = k_0$,

$$\{T = k_0\} = \{(X_1, \ldots, X_{k_0}) \in \mathbb{R}^{k_0}\} \in \sigma(X_1, \ldots, X_{k_0}).$$

(e) Since $\{T = k\} = \{(X_1, \ldots, X_k) \in B_k\}$ for some $B_k \subset \mathbb{R}^k$, we have

$$\mathbb{1}_{\{T=k\}} = \mathbb{1}_{B_k}(X_1, \ldots, X_k)$$

where $\mathbb{1}_{B_k} : \mathbb{R}^k \to R$ is defined by

$$\mathbb{1}_{B_k}(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } (x_1, \ldots, x_k) \in B_k \\ 0 & \text{otherwise} \end{cases} .$$

□

2.4 Optional Stopping Theorem

The goal of this subsection is to prove the following theorem:

Theorem. *Let* $(S_n)_{n \geq 1}$ *be a martingale and* T *a stopping time relative to* $(X_n)_{n \geq 1}$*. Suppose*

(a) $\mathbb{P}(T < \infty) = 1$,

(b) $\mathbb{E}[\,|S_T|\,] < \infty$ *and*

(c) $\lim_{n\to\infty} \mathbb{E}[S_n \,|\, T > n] \cdot \mathbb{P}(T > n) = 0$.

Then $\mathbb{E}[S_T] = \mathbb{E}[S_1]$.

To get there, we'll first look at three important lemmas that need to be proved. But before we jump into those, let us consider the following: Let $(S_n)_{n\geq 1}$ be a martingale relative to $(X_n)_{n\geq 1}$. Then $\mathbb{E}[S_n] = \mathbb{E}[S_1]$ for all $n \geq 1$. If we decide in advance to play n games, then the game is fair. However, could a clever player (who doesn't cheat) do better ? If the player decides to stop at a random time T, that is a stopping time. Would it be possible to have $\mathbb{E}[S_T] > \mathbb{E}[S_1]$?

Example 2.20. *Let X_n be i.i.d. random variables with $\mathbb{P}(X_1 = +1) = 1/2 = \mathbb{P}(X_1 = -1)$. Set $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$ for $n \geq 1$. Then $(S_n)_{n\geq 1}$ is a martingale relative to $(X_n)_{n\geq 1}$ and we have that $\mathbb{E}[S_1] = 0$. Define $T = \inf\{n \geq 1 \,:\, S_n = 1\}$. We will see that $\mathbb{P}(T < \infty) = 1$, so $S_T = 1$ and $\mathbb{E}[S_T] = 1 > 0 = \mathbb{E}[S_1]$. The stopping time T allows us to be in a situation where $\mathbb{E}[S_T] > \mathbb{E}[S_1]$, but there is a slight issue that we have overlooked until now: $\mathbb{E}[T] = \infty$.*

Let us now state and prove the lemmas we will need for the proof of Theorem 2.24.

Lemma 2.21. *Let $(S_n)_{n\geq 1}$ be a martingale (resp. supermartingale) and T a stopping time relative to $(X_n)_{n\geq 1}$. We have*

$$\mathbb{E}[S_n \cdot \mathbb{1}_{\{T=k\}}] = \mathbb{E}[S_k \cdot \mathbb{1}_{\{T=k\}}] \quad \textit{(resp. } \leq \mathbb{E}[S_k \cdot \mathbb{1}_{\{T=k\}}]\textit{)}$$

for $n \geq k \geq 1$.

Proof.

$$\begin{aligned}
\mathbb{E}[S_n \mathbb{1}_{\{T=k\}}] &\overset{(1.33)}{=} \mathbb{E}\big(\mathbb{E}[S_n \mathbb{1}_{\{T=k\}} \,|\, X_1, \ldots, X_k]\big) \\
&= \mathbb{E}\big(\mathbb{1}_{\{T=k\}} \cdot \mathbb{E}[S_n \,|\, X_1, \ldots, X_k]\big) \\
&\overset{(2.9)}{=} \mathbb{E}[S_k \cdot \mathbb{1}_{\{T=k\}}] \\
\big(\textit{resp.} &\overset{(2.14)}{\leq} \mathbb{E}[S_k \cdot \mathbb{1}_{\{T=k\}}]\big)
\end{aligned}$$

□

Lemma 2.22. *Let W be a random variable such that $\mathbb{E}[\,|W|\,] < \infty$. Let T be a stopping time relative to $(X_n)_{n\geq 1}$ such that $\mathbb{P}(T < \infty) = 1$. Then*

$$\lim_{n\to\infty} \mathbb{E}[W \cdot \mathbb{1}_{\{T>n\}}] = 0$$

and

$$\lim_{n\to\infty} \mathbb{E}[W \cdot \mathbb{1}_{\{T\leq n\}}] = \mathbb{E}[W].$$

Proof. We first prove the lemma for $|W|$.

$$\begin{aligned}
\mathbb{E}[\,|W|\, \cdot \mathbb{1}_{\{T\leq n\}}] &= \sum_{k=1}^{n} \mathbb{E}[\,|W| \cdot \mathbb{1}_{\{T=k\}}] \\
&= \sum_{k=1}^{n} \mathbb{E}[\,|W|\,|\, T = k]\mathbb{P}(T = k) \\
&\longrightarrow \sum_{k=1}^{\infty} \mathbb{E}[\,|W|\,|\, T = k]\mathbb{P}(T = k) \quad \text{as } n \to \infty \\
&= \mathbb{E}[\,|W|\,] < \infty
\end{aligned}$$

by the law of total probability (1.35) and using the fact that we are dealing with a monotone increasing sequence. Therefore

$$\lim_{n\to\infty} \mathbb{E}[\,|W| \cdot \mathbb{1}_{\{T\le n\}}] = \mathbb{E}[\,|W|\,]$$

and

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[\,|W| \cdot \mathbb{1}_{\{T>n\}}] &= \lim_{n\to\infty} \mathbb{E}[\,|W|(1 - \mathbb{1}_{\{T\le n\}})] \\
&= \mathbb{E}[\,|W|\,] - \lim_{n\to\infty} \mathbb{E}[\,|W| \cdot \mathbb{1}_{\{T\le n\}}] \\
&= \mathbb{E}[\,|W|\,] - \mathbb{E}[\,|W|\,] \\
&= 0.
\end{aligned}$$

Now, to get the result for W, observe that

$$0 \le |\mathbb{E}[W \cdot \mathbb{1}_{\{T>n\}}]| \le \mathbb{E}[\,|W| \cdot \mathbb{1}_{\{T>n\}}] \to 0, \text{ as } n \to \infty.$$

So $\lim_{n\to\infty} \mathbb{E}[W \cdot \mathbb{1}_{\{T>n\}}] = 0$ and as a consequence

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[W \cdot \mathbb{1}_{\{T\le n\}}] &= \lim_{n\to\infty} \mathbb{E}[W \cdot (1 - \mathbb{1}_{\{T>n\}})] \\
&= \mathbb{E}[W] - \lim_{n\to\infty} \mathbb{E}[W \cdot \mathbb{1}_{\{T>n\}}] \\
&= \mathbb{E}[W].
\end{aligned}$$

□

Lemma 2.23. *Let T be a stopping time relative to $(X_n)_{n\ge1}$. Then*

$$\mathbb{E}[S_{T\wedge n}] = \mathbb{E}[S_1].$$

Proof.

$$\begin{aligned}
\mathbb{E}[S_1] &\overset{(2.9)}{=} \mathbb{E}[S_n] \\
&= \sum_{k=1}^{n} \mathbb{E}[S_n \cdot \mathbb{1}_{\{T=k\}}] + \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] \\
&\overset{(2.21)}{=} \sum_{k=1}^{n} \mathbb{E}[S_k \cdot \mathbb{1}_{\{T=k\}}] + \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] \\
&= \sum_{k=1}^{n} \mathbb{E}[S_T \cdot \mathbb{1}_{\{T=k\}}] + \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] \\
&= \mathbb{E}[S_T \cdot \mathbb{1}_{\{T\le n\}}] + \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] \\
&= \mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}_{\{T\le n\}}] + \mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}_{\{T>n\}}] \\
&= \mathbb{E}[S_{T\wedge n}].
\end{aligned}$$

□

We are now ready to prove the optional stopping theorem:

Theorem 2.24. *Let $(S_n)_{n\ge1}$ be a martingale and T a stopping time relative to $(X_n)_{n\ge1}$. Suppose*

(a) $\mathbb{P}(T < \infty) = 1$,

(b) $\mathbb{E}[\,|S_T|\,] < \infty$ *and*

(c) $\lim_{n\to\infty} \mathbb{E}[S_n \,|\, T > n] \cdot \mathbb{P}(T > n) = 0.$

Then $\mathbb{E}[S_T] = \mathbb{E}[S_1]$.

Proof. For all $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}[S_T] &= \mathbb{E}[S_T \cdot \mathbb{1}_{\{T\le n\}}] + \mathbb{E}[S_T \cdot \mathbb{1}_{\{T>n\}}] \\ &= \mathbb{E}[S_{T\wedge n} \cdot (1 - \mathbb{1}_{\{T>n\}})] + \mathbb{E}[S_T \cdot \mathbb{1}_{\{T>n\}}] \\ &\overset{(2.23)}{=} \mathbb{E}[S_1] - \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] + \mathbb{E}[S_T \cdot \mathbb{1}_{\{T>n\}}].\end{aligned}$$

By assumption (c), we have

$$\mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] = \mathbb{E}[S_n \,|\, T > n]\mathbb{P}(T > n) \to 0$$

as $n \to \infty$. We can apply Lemma 2.22 to S_T since we assume (a) and have $\mathbb{E}[\,|S_T|\,] < \infty$ by assumption (b). This gives

$$\mathbb{E}[S_T \cdot \mathbb{1}_{\{T>n\}}] \to 0$$

as $n \to \infty$. Taking

$$\mathbb{E}[S_T] = \mathbb{E}[S_1] - \mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}] + \mathbb{E}[S_T \cdot \mathbb{1}_{\{T>n\}}]$$

and letting n go to infinity gives the desired result. □

Remark 2.25. *If* $\mathbb{P}(T = k_0) = 1$, *then* $E[S_T] = \mathbb{E}[S_{k_0}] = \mathbb{E}[S_1]$ *and the assumptions also hold.*

Remark 2.26. *If* $\mathbb{P}(T = \infty) > 0$, *then* $S_T = S_\infty$ *on* $\{T = \infty\}$, *which is not defined here.*

Remark 2.27. $\mathbb{E}[S_T]$ *has a meaning by assumption* (b).

Example 2.28. *Let* $(X_n)_{n\ge1}$ *be i.i.d. random variables with* $\mathbb{P}(X_1 = -1) = 1/2 = \mathbb{P}(X_1 = +1)$. *The martingale* $(S_n)_{n\ge0}$ *relative to* $(X_n)_{n\ge1}$ *defined by* $S_0 = 0$, $S_n = X_1 + \ldots + X_n$ *for* $n \ge 1$ *is called a simple random walk (r.w.) on* $\mathbb{Z}$.

Fix positive integers a *and* b. *What is the probability that the random walk visits* $-a$ *before* b *?*

Let $T = T_{-a,b} = \inf\{n \in \mathbb{N} \,:\, S_n = -a \text{ or } S_n = b\}$. T *is a stopping time since*

$$\{T = k\} = \{S_1 \notin \{-a, b\}, \ldots, S_{k-1} \notin \{-a, b\}, S_k \in \{-a, b\}\} \in \sigma(X_1, \ldots, X_k).$$

We now check the assumptions of Theorem 2.24.

(a) Consider the event

$$\{X_1 = 1, X_2 = 1, \ldots, X_{a+b} = 1\}$$

and notice that no matter where the random walk starts in the interval, we have that

$$\mathbb{P}(X_1 = 1, \ldots, X_{a+b} = 1) = 2^{-(a+b)} \le \mathbb{P}(T \le a + b)$$

since

$$\{X_1 = 1, \ldots, X_{a+b} = 1\} \subset \{T \le a + b\}.$$

Therefore, the probability of not leaving $[-a, b]$ *in* $(a + b)$ *units of time is such that*

$$\mathbb{P}(T > a + b) = 1 - \mathbb{P}(T \le a + b) \le 1 - 2^{-(a+b)}.$$

Now, it is relatively easy to see that the events

$$\{\textit{Not leaving } [-a,b] \textit{ in } 2(a+b) \textit{ time units.}\}$$

and

$$\{\textit{Not leaving the interval } [-a,b] \textit{ during the first } (a+b) \textit{ units of time, and again, starting from where we ended up at time } (a+b), \textit{ stay in } [-a,b] \textit{ during the remaining } (a+b) \textit{ time units.}\}$$

are the same. As one would expect, we have

$$\begin{aligned}
\mathbb{P}(T > 2(a+b)) &= \mathbb{E}[\mathbb{1}_{\{T>2(a+b)\}}] \\
&= \mathbb{E}[\mathbb{1}_{\{T>(a+b)\}} \cdot \mathbb{1}_{\{T>2(a+b)\}}] \\
&= \mathbb{E}\big(\mathbb{E}[\mathbb{1}_{\{T>(a+b)\}} \cdot \mathbb{1}_{\{T>2(a+b)\}} \,|\, X_1, \dots, X_{a+b}]\big) \\
&= \mathbb{E}\big(\mathbb{1}_{\{T>(a+b)\}} \cdot \mathbb{E}[\mathbb{1}_{\{T>2(a+b)\}} \,|\, X_1, \dots, X_{a+b}]\big) \\
&= \mathbb{E}[\mathbb{1}_{\{T>(a+b)\}} \cdot \mathbb{P}(T > 2(a+b) \,|\, X_1, \dots, X_{a+b})] \\
&\leq \mathbb{E}[\mathbb{1}_{\{T>(a+b)\}}] \cdot (1 - 2^{-(a+b)}) \\
&\leq (1 - 2^{-(a+b)})^2.
\end{aligned}$$

We can repeat this process and find that

$$\mathbb{P}(T > k(a+b)) \leq (1 - 2^{-(a+b)})^k$$

for $k \geq 1$. Now, since $(1 - 2^{-(a+b)}) < 1$, letting $k \to \infty$ gives

$$0 \leq \mathbb{P}(T = \infty) \leq 0$$

which concludes the verification of assumption (a).

(b) *By definition of T, $S_T \in \{-a, b\}$, which implies $|S_T| \leq \max\{a,b\}$. Condition (b) follows immediately:*

$$\mathbb{E}[\,|S_T|\,] \leq \mathbb{E}[\max\{a,b\}] = \max\{a,b\} < \infty.$$

(c) *Since $|S_n| \leq \max\{a,b\}$ on $\{T > n\}$, we have*

$$\begin{aligned}
\lim_{n\to\infty} |\mathbb{E}[S_n \,|\, T > n] \cdot \mathbb{P}(T > n)| &= \lim_{n\to\infty} |\mathbb{E}[S_n \cdot \mathbb{1}_{\{T>n\}}]| \\
&\leq \lim_{n\to\infty} \mathbb{E}[\,|S_n| \cdot \mathbb{1}_{\{T>n\}}] \\
&\leq \lim_{n\to\infty} \max\{a,b\} \cdot \mathbb{P}(T > n) \\
&= 0
\end{aligned}$$

as $P(T < \infty) = 1$.

Now that we have checked the assumptions of the optional stopping theorem, we conclude that

$$\mathbb{E}[S_T] = \mathbb{E}[S_1] = \mathbb{E}[X_1] = 0.$$

But we can go even further. By definition of the expectation and using the fact that $S_T \in \{-a, b\}$, we get

$$\mathbb{E}[S_T] = -a \cdot \mathbb{P}(S_T = a) + b \cdot \mathbb{P}(S_T = b) = 0.$$

Let $p_{-a,b} = \mathbb{P}(S_T = -a) = 1 - \mathbb{P}(S_T = b)$. We can write the previous equation with $p_{-a,b}$

$$-ap_{-a,b} + b(1 - p_{-a,b}) = 0.$$

From this, we find that the probability of the random walk visiting $-a$ before b is given by

$$p_{-a,b} = \frac{b}{a+b}.$$

As a direct consequence, we have that the probability of visiting b before $-a$ is $b/(a+b)$.

Example 2.29. *Keeping the notations we used in the example above, we now want to learn more about the first visit to b. To do so, define the stopping time $T_b = \inf\{n \geq 0 : S_n = b\}$ with $b > 0$. Notice that*

$$\{S_{T_{-a,b}} = b\} = \{T_b \leq T_{-a},\, T_b < \infty\}$$

and

$$\{T_b < \infty\} \subset \left\{ \min_{n \in \{0,\dots,T_b\}} S_n > -\infty \right\} \subset \bigcup_{a \in \mathbb{N}} \{T_b = T_{-a,b}\}.$$

In particular,

$$\begin{aligned}
\mathbb{P}(T_b < \infty) &= \mathbb{P}\left(\bigcup_{a \in \mathbb{N}} \{T_b = T_{-a,b}\} \cap \{T_b < \infty\} \right) \\
&= \mathbb{P}\left(\bigcup_{a \in \mathbb{N}} \{T_b \leq T_{-a},\, T_b < \infty\} \right) \\
&\overset{(*)}{=} \lim_{a \to \infty} \mathbb{P}(T_b \leq T_{-a}, T_b < \infty) \\
&= \lim_{a \to \infty} \mathbb{P}(S_{T_{-a,b}} = b) \\
&= \lim_{a \to \infty} \frac{a}{a+b} = 1
\end{aligned}$$

where $(*)$ is justified by the fact that $\{T_b \leq T_{-a},\, T_b < \infty\}$ is an increasing sequence of events in $a \in \mathbb{N}$. We now want to determine $\mathbb{E}[T_b]$. Let $(Y_n) = (S_n^2 - n)$ be the martingale defined in Example 2.5. Let's apply Theorem 2.24 to (Y_n) and $T := T_{-a,b}$. We check the hypotheses:

(a) *We already know $\mathbb{P}(T < \infty) = 1$.*

(b) *We first observe that*

$$\mathbb{E}[\,|Y_T|\,] = \mathbb{E}[\,|S_T^2 - T|\,] \leq \mathbb{E}[S_T^2] + \mathbb{E}[T] \leq \max(a^2, b^2) + \mathbb{E}[T],$$

so the only thing left to verify is wether $\mathbb{E}[T] < \infty$. To do so, let us first recall the following results:

Lemma 2.30. *If $X \geq 0$ is a random variable, then*

$$\mathbb{E}[X] \leq 1 + \sum_{k=1}^{\infty} \mathbb{P}(X \geq k).$$

Moreover, if $\mathbb{P}(X \in \mathbb{N}) = 1$, we have

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k).$$

Proof. To get the first result, simply notice that $X \le 1 + \sum_{k=1}^{\infty} \mathbb{1}_{\{X \ge k\}}$ and take the expectation. For the second result, we can write $X = \sum_{k=1}^{\infty} \mathbb{1}_{\{X \ge k\}}$ since we assumed $\mathbb{P}(X \in \mathbb{N}) = 1$ and again taking the expectation gives the desired equality. □

From this, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{T}{a+b}\right] &\le 1 + \sum_{k=1}^{\infty} \mathbb{P}\left(\frac{T}{a+b} \ge k\right) \\
&\le 1 + \sum_{k=1}^{\infty} \mathbb{P}\left(\frac{T}{a+b} > k-1\right) \\
&= 1 + \sum_{k=1}^{\infty} \mathbb{P}\big(T > (k-1)(a+b)\big) \\
&\le 1 + \sum_{k=1}^{\infty} (1 - 2^{-(a+b)})^{k-1} < \infty.
\end{aligned}$$

So $\mathbb{E}[T] < \infty$ *and this concludes the check of* (b).

(c) We first split the starting term in two:

$$\begin{aligned}
\mathbb{E}[Y_n \cdot \mathbb{1}_{\{T>n\}}] &= \mathbb{E}[(S_n^2 - n) \cdot \mathbb{1}_{\{T>n\}}] \\
&= \mathbb{E}[S_n^2 \cdot \mathbb{1}_{\{T>n\}}] - \mathbb{E}[n \cdot \mathbb{1}_{\{T>n\}}].
\end{aligned}$$

Now, since $S_n^2 \le \max(a^2, b^2)$ *on the set* $\{T > n\}$*, we get*

$$\mathbb{E}[S_n^2 \cdot \mathbb{1}_{\{T>n\}}] \le \max(a^2, b^2) \cdot \mathbb{P}(T > n) \longrightarrow 0$$

as $n \to \infty$*, since* $\mathbb{P}(T < \infty) = 1$*. For the second term, we have*

$$\mathbb{E}[n \cdot \mathbb{1}_{\{T>n\}}] \le \mathbb{E}[T \cdot \mathbb{1}_{\{T>n\}}] \overset{n\to\infty}{\longrightarrow} 0$$

thanks to Lemma 2.22.

All three conditions hold, so we can apply Theorem 2.24 *and we get*

$$\mathbb{E}[S_T^2 - T] = \mathbb{E}[Y_T] = \mathbb{E}[Y_1] = \mathbb{E}[S_1^2 - 1] = \mathbb{E}[X_1^2 - 1] = 0.$$

Therefore,

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[S_T^2] \\
&= a^2 \mathbb{P}(S_{T_{-a,b}} = -a) + b^2 \mathbb{P}(S_{T_{-a,b}} = b) \\
&= a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} \\
&= ab\left(\frac{a}{a+b} + \frac{b}{a+b}\right) \\
&= ab.
\end{aligned}$$

If $a = b$*,* $\mathbb{E}[T_{-a,a}] = a^2$ *is the mean time it takes to move* a *steps from the starting point. From this, we can actually calculate* $\mathbb{E}[T_b]$ *using what we have done until now.*

For all $a \in \mathbb{N}^*$*, we have that* $T_b \ge T_{-a,b}$*. Indeed, there are only two possible situations, either we visit* $-a$ *first, which means* $T_{-a,b} < T_b$*, or we visit* b *first and in this case have* $T_{-a,b} = T_b$*. From this and what we saw earlier, we get*

$$\mathbb{E}[T_b] \ge \mathbb{E}[T_{-a,b}] = ab$$

for all $a \in \mathbb{N}$*. In particular,* $\mathbb{E}[T_b] \ge \sup_{a \in \mathbb{N}} ab = +\infty$ *since* $b \ge 1$*. So* $\mathbb{E}[T_b] = \infty$ *for all* $b \ne 0$*.*

2.5 Optional Stopping Theorems for supermartingales

Before giving the optional stopping time theorems for supermartingales, we first look at two very useful results.

Theorem 2.31. ***Doob's decomposition theorem.*** *Let $(S_n)_{n\geq 1}$ be a supermartingale relative to $(X_n)_{n\geq 1}$. There exists a martingale $(M_n)_{n\geq 1}$ and a process $(A_n)_{n\geq 1}$ such that*

(1) $n \mapsto A_n$ is nondecreasing ($P(A_n \leq A_{n+1})$)= 1,

(2) $A_1 = 0$,

(3) A_{n+1} is a function of $(X_1, \dots, X_n)$ (A_n is predictable) and

(4) $S_n = M_n - A_n$.

Proof. See Problem Set 4. □

Lemma 2.32. *Let (S_n) be a supermartingale and T a stopping time relative to (X_n). Then*

$$\mathbb{E}[S_{T\wedge n}] \leq \mathbb{E}[S_1] \quad , \textit{for all } n \geq 1.$$

Proof. Let $S_n = M_n - A_n$ be the Doob decomposition of (S_n). We have

$$S_{T\wedge}n = M_{T\wedge n} - A_{T\wedge n},$$

so

$$\mathbb{E}[S_{T\wedge n}] = \mathbb{E}[M_{T\wedge n}] - \mathbb{E}[A_{T\wedge n}] \leq \mathbb{E}[M_{T\wedge n}] \stackrel{(2.23)}{=} \mathbb{E}[M_1] = \mathbb{E}[S_1],$$

where the inequality comes from the fact that $A_{T\wedge n} \geq 0$ and the last equality is a consequence from one of our assumptions, $A_1 = 0$. □

Lemma 2.33. *Let (S_n) be a supermartingale and $\varphi : \mathbb{R} \to \mathbb{R}$ a concave and nondecreasing function such that $\mathbb{E}[\,|\varphi(S_n)|\,] < \infty$ for all n. Then $(\varphi(S_n))$ is a supermartingale.*

Proof. We check the conditions of Definition 2.10.

(a) : We know by assumption that $\mathbb{E}[\,|\varphi(S_n)|\,] < \infty$.

(b) : The function φ is concave by assumption, meaning $-\varphi$ is convex. Property (G) of conditional expectation gives

$$\mathbb{E}[-\varphi(S_{n+1}) \,|\, X_1, \dots, X_n] \geq -\varphi(\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n]).$$

From this, we get

$$\mathbb{E}[\varphi(S_{n+1}) \,|\, X_1, \dots, X_n] \leq \varphi(\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n]) \leq \varphi(S_n).$$

The last inequality follows from $\mathbb{E}[S_{n+1} \,|\, X_1, \dots, X_n] \leq S_n$, since S_n is a supermartingale and φ is nondecreasing.

(c) : Since (S_n) is a supermartingale, S_n is a function of $X_1, \dots, X_n$, meaning $\varphi(S_n)$ is too.

□

Theorem 2.34. *Let (S_n) be a supermartingale and T a stopping time relative to (X_n). Suppose that $\mathbb{P}(T < \infty) = 1$ and there is a random variable $W \geq 0$ with $\mathbb{E}[W] < \infty$ and $S_{T\wedge n} \geq -W$, for all $n \in \mathbb{N}$. Then $\mathbb{E}[S_T] \leq \mathbb{E}[S_1]$.*

Remark 2.35. *We often use this theorem when (S_n) is nonnegative and take $W = 0$.*

Proof. Fix $b > 0$, and $S_n^b := \min(b, S_n)$ $n \geq 1$. Then we know that (S_n^b) is also a supermartingale, as the function $x \mapsto \min(b, x)$ is concave and nondecreasing, meaning we can apply Lemma 2.33. Using Lemma 2.32 now gives

$$\mathbb{E}[S_1^b] \geq \mathbb{E}[S_{T\wedge n}^b] \tag{$*1$}$$

for $n \geq 1$. Noticing that $S_n^n, S_T^b \in [-W, b]$, we have

$$\begin{aligned}
0 &\leq |\mathbb{E}[S_{T\wedge n}^b] - \mathbb{E}[S_T^b]| \\
&= |\mathbb{E}[S_{T\wedge n}^b - S_T^b]| \\
&= |\mathbb{E}[(S_n^b - S_T^b) \cdot \mathbb{1}_{\{T>n\}}]| \\
&\leq \mathbb{E}[(b+W) \cdot \mathbb{1}_{\{T>n\}}] \longrightarrow 0
\end{aligned}$$

as $n \to \infty$ by Lemma 2.22 applied to $(b + W)$. Therefore $\mathbb{E}[S_{T\wedge n}^b] \to \mathbb{E}[S_T^b]$ as $n \to \infty$. Letting $n \to \infty$ in $(*1)$ gives

$$\mathbb{E}[S_1^b] \geq \mathbb{E}[S_T^b]. \tag{$*2$}$$

As $b \nearrow +\infty$, we have both $S_n^b \nearrow S_n$ and $S_T^b \nearrow S_T$, applying the monotone convergence theorem gives us

$$\lim_{b\to\infty} \mathbb{E}[S_1^b] = \mathbb{E}[S_1] \quad \text{and} \quad \lim_{b\to\infty} \mathbb{E}[S_T^b] = \mathbb{E}[S_T] \tag{$*3$}$$

Combining $(*2)$ and $(*3)$ gives the desired result: $\mathbb{E}[S_1] \geq \mathbb{E}[S_T]$. □

Proposition 2.36. *Let (S_n) be a supermartingale and T a stopping time relative to (X_n). Suppose that $\mathbb{P}(S_n \geq 0) = 1$ for all $n \in \mathbb{N}^*$. Then $\mathbb{E}[S_1] \geq \mathbb{E}[S_T \cdot \mathbb{1}_{\{T<\infty\}}]$.*

Remark 2.37. *By definition,*

$$S_T \cdot \mathbb{1}_{\{T<\infty\}} = \begin{cases} S_T & \text{if } T < \infty, \\ 0 & \text{if } T = \infty \end{cases} .$$

Proof. We first observe that

$$\mathbb{E}[S_1] \overset{(2.32)}{\geq} \mathbb{E}[S_{T\wedge n}] \overset{S_n \geq 0}{\geq} \mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}_{\{T\leq n\}}] = \mathbb{E}[S_T \cdot \mathbb{1}_{\{T\leq n\}}].$$

<u>On $\{T < \infty\}$:</u>

$$S_T \cdot \mathbb{1}_{\{T\leq n\}} = \begin{cases} 0 & \text{for } n < T, \\ S_T & \text{for } n \geq T \end{cases} .$$

The function $n \mapsto S_T \cdot \mathbb{1}_{\{T\leq n\}}$ is nondecreasing and it converges to $S_T = S_T \cdot \mathbb{1}_{\{T<\infty\}}$. Moreover, $S_T \cdot \mathbb{1}_{\{T<\infty\}} = S_T$ on $\{T < \infty\}$.

<u>On $\{T = \infty\}$:</u>

$$S_T \cdot \mathbb{1}_{\{T\leq n\}} = 0 = S_T \cdot \mathbb{1}_{\{T<\infty\}}$$

In both cases,

$$S_T \cdot \mathbb{1}_{\{T\leq n\}} \nearrow S_T \cdot \mathbb{1}_{\{T<\infty\}}.$$

The monotone convergence theorem and our first observation give the desired result:

$$\mathbb{E}[S_T \cdot \mathbb{1}_{\{T<\infty\}}] = \lim_{n\to\infty} \mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}_{\{T\leq n\}}] \leq \mathbb{E}[S_1].$$

□

Theorem 2.38. *Let (S_n) be a supermartingale and U, T be two stopping times relative to (X_n). Suppose there is $N \in \mathbb{N}$ such that $\mathbb{P}(1 \leq U \leq T \leq N) = 1$, then $\mathbb{E}[S_T] \leq \mathbb{E}[S_U]$.*

Proof. Step 1: Fix $k < N$. For $k \leq n \leq N$,

$$\begin{aligned}\mathbb{E}[S_{n+1} \cdot \mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\}] &= \mathbb{E}\big(\mathbb{E}[S_{n+1} \cdot \mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\} \,|\, X_1, \ldots, X_n]\big) \\ &= \mathbb{E}\big(\mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\} \cdot \mathbb{E}[S_{n+1} \,|\, X_1, \ldots, X_n]\big) \\ &\leq \mathbb{E}[\big(\mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\} \cdot S_n].\end{aligned}$$

Step 2: The function $n \mapsto \mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}\{U = k\}]$ is nonincreasing on $[k, N]$. Indeed,

$$\begin{aligned}\mathbb{E}[S_{T\wedge n} \cdot \mathbb{1}\{U = k\}] &= \mathbb{E}[S_T \cdot \mathbb{1}\{T \leq n\} \cdot \mathbb{1}\{U = k\}] + \mathbb{E}[S_n \cdot \mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\}] \\ &\geq \mathbb{E}[S_T \cdot \mathbb{1}\{T \leq n\} \cdot \mathbb{1}\{U = k\}] + \mathbb{E}[S_{n+1} \cdot \mathbb{1}\{T > n\} \cdot \mathbb{1}\{U = k\}] \\ &= \mathbb{E}[S_{T\wedge(n+1)} \cdot \mathbb{1}_{U=k}].\end{aligned}$$

From this we get

$$\mathbb{E}[S_k \cdot \mathbb{1}_{\{U=k\}}] = \mathbb{E}[S_{T\wedge k} \cdot \mathbb{1}_{\{U=k\}}] \geq \mathbb{E}[S_{T\wedge N} \cdot \mathbb{1}_{\{U=k\}}] = \mathbb{E}[S_T \cdot \mathbb{1}_{\{U=k\}}],$$

where the first equality follows from $T \geq U = k$ and the last from our assumption $P(1 \leq U \leq T \leq N) = 1$.

Step 3: Finally,

$$\mathbb{E}[S_U] = \sum_{k=1}^{N} \mathbb{E}[S_k \cdot \mathbb{1}_{\{U=k\}}] \geq \sum_{k=1}^{N} \mathbb{E}[S_T \cdot \mathbb{1}_{\{U=k\}}] = \mathbb{E}\big[S_T \sum_{k=1}^{N} \mathbb{1}_{\{U=k\}}\big] = \mathbb{E}[S_T].$$

□

Note that we have an analogous result for submartingales:

Theorem 2.39. *Let (S_n) be a submartingale and U, T be two stopping times relative to (X_n). Suppose there is $N \in \mathbb{N}$ such that $\mathbb{P}(1 \leq U \leq T \leq N) = 1$, then $\mathbb{E}[S_T] \geq \mathbb{E}[S_U]$.*

2.6 Martingales convergence theorems

Definition 2.40. *We say that a sequence (S_n) of random variables converges* ***in probability*** *to a random variable S if:*

$$\forall\, \epsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(|S_n - S| > \epsilon) = 0.$$

We often write $S_n \overset{p}{\to} S$ to say that the sequence S_n converges in probability to S.

Definition 2.41. *A sequence of random variables (S_n) converges* ***almost-surely*** *to a random variable S if:*

$$\mathbb{P}\big(\{\omega \in \Omega : \lim_{n\to\infty} S_n(\omega) = S(\omega)\}\big) = 1.$$

We write $\lim_{n\to\infty} S_n = S$ a.s. or $S_n \overset{a.s.}{\to} S$ to denote almost-sure convergence.

Theorem 2.42. ***Monotone convergence theorem (MCT).*** *If $\mathbb{P}(S_n \geq 0) = 1$ and $\mathbb{P}(S_n \leq S_{n+1}) = 1$ for all $n \in \mathbb{N}$ and $S_n \overset{a.s.}{\to} S$, then*

$$\lim_{n\to\infty} \mathbb{E}[S_n] = \mathbb{E}[\lim_{n\to\infty} S_n] = \mathbb{E}[S].$$

Theorem 2.43. ***Dominated convergence theorem (DCT).*** *If $S_n \to S$ a.s. and if there is a random variable $W \geq 0$ such that $\mathbb{E}[W] < \infty$ and $\mathbb{P}(|Sn| \leq W) = 1$ for all $n \in \mathbb{N}$, then*

$$\lim_{n\to\infty} \mathbb{E}[S_n] = \mathbb{E}[\lim_{n\to\infty} S_n] = \mathbb{E}[S].$$

Proposition 2.44. *Let (X_n) be a sequence of random variables and X a random variable. Define*

$$A_n(\epsilon) = \{|X_n - X| > \epsilon\} \quad \text{and} \quad B_m(\epsilon) = \bigcup_{n\geq m} A_n(\epsilon).$$

(a) We have the following sufficient and necessary conditions for almost-sure convergence:

$$X_n \overset{a.s.}{\to} X \iff \forall\, \epsilon > 0\,,\, \lim_{m\to\infty} \mathbb{P}(B_m(\epsilon)) = 0 \iff \forall\, \epsilon > 0,\, \lim_{m\to\infty} \mathbb{P}(\sup_{n\geq m} |X_n - X| > \epsilon) = 0.$$

(b) If $\sum_{n=1}^{\infty} \mathbb{P}(A_n(\epsilon)) < \infty$ for all $\epsilon > 0$, then $X_n \overset{a.s.}{\to} X$.

(c) If $X_n \overset{a.s.}{\to} X$, then $X_n \overset{p}{\to} X$.

Remark 2.45. *We often write $\{A\}$ as a shorthand for $\{\omega \in \Omega : \omega \in A\}$. Here, for example, we wrote $\{|X_n - X| > \epsilon\}$ instead of the more detailed expression $\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}$.*

Proof. (a) Observe that

$$\Big\{\lim_{n\to\infty} X_n = X\Big\} = \bigcap_{k\in\mathbb{N}^*} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \big\{|X_n - X| \leq 1/k\big\}.$$

To see this, it can be useful to know that

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \iff \forall k \in \mathbb{N}^*, \exists N \in \mathbb{N}, n \geq N : |X_n(\omega) - X(\omega)| \leq 1/k.$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1 &\iff \mathbb{P}\left(\Big\{\lim_{n\to\infty} X_n = X\Big\}^c\right) = 0 \\ &\iff \mathbb{P}\Big(\bigcup_{k\in\mathbb{N}^*} \bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} \big\{|X_n - X| > 1/k\big\}\Big) = 0.\end{aligned}$$

Notice that $\cup_{n\geq N}\{|X_n - X| > 1/k\} = B_N(1/k)$ and define $C_k = \cap_{N\in\mathbb{N}} B_N(1/k)$. We have that $C_k \subset C_{k+1}$ and our previous expression simplifies to $\mathbb{P}(\cup_{k\in\mathbb{N}^*} C_k) = 0$. However, this holds if and only if $\mathbb{P}(C_k) = 0$ for all $k \in \mathbb{N}^*$, which implies

$$0 = \mathbb{P}(C_k) = \mathbb{P}\Big(\bigcap_{N\in\mathbb{N}} B_N(1/k)\Big) = \lim_{N\to\infty} \mathbb{P}(B_N(1/k)),$$

since $B_{N+1}(1/k) \subset B_N(1/k)$. From this we get

$$\mathbb{P}\big(\lim_{n\to\infty} X_n = X\big) = 1 \iff \forall k \in \mathbb{N}^*,\ \lim_{N\to\infty} \mathbb{P}(B_N(1/k)) = 0. \tag{1}$$

It remains to check that this is equivalent to

$$\forall \epsilon > 0, \lim_{N\to\infty} \mathbb{P}(B_N(\epsilon)) = 0. \tag{2}$$

Clearly, (2) $\Rightarrow$ (1), so it remains to check (1) $\Rightarrow$ (2). For $\epsilon < \epsilon'$, $A_n(\epsilon') \subset A_n(\epsilon)$, so $B_m(\epsilon') \subset B_m(\epsilon)$. If

$$\lim_{N\to\infty} \mathbb{P}(B_N(\epsilon)) = 0$$

then

$$\lim_{N\to\infty} \mathbb{P}(B_N(\epsilon')) = 0$$

and (1) $\Rightarrow$ (2) follows.

(b) First note that $\mathbb{P}(B_N(\epsilon)) \leq \sum_{n\geq N} \mathbb{P}(A_n(\epsilon))$. Since $\sum_{n=1}^{\infty} \mathbb{P}(A_n(\epsilon)) < \infty$ by assumption, we have

$$0 \leq \lim_{N\to\infty} \mathbb{P}(B_N(\epsilon)) \leq \lim_{N\to\infty} \sum_{n\geq N} \mathbb{P}(A_n(\epsilon)) = 0.$$

(c) We have $X_n \xrightarrow{p} X$ if and only if $\forall \epsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(A_n(\epsilon)) = 0$. If $X_n \to X$ a.s., then

$$0 \leq \lim_{n\to\infty} \mathbb{P}(A_n(\epsilon)) \leq \lim_{n\to\infty} \mathbb{P}(B_n(\epsilon)) = 0.$$

□

Proposition 2.46. ***Cauchy criterion for almost-sure convergence.*** *Let (S_n) be a sequence of random variables such that*

$$\forall \epsilon > 0,\ \lim_{m\to\infty} \mathbb{P}\left(\sup_{i\geq m} |S_i - S_m| > \epsilon\right) = 0.$$

Then there exists a random variable S such that $S_n \to S$ a.s.

Proof. See Problem Set 5. □

Proposition 2.47. ***The Doob-Kolmogorov inequality.*** *Let (S_n) be a submartingale relative to (X_n) such that $\mathbb{P}(S_n \geq 0) = 1$ for all $n \in \mathbb{N}$. Then for all $\lambda > 0$,*

$$\mathbb{P}\left(\max_{i=1,\ldots,n} S_i \geq \lambda\right) \leq \frac{\mathbb{E}[S_n]}{\lambda}.$$

Note that this result also holds if we have the strict inequality: $\max_{i=1,\ldots,n} S_i > \lambda$.

Proof. Set

$$T = \begin{cases} \inf\{k \geq 0 :\ S_k \geq \lambda\} & ,\text{if } \max_{i=1,\ldots,n}\ S_i \geq \lambda \\ n & ,\text{if } \max_{i=1,\ldots,n}\ S_i < \lambda \end{cases}.$$

It is easy to see that T is a stopping time relative to (X_n). For $k < n$,

$$\{T = k\} = \{S_1 < \lambda, \ldots, S_{k-1} < \lambda, S_k \geq \lambda\}$$

which is determined by $X_1, \ldots, X_k$. For $k = n$, the set

$$\{T = n\} = \{S_1 < \lambda, \ldots, S_{n-1} < \lambda\}$$

is determined by $X_1, \ldots, X_n$. By construction of T, we have that $\mathbb{P}(0 \leq T \leq n \leq n) = 1$. We can therefore apply the optional stopping theorem for bounded stopping times (Theorem 2.39), which gives

$$\mathbb{E}[S_n] \geq \mathbb{E}[S_T].$$

Moreover, thanks to our assumption $\mathbb{P}(S_n \geq 0) = 1$ and the fact that $S_T \geq \lambda$ on $\{\max_{i=1,\dots,n} \geq \lambda\}$, we get the desired result

$$\mathbb{E}[S_n] \geq \mathbb{E}[S_T] \geq \mathbb{E}\left[S_T \cdot \mathbb{1}\{\max_{i=1,\dots,n} S_i \geq \lambda\}\right] \geq \mathbb{E}\left[\lambda \cdot \mathbb{1}\{\max_{i=1,\dots,n} S_i \geq \lambda\}\right] = \lambda \cdot \mathbb{P}\left(\max_{i=1,\dots,n} S_i \geq \lambda\right).$$

□

We now give a variant of the above inequality.

Proposition 2.48. *Fix $p > 1$. If $\mathbb{E}[\,|S_i|^p\,] < \infty$ for $i = 1, \dots, n$, then*

$$\mathbb{E}\left[\max_{i=1,\dots,n} |S_i|^p\right] \leq \left(\frac{p}{p-1}\right)^p \cdot \mathbb{E}[\,|S_n|^p\,].$$

Proof. See Problem Set 6. □

Remark 2.49. *A special case of the above inequality that might be useful to keep in mind is when we take $p = 2$:*

$$\mathbb{E}\left[\max_{i=1,\dots,n} S_i^2\right] \leq 4 \cdot \mathbb{E}[S_n^2].$$

We are now ready to state and prove the main result of this subsection:

Theorem 2.50. *Let (S_n) be a martingale relative to (X_n). Suppose that there is $M < \infty$ such that*

$$\mathbb{E}[S_n^2] \leq M$$

for all $n \in \mathbb{N}$. Then there is a random variable S such that $\lim_{n\to\infty} S_n = S$ a.s.

Remark 2.51. *Note that a martingale that satisfies the assumption in the above theorem is called an L^2-bounded martingale.*

Proof. First observe that the function $n \mapsto \mathbb{E}[S_n^2]$ is nondecreasing, as the square of a martingale is a submartingale (Lemma 2.13) and we can then refer to Remark 2.14. In addition:

$$\begin{aligned}\mathbb{E}[S_{n+m}^2] &= \mathbb{E}[(S_m + S_{n+m} - S_m)^2] \\ &= \mathbb{E}[S_m^2] + 2\mathbb{E}[S_m(S_{n+m} - S_m)] + \mathbb{E}[(S_{n+m} - S_m)^2] \\ &= \mathbb{E}[S_m^2] + \mathbb{E}[(S_{n+m} - S_m)^2]\end{aligned}$$

since

$$\begin{aligned}\mathbb{E}[S_m(S_{n+m} - S_m)] &= \mathbb{E}\big(\,\mathbb{E}[S_m(S_{n+m} - S_m) \,|\, X_1, \dots X_m]\,\big) \\ &= \mathbb{E}\big(\,S_m\mathbb{E}[(S_{n+m} - S_m) \,|\, X_1, \dots X_m]\big) \\ &= \mathbb{E}\big(\,S_m(\mathbb{E}[S_{n+m} \,|\, X_1, \dots X_m] - \S_m)\big) \\ &\overset{(2.9)}{=} \mathbb{E}[S_m(S_m - S_m)] \\ &= 0.\end{aligned}$$

As $m \mapsto \mathbb{E}[S_m^2]$ is nondecreasing and bounded by M, we can set

$$M_0 = \lim_{m\to\infty} \mathbb{E}[S_m^2] \leq M.$$

For $m \in \mathbb{N}$, define $Y_n = S_{m+n} - S_m$. Then (Y_n) is a martingale relative to (Y_n). Indeed,

$$\begin{aligned}
\mathbb{E}[Y_{n+1} \,|\, Y_1, \ldots, Y_n] &= \mathbb{E}\left(\mathbb{E}[Y_{n+1} \,|\, X_1, \ldots, X_{m+n}] \,|\, Y_1, \ldots, Y_n \right) \\
&= \mathbb{E}\left(\mathbb{E}[S_{m+(n+1)} - S_m \,|\, X_1, \ldots, X_{m+n}] \,|\, Y_1, \ldots, Y_n \right) \\
&= \mathbb{E}\left(\mathbb{E}[S_{m+(n+1)} \,|\, X_1, \ldots, X_{m+n}] - \mathbb{E}[S_m \,|\, X_1, \ldots, X_{m+n}] \,|\, Y_1, \ldots, Y_n \right) \\
&= \mathbb{E}[S_{m+n} - S_m \,|\, Y_1, \ldots Y_n] \\
&= \mathbb{E}[Y_n \,|\, Y_1, \ldots, Y_n] \\
&= Y_n.
\end{aligned}$$

Applying the Doob-Kolmogorov inequality to the submartingale (Y_n^2) gives

$$\mathbb{P}\left(\max_{i=1,\ldots,n} |Y_i| > \epsilon \right) = \mathbb{P}\left(\max_{i=1,\ldots,n} Y_i^2 > \epsilon^2 \right) \leq \frac{1}{\epsilon^2} \mathbb{E}[Y_n^2]$$

that is

$$\mathbb{P}\left(\max_{i=m,\ldots,m+n} |S_i - S_m| > \epsilon \right) \leq \frac{1}{\epsilon^2} \mathbb{E}[(S_{m+n} - S_m)^2] = \frac{1}{\epsilon^2} (\mathbb{E}[S_{m+n}^2] - \mathbb{E}[S_m^2]).$$

We now want to let $n \to \infty$:

$$\bigcup_{n=1}^{\infty} \left\{ \max_{i=m,\ldots,m+n} |S_i - S_m| > \epsilon \right\} = \left\{ \sup_{i \geq m} |S_i - S_m| > \epsilon \right\}.$$

Therefore,

$$\mathbb{P}\left(\sup_{i \geq m} |S_i - S_m| > \epsilon \right) \leq \frac{1}{\epsilon^2} (M_0 - \mathbb{E}[S_m^2]).$$

Letting $m \to \infty$ gives

$$0 \leq \lim_{m \to \infty} \mathbb{P}\left(\sup_{i \geq m} |S_i - S_m| > \epsilon \right) \leq \frac{1}{\epsilon^2} (M_0 - M_0) = 0.$$

The desired result now follows from the Cauchy criterion for almost-sure convergence. □

Definition 2.52. *Let (S_n) be a submartingale. For $a < b$, $N \in \mathbb{N}^*$, let $V_{a,b,N}$ be the number of couples (i, j) with $1 \leq i < j \leq N$ such that*

$$S_i \leq a, \quad a < S_i < b \quad \text{for } i < k < j \quad \text{and } b \leq S_j.$$

$V_{a,b,N}$ is the number of ***upcrossings*** *of $[a, b]$ between times 1 and N. Below is a graphic illustrating an example of an upcrossing:*

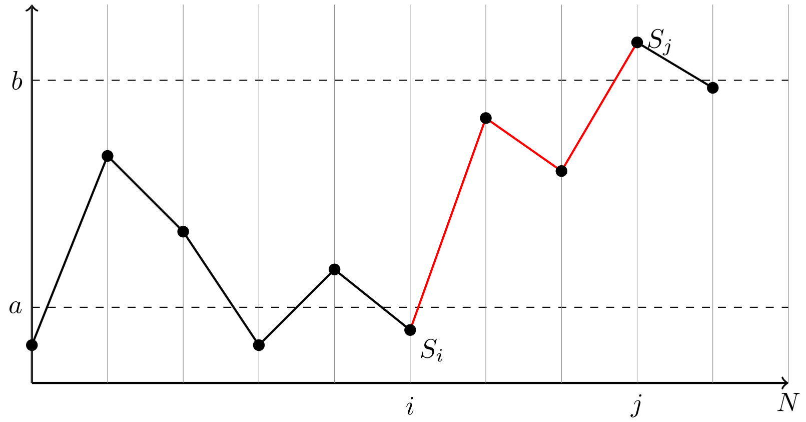


Figure 1: Here is an example of an upcrossing of $[a, b]$.

Remark 2.53. *If $f : \mathbb{N} \to \mathbb{R}$ is nondecreasing, then there is at most one upcrossing of $[a, b]$ during $\{1, \dots, N\}$. A submartingale is a stochastic analogue of this, so there should not be too many upcrossings of $[a, b]$ during $\{1, \dots, N\}$*

Lemma 2.54. ***The upcrossings inequality:*** *Keeping the notation introduced above, we have the following inequality:*

$$\mathbb{E}[V_{a,b,N}] \leq \frac{1}{b-a}\Big(\mathbb{E}[(S_N - a)_+] - \mathbb{E}[(S_1 - a)_+]\Big).$$

Proof. Set $X_n = (S_n - a)_+$. Since the function $x \mapsto (x-a)_+$ is nondecreasing and convex, Lemma 2.13 ensures that (X_n) is also a submartingale. Set $T_0 = 0$ and for $k = 1, \dots, N$:

If k is odd:

$$T_k = \begin{cases} N & \text{if } X_j > 0 \text{ for } j \in [T_{k-1}+1, N] \\ \min\{j \in [T_{k-1}+1, N],\, X_j = 0\} & \text{otherwise} \end{cases}.$$

If k is even:

$$T_k = \begin{cases} N & \text{if } X_j < b-a \text{ for } j \in [T_{k-1}+1, N] \\ \min\{j \in [T_{k-1}+1, N],\, X_j \geq b-a\} & \text{otherwise} \end{cases}.$$

Lastly, set $T_{N+1} = N$.

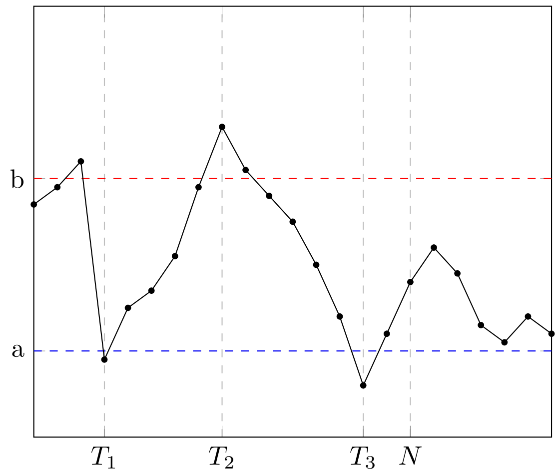


Figure 2: Notice that in this particular example, $N = T_4 = T_5 = \ldots = T_N = T_{N+1}$.

The T_k are stopping times and are defined such that $1 \le T_k \le T_{k+1} \le N$. They are the right-endpoints of an upcrossing or a downcrossing. If for some $k \le N$, $T_k = N$, then $T_{k+1} = N$. We get, applying the optional stopping theorem for bounded stopping times (Theorem 2.39),

$$\mathbb{E}[X_{T_k}] \le \mathbb{E}[X_{T_{k+1}}]. \tag{$*$}$$

Notice that

$$\begin{aligned} X_N - X_1 &= \sum_{k=1}^{N} \left(X_{T_{k+1}} - X_{T_k}\right) + X_{T_1} - X_1 \\ &= \sum_{k \text{ odd}} \left(X_{T_{k+1}} - X_{T_k}\right) + \sum_{k \text{ even}} \left(X_{T_{k+1}} - X_{T_k}\right) + X_{T_1} - X_1 \\ &\ge (b-a)V_{a,b,N} + \sum_{k \text{ even}} \left(X_{T_{k+1}} - X_{T_k}\right) + X_{T_1} - X_1, \end{aligned}$$

where the last inequality comes from the fact that for k odd (T_k, T_{k+1}) is an upcrossing and $(X_{T_{k+1}} - X_{T_k}) \ge (b-a)$. Taking the expectation gives

$$\begin{aligned} \mathbb{E}[X_N - X_1] &\ge (b-a)\mathbb{E}[V_{a,b,N}] + \sum_{k \text{ even}} \Big(\mathbb{E}[X_{T_{k+1}}] - \mathbb{E}[X_{T_k}]\Big) + \Big(\mathbb{E}[X_{T_1}] - \mathbb{E}[X_1]\Big) \\ &\overset{(*)}{\ge} (b-a)\mathbb{E}[V_{a,b,N}]. \end{aligned}$$

Going back to our definition of (X_n), we get the inequality we were looking for:

$$\mathbb{E}[V_{a,b,N}] \le \frac{1}{b-a}\mathbb{E}[(S_N - a)_+ - (S_1 - a)_+].$$

□

Statement 2.55. ***Fatou's lemma.*** *Let $X_n \ge 0$ a.s., then*

$$\mathbb{E}[\liminf_{n\to\infty} X_n] \le \liminf_{n\to\infty} \mathbb{E}[X_n].$$

Theorem 2.56. ***A.s. convergence of submartingales.*** *Let (S_n) be a submartingale relative to (X_n), such that $\sup_{n\in\mathbb{N}} \mathbb{E}[S_n^+] < \infty$. Then there is a random variable S such that*

$$S_n \xrightarrow{a.s.} S \quad \text{and } \mathbb{E}[\,|S|\,] < \infty.$$

Proof. Fix $a < b$, $N \in \mathbb{N}$ and $V_{a,b,N}$ as before. Then

$$\mathbb{E}[V_{a,b,N}] \le \frac{1}{b-a}\,\mathbb{E}[(S_N - a)_+] \le \frac{1}{b-a}\Big(\mathbb{E}[S_N^+] + |a|\Big) \le \frac{1}{b-a}\Big(\sup_{n\in\mathbb{N}} \mathbb{E}[S_n^+] + |a|\Big) < \infty.$$

For $N \nearrow \infty$, we have

$$V_{a,b,N} \nearrow V_{a,b}$$

where $V_{a,b}$ is the number of upcrossings of $[a,b]$ by $(S_n : n \in \mathbb{N})$. The monotone convergence theorem gives

$$\mathbb{E}[V_{a,b}] = \lim_{N\to\infty} \mathbb{E}[V_{a,b,N}] \leq \frac{1}{b-a}\big(\sup_{N\in\mathbb{N}} \mathbb{E}[S_N^+] + |a|\big)$$

and as a consequence $\mathbb{P}(V_{a,b} < \infty) = 1$. Now consider the event

$$\{\liminf_{n\to\infty} S_n < a < b < \limsup_{n\to\infty} S_n\} = \{V_{a,b} = \infty\}.$$

Therefore $\mathbb{P}(\liminf_{n\to\infty} S_n < a < b < \limsup_{n\to\infty} S_n) = 0$ and

$$\mathbb{P}\left(\bigcup_{\substack{a<b\\ a,b\in\mathbb{Q}}} \{\liminf_{n\to\infty} S_n < a < b < \limsup_{n\to\infty} S_n\}\right) = 0.$$

Equivalently, $\mathbb{P}(\liminf_{n\to\infty} S_n < \limsup_{n\to\infty} S_n) = 0$, meaning $\mathbb{P}(\liminf_{n\to\infty} S_n = \limsup_{n\to\infty} S_n) = 1$. Therefore, (S_n) converges a.s. and $S = \lim_{n\to\infty} S_n$ is the desired random variable.

A priori, $S = \pm\infty$ is possible. We now want to exclude this possibility. Thanks to Fatou's lemma, we have

$$\mathbb{E}[S^+] = \mathbb{E}[\lim_{n\to\infty} S_n^+] = \mathbb{E}[\liminf_{n\to\infty} S_n^+] \leq \liminf_{n\to\infty} \mathbb{E}[S_n^+] \leq \sup_{n\in\mathbb{N}} \mathbb{E}[S_n^+] < \infty.$$

So $\mathbb{E}[S^+] < \infty$, and since $S_n = S_n^+ - S_n^-$, we have $\mathbb{E}[S_n^-] = \mathbb{E}[S_n^+] - \mathbb{E}[S_n] \leq \mathbb{E}[S_n^+] - \mathbb{E}[S_1]$ thanks to Remark 2.14. Using Fatou's lemma once again gives

$$\mathbb{E}[S^-] = \mathbb{E}[\liminf_{n\to\infty} S_n^-] \leq \liminf_{n\to\infty} \mathbb{E}[S_n^-] \leq \liminf_{n\to\infty} \mathbb{E}[S_n^+] - \mathbb{E}[S_1] \leq \sup_{n\in\mathbb{N}} \mathbb{E}[S_n^+] - \mathbb{E}[S_1] < \infty.$$

Finally, $\mathbb{E}[\,|S|\,] = \mathbb{E}[S^+] + \mathbb{E}[S^-] < \infty$.

□

Theorem 2.57. ***L^2-convergence of L^2-bounded martingales.*** *Let (S_n) be a martingale relative to (X_n) such that $\sup_{n\in\mathbb{N}} \mathbb{E}[S_n^2] < \infty$. Then there is a random variable S such that*

$$\lim_{n\to\infty} S_n = S \quad a.s. \quad and \quad \lim_{n\to\infty} \mathbb{E}[(S_n - S)^2] = \lim_{n\to\infty} ||S_n - S||_{L^2}^2 = 0.$$

Proof. We have already shown that there exists a random variable S such that $\lim_{n\to\infty} S_n = S$ a.s. It remains to prove $\lim_{n\to\infty} \mathbb{E}[(S_n - S)^2] = 0$. Using the variant of the Doob-Kolmogorov inequality (Proposition 2.48) with $p = 2$, we have

$$\mathbb{E}[\max_{i\in\{1,\dots n\}} S_i^2] \leq 4 \cdot \mathbb{E}[S_n^2]. \qquad (*)$$

Noticing $\max_{i\in\{1,\dots n\}} S_i^2 \nearrow \sup_{i\in\mathbb{N}} S_i^2$ as $n \to \infty$, i.e., $(\max_{i\in\{1,\dots,n\}} S_i^2)_n$ is a sequence of nonnegative random variables increasing to $\sup_{i\in\mathbb{N}} S_i^2$, we can use the monotone convergence theorem:

$$\mathbb{E}[\sup_{i\in\mathbb{N}} S_i^2] = \mathbb{E}[\lim_{n\to\infty} \max_{i\in\{1,\dots,n\}} S_i^2] \overset{(MCT)}{=} \lim_{n\to\infty} \mathbb{E}[\max_{i\in\{1,\dots,n\}} S_i^2].$$

And now $(*)$ gives

$$\mathbb{E}[\sup_{i\in\mathbb{N}} S_i^2] \leq 4 \cdot \lim_{n\to\infty} \mathbb{E}[S_n^2] \leq 4 \cdot \sup_{n\in\mathbb{N}} \mathbb{E}[S_n^2] < \infty,$$

where the last inequality holds by assumption.

Before we can conclude this proof, observe that

$$\begin{aligned}(S_n - S)^2 &= |S_n - S|^2 \\ &\leq (|S_n| + |S|)^2 \\ &\leq (2\sup_{i\in\mathbb{N}} |S_i|)^2 \\ &= 4\sup_{i\in\mathbb{N}} S_i^2 < \infty.\end{aligned}$$

Let $W = 4\sup_{i\in\mathbb{N}} S_i^2$. We have $\mathbb{E}[W] < \infty$, $(S_n - S)^2 \to 0$ a.s. as $n \to \infty$ and $(S_n - S)^2 \leq W$ for all $n \in \mathbb{N}$ and we can therefore apply the dominated convergence theorem

$$\lim_{n\to\infty} \mathbb{E}[(S_n - S)^2] = \mathbb{E}[\lim_{n\to\infty} (S_n - S)^2] = \mathbb{E}[0] = 0,$$

which concludes the proof. □

Remark 2.58. *As a side note, to better understand the steps of this proof (as well as many others), retracing the steps in reverse is often beneficial. Here, for example, we want to show that* $\lim_{n\to\infty} \mathbb{E}[(S_n - S)^2] = 0$, *but we already know from a previous result that* $S_n \to S$ *almost surely for some random variable* S. *From this, we infer that* $(S_n - S)^2 \to 0$ *almost surely, and we now just need to find a way to show* $\mathbb{E}[(S_n - S)^2] \to 0$ *as* $n \to \infty$. *Multiple theorems and results could be helpful to show this; the only thing left to do is to find which one fits our situation best. In this case, the monotone convergence theorem won't be of any help since we do not know whether the sequence is increasing in* n *or not. Another possibility would be the bounded convergence theorem (essentially the dominated convergence theorem but using a constant to bound the sequence of random variables), but again, it seems quite difficult to find a constant that would bound our sequence* $(S_n - S)^2$ *since we know so little about it. Finally, we can try our luck with the dominated convergence theorem. In this particular case, the random variable used as a bound comes out naturally, but most of the time, more work will be needed to find a suitable one.*

3 Branching Processes

In probability theory, a *branching process* is a type of stochastic process, which is a collection of random variables indexed by a set, typically the natural numbers or non-negative real numbers. Originally, branching processes were developed to model populations where each individual in generation n produces a random number of individuals in generation $n+1$. This reproduction is often modeled using a fixed probability distribution that is consistent across all individuals. We will primarily consider branching processes in this context.

Branching processes are frequently employed to model biological reproduction scenarios. For example, consider bacteria where each bacterium can produce 0, 1, or 2 offspring with certain probabilities within a single time unit. Beyond biological applications, branching processes can also be used to model other dynamic systems, such as the spread of surnames in genealogy or the propagation of neutrons in a nuclear reactor.

Hypothesis A: The numbers of offspring from different individuals are independent and identically distributed (i.i.d.) random variables, each following the distribution of a random variable Z. We will write $p_j = \mathbb{P}(Z = j)$ for $j \in \mathbb{N}$.

Remark 3.1. $\mathbb{E}[Z]$ *is the expected number of offspring per individual. If* $\mathbb{E}[Z] > 1$*, the population should grow and in case* $\mathbb{E}[Z] < 1$*, we would expect the population to decrease. If* $\mathbb{E}[Z] = 1$*, what happens ?*

Our first objective will be to determine the **extinction probability**, i.e., the probability that there is an $n \in \mathbb{N}$ such that $X_n = 0$. Note that this implies $X_m = 0$ for $m \geq n$. With this goal in mind, we define the **generating function** of the random variable Z, denoted by $g_Z(\cdot)$, as follows:

$$
\begin{aligned}
g_Z(s) &= \mathbb{P}(Z=0) + s\mathbb{P}(Z=1) + s^2\mathbb{P}(Z=2) + \dots \\
&= \sum_{j=0}^{\infty} s^j \mathbb{P}(Z=j) \\
&= \sum_{j=0}^{\infty} p_j s^j .
\end{aligned}
$$

And we give some properties that follow immediately from the definition.

Proposition 3.2. *(a) The series* $\sum_{j=0}^{\infty} p_j s^j$ *converges uniformly on* $[-1,1]$*;*

(b) $g_Z(s) = \mathbb{E}[s^Z]$*;*

(c) If Z *and* Z' *are two independent random variables taking values in* $\mathbb{N}$*, then* $g_{Z+Z'}(s) = g_Z(s) \cdot g_{Z'}(s)$*;*

(d) If $\sum_{j=0}^{\infty} j\, p_j < \infty$*, then* $g'_Z(1) = \mathbb{E}[Z]$*, and if* $\sum_{j=0}^{\infty} j^2\, p_j < \infty$*, then* $Var(Z) = g''_Z(1) + \mathbb{E}[Z] - \mathbb{E}[Z]^2$*.*

Proof. See Problem Set 7. □

Remark 3.3. *Recall that the moment generating function of the random variable* Z *is given by* $M_Z(t) = \mathbb{E}[e^{tZ}]$*, provided this expectation exists for* t *in some open neighborhood of zero. So,* $g_Z(s) = M_Z(\ln s)$*.*

Let us now explore in more details the generating function of X_n, when we set $X_0 = 1$. By definition,

$$
g_{X_{n+1}}(s) = \mathbb{E}[s^{X_{n+1}}] = \mathbb{E}\left(\mathbb{E}[s^{X_{n+1}} \mid X_n]\right) = \mathbb{E}[\psi(X_n)],
$$

where $\psi(j) = \mathbb{E}[s^{X_{n+1}} \mid X_n = j]$. Observe that

$$\mathbb{E}[s^{X_{n+1}} \mid X_n = j] = \mathbb{E}[s^{Z_1^n + \ldots + Z_j^n} \mid X_n = j]$$

with Z_i^n the number of offspring of individual i of generation n. Noticing that $\{X_n = j\}$ is determined by the Z_i^k, where $k < n$, Hypothesis A gives

$$\mathbb{E}[s^{X_{n+1}} \mid X_n = j] = \mathbb{E}[s^{Z_1^n + \ldots + Z_j^n}] = \mathbb{E}[s^{Z_1^n}] \cdot \ldots \cdot \mathbb{E}[s^{Z_j^n}] = \big(\mathbb{E}[s^Z]\big)^j = \big(g_Z(s)\big)^j,$$

as the event $\{X_n = j\}$ is independent of the Z_i^k, $k < n$. Therefore,

$$\mathbb{E}[s^{X_{n+1}} \mid X_n] = \psi(X_n) = \big(g_Z(s)\big)^{X_n}$$

and

$$g_{X_{n+1}}(s) = \mathbb{E}[(g_Z(s))^{X_n}] = g_{X_n}\big(g_Z(s)\big).$$

That is,

$$\begin{aligned} g_{X_{n+1}} &= g_{X_n} \circ g_Z = (g_{X_{n-1}} \circ g_Z) \circ g_Z \\ &= g_{X_{n-1}} \circ g_Z^{(2)} \\ &= \ldots \\ &= g_{X_1} \circ g_Z^{(n)} \\ &= g_Z^{(n+1)}. \end{aligned}$$

The last equality follows from the fact that $g_{X_1} = g_Z$, since $X_0 = 1$. This gives the following formula for g_{X_n}:

$$g_{X_n} = g_Z^{(n)} = \underbrace{g_Z \circ g_Z \circ \ldots \circ g_Z}_{n \text{ times}}.$$

Let's return to our main objective: determining the extinction probability. Let $F = \bigcup_{n=1}^{\infty} \{X_n = 0\} = \{\text{ the population becomes extinct }\}$. We want to find $\mathbb{P}(F)$. Observe that $\{X_n = 0\} \subseteq \{X_{n+1} = 0\}$, so $\mathbb{P}(F) = \lim_{n\to\infty} \mathbb{P}(X_n = 0)$.

Theorem 3.4. *Suppose $X_0 = 1$ and Hypothesis A holds. Then $\mathbb{P}(F)$ is the smallest number $\alpha \geq 0$ such that*

$$\alpha = \sum_{j=0}^{\infty} \alpha^j \, p_j = g_Z(\alpha).$$

Remark 3.5. *(a) If the probability that an individual has no offspring is zero, i.e., $p_0 = \mathbb{P}(Z = 0) = 0$, then $0 = g_Z(0)$ and $\mathbb{P}(F) = 0$.*

(b) Case of male descendants in the U.S.A (Lotka 1931 [3]). Statistical methods showed that

$$g_Z(s) = \frac{0.482 - 0.041s}{1 - 0.559s}.$$

So $\alpha = g_Z(\alpha)$ gives a 2^{nd} degree polynomial equation with two solutions: 1, which is always a solution, and $\alpha = 0.86$, which was interpreted as the probability of extinction of a family name.

Proof. Recall $\mathbb{P}(F) = \lim_{n\to\infty} \mathbb{P}(X_n = 0)$. Set $\alpha_n = \mathbb{P}(X_n = 0)$, we know that $\alpha = \lim_{n\to\infty} \alpha_n = \mathbb{P}(F)$ exists. Therefore,

$$\alpha_n = \mathbb{P}(X_n = 0) = g_{X_n}(0) = g_Z^{(n)}(0) = g_Z\big(g_Z^{(n-1)}(0)\big) = g_Z\big(g_{X_{n-1}}(0)\big) = g_Z(\alpha_{n-1}).$$

If we now let $n \to \infty$, we get

$$\alpha = \lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} g_Z(\alpha_{n-1}) \overset{(*)}{=} g_Z\big(\lim_{n\to\infty} \alpha_{n-1}\big) = g_Z(\alpha),$$

where we used the continuity of g_Z in $(*)$. Let us now prove the second part of the claim. Let $\beta \geq 0$ be another solution, i.e., $\beta = g_Z(\beta)$. Since g_Z in nondecreasing on $[0,1]$ and $0 \leq \beta$, we have

$$g_Z(0) \leq g_Z(\beta) = \beta.$$

Again,

$$g_Z^{(2)}(0) = g_Z\big(g_Z(0)\big) \leq g_Z(g_Z(\beta)) = g_Z(\beta) = \beta.$$

Recall that $\alpha_n = g_Z^{(n)}(0)$. Repeating what we did above, we will eventually get

$$\alpha_n = g_{X_n}(0) = g_Z^{(n)}(0) \leq g_Z^{(n)}(\beta) = \beta.$$

Since this holds for all $n \in \mathbb{N}$, letting $n \to \infty$ gives the desired result: $\alpha \leq \beta$. □

Hypothesis B:

(a) We will assume $p_j < 1$ for all j and $p_0 + p_1 < 1$. If $p_0 + p_1 = 1$, then $p_j = 0$ for $j \geq 2$ and each individual has 0 or 1 offspring, nothing interesting happens.

(b) $\mathbb{E}[Z] = \sum_{j=1}^{\infty} j\, p_j < \infty$.

Proposition 3.6. *Under Hypotheses A and B, we have*

1. *Extinction probability* $= 1 \iff \mathbb{E}[Z] \leq 1$.
2. *Extinction probability* $= 0 \iff p_0 = \mathbb{P}(Z = 0) = 0$.

Proof. Follows from the fact that $g_Z(1) = 1$ and $s \mapsto g_Z(s)$ is strictly convex and nondecreasing on $[0,1]$. □

Let us now study the asymptotic behaviour of X_n as $n \to \infty$.

Lemma 3.7. *Set* $m = \mathbb{E}[Z]$, $\sigma^2 = Var(Z) < \infty$. *Then*

$$\mathbb{E}[X_n] = m^n$$

and

$$Var(X_n) = \begin{cases} \sigma^2 \dfrac{m^n(m^n - 1)}{m^2 - m} & , \textit{ for } m \neq 1 \\ n\sigma^2 & , \textit{ for } m = 1 \end{cases}.$$

Proof. First notice that

$$\mathbb{E}[X_n] = g'_{X_n}(1) = \big(g_{X_{n-1}} \circ g_Z\big)'(1) = g'_{X_{n-1}}\big(g_Z(1)\big) \cdot g'_Z(1) = g'_{X_{n-1}}(1) \cdot m,$$

meaning $\mathbb{E}[X_n] = m\mathbb{E}[X_{n-1}]$. By iterating, we get the desired result: $\mathbb{E}[X_n] = m^n$. For the second part of the claim, we have

$$g'_{X_{n+1}}(s) = g'_{X_n}\big(g_Z(s)\big) \cdot g'_z(s)$$

and

$$\begin{aligned} g''_{X_{n+1}}(s) &= g''_{X_n}\big(g_Z(s)\big) \cdot g'_Z(s) \cdot g'_Z(s) + g'_{X_n}\big(g_Z(s)\big) \cdot g''_Z(s) \\ &= g''_{X_n}\big(g_Z(s)\big) \cdot g'_Z(s)^2 + g'_{X_n}\big(g_Z(s)\big) \cdot g''_Z(s). \end{aligned}$$

Recall point (d) of Proposition 3.2

$$g''_{X_n}(1) = \mathrm{Var}(X_n) - \mathbb{E}[X_n] + \big(\mathbb{E}[X_n]\big)^2.$$

From this, we get

$$\begin{aligned} \mathrm{Var}(X_{n+1}) &= g''_{X_{n+1}}(1) + m^{n+1} - m^{2(n+1)} \\ &= g''_{X_n}(\underbrace{g_Z(1)}_{=1}) \cdot \big(\underbrace{g'_Z(1)}_{=m}\big)^2 + g'_{X_n}\big(\underbrace{g'_Z(1)}_{=1}\big) \cdot g''_Z(1) + m^{n+1} - m^{2(n+1)} \\ &= \big(\mathrm{Var}(X_n) - \mathbb{E}[X_n] + \mathbb{E}[X_n]^2\big) \cdot m^2 + \mathbb{E}[X_n] \cdot (\sigma^2 - m + m^2) + m^{n+1} - m^{2(n+1)} \\ &= \mathrm{Var}(X_n) \cdot m^2 - m^{n+2} + m^{2n+2} + m^n\sigma^2 - m^{n+1} + m^{n+2} + m^{n+1} - m^{2n+2} \\ &= \mathrm{Var}(X_n) \cdot m^2 + m^n\sigma^2. \end{aligned}$$

Finally,

$$\begin{aligned} \mathrm{Var}(X_{n+1}) &= m^n\sigma^2 + m^2\mathrm{Var}(X_n) \\ &= m^n\sigma^2 + m^2\big(m^{n-1}\sigma^2 + m^2\mathrm{Var}(X_n)\big) \\ &= (m^n + m^{n+1})\sigma^2 + m^4\mathrm{Var}(X_{n-1}) \\ &= \dots \\ &= (m^n + m^{n+1} + \dots + m^{2n})\sigma^2 + m^{2(n+1)}\underbrace{\mathrm{Var}(X_0)}_{=0} \\ &= \begin{cases} (n+1)\sigma^2 & \text{, if } m = 1 \\ \dfrac{m^{2n+1} - m^n}{m-1}\sigma^2 & \text{, if } m \neq 1 \end{cases}. \end{aligned}$$

We can now simply rewrite the case $m \neq 1$ and get

$$\mathrm{Var}(X_{n+1}) = \begin{cases} (n+1)\sigma^2 & \text{, if } m = 1 \\ \dfrac{m^{n+1}(m^{n+1} - 1)}{m^2 - m}\sigma^2 & \text{, if } m \neq 1 \end{cases}.$$

□

Theorem 3.8. *Let $S_n = m^{-n}X_n$. Then (S_n) is a martingale relative to (X_n). If $m > 1$, there exists a random variable S such that $S_n \to S$ a.s. and $\mathbb{E}[(S_n - S)^2] \to 0$ as $n \to \infty$.*

Proof. We first check $\mathbb{E}[\,|S_n|\,]$ is finite for all $n \geq 1$:

$$\mathbb{E}[\,|S_n|] = \mathbb{E}[S_n] = \frac{1}{m^n}\mathbb{E}[X_n] = 1 < \infty.$$

For the second condition, we have

$$\mathbb{E}[S_{n+1}\,|\,X_1, \ldots, X_n] = \mathbb{E}\left[\frac{X_{n+1}}{m^{n+1}}\,|\,X_1, \ldots, X_n\right] = \frac{1}{m^{n+1}}\mathbb{E}[X_{n+1}\,|\,X-1, \ldots, X_n].$$

Observe that

$$\begin{aligned}\mathbb{E}[X_{n+1}\,|\,X_n = j_n, \ldots, X_1 = j_1] &\overset{\text{(Hyp. A)}}{=} \mathbb{E}[Z_1^n + \ldots + Z_{j_n}^n] \\ &= j_n \cdot \mathbb{E}[Z] \\ &= j_n \cdot m,\end{aligned}$$

where Z_k^n is the number of offspring of individual k in the generation n. From this, we get

$$\mathbb{E}[X_{n+1}\,|\,X_1, \ldots, X_n] = mX_n$$

and therefore

$$\mathbb{E}[S_{n+1}\,|\,X_1, \ldots, X_n] = \frac{1}{m^{n+1}} \cdot m \cdot X_n = \frac{X_n}{m^n} = S_n.$$

Now, suppose $m > 1$. We want to use the Martingale convergence theorem for L^2-bounded martingales, to do so, we first need to check that (S_n) is an L^2-bounded martingale:

$$\begin{aligned}\mathbb{E}[S_n^2] &= \mathrm{Var}(S_n) + \big(\mathbb{E}[S_n]\big)^2 \\ &= \frac{1}{m^{2n}}\mathrm{Var}(X_n) + 1 \\ &= \frac{\sigma^2}{m^{2n}} \cdot \frac{m^n(m^n-1)}{m^2-m} + 1 \\ &= \frac{\sigma^2}{m^2-m}(1 - \frac{1}{m^n}) + 1 \\ &\leq \frac{\sigma^2}{m^2-m} + 1 < \infty.\end{aligned}$$

We can apply our theorem, which guarantees the existence of a random variable S such that $S_n \to S$ a.s. and $\mathbb{E}[(S_n - S)^2] \to 0$ as $n \to \infty$. $\square$

Conclusions:

(a) If $m \leq 1$, then $\mathbb{P}(\exists n : X_n = 0) = 1$ (See Proposition 3.6).

(b) If $m > 1$, then for large n, $X_n \sim m^n S$.

First note that $\mathbb{P}(S \geq 0) = 1$. If $S > 0$, the population grows exponentially. In case $S = 0$, we have that $m^{-n}X_n \to 0$, the population "grows" more slowly than $n \mapsto m^n$. In fact, we will show that the population becomes extinct in this case. From this, we get that there is two possibilities for the population of interest: "either extinction or explosion".

Remark 3.9. *Since* $\mathbb{E}[(S_n - S)^2] \to 0$*, we have*

$$\underbrace{\mathbb{E}[S_n]}_{=1} \to \mathbb{E}[S] = 1, \quad Var(S_n) \to Var(S) = \frac{\sigma^2}{m^2 - m} > 0.$$

Proposition 3.10. *If* $m > 1$.

(1) $\mathbb{P}(S = 0) = \mathbb{P}(\{\textit{extinction}\})$.

(2) Var($S \mid S > 0$*)*> 0 *: given* $S > 0$ *(no extinction),* S *is still random, it is not a constant.*

Proof. We will only give a proof for (1). Clearly, {extinction} $\subseteq \{S = 0\}$, because $S_n = m^{-n} X_n \to S$ a.s., so $\mathbb{P}(\{\text{extinction}\}) \leq \mathbb{P}(S = 0)$.

To verify that this inequality holds as an equality, let $q = \mathbb{P}(S = 0)$ and note that

$$\{S = 0\} = \left\{ \lim_{n\to\infty} S_n = 0 \right\}.$$

Now, observe that

$$\begin{aligned}
\mathbb{P}\left(\lim_{n\to\infty} S_n = 0 \mid X_1 = k \right) &= \mathbb{P}\left(\lim_{n\to\infty} \frac{1}{m^n} \sum_{l=1}^{k} X_n^{(k)} = 0 \right) \\
&= \mathbb{P}\left(\bigcap_{l=1}^{k} \{ \lim_{n\to\infty} m^{-n} X_n^{(l)} = 0\} \mid X_1 = k \right) \\
&\overset{(*)}{=} \prod_{l=1}^{k} \mathbb{P}\left(\lim_{n\to\infty} m^{-n} X_n^{(l)} \mid X_1 = k \right) \\
&= \prod_{l=1}^{k} \mathbb{P}\left(\lim_{n\to\infty} m^{-n} X_n = 0 \right) \\
&= \left(\mathbb{P}(S = 0) \right)^k.
\end{aligned}$$

Therefore,

$$\begin{aligned}
q = \mathbb{P}(S = 0) &= \sum_{k=0}^{\infty} \underbrace{\mathbb{P}(S = 0 \mid X_1 = k)}_{q^k} \cdot \underbrace{\mathbb{P}(X_1 = k)}_{p_k} \\
&= \sum_{k=0}^{\infty} q^k \cdot p_k \\
&= g_Z(q).
\end{aligned}$$

That is: $q = g_Z(q)$. In addition, $q = \mathbb{P}(S = 0) < 1$ because $\mathrm{Var}(S) > 0$. We know that $s = g_Z(s)$ has exactly two nonnegative solutions (when $m > 1$), namely 1 and $\mathbb{P}(\{\text{extinction}\})$, therefore $q = \mathbb{P}(\{\text{extinction}\})$.

□

4 Brownian motion

Definition 4.1. *A **Brownian motion** (BM) is a stochastic process* $(B_t : t \in \mathbb{R}_+)$ *with two properties:*

(a) For all $s, t \in \mathbb{R}_+$*, we have*

$$B_{t+s} - B_s \sim \mathcal{N}(0, \sigma^2 t),$$

with $\sigma > 0$ *a fixed parameter.*

(b) For $t_1 < t_2 \leq t_3 < t_4$*, the increments* $(B_{t_4} - B_{t_3})$ *and* $(B_{t_2} - B_{t_1})$ *are independent. The same holds for increments over* $n \in \mathbb{N}$ *nonoverlapping intervals.*

Remark 4.2.

1. *The law of* B_0 *is not specified.*
2. *We will often write* $B(t)$ *instead of* B_t *for clarity.*
3. *Unless stated otherwise, we will consider the case* $\sigma = 1$ *and* $B_0 = 0$*. We will designate this specific instance as a standard Brownian motion. We have in this case* $B_t \sim \mathcal{N}(0, t)$*,* $t \in \mathbb{R}_+$*.*

We now want to find an expression for the joint probability distribution of $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$, where $0 < t_1 < \dots < t_n$. Define

$$p(x, t) = \frac{1}{\sqrt{2\pi t}} \cdot \exp\left\{-\frac{x^2}{2t}\right\},$$

which is the probability density function of a $\mathcal{N}(0, t)$.

Proposition 4.3. *Using the function* $p(x, t)$ *defined above, we have*

$$f_{(B_{t_1}, \dots, B_{t_n})}(x_1, \dots, x_n) = p(x_1, t_1) p(x_2 - x_1, t_2 - t_1) \dots p(x_n - x_{n-1}, t_n - t_{n-1}).$$

Proof. By property (b) of our definition, the probability density function of $Y = (B_{t_1} - B_0, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$, where $B_0 = 0$, is given by

$$g(y_1, \dots, y_n) = p(y_1, t_1 - t_0) p(y_2, t_2 - t_1) \dots p(y_n, t_n - t_{n-1}).$$

Define $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(x_1, \dots, x_n) = (x_1, x_2 - x_1, \dots, x_n - x_{n-1})$. Then, $T(B_{t_1}, \dots, B_{t_n}) = Y$. Recall that if $Y = T(X)$ with T bijective and $\mathcal{C}^0$, then

$$f_X(x) = f_Y\big(T(x)\big) \cdot |det(J)|$$

where $J = (\partial T_i / \partial x_j)$ is the Jacobian matrix. In our case, the Jacobian matrix is of the form

$$J = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix},$$

meaning $\det(J) = 1$. Therefore,

$$f_{(B_{t_1}, \dots, B_{t_n})}(x_1, \dots, x_n) = p(x_1, t_1) p(x_2 - x_1, t_2 - t_1) \dots p(x_n - x_{n-1}, t_n - t_{n-1}).$$

□

Lemma 4.4. ***Translation invariance.*** *For $s < t$, the conditional law of B_t given $B_s = y$ has the following property:*

$$\mathbb{P}(B_t \le x \,|\, B_s = y) = \mathbb{P}(B_t \le x - y \,|\, B_s = 0).$$

Proof. We know that the joint density of (B_s, B_t) is:

$$f(y, x) = p(y, s)p(x - y, t - s).$$

The conditional density of B_t given $B_s = y$ is:

$$f_{B_t \,|\, B_s}(x \,|\, y) = \frac{f(y,\, s)}{f_{B_s}(y)} = \frac{p(y,\, s)p(x - y,\, t - s)}{p(y,\, s)} = p(x - y,\, t - s).$$

It follows that

$$\begin{aligned}
\mathbb{P}(B_t \le x \,|\, B_s = y) \quad &= \quad \int_{-\infty}^{x} p(v - y, t - s)dv \\
&\overset{(u=v-y)}{=} \int_{-\infty}^{x-y} p(u, t - s)du \\
&= \quad \mathbb{P}(B_t \le x - y \,|\, B_s = 0).
\end{aligned}$$

□

Statement 4.5. ***Continuity of sample paths.*** *Fix $\omega \in \Omega$. With probability one,*

$$\begin{aligned}
\mathbb{R}_+ &\to \mathbb{R} \\
t &\mapsto B_t(\omega)
\end{aligned}$$

is continuous.

Theorem 4.6. ***Markov property.*** *For $t_1 < \ldots < t_n < t$,*

$$\mathbb{P}(B_t \le x \,|\, B_{t_1} = x_1, \ldots, B_{t_n} = x_n) = \mathbb{P}(B_t \le x \,|\, B_{t_n} = x_n).$$

Proof. See Problem Set 8. □

Proposition 4.7. *The covariance of B_s and B_t is given by*

$$Cov(B_s, B_t) = \mathbb{E}[B_s B_t] = s \wedge t.$$

Proof. Without loss of generality, suppose $s \le t$. Using the independence of increments and the fact that B_t follows a $\mathcal{N}(0, t)$ distribution, we have

$$\begin{aligned}
\mathbb{E}[B_s B_t] &= \mathbb{E}\left[B_s\big((B_t - B_s) + B_s\big)\right] \\
&= \mathbb{E}[B_s(B_t - B_s)] + \mathbb{E}[B_s^2] \\
&= \mathbb{E}[B_s] \cdot \mathbb{E}[B_t - B_s] + s \\
&= 0 + s \\
&= s \wedge t.
\end{aligned}$$

□

4.1 Probabilities of behaviors of Brownian motion

Our goal is now to derive expressions for the probabilities of behaviors defined by specified values at particular times. Let $0 < t_1 < \ldots < t_n$ be specific time points and $a_i, b_i \in \mathbb{R}$ such that $a_i < b_i$, with $i \in \{1, \ldots, n\}$. The probability for B_{t_i} to be in the interval $[a_i, b_i]$ for all $i \in \{1, \ldots, n\}$ is

$$\mathbb{P}(B_{t_1} \in [a_1, b_1], \ldots, B_{t_n} \in [a_n, b_n]) = \int_{a_1}^{b_1} dx_1 \ldots \int_{a_n}^{b_n} dx_n \, p(x_1, t_1) p(x_2 - x_1, t_2 - t_1) \\ \ldots p(x_n - x_{n-1}, t_n - t_{n-1}).$$

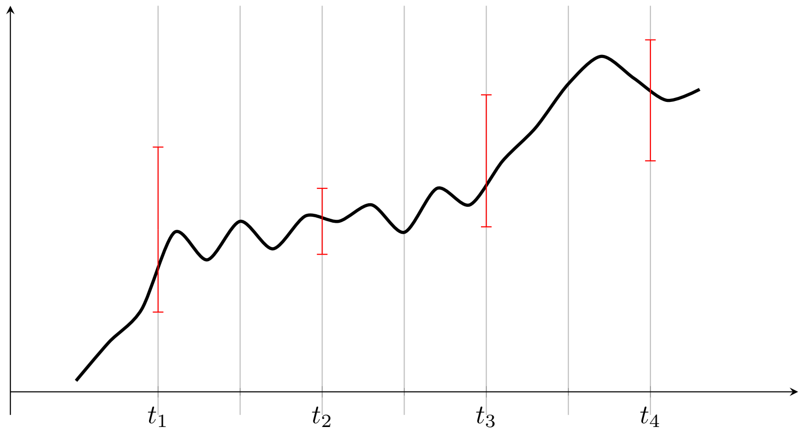


Figure 3: One example of a sample path that satisfies the constraints at times t_1, t_2, t_3 and t_4.

Remark 4.8. *Many interesting events are not of this kind, for example:*

$$\mathbb{P}\left(\max_{u \in [0,t]} B_u > a\right) = \mathbb{P}\left(\bigcup_{r \in \mathbb{Q} \cap [0,t]} \{B_r > a\}\right).$$

Proposition 4.9. ***Reflection principle (Bachelier).*** *For $a \geq 0$,*

$$\mathbb{P}\left(\max_{u \in [0,t]} B_u \geq a\right) = 2\mathbb{P}(B_t \geq a).$$

The formal proof of the above proposition requires the use of the strong Markov property (Theorem 4.53), which we have not yet introduced. Therefore, we will provide an informal explanation of the result and indicate the points at which the strong Markov property would be applied in a rigorous proof. Let

$$\tau = \begin{cases} \inf\{a \geq 0 : B_u = a\} & \text{, if } \{u \geq 0 : B_u = a\} \neq \emptyset \\ +\infty & \text{, otherwise} \end{cases}.$$

Define

$$\tilde{B}_t = \begin{cases} B_t & \text{, if } t < \tau \\ a - (B_t - a) & \text{, if } t \geq \tau \end{cases}.$$

Intuitively, $(\tilde{B}_t)$ is also a BM by symmetry. However, this is where one would need to use the strong Markov property in order to get a rigorous justification. Consider the event

$$\left\{\max_{u\in[0,t]} B_u \geq a,\, B_t \geq a\right\} = \left\{\max_{u\in[0,t]} \tilde{B}_u \geq a,\, \tilde{B}_t \leq a\right\}.$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a,\, B_t \geq a\right) &= \mathbb{P}\left(\max_{u\in[0,t]} \tilde{B}_u \geq a,\, \tilde{B}_t \leq a\right) \\ &\overset{(*)}{=} \mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a,\, B_t \leq a\right),\end{aligned}$$

where the equality marked by $(*)$ follows form the fact that $\tilde{B}_t$ and B_t are both BMs. From this, and recalling that $\mathbb{P}(B_t = a) = 0$ for the equality denoted by $(\star)$, we get the desired result

$$\begin{aligned}\mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a\right) &\overset{(\star)}{=} \mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a,\, B_t \geq a\right) + \mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a,\, B_t \leq a\right) \\ &= 2\mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a,\, B_t \geq a\right) \\ &= 2\mathbb{P}\left(B_t \geq a\right).\end{aligned}$$

Proposition 4.10. *Fix $a \neq 0$. We define the first hitting time of level a by*

$$T_a = \begin{cases} \inf\{t \geq 0 : B_t = a\} & ,\ \text{if } \{t \geq 0 : B_t = a\} \neq \emptyset \\ +\infty & ,\ \text{otherwise} \end{cases}.$$

The probability density function of T_a is

$$f_{T_a}(t) = \frac{|a|}{\sqrt{2a}} \cdot \frac{1}{\sqrt{t^3}} \exp\left\{\frac{-a^2}{2t}\right\},$$

for $t > 0$.

Proof. Suppose $a > 0$. Then

$$\{T_a \leq t\} = \left\{\max_{u\in[0,t]} B_u \geq a\right\}$$

and from this

$$\begin{aligned}\mathbb{P}(T_a \leq t) &= \mathbb{P}\left(\max_{u\in[0,t]} B_u \geq a\right) \\ &= 2\mathbb{P}(B_t \geq a) \\ &= 2\int_a^\infty \frac{1}{\sqrt{2\pi t}} \exp\left\{\frac{-x^2}{2t}\right\} dx \\ &\overset{(*)}{=} \frac{2}{\sqrt{2\pi t}} \int_{a/\sqrt{t}}^\infty \exp\left\{\frac{-y^2}{2}\right\} \sqrt{t}\, dy \\ &= \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^\infty \exp\left\{\frac{-y^2}{2}\right\} dy\end{aligned}$$

where we used the change of variable $x = y\sqrt{t}$ in $(*)$. Finally,

$$f_{T_a}(t) = \frac{d}{dt}\mathbb{P}(T_a \le t) = \frac{2}{\sqrt{2\pi}}\left(-\exp\left\{\frac{-a^2}{2t}\right\}\right)\left(\frac{-a}{2\sqrt{t^3}}\right) = \frac{a}{\sqrt{2\pi}}\frac{1}{\sqrt{t^3}}\exp\left\{\frac{-a^2}{2t}\right\}.$$

The case $a < 0$ can be treated analogously. □

Remark 4.11. *(a) The stopping time T_a is a.s. finite:*

$$\mathbb{P}(T_a < \infty) = \lim_{t\to\infty}\mathbb{P}(T_a \le t) = \lim_{t\to\infty}\frac{2}{\sqrt{2\pi}}\int_{a/\sqrt{t}}^{\infty}\exp\left\{\frac{-y^2}{2}\right\}dy = \frac{2}{\sqrt{2\pi}}\int_0^{\infty}\exp\left\{\frac{-y^2}{2}\right\}dy = 1.$$

(b) The stopping time T_a does not have a finite expectation:

$$\begin{aligned}\mathbb{E}[T_a] = \int_0^{\infty} t\, f_{T_a}(t)dt &= \int_0^{\infty} t\frac{a}{\sqrt{2\pi}}\frac{1}{\sqrt{t^3}}\exp\left\{\frac{-a^2}{2t}\right\}dt \ge \int_1^{\infty}\frac{a}{\sqrt{2\pi}}\frac{1}{\sqrt{t}}\exp\left\{\frac{-a^2}{2t}\right\}dt\\ &\ge \int_1^{\infty}\frac{a}{\sqrt{2\pi}}\frac{1}{\sqrt{t}}\exp\left\{\frac{-a^2}{2}\right\}dt = \infty.\end{aligned}$$

Proposition 4.12. *The probability $\mathbb{P}(\min_{u\in[0,t]} B_u \le 0 \,|\, B_0 = a)$, i.e., the probability of moving down a units is given by*

$$\mathbb{P}\left(\min_{u\in[0,t]} B_u \le 0 \,|\, B_0 = a\right) = \frac{a}{\sqrt{2\pi}}\int_0^t u^{-3/2}\exp\left\{\frac{-a^2}{2u}\right\}du.$$

Proof. By symmetry and invariance under translation in space, we have

$$\begin{aligned}\mathbb{P}\left(\min_{u\in[0,t]} B_u \le 0 \,|\, B_0 = a\right) &= \mathbb{P}\left(\max_{u\in[0,t]} B_u \ge 2a \,|\, B_0 = a\right)\\ &= \mathbb{P}\left(\max_{u\in[0,t]} B_u \ge a \,|\, B_0 = 0\right)\\ &= \mathbb{P}(T_a \le t \,|\, B_0 = 0)\\ &= \frac{a}{\sqrt{2\pi}}\int_0^t u^{-3/2}\exp\left\{\frac{-a^2}{2u}\right\}du.\end{aligned}$$

□

Proposition 4.13. *As a direct consequence of Proposition 4.12, we get that the probability of visiting zero, starting at $B_{t_0} = a$, during $[t_0, t_1]$, for $0 < t_0 < t_1$, is*

$$\mathbb{P}\left(\min_{u\in[t_0,t_1]} B_u \le 0 \,|\, B_{t_0} = a\right) = \mathbb{P}\left(\min_{u\in[0,t_1-t_0]} B_u \le 0 \,|\, B_0 = a\right) = \frac{a}{\sqrt{2\pi}}\int_0^{t_1-t_0} u^{-3/2}\exp\left\{\frac{-a^2}{2u}\right\}du.$$

Proposition 4.14. *Let $0 < t_0 < t_1$. The probability α of visiting zero in the interval (t_0, t_1) is*

$$\alpha = \mathbb{P}(\exists t \in (t_0, t_1)\, B_t = 0 \,|\, B_0 = 0) = \frac{2}{\pi} \arccos\left(\sqrt{\frac{t_0}{t_1}}\right).$$

Proof. First recall the following formula of total probability:

$$\mathbb{P}(F) = \int_{-\infty}^{\infty} \mathbb{P}(F \,|\, X = x)\, f_X(x) dx.$$

We have

$$\begin{aligned}
\alpha &= \int_{-\infty}^{\infty} \mathbb{P}(\exists t \in (t_0, t_1) \,:\, B_t = 0 \,|\, B_{t_0} = a)\, f_{B_{t_0}}(a)\, da \\
&= 2 \int_0^{\infty} \mathbb{P}\left(\min_{u \in (t_0, t_1)} B_u \leq 0 \,|\, B_{t_0} = a \right) f_{B_{t_0}}(a)\, da \\
&= 2 \int_0^{\infty} da\, f_{B_{t_0}}(a) \int_0^{t_1 - t_0} du \frac{a}{\sqrt{2\pi}} u^{-3/2} \exp\left\{ \frac{-a^2}{2u} \right\} \\
&\overset{(1)}{=} \frac{1}{\pi \sqrt{t_0}} \int_0^{t_1 - t_0} du\, u^{-3/2} \int_0^{\infty} da\, a \exp\left\{ -\frac{a^2}{2} \left(\frac{1}{t_0} + \frac{1}{u} \right) \right\} \\
&\overset{(2)}{=} \frac{\sqrt{t_0}}{\pi} \int_0^{t_1 - t_0} \frac{du}{(t_0 + u)\sqrt{u}} \\
&\overset{(3)}{=} \frac{2}{\pi} \int_0^{\sqrt{(t_1 - t_0)/t_0}} \frac{dv}{1 + v^2} \\
&= \frac{2}{\pi} \arctan\left\{ \sqrt{\frac{t_1 - t_0}{t_0}} \right\}.
\end{aligned}$$

We switched the integrals in (1), used the explicit antiderivative of $a \exp\{-(a^2/2)(1/t_0 + 1/u)\}$ given by

$$-\exp\left\{ \frac{-a^2}{2} \left(\frac{1}{t_0} + \frac{1}{u} \right) \right\} \frac{t_0 u}{t_0 + u}$$

in (2) and did the change of variables $u = t_0\, v^2$ in (3). It follows that

$$\begin{aligned}
\tan^2\left(\frac{\pi}{2}\alpha\right) &= \frac{t_1}{t_0} - 1 \\
\Rightarrow \frac{t_1}{t_0} &= 1 + \tan^2\left(\frac{\pi}{2}\alpha\right) = \frac{1}{\cos^2((\alpha\pi)/2)} \\
\Rightarrow \sqrt{\frac{t_0}{t_1}} &= \cos\left(\frac{\pi\alpha}{2}\right).
\end{aligned}$$

From this, we get the desired result

$$\alpha = \mathbb{P}(\exists t \in (t_0, t_1)\, B_t = 0 \,|\, B_0 = 0) = \frac{2}{\pi} \arccos\left(\sqrt{\frac{t_0}{t_1}}\right).$$

□

Proposition 4.15. ***The arcsin law.*** *Define $L_1 := \sup\{t \leq 1 : B_t = 0\}$, the time of the last visit to 0 before time 1. Then*

$$\mathbb{P}(L_1 \leq s) = \frac{2}{\pi} \arcsin(\sqrt{s}), \quad s \in [0,1].$$

Remark 4.16. *Note that $\mathbb{P}(B_1 = 0) = 0$, so $\mathbb{P}(L_1 < 1) = 1$.*

Proof. First observe that

$$\{L_1 \leq s\} = \{B_t \neq 0 \text{ for all } t \in [s,1]\},$$

meaning

$$\begin{aligned}\mathbb{P}(L_1 \leq s) &= 1 - \mathbb{P}(\exists t \in [s,1] : B_t = 0 \,|\, B_0 = 0) \\ &= 1 - \frac{2}{\pi} \arccos(\sqrt{s}).\end{aligned}$$

Now recalling the following identity:

$$\arccos(x) + \arcsin(x) = \frac{\pi}{2},$$

we find the desired result

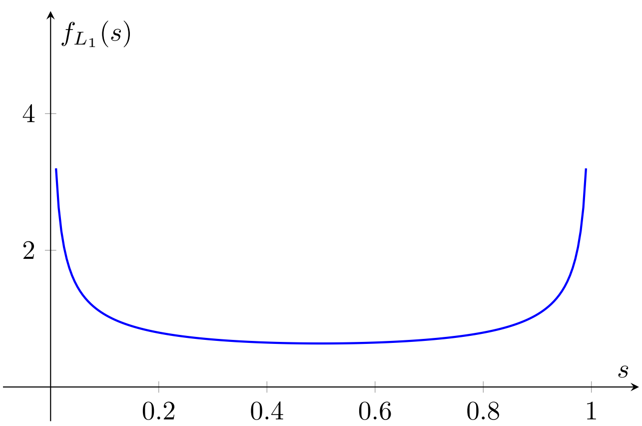
$$\mathbb{P}(L_1 \leq s) = 1 - \frac{2}{\pi}\left(\frac{\pi}{2} - \arcsin(\sqrt{s})\right) = \frac{2}{\pi}\arcsin(\sqrt{s}).$$

□

Remark 4.17. *The probability density function of L_1 is given by:*

$$f_{L_1}(s) = \frac{2}{\pi} \frac{1}{\sqrt{1-s}} \frac{1}{2\sqrt{s}} = \frac{1}{\pi} \frac{1}{\sqrt{s(1-s)}},$$

for $s \in (0,1)$. The density function of L_1 has the following plot:



So we would typically expect L_1 to be near 0 or 1 and that it would be less likely for it to be near $1/2$.

Remark 4.18. *We will use the following notation: $\mathbb{P}_a(\ldots) = \mathbb{P}(\ldots \,|\, B_0 = a)$.*

4.2 Three invariant properties of Brownian motion

Proposition 4.19. ***Scaling invariance.*** *Let $(B_t,\, t \in \mathbb{R}_+)$ be a standard BM. Fix $a > 0$ and define $\tilde{B}_t = \sqrt{a}^{-1} B_{at}$. Then $(\tilde{B}_t)$ is a standard BM.*

Proof. We check the two defining properties:

(a)

$$\begin{aligned}
\tilde{B}_{t+s} - \tilde{B}_s &= \frac{1}{\sqrt{a}} B_{a(t+s)} - \frac{1}{\sqrt{a}} B_{as} \\
&= \frac{1}{\sqrt{a}} (B_{as+at} - B_{as}) \\
&= \frac{1}{\sqrt{a}} \mathcal{N}(0, at) \\
&= \mathcal{N}(0, t).
\end{aligned}$$

(b) Let $t_1 < t_2 \le t_3 < t_4$. The increments

$$\tilde{B}_{t_4} - \tilde{B}_{t_3} = \frac{1}{\sqrt{a}} (B_{at_4} - B_{at_3})$$

and

$$\tilde{B}_{t_2} - \tilde{B}_{t_1} = \frac{1}{\sqrt{a}} (B_{at_2} - B_{at_1})$$

are independent as

$$t_1 < t_2 \le t_3 < t_4 \iff at_1 < at_2 \le at_3 < at_4.$$

□

Proposition 4.20. ***Time inversion.*** *Let (B_t) be a standard BM and define $\tilde{B}_t = tB_{1/t}$ if $t > 0$ and $\tilde{B}_0 = 0$. Then $(\tilde{B}_t)$ is a standard BM.*

Proof. We check the two defining properties:

(a)

$$\begin{aligned}
\tilde{B}_{t+s} - \tilde{B}_s &= (t+s) B_{1/(t+s)} - sB_{1/s} \\
&= tB_{1/(t+s)} - s(B_{1/s} - B_{1/(t+s)}) \\
&= t\mathcal{N}\left(0, \frac{1}{t+s}\right) - s\mathcal{N}\left(0, \frac{1}{s} - \frac{1}{t+s}\right) \\
&\overset{(*)}{=} \mathcal{N}\left(0, t^2 \frac{1}{t+s} + s^2 \left(\frac{1}{s} - \frac{1}{t+s}\right)\right) \\
&= \mathcal{N}(0, t),
\end{aligned}$$

where $(*)$ holds by independence.

(b) Let $t_1 < t_2 \leq t_3 < t_4$ and recall that for a standard BM:

$$\mathrm{Cov}(B_u, B_v) = \mathbb{E}[B_u B_v] = u \wedge v.$$

To check independence, it suffices to show that the covariance is zero:

$$\begin{aligned}
\mathbb{E}\left[\left(\tilde{B}_{t_4} - \tilde{B}_{t_3}\right)\left(\tilde{B}_{t_2} - \tilde{B}_{t_1}\right)\right] &= \mathbb{E}\left[\left(t_4 B_{1/t_4} - t_3 B_{1/t_3}\right)\left(t_2 B_{1/t_2} - t_1 B_{1/t_1}\right)\right] \\
&= t_4 t_2 \mathbb{E}[B_{1/t_4} B_{1/t_2}] - t_4 t_1 \mathbb{E}[B_{1/t_4} B_{1/t_1}] \\
&\quad - t_3 t_2 \mathbb{E}[B_{1/t_3} B_{1/t_2}] + t_3 t_1 \mathbb{E}[B_{1/t_3} B_{1/t_1}] \\
&= t_4 t_2 \cdot \frac{1}{t_4} - t_4 t_1 \cdot \frac{1}{t_4} - t_2 t_3 \cdot \frac{1}{t_3} + t_3 t_1 \cdot \frac{1}{t_3} \\
&= 0.
\end{aligned}$$

□

Remark 4.21. *The continuity of $t \mapsto \tilde{B}_t(\omega)$ at $t = 0$ is obvious. It is suggested by the fact that*

$$\mathbb{E}[(\tilde{B}_t)^2] = t^2 \mathbb{E}[B_{1/t}^2] = t^2 \cdot \frac{1}{t} = t \to 0$$

as $t \searrow 0$.

Proposition 4.22. ***Invariance under translation of time.*** *Fix $h > 0$, define $\tilde{B}_t = B_{t+h} - B_h$. Then $(\tilde{B}_t)$ is a standard BM.*

Proof. We check the two defining properties:

(a)

$$\begin{aligned}
\tilde{B}_{t+s} - \tilde{B}_s &= (B_{t+s+h} - B_h) - (B_{s+h} - B_h) \\
&= B_{t+s+h} - B_{s+h} \\
&\overset{\text{law}}{=} \mathcal{N}(0, t)
\end{aligned}$$

(b) Let $t_1 < t_2 \leq t_3 < t_4$. The increments

$$\tilde{B}_{t_4} - \tilde{B}_{t_3} = B_{t_4+h} - B_{t_3+h}$$

and

$$\tilde{B}_{t_2} - \tilde{B}_{t_1} = B_{t_2+h} - B_{t_1+h}$$

are independent because $t_1 + h < t_2 + h \leq t_3 + h < t_4 + h$ since $t_1 < t_2 \leq t_3 < t_4$.

□

4.3 Some transformations of Brownian motion

Definition 4.23. *A **reflected Brownian motion** is the process $(X_t,\, t \in \mathbb{R}_+)$ defined by $X_t = |B_t|$, where (B_t) is a standard BM.*

Proposition 4.24. *The reflected BM has the Markov property. In addition, for $0 \leq s < t$,*

$$f_{X_t}(y \,|\, X_s = x) = p(y - x,\, t - s) + p(y + x,\, t - s),$$

where $y, x \geq 0$ and where $p(u, r) = (\sqrt{2\pi r})^{-1} \exp\{-u^2/(2r)\}$.

Proof. Fix $0 \leq t_1 < t_2 < \ldots < t_n < t$ and $x_1, \ldots, x_n, x \geq 0$. Then

$$\begin{aligned}
\mathbb{P}(X_t \leq x \,|\, X_{t_n} = x_n, \ldots, X_{t_1} = x_1) &= \mathbb{P}(-x \leq B_t \leq x \,|\, B_{t_n} = \pm x_n, \ldots, B_{t_1} = \pm x_1) \\
&= \mathbb{P}(-x \leq B_t \leq x \,|\, (B_{t_n} = +x_n,\ \text{past}) \text{ or } (B_{t_n} = -x_n,\ \text{past})) \\
&\overset{(1)}{=} \mathbb{P}(-x \leq B_t \leq x \,|\, B_{t_n} = x_n,\ \text{past}) \\
&\overset{(2)}{=} \mathbb{P}(-x \leq B_t \leq x \,|\, B_{t_n} = x_n) \\
&= \int_{-x}^{x} p(y - x_n,\, t - t_n) dy \\
&\overset{(3)}{=} \mathbb{P}(X_t \leq x \,|\, X_{t_n} = x_n),
\end{aligned}$$

where (1) follows by symmetry, (2) holds since a standard BM has the Markov property and lastly, we can see in (3) that the "past" positions did not affect our calculation.

From this, we get that the conditional density of X_t given $X_{t_n} = x_n$ is:

$$\begin{aligned}
\frac{d}{dx}\mathbb{P}(X_t \leq x \,|\, X_{t_n} = x_n) &= p(x - x_n,\, t - t_n) - p(-x - x_n,\, t - t_n)(-1) \\
&= p(x - x_n,\, t - t_n) + p(x + x_n,\, t - t_n),
\end{aligned}$$

which is exactly what we wanted. □

Definition 4.25. ***Absorbed Brownian motion.*** *Let (B_t) be a BM with $B_0 = a \neq 0$. We define the BM absorbed at 0 to be the process $(Y_t,\, t \in \mathbb{R}_+)$ defined by*

$$Y_t = \begin{cases} B_t & \text{, if } t \leq T_0 = \inf\{t \geq 0 : B_t = 0\} \\ 0 & \text{, if } t > T_0 \end{cases}.$$

Proposition 4.26. *The BM absorbed at 0 has the Markov property. In addition, for $0 \leq s < t$,*

$$f_{Y_t}(y \,|\, Y_s) = p(y - x,\, t - s) - p(y + x,\, t - s),$$

with $y, x > 0$.

Proof. We prove the case $B_0 = a > 0$. Let $0 \le t_1 < \ldots < t_n < t$ and $y, x, x_i \ge 0$ for $i \in \{1, \ldots, n-1\}$. We will check that

$$\mathbb{P}(Y_t > y \,|\, Y_{t_n} = x, Y_{t_{n-1}} = x_{n-1}, \ldots, Y_{t_1} = x_1) = \mathbb{P}(Y_t > y \,|\, Y_{t_n} = x).$$

If $x = 0$, then both sides are equal to 0. If $x > 0$ and $x_i > 0$ for $i \in \{1, \ldots, n-1\}$, so that the path is possible, then we have

$$\begin{aligned}
\text{LHS} &= \mathbb{P}\left(Y_t > y \,|\, \min_{u\in[0,t_n]} B_u > 0, Y_{t_n} = x, \ldots, Y_{t_1} = x_1\right) \\
&= \mathbb{P}\left(B_t > y, \min_{u\in[t_n,t]} B_u > 0 \,|\, B_{t_n} = x, \underbrace{\min_{u\in[0,t_n]} B_u > 0, B_{t_{n-1}} = x_{n-1}, \ldots, B_{t_1} = x_1}_{\text{past}}\right) \\
&\overset{(1)}{=} \mathbb{P}\left(B_t > y, \min_{u\in[t_n,t]} B_u > 0 \,|\, B_{t_n} = x\right) \\
&\overset{(2)}{=} \mathbb{P}(Y_t > y \,|\, Y_{t_n} = x),
\end{aligned}$$

where (1) follows from the Markov property applied to the BM (B_t) and we can see in (2) that the values of Y_t at earlier times played no role.

Let us now compute $(\star) = \mathbb{P}\left(B_t > y, \min_{u\in[t_n,t]} B_u > 0 \,|\, B_{t_n} = x\right)$:

$$\begin{aligned}
(\star) &= \mathbb{P}(B_t < 2x - y, \max_{u\in[t_n,t]} B_u \le 2x \,|\, B_{t_n} = x) \\
&\overset{(1)}{=} \mathbb{P}(B_{t-t_n} < 2x - y, \max_{u\in[0,t-t_n]} B_u \le 2x \,|\, B_0 = x) \\
&\overset{(2)}{=} \mathbb{P}(B_{t-t_n} < x - y, \max_{u\in[0,t-t_n]} B_u < x \,|\, B_0 = 0) \\
&= \mathbb{P}_0(B_{t-t_n} \le x - y) - \mathbb{P}_0(B_{t-t_n} \le x - y, \max_{u\in[0,t-t_n]} B_u > x) \\
&\overset{(3)}{=} \mathbb{P}_0(B_{t-t_n} \le x - y) - \mathbb{P}_0(B_{t-t_n} > x + y) \\
&= \mathbb{P}_0(B_{t-t_n} < x - y) - \mathbb{P}_0(B_{t-t_n} < -x - y) \\
&= \mathbb{P}_0(-x - y < B_{t-t_n} < x - y) \\
&= \int_{-x-y}^{x-y} p(u,\, t - t_n) du,
\end{aligned}$$

where we used the translation of time property in (1) and of space in (2). The equality denoted by (3) follows from the reflection principle. Finally, we get the desired density function

$$f_{Y_t}(y \,|\, Y_{t_n} = x) = -\Big(p(x-y,\, t-t_n)(-1) - p(-x-y,\, t-t_n)(-1)\Big) = p(x-y,\, t-t_n) - p(x+y, t-t_n).$$

□

Definition 4.27. *A **Brownian motion with drift** is the process $(Z_t,\, t \in \mathbb{R}_+)$ defined by $Z_t = B_t + \mu t$, where (B_t) is a standard BM, and $\mu \ne 0$.*

Proposition 4.28. *For $0 \le t_1 < t_2 \le t_3 < t_4$, the increments*

$$Z_{t_4} - Z_{t_3} = B_{t_4} - B_{t_3} + \mu(t_4 - t_3) \quad ; \quad Z_{t_2} - Z_{t_1} = B_{t_2} - B_{t_1} + \mu(t_2 - t_1)$$

are independent. Therefore,

$$f_{Z_{t_2}-Z_{t_1},\, Z_{t_4}-Z_{t_3}}(x,\, y) = f_{B_{t_2}-B_{t_1}}\Big(x-\mu(t_2-t_1)\Big) f_{B_{t_4}-B_{t_3}}\Big(y-\mu(t_4-t_3)\Big).$$

Proof. Directly follows from Definition 4.27. □

Proposition 4.29. *A BM with drift has the Markov property.*

Proof. Fix $0 \le t_1 < \ldots < t_n < t$. Then,

$$\begin{aligned}
&\mathbb{P}_0(Z_t \le x \,|\, Z_{t_n} = x_n, \ldots, Z_{t_1} = x_1) \\
&= \mathbb{P}(Z_t - Z_{t_n} \le x - x_n \,|\, Z_{t_n} - Z_{t_{n-1}} = x_n - x_{n-1}, \ldots, Z_{t_2} - Z_{t_1} = x_2 - x_1, Z_{t_1} - Z_0 = x_1 - 0) \\
&\overset{(\star)}{=} \mathbb{P}(Z_t - Z_{t_n} \le x - x_n) \\
&= \mathbb{P}(Z_t - Z_{t_n} \le x - x_n \,|\, Z_{t_n} - Z_0 = x_n - 0) \\
&= \mathbb{P}(Z_t \le x \,|\, Z_{t_n} = x_n),
\end{aligned}$$

where $(\star)$ is a consequence of Proposition 4.28. □

Let us now study the time it takes for a standard BM to exit an interval. For $a \in \mathbb{R}$, we defined $\tau_a = \inf\{t \ge 0 : B_t = a\}$, the first time we hit the level a. We already saw that $\mathbb{P}(\tau_a < \infty) = 1$. If $B_0 \in [a, b]$, $a < b$ and $a, b \in \mathbb{R}$, then $T_{a,b} = \tau_a \wedge \tau_b$ is the first exit time of the interval $[a, b]$. Let $x \in (a, b)$ and fix $h > 0$, small, so that $[x-h, x+h] \subset (a, b)$. The BM starting at x must exit the interval $[x-h, x+h]$ before exiting $[a, b]$. By symmetry,

$$\mathbb{P}_x(B_{T_{x-h,x+h}} = x-h) = \frac{1}{2} = \mathbb{P}_x(B_{T_{x-h,x+h}} = x+h).$$

Therefore,

$$\begin{aligned}
\mathbb{P}_x(B_{T_{a,b}} = b) =&\mathbb{P}_x(B_{T_{a,b}} = b \,|\, B_{T_{x-h,x+h}} = x-h) \cdot \mathbb{P}_x(B_{T_{x-h,x+h}} = x-h) \\
&+ \mathbb{P}_x(B_{T_{a,b}} = b \,|\, B_{T_{x-h,x+h}} = x+h) \cdot \mathbb{P}_x(B_{T_{x-h,x+h}} = x+h).
\end{aligned}$$

Let $f(x) = \mathbb{P}_x(T_{a,b} = b)$, then

$$f(x) = \frac{1}{2} f(x-h) + \frac{1}{2} f(x+h).$$

This implies that f is affine, i.e., of the form $f(x) = \alpha x + \beta$, $\alpha, \beta \in \mathbb{R}$. If $f \in \mathcal{C}^2$, write

$$0 = \frac{1}{2}\Big(f(x+h) + 2f(x) + 2f(x-h)\Big)$$

and divide by h^2

$$0 = \frac{1}{2} \frac{f(x+h) + 2f(x) + f(x-h)}{h^2}.$$

Finally, letting $h \searrow 0$ gives

$$0 = \frac{1}{2} f''(x),$$

confirming that $f(x) = \alpha x + \beta$ for some $\alpha, \beta \in \mathbb{R}$. We have the following two constraints: $f(a) = 1$ and $f(b) = 1$. From this, we get

$$f(x) = \frac{x-a}{b-a}.$$

Proposition 4.30. *For $x \in [a,b]$, we have*

$$\mathbb{P}_x(B_{T_{a,b}} = b) = \frac{x-a}{b-a}.$$

Our next goal will be to determine $\mathbb{E}_x[T_{a,b}]$, $x \in [a,b]$, i.e., find the mean exit time of $[a,b]$. As defined above, we have $T_{a,b} = \tau_a \wedge \tau_b$, with $\mathbb{E}_x[\tau_a] = +\infty = \mathbb{E}_x[\tau_b]$. Fix $h > 0$. We aim to first find an expression for $\mathbb{E}_x[T_{x-h,x+h}] = \mathbb{E}_0[T_{-h,h}]$, where $T_{-h,h} = \inf\{u \geq 0 : |B_u| = h\}$. By the scaling invariance property, $(\tilde{B}_u = hB_{u/h^2}\, u \in \mathbb{R}_+)$ is also a standard BM. Therefore, $T_{-h,h}$ has the same law as $\tilde{T}_{-h,h} = \inf\{u \geq 0\,; |\tilde{B}_u| = h\}$. By noticing that

$$\begin{aligned}
\tilde{T}_{-h,h} &= \inf\{u \geq 0 \,:\, |\tilde{B}_u| = h\} \\
&= \inf\{u \geq 0 \,:\, |hB_{u/h^2}| = h\} \\
&= \inf\{u \geq 0 : |B_{u/h^2}| = 1\} \\
&= \inf\{vh^2 \geq 0 \,:\, |B_v| = 1\} \\
&= h^2 \inf\{v \geq 0 \,:\, |B_v| = 1\} \\
&= h^2 T_{-1,1},
\end{aligned}$$

we find $\mathbb{E}_0[T_{-h,h}] = \mathbb{E}_0[\tilde{T}_{-h,h}] = h^2 \mathbb{E}_0[T_{-1,1}] = h^2 \cdot c_0$, with $c_0 = \mathbb{E}_0[T_{-1,1}] > 0$. We will follow the same reasoning used to determine $\mathbb{P}_x(T_{a,b} = b)$ in order to derive an expression for $\mathbb{E}_x[T_{a,b}]$, leaving c_0 as the only remaining unknown.

Fix $h > 0$ small. A Brownian motion (BM) starting from x must first exit the interval $[x-h, x+h]$ before it can exit the larger interval $[a,b] \supset [x-h, x+h]$. This first exit takes $h^2 c_0$ units of time. Then, starting from either $x-h$ or $x+h$, it must exit $[a,b]$, so

$$\mathbb{E}_x[T_{a,b}] = \left(c_0 h^2 + \mathbb{E}_{x-h}[T_{a,b}]\right) \cdot \frac{1}{2} + \left(c_0 h^2 + \mathbb{E}_{x+h}[T_{a,b}]\right) \cdot \frac{1}{2}.$$

Let $g(x) = \mathbb{E}_x[T_{a,b}]$. Then,

$$g(x) = \left(c_0 h^2 + g(x-h)\right) \cdot \frac{1}{2} + \left(c_0 h^2 + g(x+h)\right) \cdot \frac{1}{2}.$$

Expanding and rearranging gives

$$0 = c_0 h^2 + \frac{1}{2} \left(g(x-h) - 2g(x) + g(x+h)\right).$$

Dividing by h^2 yields

$$0 = c_0 + \frac{1}{2} \cdot \frac{g(x-h) - 2g(x) + g(x+h)}{h^2},$$

and taking the limit as $h \searrow 0$ gives

$$0 = c_0 + \frac{1}{2}g''(x).$$

Hence, $g''(x) = -2c_0$, implying that g is a second-degree polynomial. Using the boundary conditions $g(a) = 0 = g(b)$, we find the explicit form

$$g(x) = c_0(b-x)(x-a).$$

Proposition 4.31. *For $x \in [a,b]$,*

$$\mathbb{E}_x[T_{a,b}] = (b-x)(x-a),$$

where we have used the fact that $c_0 = \mathbb{E}_0[T_{-1,1}] = 1$.

We now consider a slightly more computationally involved problem. Specifically, we aim to compute the probability that a BM with drift exits the interval $[a,b]$ through the point b before reaching a. Let $Z_t = B_t + \mu t$, where $\mu \neq 0$ and (B_t) is a standard BM, so that (Z_t) is a Brownian motion with drift μ. Define $T_{a,b} = \inf\{u \geq 0 : Z_u \in \{a,b\}\}$. Our goal is to compute $\mathbb{P}_x(Z_{T_{a,b}} = b)$, the probability that the process exits through b when started from $x \in [a,b]$.

Fix $h > 0$, small. During h^2 units of time, the BM moves about $\pm h$ units ($B_t \sim \mathcal{N}(0,t)$). The drift contributes for μh^2 units, and $\mu h^2 << h$ since we consider a small value of h. We must have

$$\mathbb{P}_x(Z_{T_{a,b}} = b) = \frac{1}{2}\mathbb{P}_{x+\mu h^2+h}(Z_{T_{a,b}} = b) + \frac{1}{2}\mathbb{P}_{x+\mu h^2-h}(Z_{T_{a,b}} = b).$$

Let $f(x) = \mathbb{P}_x(Z_{T_{a,b}} = b)$. We have

$$f(x) = \frac{1}{2}f(x+\mu h^2+h) + \frac{1}{2}f(x+\mu h^2-h)$$

and

$$\begin{aligned} 0 =& \frac{1}{2}\Big[f(x+\mu h^2+h) - f(x-\mu h^2+h)\Big] \\ &+ \frac{1}{2}\Big[\underbrace{f\big(x+(h-\mu h^2)\big) - 2f(x) + f\big(x-(h-\mu h^2)\big)}_{=c}\Big]. \end{aligned}$$

Using the Taylor expansions:

$$0 = \frac{1}{2}\left\{f'(x)\Big[(\mu h^2+h)-(h-\mu h^2)\Big] + \frac{1}{2}f''(x)\Big[(\mu h^2+h)^2-(h-\mu h^2)^2\Big]\right\} + \frac{1}{2}c$$

so

$$0 = \frac{1}{2}\left[f'(x)2\mu h^2 + o(h^2)\right] + \frac{1}{2}c.$$

Dividing by h^2 gives

$$0 = \mu f'(x) + \frac{1}{2}\frac{c}{(h-\mu h^2)} \cdot \frac{(h-\mu h^2)^2}{h^2} \cdot \frac{o(h^2)}{h^2},$$

and letting $h \searrow 0$ yields

$$0 = \mu f'(x) + \frac{1}{2}f''(x),$$

since

$$\frac{c}{(h-\mu h^2)} \to f''(x) \quad ; \quad \frac{(h-\mu h^2)^2}{h^2} \to 1 \quad ; \quad \frac{o(h^2)}{h^2} \to 0$$

as $h \searrow 0$.

Therefore,

$$f'(x) = c_1 \exp\{-2\mu x\},$$
$$f(x) = \tilde{c}_1 \exp\{-2\mu x\} + \tilde{c}_2.$$

Using the boundary conditions $f(a) = 0$, $f(b) = 1$, we get

$$f(x) = \frac{\exp\{-2\mu x\} - \exp\{-2\mu a\}}{\exp\{-2\mu b\} - \exp\{-2\mu a\}}.$$

Proposition 4.32. *For a BM with drift $\mu \neq 0$,*

$$\mathbb{P}_x(Z_{T_{a,b}} = b) = \frac{\exp\{-2\mu x\} - \exp\{-2\mu a\}}{\exp\{-2\mu b\} - \exp\{-2\mu a\}},$$

for $x \in [a, b]$.

Corollary 4.33. *Let $Z_t = B_t + \mu t$, $\mu < 0$ be a BM with drift and set $M := \sup_{t\in\mathbb{R}_+} Z_t$. For $m \geq 0$,*

$$\mathbb{P}_0(M \geq m) = \exp\{2\mu m\}.$$

Therefore, $\mathbb{P}_0(M = +\infty) = 0$ and

$$\mathbb{E}_0[M] = \int_0^\infty \mathbb{P}_0(M \geq m)dm = \frac{-1}{2\mu} > 0.$$

Proof. Set $\tau_a = \inf\{t \geq 0 : Z_t = a\}$. By Proposition 4.32,

$$\mathbb{P}_0(\tau_b < \tau_{-a}) = \frac{\exp\{-2\mu \cdot 0\} - \exp\{2\mu a\}}{\exp\{-2\mu b\} - \exp\{2\mu a\}} = \frac{1 - \exp\{2\mu a\}}{\exp\{-2\mu b\} - \exp\{2\mu a\}} \tag{$\star$}$$

for $-a < 0 < b$ with $a, b > 0$. Now suppose that

$$\mathbb{P}_0\left(\lim_{a\to+\infty} \tau_{-a} = +\infty\right) = 1. \tag{$\star\star$}$$

If we let $a \nearrow +\infty$ in ($\star$), we find

$$\mathbb{P}_0(M \geq b) = \mathbb{P}_0(\tau_b < +\infty) = \frac{1-0}{\exp\{-2\mu b\} - 0} = \exp\{2\mu b\}.$$

It remains to be shown that ($\star\star$) holds. First, notice that the map $a \mapsto \tau_{-a}$ is nondecreasing. Indeed, if $a_1 < a_2$, then (Z_t) must hit level $-a_1$ before it can reach level $-a_2$, so $\tau_{-a_1} < \tau_{-a_2}$. In particular,

$$\lim_{a\to+\infty} \tau_{-a} = \sup_{a\in\mathbb{R}_+} \tau_{-a}.$$

For $t > 0$,

$$\Big\{ \sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t \Big\} = \bigcap_{a \in \mathbb{N}} \{\tau_{-a} \leq t\} = \Big\{ \inf_{s \in [0,t]} Z_s = -\infty \Big\}.$$

From this and the fact that $s \mapsto Z_s(\omega)$ is continuous, we must have

$$\mathbb{P}_0 \left(\sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t \right) = \mathbb{P}_0 \left(\inf_{s \in [0,t]} Z_s = -\infty \right) = 0.$$

Moreover, the event $\{\sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t\}$ is increasing in t, i.e., if $t_1 \leq t_2$, then $\{\sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t_1\} \subseteq \{\sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t_2\}$. Therefore,

$$\mathbb{P}_0 \left(\sup_{a \in \mathbb{R}_+} \tau_{-a} < +\infty \right) = \mathbb{P}_0 \left(\bigcup_{t \geq 0\,;\, t \in \mathbb{N}} \Big\{ \sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t \Big\} \right) = \lim_{t \to \infty\,;\, t \in \mathbb{N}} \mathbb{P}_0 \left(\sup_{a \in \mathbb{R}_+} \tau_{-a} \leq t \right) = 0,$$

and $\mathbb{P}_0(\sup_{a \in \mathbb{R}_+} \tau_{-a} = +\infty) = 1$. □

Let us now do a quick overview of the asymptotic behavior of $t \mapsto Z_t$, with $\mu < 0$. Recall our definition of a BM with drift $Z_t = B_t + \mu t$. Since $\mathbb{E}[B_t^2] = t$, B_t has order of magnitude $\sqrt{t}$, i.e. $B_t \sim \sqrt{t}$ and for Z_t we have:

(1) For t small ($t \sim 0$), $\sqrt{t} >> t$, meaning Z_t behaves like B_t.

(2) For t large ($t \to +\infty$), $\sqrt{t} << t$, so Z_t behaves like μt.

4.4 Zero set of Brownian motion and its Hausdorff dimension

Definition 4.34. *Let $(B_t\,, t \in \mathbb{R}_+)$ be a standard BM. We call the set*

$$Z(\omega) := \{t \in \mathbb{R}_+ \,:\, B_t(\omega) = 0\}\,,$$

the zero set of the standard BM (B_t).

Proposition 4.35.

$$\mathbb{P}_0 \left(\bigcap_{n \in \mathbb{N}^*} \Big\{ \exists t \in \big(0, n^{-1}\big) \,:\, B_t = 0 \Big\} \right) = 1$$

Therefore, there exists a sequence $(t_n)_{n \in \mathbb{N}}$ such that $t_n > t_{n+1} > 0$, $\lim_{n \to +\infty} t_n = 0$ and $B_{t_n} = 0$ for all n.

Proof. Let $n \in \mathbb{N}^*$. Then, for all $m > n$:

$$\begin{aligned} \mathbb{P}_0 \left(\exists t \in (0, n^{-1}) \,:\, B_t = 0\right) &\geq \mathbb{P}_0 \left(\exists t \in (m^{-1}, n^{-1}) \,:\, B_t = 0\right) \\ &= \frac{2}{\pi} \arccos \left(\sqrt{\frac{n}{m}} \right). \end{aligned}$$

By letting $m \to +\infty$,

$$\mathbb{P}_0\left(\exists t \in (0, n^{-1}) : B_t = 0\right) \geq \frac{2}{\pi}\arccos(0) = 1.$$

Set $F_n = \{\exists t \in (0, n^{-1}) : B_t = 0\}$. Then, $F_n \supseteq F_{n+1}$ and therefore:

$$\mathbb{P}_0\left(\bigcap_{n\in\mathbb{N}^*}\left\{\exists t \in \left(0, n^{-1}\right) : B_t = 0\right\}\right) = \mathbb{P}_0\left(\bigcap_{n\in\mathbb{N}^*} F_n\right) = \lim_{n\to+\infty}\mathbb{P}_0(F_n) = 1,$$

since $\mathbb{P}_0(F_n) = 1$ for all $n \in \mathbb{N}^*$. □

Before stating and proving the next result, let us recall a simple and useful fact:

Lemma 4.36. *Let Y be a random variable. If $\mathbb{E}[Y] = 0$ and $\mathbb{P}(Y \geq 0) = 1$, then $\mathbb{P}(Y > 0) = 0$.*

Proof. For any $n \in \mathbb{N}^*$, define

$$E_n = \left\{Y > n^{-1}\right\}.$$

Since Y is nonnegative,

$$Y \geq Y \cdot \mathbb{1}_{E_n} \geq n^{-1}\mathbb{1}_{E_n},$$

by definition of E_n. Taking the expectation gives

$$0 = \mathbb{E}[Y] \geq n^{-1}\mathbb{P}(E_n) \geq 0,$$

and therefore $\mathbb{P}(E_n) = 0$. From this, we get the desired result

$$0 \leq \mathbb{P}(Y > 0) = \mathbb{P}\left(\cup_{n\in\mathbb{N}^*} E_n\right) = \lim_{n\to\infty}\mathbb{P}(E_n) = 0.$$

□

Remark 4.37. *When the BM starts at $x_0 \neq 0$, we can set $\tau_0 = \inf\{t \geq 0 : B_t = 0\}$ and we already know that $\mathbb{P}_{x_0}(\tau_0 < +\infty) = 1$ (see Remark 4.11). From this and what we have seen above, there will be an infinite number of visits to zero in the interval $[\tau_0, \tau_0 + \epsilon]$, for $\epsilon > 0$.*

Proposition 4.38.

$$\mathbb{P}_0\Big(\exists 0 \leq t_0 < t_1 : B_t = 0 \textit{ for all } t \in [t_0, t_1]\Big) = 0.$$

Proof. For $N \in \mathbb{N}$, define

$$F_N = \Big\{\exists 0 \leq t_0 < t_1 \leq N : B_t = 0 \text{ for all } t \in [t_0, t_1]\Big\}.$$

Then, since $F_N \subseteq F_{N+1}$ and $\cup_{N\in\mathbb{N}} F_N = \{\exists 0 \leq t_0 < t_1 : B_t = 0 \text{ for all } t \in [t_0, t_1]\}$, we have $\mathbb{P}_0(\cup_{N\in\mathbb{N}} F_N) = \lim_{N\to+\infty}\mathbb{P}_0(F_N)$ and it suffices to show $\mathbb{P}_0(F_N) = 0$ for all $N \in \mathbb{N}$ to get the desired result.

If F_N occurs, then

$$\int_0^N \mathbb{1}\{B_t = 0\}dt \geq \int_{t_0}^{t_1} 1 \cdot dt = t_1 - t_0 > 0.$$

Set $Y = \int_0^N \mathbb{1}\{B_t = 0\}dt \geq 0$. Thanks to Fubini's theorem, we find

$$\mathbb{E}_0[Y] = \mathbb{E}_0\left[\int_0^N \mathbb{1}\{B_t = 0\}dt\right] = \int_0^N \mathbb{E}_0[\mathbb{1}\{B_t = 0\}]dt = \int_0^N \underbrace{\mathbb{P}_0(B_t = 0)}_{=0} dt = 0.$$

Combining $\mathbb{E}_0[Y] = 0$ and $\mathbb{P}_0(Y \geq 0) = 1$, we conclude $\mathbb{P}_0(Y > 0) = 0$. Finally, since $\{Y > 0\} \supseteq F_N$, we get $\mathbb{P}_0(F_N) = 0$ for all $N \in \mathbb{N}$. □

Remark 4.39. *Fix $N \in \mathbb{N}^*$. Define*

$$Z_N(\omega) = \{t \in (0, N) : B_t(\omega) = 0\}.$$

Then, $Z_N(\omega)$ is a closed set in $(0, N)$, meaning $Z_N^c(\omega)$ is an open set in $(0, N)$ and

$$Z_N^c(\omega) = \bigcup_{i=1}^{\infty} O_i,$$

where O_i, $i \in \mathbb{N}$ are disjoint open intervals. Then we have $Z_N(\omega) \cup \mathbb{Z}_N^c(\omega) = (0, N)$ and

$$\mathbb{1}_{Z_N(\omega)}(t) + \mathbb{1}_{Z_N^c(\omega)}(t) = 1$$

for $t \in (0, N)$. Therefore,

$$\int_0^N \mathbb{1}_{Z_N(\omega)}(t)dt + \int_0^N \mathbb{1}_{Z_N^c(\omega)}(t)dt = N.$$

Since

$$\int_0^N \mathbb{1}_{Z_N(\omega)}(t)dt = 0,$$

we have that in some sense $Z_N(\omega)$ has length zero, and $Z_N^c(\omega)$ has length N.

Our main goal now will be to find the Hausdorff dimension of $Z(\omega)$, but before that, we have to go over some definitions and results.

Definition 4.40. *Fix $E \subset \mathbb{R}$.*

(a) A family $(U_i,\, i \in \mathbb{N})$ of intervals is a cover of E if $E \subset \cup_{i \in \mathbb{N}} U_i$.

(b) Let $L(U_i)$ denote the length of U_i, where $L([a, b]) = b - a$ for $a \leq b$. Fix $\delta > 0$, (U_i) is a δ-cover of E if it is a cover of E and $L(U_i) \leq \delta$ for all i.

(c) Fix $\alpha > 0$. Define

$$\mu_\alpha(E) = \lim_{\delta \searrow 0} \inf_{(U_i) \text{ a } \delta\text{-cover of } E} \sum_{i=1}^{\infty} \Big(L(U_i)\Big)^\alpha.$$

Note that $0 \leq \mu_\alpha(E) \leq +\infty$.

Lemma 4.41. *(a) Fix $\alpha > 0$ and suppose that $\mu_\alpha(E) < +\infty$. Then for $\beta > \alpha$, $\mu_\beta(E) = 0$.*

(b) Fix $\alpha > 0$ and suppose that $\mu_\alpha(E) > 0$. Then for $\beta < \alpha$, $\mu_\beta(E) = +\infty$.

Proof. (a) Fix $\beta > \alpha$. Then,

$$\begin{aligned} 0 \le \mu_\beta(E) &= \lim_{\delta \searrow 0} \inf_{C_\delta(E)} \sum_{i=1}^{\infty} \Big(L(U_i)\Big)^\beta \\ &= \lim_{\delta \searrow 0} \inf_{C_\delta(E)} \sum_{i=1}^{\infty} \Big(L(U_i)\Big)^{\beta-\alpha} \cdot \Big(L(U_i)\Big)^\alpha \\ &\le \lim_{\delta \searrow 0} \delta^{\beta-\alpha} \inf_{C_\delta(E)} \sum_{i=1}^{\infty} \Big(L(U_i)\Big)^\alpha, \end{aligned}$$

where we write $C_\delta(E)$ to denote the fact that we take the infimum over the δ-covers (U_i) of E. Fix $\delta_0 > 0$, we now have

$$\mu_\beta(E) \le \delta_0^{\beta-\alpha} \lim_{\delta \searrow 0} \inf_{C_\delta(E)} \sum_{i=1}^{\infty} \Big(L(U_i)\Big)^\alpha = \delta_0^{\beta-\alpha} \mu_\alpha(E).$$

Since $\beta > \alpha$, $\mu_\alpha(E) < +\infty$ and the above holds for all $\delta_0 > 0$, we have that $\mu_\beta(E) = 0$.

(b) Suppose $\mu_\alpha(E) > 0$. Fix $\beta < \alpha$. Suppose by contradiction that $\mu_\beta(E) < +\infty$. By (a), we would have $\mu_\alpha(E) = 0$, which gives a contradiction. Hence, $\mu_\beta(E) = +\infty$.

□

From Lemma 4.41, we conclude that there exists α_0 such that if $\beta > \alpha_0$, then $\mu_\beta(E) = 0$ and if $\beta < \alpha_0$, then $\mu_\beta(E) = +\infty$. The value $\mu_{\alpha_0}(E)$ can be 0, $+\infty$ or in $(0, +\infty)$.

Definition 4.42. *The value α_0 mentioned above is called the Hausdorff dimension of E.*

The Hausdorff dimension of a set E in $\mathbb{R}^d$ is defined similarly:

Definition 4.43. *(a) For $E \subset \mathbb{R}^d$, $(U_i)_{i\in\mathbb{N}}$ is a δ-cover of E if $E \subset \cup_i U_i$, and each U_i is a ball with radius $\epsilon_i > 0$ and diam$(U_i) = 2\epsilon_i \le \delta$.*

(b) For $\alpha > 0$,

$$\mu_\alpha(E) = \lim_{\delta \searrow 0} \inf_{C_\delta(E)} \sum_{i=1}^{\infty} \Big(diam(U_i)\Big)^\alpha.$$

(c) $\dim_{\mathcal{H}}(E) = \inf\{\alpha > 0 : \mu_\alpha(E) = 0\} = \sup\{a > 0 : \mu_\alpha(E) = +\infty\}$

Remark 4.44. *To show that $\dim_{\mathcal{H}}(E) \le \alpha_0$, it suffices, for each $\alpha > \alpha_0$ and $n \in \mathbb{N}^*$, to find a $(1/n)$-cover (U_i^n) of E such that*

$$\liminf_{n\to+\infty} \sum_i \Big(diam(U_i^n)\Big)^\alpha = 0.$$

Indeed,

$$\mu_\alpha(E) = \lim_{\delta \searrow 0} \inf_{C_\delta(E)} \sum_i \Big(diam(U_i)\Big)^\alpha$$
$$\leq \lim_{n\to+\infty} \inf_{C_{1/n}(E)} \sum_i \Big(diam(U_i)\Big)^\alpha$$
$$\leq \liminf_{n\to+\infty} \sum_i \Big(diam(U_i^n)\Big)^\alpha = 0.$$

Example 4.45. *(a) Let $d = 2$ and $E = [0,1] \times \{0\}$. Define the sets*

$$U_i^n := \left[\frac{i}{n}, \frac{i+1}{n}\right] \times \left[-\frac{1}{2n}, \frac{1}{2n}\right], \quad i = 0, \dots, n-1.$$

The sets U_i^n are balls of diameter n^{-1} for the norm $||(x_1, x_2)|| = \max(|x_1|, |x_2|)$. In order to cover E, we need n balls with diameter n^{-1}. Since

$$\sum_{i=1}^n \Big(diam(U_i^n)\Big)^\alpha = n\left(\frac{1}{n}\right)^\alpha = n^{1-\alpha} \overset{n\to+\infty}{\longrightarrow} \begin{cases} +\infty & ,\alpha < 1 \\ 0 & ,\alpha > 1 \end{cases}$$

suggesting $dim_{\mathcal{H}}(E) = 1$. In fact, it proves that $dim_{\mathcal{H}}(E) \leq 1$.

(b) Let $d = 2$ and $E = [0,1]^2$. Define

$$U_{i,j}^n = \left[\frac{i}{n}, \frac{i+1}{n}\right] \times \left[\frac{j}{n}, \frac{j+1}{n}\right], \quad i, j \in \{0, \dots, n-1\}.$$

This time, we need n^2 balls with diameter n^{-1} in order to cover E:

$$\sum_{i,j=1}^n \Big(diam(U_{i,j}^n)\Big)^\alpha = n^2\left(\frac{1}{n}\right)^\alpha = n^{2-\alpha} \overset{n\to+\infty}{\longrightarrow} \begin{cases} +\infty & ,\alpha < 2 \\ 0 & ,\alpha > 2 \end{cases},$$

which suggests $dim_{\mathcal{H}}(E) = 2$ and ensures $dim_{\mathcal{H}}(E) \leq 2$.

(c) Let $\mathcal{C}$ be the Cantor set. The Cantor set is defined inductively as follows. As a base case, we let

$$C_0 := [0,1];$$
$$C_1 := [0,1/3] \cup [2/3,1];$$
$$C_2 := [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1].$$

For $n \in \mathbb{N}$, the set C_n is the union of 2^n disjoint closed interval of length 3^{-n}. To get C_{n+1} from C_n, we remove the open middle third from each of the intervals in C_n. The Cantor set is defined as

$$\mathcal{C} := \bigcap_{n=0}^\infty C_n.$$

At stage n, there are 2^n intervals with length 3^{-n} needed to cover the Cantor set,

$$\sum_{i=1}^{2^n} \Big(diam(U_i^n)\Big)^\alpha = 2^n\left(\frac{1}{3}\right)^{\alpha\cdot n} = \exp\{n[\log(2) - \log(3)]\} \overset{n\to+\infty}{\longrightarrow} \begin{cases} +\infty & ,\alpha < [\log(2)/\log(3)] \\ 0 & ,\alpha > [\log(2)/\log(3)] \end{cases},$$

which suggests $dim_{\mathcal{H}}(\mathcal{C}) = \log(2)/\log(3)$ and ensures $dim_{\mathcal{H}}(\mathcal{C}) \leq \log(2)/\log(3)$.

Theorem 4.46. *Let $(B_t)_{t\geq 0}$ be a standard BM. Then,*

$$\mathbb{P}_0\left(\omega\,:\dim_{\mathcal{H}}\{t\in\mathbb{R}_+\,:B_t(\omega)=0\}=\frac{1}{2}\right)=1.$$

We will not prove this theorem in full; instead, we will prove the following result:

Proposition 4.47.

$$\mathbb{P}_0\left(\omega\,:\dim_{\mathcal{H}}\{t\in\mathbb{R}_+\,:B_t(\omega)=0\}\leq\frac{1}{2}\right)=1.$$

To do so, we will need the two following lemmas.

Lemma 4.48. *There exists $0<c_0<c_1<+\infty$ and $s_0>0$ such that for all $t\in[1,2]$ and $0\leq s\leq s_0$,*

$$c_0\sqrt{s}\leq\arccos\left(\sqrt{\frac{t}{t+s}}\right)\leq c_1\sqrt{s}.$$

Proof. Setting $y=(t+s)/t$, we get

$$\frac{\arccos\left(\sqrt{\frac{t}{t+s}}\right)}{\sqrt{s}}=\frac{\arccos\left(\frac{1}{\sqrt{y}}\right)}{\sqrt{t}\sqrt{y-1}}$$

and it suffices to show that

$$\lim_{y\searrow 1}\frac{\arccos\left(\frac{1}{\sqrt{y}}\right)}{\sqrt{y-1}}=1.$$

Applying B-H, we get

$$\lim_{y\searrow 1}\frac{\frac{-1}{\sqrt{1-1/y}}\cdot\frac{-1}{2}y^{-3/2}}{\frac{1}{2\sqrt{y-1}}}=\lim_{y\searrow 1}\sqrt{\frac{y}{y-1}}\cdot\frac{\sqrt{y-1}}{y^{3/2}}=1.$$

□

Lemma 4.49. *For $I=[t_0,t_1]\subset[1,2]$, $t_0<t_1$ and t_1-t_0 small enough, we have*

$$\mathbb{P}_0(\exists t\in I\,:B_t=0)\leq\frac{2}{\pi}c_1\sqrt{L(I)}.$$

Proof. Using the previous lemma,

$$\mathbb{P}_0(\exists t\in[t_0,t_1]\,:B_t=0)=\frac{2}{\pi}\arccos\left(\sqrt{\frac{t_0}{t_0+(t_1-t_0)}}\right)\leq\frac{2}{\pi}c_1\sqrt{t_1-t_0},$$

provided $t_1-t_0\leq s_0$. □

Proof. of Proposition 4.47. Let $Z(\omega) = \{t \in [1,2] : B_t(\omega) = 0\}$. Define

$$U_i^n(\omega) = \begin{cases} \left[1+\frac{i}{n}, 1+\frac{i+1}{n}\right] & , \text{ if } Z(\omega) \cap \left[1+\frac{i}{n}, 1+\frac{i+1}{n}\right] \neq \emptyset \\ \emptyset & , \text{ otherwise} \end{cases}.$$

Clearly, $Z(\omega) \subset \cup_{i=0}^{n-1} U_i^n$, meaning $(U_i^n,\, i = 0, \ldots, n-1)$ is a n^{-1}-cover of $Z(\omega)$. Fix $\alpha > 2^{-1}$. We will show that $\mathbb{P}(\mu_\alpha(Z) = 0) = 1$. Since $\mu_\alpha(Z) \geq 0$, we have

$$\begin{aligned}
0 &\leq \mathbb{E}_0[\mu_\alpha(Z)] \\
&\leq \mathbb{E}_0\left[\liminf_{n\to\infty} \sum_{i=0}^{n-1} \Big(L(U_i^n)\Big)^\alpha\right] \\
&\overset{(1)}{\leq} \liminf_{n\to\infty} \mathbb{E}_0\left[\sum_{i=0}^{n-1} \Big(L(U_i^n)\Big)^\alpha\right] \\
&= \liminf_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}_0\left[\Big(L(U_i^n)\Big)^\alpha\right] \\
&= \liminf_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}_0\left[\left(\frac{1}{n}\right)^\alpha \mathbb{1}\{U_i^n \neq \emptyset\} + 0^\alpha \mathbb{1}\{U_i^n = \emptyset\}\right] \\
&= \liminf_{n\to\infty} \sum_{i=0}^{n-1} \left(\frac{1}{n}\right)^\alpha \mathbb{P}_0\left(\exists t \in \left[1+\frac{i}{n}, 1+\frac{i+1}{n}\right] : B_t = 0\right) \\
&\overset{(2)}{\leq} \liminf_{n\to\infty} \sum_{i=0}^{n-1} \left(\frac{1}{n}\right)^\alpha \cdot c\sqrt{\frac{1}{n}} \\
&= \liminf_{n\to\infty} n^{1-\alpha-(1/2)} \\
&= \liminf_{n\to\infty} n^{1/2-\alpha} = 0
\end{aligned}$$

as $\alpha > 1/2$. We used Fatou's lemma (Lemma 2.55) in (1) and Lemma 4.49 in (2).

Finally, combining $\mathbb{E}_0[\mu_\alpha(Z)] = 0$ and $\mathbb{P}(\mu_\alpha(Z) \geq 0) = 1$, we get $\mathbb{P}_0(\mu_\alpha(Z) = 0) = 1$, which in turn implies $\mathbb{P}_0(\dim_{\mathcal{H}}(Z) \leq 1/2) = 1$. The only step left to do is extend this result to $\mathbb{R}_+$.

Now, let $Z(\omega) = \{t \in \mathbb{R}_+ : B_t(\omega) = 0\}$. We have just proved that

$$\mathbb{P}_0\Big(\mu_\alpha(Z \cap [1,2]) = 0\Big) = 1$$

for $\alpha > 1/2$. Similarly,

$$\mathbb{P}_0\Big(\mu_\alpha(Z \cap [n, n+1]) = 0\Big) = 1 \quad \text{and} \quad \mathbb{P}_0\left(\mu_\alpha\left(Z \cap \left[\frac{1}{n+1}, \frac{1}{n}\right]\right) = 0\right) = 1$$

for all $n \geq 1$. Therefore,

$$\begin{aligned}
\mu_\alpha(Z) &= \mu_\alpha\left(\{0\} \cup \left[\bigcup_{n=1}^\infty \left\{Z \cap \left(\frac{1}{n+1}, \frac{1}{n}\right]\right\}\right] \cup \left[\bigcup_{n=1}^\infty \{Z \cap [n, n+1)\}\right]\right) \\
&\overset{(\star)}{=} \mu_\alpha(\{0\}) + \sum_{i=1}^\infty \mu_\alpha\left(Z \cap \left(\frac{1}{n+1}, \frac{1}{n}\right]\right) + \sum_{i=1}^\infty \mu_\alpha\left(Z \cap [n, n+1)\right) \\
&= 0,
\end{aligned}$$

with probability one. The equality denoted by $(\star)$ follows from the fact that μ_α is a measure and all sets involved are disjoint. $\square$

4.5 Continuity of sample paths

The objective of this subsection is to construct a random continuous function $t \mapsto B_t(\omega)$ satisfying properties (a) and (b) of Definition 4.1. To achieve this, we will follow Paul Lévy's construction. The idea is as follows:

Let (X_n) be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables, and define $S_0 = 0$, $S_n = X_1 + \ldots + X_n$ for $n \geq 1$. The process (S_n) is a Gaussian random walk on $\mathbb{R}$. We view (S_n) as the restriction to $\mathbb{N}$ of a BM (B_t) on $\mathbb{R}_+$, such that $B_n = S_n$ for all $n \in \mathbb{N}$. Thus, we have defined the BM at integer times, but it remains to define B_t for $t \in \mathbb{R}_+ \setminus \mathbb{N}$.

<u>Step 1: Linear interpolation.</u>

Define

$$B_t^{(1)} = \begin{cases} B_t & \text{, if } t \in \mathbb{N} \\ B_n + (t-n)(B_{n+1} - B_n) & \text{, if } t \in (n, n+1), \quad n \in \mathbb{N} \end{cases}$$

and observe that we have the equality

$$B_{n+1/2}^{(1)} = \frac{B_n^{(1)} + B_{n+1}^{(1)}}{2} = \mathbb{E}[B_{n+1/2} \,|\, B_n, B_{n+1}]$$

as, given B_n, B_{n+1}, the law of $B_{n+1/2}$ is $\mathcal{N}\{(B_n + B_{n+1})/2, 1/4\}$. See Exercise 3 of Problem Set 8.

<u>Step 2: Refine the time step.</u>

Set

$$\begin{aligned} B_n^{(2)} &= B_n^{(1)} = S_n & &, n \in \mathbb{N}; \\ B_{n+(1/2)}^{(2)} &= \frac{1}{2}\left(B_n^{(1)} + B_{n+1}^{(1)}\right) + \frac{1}{2} Z_n^{(1)} & &, n \in \mathbb{N}, \end{aligned}$$

where $\left(Z_n^{(1)}, n \in \mathbb{N}\right)$ are i.i.d. $\mathcal{N}(0,1)$, and independent of (X_n).

We then repeat these two steps. It remains to verify that the process defined in the limit is indeed continuous and satisfies properties (a) and (b) of BM.

While this might seem confusing at first, it is the core principle behind the rigorous construction of Brownian motion, which we now begin. Hopefully, the formal mathematical framework will clarify the somewhat messy ideas introduced above. We restrict ourselves to the construction of the BM on the interval $[0,1]$, the generalization to $\mathbb{R}_+$ follows naturally and is given as an exercise, see Exercise 1 of Problem Set 13.

For $n \in \mathbb{N}$, let

$$D_n = \left\{ \frac{k}{2^n}, \, k = 0, 1, \ldots, 2^n \right\},$$

the dyadics of order n in $[0,1]$. The sets $\{D_n\}_{n \in \mathbb{N}}$ are increasing in $n \in \mathbb{N}$ and $D = \cup_{n \in \mathbb{N}} D_n$ is the set of dyadics in $[0,1]$. Moreover, the set D is countable and dense in $[0,1]$.

Let $(Z_d, d \in D)$ be i.i.d. $\mathcal{N}(0,1)$ random variables. We begin by defining B_t for $t \in D$. We start with the base case $n = 0$: Set $B_0 = 0$ and $B_1 = Z_1$. This defines B_t for $t \in D_0$.

For $n \geq 1$, we will define B_t for $t \in D_n$ inductively, so that

(i) for $r < s < t$ in D_n, $B_t - B_s \sim \mathcal{N}(0, t-s)$ and is independent of $B_s - B_r$;

(ii) for $t \in D_n$, B_t is determined by the Z_e with $e \in D_n$ and is therefore independent of Z_f, where $f \in D \setminus D_n$.

For the base case $n = 0$, properties (i) and (ii) hold. For $n \geq 1$, suppose, by induction, that $(B_t,\, t \in D_{n-1})$ has been defined in such a way that (i) and (ii) hold for $n-1$. The goal is now to define B_t, for $t \in D_n \setminus D_{n-1}$ using $(B_t,\, t \in D_{n-1})$ so that (i) and (ii) are satisfied.

For $d \in D_n \setminus D_{n-1}$, let

$$d_n^- = d - 2^{-n}\,; \quad d_n^+ = d + 2^{-n},$$

so that $d_n^-, d_n^+ \in D_{n-1}$. Now, set

$$B_d = \frac{1}{2}\Big(B_{d_n^-} + B_{d_n^+}\Big) + \frac{1}{2^{(n+1)/2}} Z_d, \tag{$\star$1}$$

and we define, for $d \in D_n$,

$$B_d = \begin{cases} B_d & \text{, if } d \in D_{n-1} \\ \frac{1}{2}\Big(B_{d_n^-} + B_{d_n^+}\Big) + \frac{1}{2^{(n+1)/2}} Z_d & \text{, if } d \in D_n \setminus D_{n-1} \end{cases}.$$

Now that we have defined $(B_t,\, t \in D_n)$, it remains to check that (i) and (ii) are satisfied. Observe that

$$X_1 = \frac{1}{2}\Big(B_{d_n^+} - B_{d_n^-}\Big) \sim \mathcal{N}\left(0, \frac{1}{2^{n+1}}\right)$$

by induction and (i) for the case $(n-1)$, and

$$X_2 = \frac{1}{2^{(n+1)/2}} Z_d \sim \mathcal{N}\left(\frac{1}{2^{n+1}}\right).$$

Property (ii) for $(n-1)$, ensures that X_1 and X_2 are independent. Therefore, their sum and difference are independent and $\mathcal{N}(0, 2^{-n})$. This observation follows from the fact that $X_1, X_2 \sim \mathcal{N}(0, 2^{-(n+1)})$ combined with

$$\begin{aligned} \mathrm{Cov}(X_1 + X_2,\, X_1 - X_2) &= \mathbb{E}[(X_1 + X_2)(X_1 - X_2)] \\ &= \mathbb{E}[X_1^2] - \mathbb{E}[X_2^2] \\ &= 0. \end{aligned}$$

Note that

$$X_1 + X_2 = B_d - B_{d_n^-} \quad ; \quad X_1 - X_2 = B_{d_n^+} - B_d. \tag{$\star$2}$$

Consequently, for $d, f \in D_n \setminus D_{n-1}$, $d < f$, the increments

$$B_{d_n^+} - B_d, \quad B_d - B_{d_n^-} \quad ; \quad B_{f_n^+} - B_f, \quad B_f - B_{f_n^-}$$

are independent. Indeed, we have just seen that the first two are independent. Same holds for the second two. The remaining independencies are easy to check. Take the increments

$$B_{d_n^+} - B_d \quad \text{and} \quad B_f - B_{d_n^-}$$

for example. Using ($\star$2), we see that the first is determined by $(B_{d_n^+} - B_{d_n^-})$ and Z_d, while the second is determined by $(B_{f_n^+} - B_{f_n^-})$ and Z_f. Their independence then follows from (i), (ii) for $(n-1)$ and the equivalence between "pairwise independence" and "mutual independence" for Gaussians.

The family $(B_t,\, t \in D)$ we just defined satisfies the properties (i) and (ii). Property (i) holds as an increment over a long interval can be expressed as the sum of increments over intervals of lengths 2^{-n}. Property (ii) holds by construction, in particular by ($\star$1). We see immediately that $(B_t,\, t \in D)$ has the properties (a) and (b) of a BM (Definition 4.1).

The final step in our construction is to extend the family $(B_t,\, t \in D)$ to all of $[0, 1]$. To achieve this, we will introduce a new approach that may initially appear unrelated to what we have done until now. However, the connection will soon become apparent, and we will see how this new perspective simplifies the process. In particular, it will clarify both the extension from D to $[0, 1]$ and the proof of the continuity of sample paths.

Define $F_0(t) = tZ_1$, $t \in [0, 1]$ for $n = 0$. For $n \geq 1$, let

$$F_n(t) = \begin{cases} 2^{-(n+1)/2} Z_t & \text{, if } t \in D_n \setminus D_{n-1} \\ 0 & \text{, if } t \in D_{n-1} \\ \text{linear (affine)} & \text{, between consecutive elements of } D_n \end{cases}$$

Each F_n is a continuous function on $[0, 1]$. Define $B_t^{(n)} = \sum_{i=0}^{n} F_i(t)$. This function is continuous and affine between consecutive elements of D_n. From the definition of the F_n's, we see that $B_t^{(n)} = \sum_{i=0}^{n} F_i(t) = \sum_{i=0}^{\infty} F_i(t)$ for $t \in D_n$.

Definition 4.50. *For $t \in D = \cup_{n=0}^{\infty} D_n$, let*

$$\tilde{B}_t = \sum_{i=0}^{\infty} F_i(t).$$

As mentioned above, $\tilde{B}_t = B_t^{(n)}$ for $t \in D_n$.

The next proposition will make the link between Definition 4.50 and the construction used in ($\star$1).

Proposition 4.51. *For $t = d \in D_n$, $\tilde{B}_t = B_d$, where B_d denotes the construction in ($\star$1).*

Proof. For $n = 0$, we only need to check for $t \in \{0, 1\}$. If $t = 0$, then $\tilde{B}_0 = 0 = B_0$. If $t = 1$, then $\tilde{B}_1 = F_0(1) = Z_1 = B_1$. Now suppose that the result holds for $(n-1)$. Then, for $t \in D_n \cap D_{n-1}$, we have $\tilde{B}_t = B_t$ by induction. In case $t = d \in D_n \setminus D_{n-1}$, we start by noticing that for $i \in \{0, 1, \ldots, n-1\}$, F_i is affine on $[d_n^-, d_n^+]$. Therefore,

$$\begin{aligned} \sum_{i=0}^{n-1} F_i(d) &= \sum_{i=0}^{n-1} \frac{1}{2}\Big(F_i(d_n^-) + F_i(d_n^+)\Big) \\ &= \frac{1}{2}\left[\sum_{i=0}^{n-1} F_i(d_n^-) + \sum_{i=0}^{n-1} F_i(d_n^+)\right] \\ &= \frac{1}{2}\Big(B_{d_n^-} + B_{d_n^+}\Big) \end{aligned}$$

where the last inequality follows by induction. Finally, using the definition of F_n, we get the desired result

$$\tilde{B}_t = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{n-1} F_i(d) + F_n(d) = \frac{1}{2}\Big(B_{d_n^-} + B_{d_n^+}\Big) + 2^{-(n+1)/2} Z_d = B_d.$$

□

We now have all the necessary results to establish the continuity of the sample paths.

Proposition 4.52. *The series of functions* $\sum_{n=0}^{\infty} F_n$ *converges uniformly on* $[0,1]$*, with probability one. Therefore,*

$$t \mapsto \tilde{B}_t = \sum_{n=0}^{\infty} F_n(t)$$

is a continuous function with probability one.

Proof. First notice that

$$\sup_{t\in[0,1]} |F_n(t)| \leq 2^{-(n+1)/2} \sup_{t\in D_n \setminus D-n-1} |Z_t|.$$

Let $c > 0$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in D_n\setminus D_{n-1}} |Z_t| \geq c\sqrt{n}\right) &\leq \mathbb{P}\left(\sup_{t\in D_n} |Z_t| \geq c\sqrt{n}\right) \\
&\leq \sum_{t\in D_n} \mathbb{P}\Big(|Z_t| \geq c\sqrt{n}\Big) \\
&\leq (2^n + 1) \exp\left\{-\frac{c^2 n}{2}\right\} \\
&\leq 2 \cdot 2^n \exp\left\{-\frac{c^2 n}{2}\right\} \\
&= 2 \exp\left\{n\left(\log(2) - \frac{c^2}{2}\right)\right\}
\end{aligned}$$

using the fact that for $Z \sim \mathcal{N}(0,1)$ and $x \geq 0$,

$$\mathbb{P}(|Z| \geq x) \leq \exp\left\{-\frac{x^2}{2}\right\}.$$

If $\log(2) - (c^2/2) < 0$, i.e., $c > \sqrt{2\log(2)}$, then

$$\sum_{n=0}^{\infty} \mathbb{P}\left(\sup_{t\in D_n\setminus D_{n-1}} |Z_t| \geq c\sqrt{n}\right) < \infty.$$

By the Borel-Cantelli lemma, there exists $N(\omega) \in \mathbb{N}$ such that for $n \geq N(\omega)$,

$$\sup_{t\in D_n\setminus D_{n-1}} |Z_t| < c\sqrt{n}$$

and

$$\sup_{t\in[0,1]} |F_n(t)| \le c\sqrt{n}2^{-(n+1)/2}.$$

Meaning

$$\sum_{n=0}^{\infty} \sup_{t\in[0,1]} |F_n(t)| \le c \sum_{n=0}^{\infty} \sqrt{n}2^{-(n+1)/2} < \infty.$$

Therefore, $\sum_{n=0}^{\infty} F_n$ converges normally, which implies uniform convergence on $[0,1]$. □

We have just shown that

$$t \mapsto \tilde{B}_t = \sum_{n=0}^{\infty} F_n(t), \quad t \in [0,1]$$

is continuous with probability one. Moreover, since $\tilde{B}_t = B_t$ for all $t \in D$, the process $(\tilde{B}_t,\, t \in D)$ satisfies properties (a) and (b) of Brownian motion. As D is dense in $[0,1]$ and the mapping $t \mapsto \tilde{B}_t$ is continuous, it follows that the process $(\tilde{B}_t,\, t \in [0,1])$ also satisfies properties (a) and (b). This completes the construction of the continuous stochastic process $t \mapsto \tilde{B}_t$ we set out to define.

4.6 The Strong Markov Property

This subsection presents a result that was not covered in the course. However, we state it without proof, as it was mentioned earlier and will be needed for the final exercise of Problem Set 13.

Theorem 4.53. ***Strong Markov Property.*** *For every almost surely finite stopping time T, the process $(B_{T+t} - B_T,\, t \in \mathbb{R}_+)$ is a standard Brownian motion independent of $(B_t,\, t \le T)$.*

We have not formally defined what a stopping time for Brownian motion is, as doing so would take us beyond the scope of this course. Intuitively, one can think of a stopping time T as a random variable taking values in $[0,\infty]$ such that, for each $t \in \mathbb{R}_+$, the event $\{T \le t\}$ is determined by $(B_s,\, s \le t)$. For readers seeking a precise definition, we refer to Chapter 8 of *Probability: Theory and Examples* by R. Durrett [2].

References

[1] Robert C Dalang and Daniel Conus. *Introduction à la théorie des probabilités*. EPFL Press, 2018.

[2] Rick Durrett. *Probability: Theory and Examples*. Cambridge university press, 2010.

[3] Alfred J Lotka. The extinction of families—i. *Journal of the Washington Academy of Sciences*, 21(16):377–380, 1931.