Stochastic processes in discrete time

Stefan Geiss

November 18, 2019

Contents

Important remark for the reading: These notes are only lecture notes which combines material from different sources, for example from the beautiful books [6] and [7]. Consequently, these lecture notes cannot replace the careful study of the literature as the mentioned books of Williams [7] and Shirjaev [6], and of Neveu [5]. Furthermore, there is a lecture notes of Hitczenko [3] which concerns the Central Limit Theorem for martingales not touched in these notes.

Chapter 1

Introduction

In the Introduction we want to motivate by two examples the main two parts of the lecture which deal with

- sums of independent random variables and
- martingale theory.

The examples are given at this stage in an intuitive way without being rigorous. In the course we will come back to the examples and treat them in a rigorous way.

Throughout this lecture we assume that the students are familiar with the basic probability theory. In particular, we assume that the material of the lecture notes [1] is known and will refer from time to time to these notes.

1.1 The gamblers ruin problem

We assume that there is a gambler starting with x_0 Euro as his starting capital at time $n = 0$. The game is as follows: we fix some $p \in (0, 1)$ as our *probability of loss* and at each time-step $n = 1, 2, ...$ the gambler

- loses with probability p the amount of 1 Euro,
- wins with probability $1 - p$ the amount of 1 Euro.

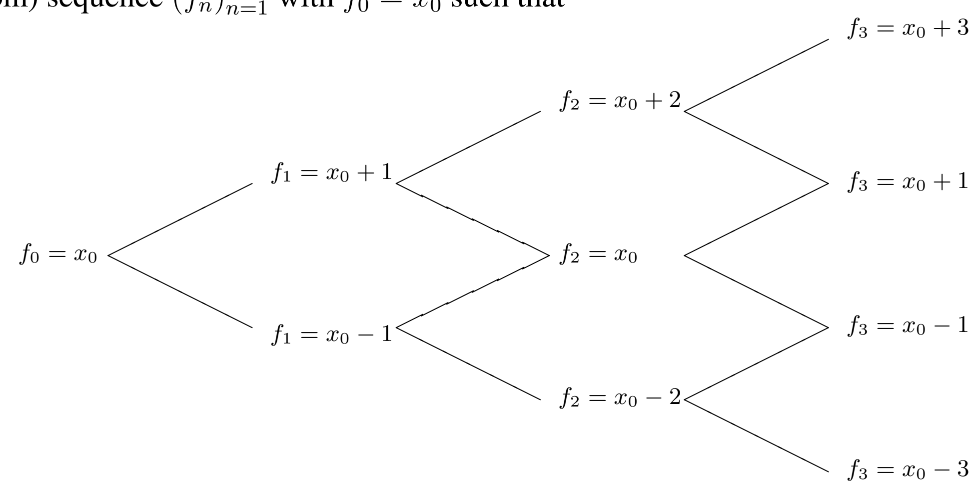
To realize this we can assume a coin which gives by flipping

- tails with probability p,

- heads with probability $1 - p$,

so that the gambler loses or wins if tails or heads, respectively, occurs. Of course, if $p \neq 1/2$, then the coin is not fair.

If f_n is the amount of capital of the gambler at time n, then we obtain a (random) sequence $(f_n)_{n=1}^{\infty}$ with $f_0 = x_0$ such that



To make our example more realistic, one would need to have at least a lower bound for f_n to bound the losses. Here we take two bounds $-\infty < A < x_0 < B < \infty$ where A and B are integers with the following interpretation:

- If the gambler reaches A, then he is ruined.
- If the gambler reaches B, then he wins.

This can also be interpreted as the random walk of a particle in a tube, where A is the lower boundary and B is the upper boundary. *Ruin* and *win* have here the meaning of touching the lower or upper boundary of the tube.

What are the questions one should ask now?

(Q1) What is the probability of ruin, that means

$$\mathbb{P}_{\text{ruin}} := \text{Probability}\left((f_n)_{n=0}^{\infty} \text{ reaches } A \text{ before } B\right)?$$

(Q2) What is the probability of win, that means

$$\mathbb{P}_{\text{win}} := \text{Probability}\left((f_n)_{n=0}^{\infty} \text{ reaches } B \text{ before } A\right)?$$

(Q3) What is the probability of an infinite length of the game, that means

$$\mathrm{Probability}\left((f_n)_{n=0}^{\infty} \text{ never reaches } A \text{ and } B\right)?$$

(Q4) What is the expected length of the game?

(Q5) Assume that we do not have bounds A and B. What is the probability of infinitely large losses, that means

$$\mathrm{Probability}\left(\inf_{n=1,2,\dots} f_n = -\infty\right).$$

(Q6) Assume that we do not have bounds A and B and that $x_0 = 0$. What is the probability that f_n reaches 0 infinitely often, that means

$$\mathrm{Probability}\left(\mathrm{card}\{n : f_n = 0\} = \infty\right)?$$

(Q7) Assume that $x_0 = 0$. Are there functions $\varphi^+, \varphi^- : [0,\infty) \to [0,\infty)$ such that

$$\mathrm{Probability}\left(\limsup_n \frac{f_n}{\varphi^+(n)} = 1\right) = 1$$

and

$$\mathrm{Probability}\left(\liminf_n \frac{f_n}{\varphi^-(n)} = -1\right) = 1.$$

Some of the above questions require deep insight into the theory of independent random variables, but some of them we can answer immediately:

Proposition 1.1.1. *Let* $0 = A < x_0 < B < \infty$, $B \in \mathbb{N}$, *and* $0 < p, q < 1$ *with* $p + q = 1$. *Then one has the following:*

(i) $$\mathbb{P}_{\mathrm{ruin}} := \begin{cases} \dfrac{\left(\frac{p}{q}\right)^B - \left(\frac{p}{q}\right)^{x_0}}{\left(\frac{p}{q}\right)^B - 1} & : \ p \neq q \\ 1 - \frac{x_0}{B} & : \ p = q = \frac{1}{2} \end{cases}.$$

(ii) $\mathbb{P}_{\mathrm{ruin}} + \mathbb{P}_{\mathrm{win}} = 1.$

This proposition answers (Q1) and (Q2) by explicit formulas and (Q3) by item (ii) which means that the game ends with probability one. The following table is an illustration:

p	q	x_0	B	$\mathbb{P}_{\text{ruin}}$
0.5	0.5	90	100	0.1
0.55	0.45	90	100	0.866

The table says that the probability of ruin can drastically go up if the coin is slightly unfair even if we start off from a very good position.

Proof of Proposition 1.1.1. (a) Let $0 \le x \le B$, where x is an integer, and let

- r_x be the probability of ruin if one starts gambling with x Euro,
- w_x is the probability of win if one starts gambling with x Euro.

We have that $r_0 = 1$, $r_B = 0$, $w_0 = 0$, and $w_B = 1$. We get the *difference equation*

$$r_x = q r_{x+1} + p r_{x-1} \quad \text{with} \quad r_0 = 1 \quad \text{and} \quad r_B = 0, \tag{1.1}$$

where $x = 1, ..., B-1$.

(b) Existence and uniqueness of the solution for $p \neq q$: Assume that we have a solution $(r_x)_{x=0}^B$, we find unique α and β such that

$$r_0 = \alpha + \beta \quad \text{and} \quad r_1 = \alpha + \beta \frac{p}{q}$$

(note the $p/q \neq 1$ gives unique α and β). By the first equation in (1.1) we can compute that $r_x = \alpha + \beta (p/q)^x$ by iteration over $x = 1, 2, ...$ Hence r_x is a solution to $r_x = q r_{x+1} + p r_{x-1}$ if and only if it is of the form

$$r_x = \alpha + \beta (p/q)^x.$$

To adjust the boundary conditions we re-compute the parameters α and β and solve

$$\begin{aligned} 1 = r_0 &= \alpha + \beta, \\ 0 = r_B &= \alpha + \beta \left(\frac{p}{q} \right)^B \end{aligned}$$

so that

$$\beta = \frac{1}{1-\left(\frac{p}{q}\right)^B} \qquad \text{and} \qquad \alpha = \frac{-\left(\frac{p}{q}\right)^B}{1-\left(\frac{p}{q}\right)^B}$$

and

$$r_x = \alpha + \beta\left(\frac{p}{q}\right)^x = \frac{\left(\frac{p}{q}\right)^x - \left(\frac{p}{q}\right)^B}{1-\left(\frac{p}{q}\right)^B}.$$

(c) In the same way one gets a solution for $\mathbb{P}_{\text{win}}$ and one computes that $\mathbb{P}_{\text{win}} + \mathbb{P}_{\text{ruin}} = 1$.

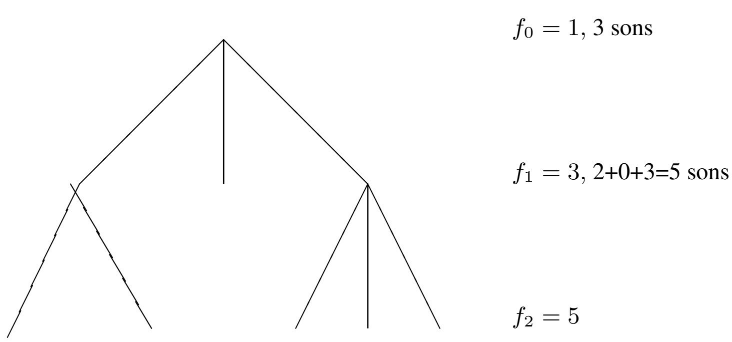
(d) The case $p = q = \frac{1}{2}$ is an exercise. □

1.2 Branching processes

Assume that at some time $n = 0$ there was exactly one family with the name HÄKKINEN in Finland. From generation n to generation $n+1$ the following may happen:

- If a family with name HÄKKINEN has a son at generation n, then the son carries this name to the next generation $n+1$.
- If there is no son then this family cannot carry over his name to the generation $n+1$.

So we get the following tree, which can be interpreted as a *branching process*:



Let f_n be the total number of families in generation n having the name HÄKKINEN. If for some n we have that $f_n = 0$, then the name has died out. What is the probability of this event and what does it depend on?

We let $0 \le q_0, ..., q_N \le 1$ such that $1 = q_0 + \cdots + q_N$ with the interpretation that q_k is the probability that a family has k sons. Obviously, we can assume that N is finite. To describe f_n we use *independent* random variables $(X_i^{(n)})_{i,n=1}^\infty$ such that

$$\mathbb{P}(X_i^{(n)} = k) = q_k.$$

In this way, $X_i^{(n)} = k$ describes the event that the i-th family in generation $n-1$ has k sons. Then we consider the system

$$\begin{aligned} f_0 &:= 1, \\ f_{n+1} &:= X_1^{(n+1)} + X_2^{(n+1)} + \cdots + X_{f_n}^{(n+1)}. \end{aligned}$$

This gives a sum of independent random variables which is (this is the point) *randomly* stopped. To treat sums like this we shall need martingale theory treated in the second part of our lecture. In particular, we will be able to prove the following

Proposition 1.2.1. *Let* $\mu := \sum_{k=0}^N k q_k > 0$, $q_0 \in (0,1)$, *and* $M_n := \frac{f_n}{\mu^n}$. *Let* $(\mathcal{F}_n)_{n=0}^\infty$ *be the filtration generated by* $(f_n)_{n=0}^\infty$. *Then one has the following:*

(i) $(M_n)_{n=0}^\infty$ *is a martingale with respect to the filtration* $(\mathcal{F}_n)_{n=0}^\infty$, *that means* $\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n$ *a.s.,*

(ii) $M_\infty := \lim_{n\to\infty} M_n$ *exists almost surely.*

(iii) *If* $\mu \le 1$, *then* $M_\infty = 0$ *a.s. and* $\lim_n \mathbb{P}(M_n = 0) = 1$.

(iv) *If* $\mu > 1$, *then* $\mathbb{E}M_\infty = 1$ *and* $\lim_n \mathbb{P}(M_n = 0) \in (0,1)$.

The interpretation of μ is the expected number of sons in each generation of a single family. Then, μ^n can be seen as some kind of average number of families with name HÄKKINEN in the n-th generation. So f_n/μ^n should be the right quantity which stabilizes after some time.

1.3 Some notation

The symbol $\mathbb{N}$ stands for the natural numbers $\{1, 2, ...\}$. For a sequence of real numbers $(\xi_n)_{n=1}^{\infty}$ we recall that

$$\begin{aligned}\limsup_{n\to\infty} \xi_n &:= \lim_{n\to\infty} \sup_{k\geq n} \xi_k, \\ \liminf_{n\to\infty} \xi_n &:= \lim_{n\to\infty} \inf_{k\geq n} \xi_k.\end{aligned}$$

If $a, b \in \mathbb{R}$, then we let $a \wedge b := \min\{a, b\}$ and $a^+ := \max\{a, 0\}$. Besides this notation we will need

Lemma 1.3.1. [Lemma of Borel-Cantelli] [1] [2] *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A_1, A_2, ... \in \mathcal{F}$. Then one has the following:*

(1) *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$.*

(2) *If $A_1, A_2, ...$ are assumed to be independent and $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$.*

[1] Félix Edouard Justin Émile Borel, 07/01/1871 (Saint Affrique, Aveyron, Midi-Pyrénées, France)- 03/02/1956 (Paris, France), measure theory, modern theory of functions of a real variable.

[2] Francesco Paolo Cantelli, 20/12/1875 (Palermo, Sicily, Italy) - 21/07/1966 (Rome, Italy), Italian mathematician.

Chapter 2

Sums of independent random variables

First let us recall the independence of random variables.

Definition 2.0.1. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $f_\alpha : \Omega \to \mathbb{R}$ *with* $\alpha \in I$ *be a family of random variables. The family is called independent provided that for all pairwise different* $\alpha_1, ..., \alpha_n \in I$ *and all* $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$ *one has that*

$$\mathbb{P}(f_{\alpha_1} \in B_1, ..., f_{\alpha_n} \in B_n) = \mathbb{P}(f_{\alpha_1} \in B_1) \cdots \mathbb{P}(f_{\alpha_n} \in B_n).$$

2.1 Zero-One laws

It is known that $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$ but that $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ does converge. What is going to happen if we take a random sequence of signs?

Definition 2.1.1. *For* $p \in (0,1)$ *we denote by* $\varepsilon_1^{(p)}, \varepsilon_2^{(p)}, ... : \Omega \to \mathbb{R}$ *a sequence of independent random variables such that*

$$\mathbb{P}\left(\varepsilon_n^{(p)} = -1\right) = p \quad \text{and} \quad \mathbb{P}\left(\varepsilon_n^{(p)} = 1\right) = 1 - p.$$

If $p = 1/2$ *we write* $\varepsilon_n = \varepsilon_n^{(1/2)}$ *and call this sequence a sequence of* BERNOULLI [1] *random variables.*

[1] Jacob Bernoulli, 27/12/1654 (Basel, Switzerland)- 16/08/1705 (Basel, Switzerland), Swiss mathematician.

Now we are interested in

$$\mathbb{P}\left(\sum_{n=1}^{\infty}\frac{\varepsilon_n}{n}\text{ converges}\right).$$

Remark 2.1.2. Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be random variables over $(\Omega, \mathcal{F}, \mathbb{P})$. Then

$$A := \left\{\omega \in \Omega : \sum_{n=1}^{\infty} \xi_n(\omega) \text{ converges}\right\} \in \mathcal{F}.$$

This is not difficult to verify because $\omega \in A$ if and only if

$$\omega \in \bigcap_{N=1,2,...}\ \bigcup_{n_0=1,2,...}\ \bigcap_{m>n\geq n_0} \left\{\omega \in \Omega : \left|\sum_{k=n+1}^{m} \xi_k(\omega)\right| < \frac{1}{N}\right\}.$$

What is a typical property of the set $A = \{\omega \in \Omega : \sum_{n=1}^{\infty} \xi_n(\omega) \text{ converges}\}$? The condition does not depend on the first realizations $\xi_1(\omega), ..., \xi_N(\omega)$ since the convergence of $\sum_{n=1}^{\infty} \xi_n(\omega)$ and $\sum_{n=N+1}^{\infty} \xi_n(\omega)$ are equivalent so that

$$A = \left\{\omega \in \Omega : \sum_{n=N+1}^{\infty} \xi_n(\omega) \text{ converges}\right\}.$$

We shall formulate this in a more abstract way. For this we need

Definition 2.1.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\xi_\alpha : \Omega \to \mathbb{R}$ be a family of random variables. Then $\sigma(\xi_\alpha : \alpha \in I)$ is the smallest σ-algebra which contains all sets of form*

$$\{\omega \in \Omega : \xi_\alpha(\omega) \in B\} \quad \textit{where} \quad \alpha \in I \quad \textit{and} \quad B \in \mathcal{B}(\mathbb{R}).$$

Exercise 2.1.4. Show that $\sigma(\xi_\alpha : \alpha \in I)$ is the smallest σ-algebra on Ω such that all random variables $\xi_\alpha : \Omega \to \mathbb{R}$ are measurable.

Example 2.1.5. For random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ and

$$A = \left\{\omega \in \Omega : \sum_{n=1}^{\infty} \xi_n(\omega) \text{ converges}\right\}$$

one has that $A \in \bigcap_{N=1}^{\infty} \sigma(\xi_N, \xi_{N+1}, \xi_{N+2}, ...)$.

The above example leads us straight to the famous Zero-One law of KOLMOGOROV [2]:

Proposition 2.1.6 (Zero-One law of KOLMOGOROV)**.** *Assume independent random variables* $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ *on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let*

$$\mathcal{F}_N^\infty := \sigma(\xi_N, \xi_{N+1}, ...) \quad \text{and} \quad \mathcal{F}^\infty := \bigcap \mathcal{F}_n^\infty.$$

Then $\mathbb{P}(A) \in \{0,1\}$ *for all* $A \in \mathcal{F}^\infty$.

For the proof the following lemma is needed:

Lemma 2.1.7. *Let* $\mathcal{A} \subseteq \mathcal{F}$ *be an algebra and let* $\sigma(\mathcal{A}) = \mathcal{F}$*. Then, for all* $\varepsilon > 0$ *and* $B \in \mathcal{F}$*, there is an* $A \in \mathcal{A}$ *such that*

$$\mathbb{P}(A \Delta B) < \varepsilon.$$

Proof of Proposition 2.1.6. The idea of the proof is to show that $\mathbb{P}(A) = \mathbb{P}(A)^2$. Define the algebra

$$\mathcal{A} := \bigcup_{n=1}^{\infty} \sigma(\xi_1, ..., \xi_n).$$

We have that $\mathcal{F}^\infty \subseteq \sigma(\mathcal{A})$. Hence Lemma 2.1.7 implies that for $A \in \mathcal{F}^\infty$ there are $A_n \in \sigma(\xi_1, ..., \xi_{N_n})$ such that

$$\mathbb{P}(A \Delta A_n) \to_{n\to\infty} 0.$$

We get also that

$$\mathbb{P}(A_n \cap A) \to_{n\to\infty} \mathbb{P}(A) \quad \text{and} \quad \mathbb{P}(A_n) \to_{n\to\infty} \mathbb{P}(A).$$

The first relation can be seen as follows: since

$$\mathbb{P}(A_n \cap A) + \mathbb{P}(A_n \Delta A) = \mathbb{P}(A_n \cup A) \geq \mathbb{P}(A)$$

[2] Andrey Nikolaevich Kolmogorov 25/04/1903 (Tambov, Russia) - 20/10/1987 (Moscow, Russia), one of the founders of modern probability theory, Wolf prize 1980.

we get that

$$\liminf_n \mathbb{P}(A_n \cap A) \geq \mathbb{P}(A) \geq \limsup \mathbb{P}(A_n \cap A).$$

The second relation can be also checked easily. But now we get, by independence, that

$$\mathbb{P}(A) = \lim_n \mathbb{P}(A \cap A_n) = \lim_n \mathbb{P}(A)\mathbb{P}(A_n) = \mathbb{P}(A)^2$$

so that $\mathbb{P}(A) \in \{0, 1\}$. □

Corollary 2.1.8. *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be independent random variables over $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \textit{ converges}\right) \in \{0, 1\}.$$

We also obtain an early version of the law of the iterated logarithm (LIL). But for this we need the central limit theorem (CLT) for the BERNOULLI random variables, which we will not prove here.

Proposition 2.1.9 (CLT). *Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ be independent* BERNOULLI *random variables and $c \in \mathbb{R}$. Then one has that*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\varepsilon_1 + \cdots + \varepsilon_n}{\sqrt{n}} \leq c\right) = \int_{-\infty}^{c} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

As a second application of KOLMOGOROV's Zero-One law we get

Corollary 2.1.10. *Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ be independent* BERNOULLI *random variables. Then one has that*

$$\mathbb{P}\left(\limsup_n \frac{f_n}{\sqrt{n}} = \infty, \quad \liminf_n \frac{f_n}{\sqrt{n}} = -\infty\right) = 1$$

where $f_n := \varepsilon_1 + \cdots + \varepsilon_n$.

Proof. The assertion is equivalent to

$$\mathbb{P}\left(\limsup_n \frac{f_n}{\sqrt{n}} = \infty\right) = 1 \quad \text{and} \quad \mathbb{P}\left(\liminf_n \frac{f_n}{\sqrt{n}} = -\infty\right) = 1.$$

By symmetry it is sufficient to prove the first equality only. Letting

$$\overline{A}_c := \left\{ \limsup_n \frac{f_n}{\sqrt{n}} \geq c \right\}$$

for $c \geq 0$ we get that

$$\mathbb{P}\left(\limsup_n \frac{f_n}{\sqrt{n}} = \infty \right) = \mathbb{P}\left(\bigcap_{m=1}^{\infty} \overline{A}_m \right) = \lim_m \mathbb{P}\left(\overline{A}_m \right)$$

because of $\overline{A}_m \supseteq \overline{A}_{m+1}$. Hence it would be sufficient to prove that $\mathbb{P}\left(\overline{A}_m\right) = 1$. Since

$$\begin{aligned} & \limsup_n \frac{\varepsilon_1(\omega) + \cdots + \varepsilon_n(\omega)}{\sqrt{n}} \\ = \; & \limsup_{n\to\infty, n\geq N} \left(\frac{\varepsilon_1(\omega) + \cdots + \varepsilon_{N-1}(\omega)}{\sqrt{n}} + \frac{\varepsilon_N(\omega) + \cdots + \varepsilon_n(\omega)}{\sqrt{n}} \right) \\ = \; & \limsup_{n\to\infty, n\geq N} \frac{\varepsilon_N(\omega) + \cdots + \varepsilon_n(\omega)}{\sqrt{n}} \end{aligned}$$

because of

$$\lim_{n\to\infty, n\geq N} \frac{\varepsilon_1(\omega) + \cdots + \varepsilon_{N-1}(\omega)}{\sqrt{n}} = 0,$$

where we assume that $N \geq 2$, we get that

$$\overline{A}_m \in \bigcap_{N=1}^{\infty} \sigma\left(\varepsilon_N, \varepsilon_{N+1}, ...\right).$$

Consequently, in order to prove $\mathbb{P}\left(\overline{A}_m\right) = 1$ we are in a position to apply KOLMOGOROV's Zero-One law (Proposition 2.1.6) and the only thing to verify is

$$\mathbb{P}\left(\overline{A}_m\right) > 0.$$

To get this, we first observe that

$$\overline{A}_m = \left\{ \limsup_n \frac{f_n}{\sqrt{n}} \geq m \right\} \supseteq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ \frac{f_k}{\sqrt{k}} \geq m \right\}$$

since $\omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ \frac{f_k}{\sqrt{k}} \geq m \right\}$ implies that for all $n = 1, 2, ...$ there is a $k \geq n$ such that $f_k(\omega)/\sqrt{k} \geq m$, which gives a subsequence $(k_l)_{l=1}^{\infty}$ with $f_{k_l}(\omega)/\sqrt{k_l} \geq$

m and finally $\limsup_n f_n(\omega)/\sqrt{n} \geq m$. Applying FATOU [3]'s lemma we derive from that

$$\mathbb{P}(\overline{A}_m) \geq \mathbb{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left\{\frac{f_k}{\sqrt{k}} \geq m\right\}\right) \geq \limsup_n \mathbb{P}\left(\frac{f_n}{\sqrt{n}} \geq m\right).$$

Since the central limit theorem (Proposition 2.1.9) gives

$$\lim_n \mathbb{P}\left(\frac{f_n}{\sqrt{n}} \geq m\right) = \int_m^{\infty} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} > 0$$

we are done. □

Now let us come back to question (Q6) of the ruin problem in Section 1.1. Let $p, q \in (0,1)$, $p+q=1$, and $f_n := \sum_{i=1}^n \varepsilon_i^{(p)}$ where $n = 1, 2, ...$ and the random variables $\varepsilon_i^{(p)}$ were defined in Definition 2.1.1. Consider the event

$$A := \{\omega : \#\{n : f_n(\omega) = 0\} = \infty\}.$$

In words, A is the event, that the path of the process $(f_n)_{n=1}^{\infty}$ reaches 0 infinitely many often. We would like to have a Zero-One law for this event. However, we are not able to apply KOLMOGOROV's Zero-One law (Proposition 2.1.6). But what is the typical property of A? We can rearrange *finitely* many elements of the sum $\sum_{i=1}^n \varepsilon_i^{(p)}$ and we will get the same event since from some N on, this rearrangement does not influence the sum anymore. To give a formal definition of this property we need

Definition 2.1.11. *A map $\pi : \mathbb{N} \to \mathbb{N}$ is called* finite permutation *if*

(i) *the map π is a bijection,*

(ii) *there is some $N \in \mathbb{N}$ such that $\pi(n) = n$ for all $n \geq N$.*

Moreover, we recall that $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ is the smallest σ-algebra on $\mathbb{R}^{\mathbb{N}} = \{(\xi_1, \xi_2, ...) : \xi_n \in \mathbb{R}\}$ which contains all cylinder sets Z of form

$$Z := \{(\xi_1, \xi_2, ...) : a_n < \xi_n < b_n, n = 1, 2, ...\}$$

for some $-\infty < a_n < b_n < \infty$. Now the symmetry property, needed for the next Zero-One law, looks as follows:

[3]Pierre Joseph Louis Fatou, 28/02/1878 (Lorient, France) - 10/08/1929 (Pornichet, France), French mathematician, dynamical systems, Mandelbrot set.

Definition 2.1.12. *Let* $(\xi_n)_{n=1}^\infty$, $\xi_n : \Omega \to \mathbb{R}$, *be a sequence of independent random variables over a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *Let* $B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ *and*

$$A := \{\omega \in \Omega : (\xi_1(\omega), \xi_2(\omega), ...) \in B\}.$$

The set A *is called* symmetric *if for all finite permutations* $\pi : \mathbb{N} \to \mathbb{N}$ *one has that*

$$A = \left\{\omega \in \Omega : (\xi_{\pi(1)}(\omega), \xi_{\pi(2)}(\omega), ...) \in B\right\}.$$

A typical set B which serves as *basis* for the set A is given by

Example 2.1.13. *We let* $B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ *be the set of all sequences* $(\xi_n)_{n=1}^\infty$, $\xi_n \in \{-1, 1\}$, *such that*

$$\#\{n : \xi_1 + \cdots + \xi_n = 0\} = \infty.$$

The next Zero-One law works for identically distributed random variables. Recall that a sequence of random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is *identically distributed* provided that

$$\mathbb{P}(\xi_k > \lambda) = \mathbb{P}(\xi_l > \lambda)$$

for all $k, l \in \mathbb{N}$ and $\lambda \in \mathbb{R}$.

Proposition 2.1.14 (Zero-One law of HEWITT and SAVAGE). [4] *Assume a sequence of independent and identically distributed random variables* $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ *over some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *If the event* $A \in \mathcal{F}$ *is symmetric, then* $\mathbb{P}(A) \in \{0, 1\}$.

Proof. (a) We are going to use the permutations

$$\pi_n(k) := \begin{cases} n+k & : \quad 1 \le k \le n \\ k-n & : \quad n+1 \le k \le 2n \\ k & : \quad k > 2n \end{cases}.$$

Now we approximate the set A. By Lemma 2.1.7 we find $B_n \in \mathcal{B}(\mathbb{R}^n)$ such that

$$\mathbb{P}(A_n \Delta A) \to_n 0 \quad \text{for} \quad A_n := \{\omega : (\xi_1(\omega), \ldots, \xi_n(\omega)) \in B_n\}.$$

[4] Leonard Jimmie Savage 20/11/1917 (Detroit, USA) - 1/11/1971 (New Haven, USA), American mathematician and statistician.

(b) Our goal is to show that

$$\begin{aligned} \mathbb{P}(A_n) &\to \mathbb{P}(A), \\ \mathbb{P}(\pi_n(A_n)) &\to \mathbb{P}(A), \\ \mathbb{P}(A_n \cap \pi_n(A_n)) &\to \mathbb{P}(A), \end{aligned}$$

as $n \to \infty$, where

$$\begin{aligned} \pi_n(A_n) &:= \left\{\omega : \left(\xi_{\pi_n(k)}(\omega)\right)_{k=1}^n \in B_n\right\} \\ &= \left\{\omega : (\xi_{n+1}(\omega), \ldots, \xi_{2n}(\omega)) \in B_n\right\}. \end{aligned}$$

(c) Assuming for a moment that (b) is proved we derive

$$\mathbb{P}(A_n \cap \pi_n(A_n)) = \mathbb{P}(A_n)\mathbb{P}(\pi_n(A_n))$$

since A_n is a condition on $\xi_1, ..., \xi_n$, $\pi_n(A_n)$ is a condition on $\xi_{n+1}, ..., \xi_{2n}$, and $\xi_1, ..., \xi_{2n}$ are independent. By $n \to \infty$ this equality turns into

$$\mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(A)$$

so that $\mathbb{P}(A) \in \{0, 1\}$.

(d) Now we prove (b). The convergence

$$\mathbb{P}(A_n) \to \mathbb{P}(A)$$

is obvious since $\mathbb{P}(A_n \Delta A) \to 0$. This implies

$$\mathbb{P}(\pi_n(A_n)) \to \mathbb{P}(A)$$

since $\mathbb{P}(\pi_n(A_n)) = \mathbb{P}(A_n)$ which follows from the fact that

$$(\xi_{n+1}, ..., \xi_{2n}, \xi_1, ..., \xi_n) \quad \text{and} \quad (\xi_1, ..., \xi_n, \xi_{n+1}, ..., \xi_{2n})$$

have the same law in $\mathbb{R}^{2n}$ as a consequence that we have an identically distributed sequence of random variables. Finally

$$\begin{aligned} \mathbb{P}(A \Delta A_n) &= \mathbb{P}(\pi_n(A \Delta A_n)) \\ &= \mathbb{P}(\pi_n(A) \Delta \pi_n(A_n)) \\ &= \mathbb{P}(A \Delta \pi_n(A_n)) \end{aligned}$$

where the first equality follows because the random variables $(\xi_1, \xi_2, ...)$ have the same distribution and the last equality is a consequence of the symmetry of A. Hence

$$\mathbb{P}\left(A\Delta A_n\right) \to_n 0 \quad \text{and} \quad \mathbb{P}\left(A\Delta\pi_n(A_n)\right) \to_n 0$$

which implies

$$\mathbb{P}\left(A\Delta\left(A_n \cap \pi_n(A_n)\right)\right) \to_n 0 \quad \text{and} \quad \mathbb{P}\left(A_n \cap \pi_n(A_n)\right) \to \mathbb{P}(A).$$

□

As an application we consider

Proposition 2.1.15. *Let* $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ *be independent* BERNOULLI *random variables and* $f_n = \sum_{i=1}^n \varepsilon_i$, $n = 1, 2, ...$ *Then*

$$\mathbb{P}(\#\{n : f_n = 0\} = \infty) = 1.$$

Proof. Consider the sets

$$\begin{aligned} A^+ &:= \{\omega \in \Omega : \#\{n : f_n(\omega) = 0\} < \infty\} \cap \{\omega \in \Omega : \liminf f_n(\omega) > 0\}, \\ A &:= \{\omega \in \Omega : \#\{n : f_n(\omega) = 0\} = \infty\}, \\ A_- &:= \{\omega \in \Omega : \#\{n : f_n(\omega) = 0\} < \infty\} \cap \{\omega \in \Omega : \limsup f_n(\omega) < 0\}. \end{aligned}$$

Since the random walk is symmetric we have

$$\mathbb{P}(A_+) = \mathbb{P}(A_-).$$

Moreover, A_+, A, and A_- are symmetric, so that

$$\mathbb{P}(A_+), \mathbb{P}(A), \mathbb{P}(A_-) \in \{0, 1\}$$

by the Zero-One law of HEWITT-SAVAGE. As the only solution to that we obtain

$$\mathbb{P}(A) = 1 \quad \text{and} \quad \mathbb{P}(A_+) = \mathbb{P}(A_-) = 0.$$

□

Other examples for which the Zero-One law of KOLMOGOROV or HEWITT and SAVAGE might be applied are given in

Example 2.1.16. Let us illustrate by some additional very simple examples the difference of the assumptions for the Zero-One law of KOLMOGOROV and the Zero-One law of HEWITT-SAVAGE. Assume a sequence of independent and identically distributed random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and that $\mathcal{F}^\infty = \bigcap_{n=1}^\infty \sigma(\xi_n, \xi_{n+1}, \ldots)$.

(a) Given $B \in \mathcal{B}(\mathbb{R})$ we have that $\{\limsup_n \xi_n \in B\} \in \mathcal{F}^\infty$. Here we do not need the same distribution of the random variables $\xi_1, \xi_2, ...$

(b) The set $A := \{\omega \in \Omega : \ \xi_n(\omega) = 0 \text{ for all } n = 1, 2, \ldots\}$ does not belong to $\mathcal{F}^\infty$ but is symmetric, because

$$\begin{aligned} A &= \{\omega \in \Omega : (\xi_n(\omega))_{n=1}^\infty \in B\} \\ &= \left\{\omega \in \Omega : \left(\xi_{\pi(n)}(\omega)\right)_{n=1}^\infty \in B\right\} \end{aligned}$$

for $B = \{(0,0,0,\ldots)\} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$.

(c) The set $A := \{\omega \in \Omega : \ \Sigma_{n=1}^\infty |\xi_n(\omega)| \text{ exists and } \Sigma_{n=1}^\infty \xi_n(\omega) < 1\}$ does not belong to $\mathcal{F}^\infty$ but is symmetric, because

$$\begin{aligned} A &= \{\omega \in \Omega : (\xi_n(\omega))_{n=1}^\infty \in B\} \\ &= \left\{\omega \in \Omega : \left(\xi_{\pi(n)}(\omega)\right)_{n=1}^\infty \in B\right\} \end{aligned}$$

if B is the set of all summable sequences with sum strictly less than one.

We finish by the non-symmetric random walk. As preparation we need STIRLING [5]'s formula

$$n! = \sqrt{2\pi n} \quad \left(\frac{n}{e}\right)^n \quad e^{\frac{\theta}{12n}}$$

for $n = 1, 2, ...$ and some $\theta \in (0,1)$ depending on n. This gives

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} = \frac{\sqrt{2\pi(2n)}\left(\frac{2n}{e}\right)^{2n} e^{\frac{\theta_1}{12\cdot 2n}}}{(\sqrt{2\pi n})^2 \left(\frac{n}{e}\right)^{2n} \left(e^{\frac{\theta_2}{12n}}\right)^2} \sim \frac{4^n}{\sqrt{\pi n}}.$$

[5] James Stirling May 1692 (Garden, Scotland) - 5/12/1770 (Edinburgh, Scotland), Scottish mathematician whose most important work Methodus Differentialis in 1730 is a treatise on infinite series, summation, interpolation and quadrature.

Proposition 2.1.17. *Let $p \neq 1/2$ and $f_n = \sum_{i=1}^n \varepsilon_i^{(p)}$, $n = 1, 2, \ldots$, where the random variables $\varepsilon_1^{(p)}, \varepsilon_2^{(p)}, \ldots : \Omega \to \mathbb{R}$ are given by Definition 2.1.1. Then*

$$\mathbb{P}\left(\#\left\{n : f_n = 0\right\} = \infty\right) = 0.$$

Proof. Letting $B_n := \{f_{2n} = 0\}$ we obtain

$$\mathbb{P}(B_n) = \binom{2n}{n}(pq)^n \sim \frac{(4pq)^n}{\sqrt{\pi n}}$$

by STIRLING's formula. Since $p \neq q$ gives $4pq < 1$, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(B_n) < \infty$$

so that the Lemma of BOREL-CANTELLI implies that

$$\mathbb{P}\left(\omega \in B_n \text{ infinitely often }\right) = 0.$$

□

2.2 Convergence of sums

If $\varepsilon_1, \varepsilon_2, \ldots$ are independent BERNOULLI random variables, that means $\mathbb{P}\left(\varepsilon_n = 1\right) = \mathbb{P}\left(\varepsilon_i = -1\right) = \frac{1}{2}$, then we know from Section 2.1 that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \frac{\varepsilon_n}{n} \text{ converges}\right) \in \{0, 1\}.$$

But, do we get probability zero or one? For this there is a beautiful complete answer: It consists of the Three-Series-Theorem of KOLMOGOROV which we will deduce from the Two-Series-Theorem of KOLMOGOROV. First we formulate the Two-Series-Theorem, then we give some examples and consequences (including the Three-Series-Theorem), and finally we prove the Two-Series-Theorem.

Besides the investigation of sums of independent random variables this section also provides some basic concepts going into the direction of the martingale-theory.

2.2.1 The Two-Series-Theorem and its applications

Proposition 2.2.1 (Two-Series-Theorem of KOLMOGOROV). *Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ which are assumed to be* independent.

(i) *If $\sum_{n=1}^{\infty} \mathbb{E}\xi_n$ converges and $\sum_{n=1}^{\infty} \mathbb{E}\left(\xi_n - \mathbb{E}\xi_n\right)^2 < \infty$, then*

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \text{ converges }\right) = 1.$$

(ii) *Let $|\xi_n(\omega)| \leq d$ for all $n = 1, 2, ...$ and all $\omega \in \Omega$, where $d > 0$ is some constant. Then the following assertions are equivalent:*

  (a) $\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \text{ converges }\right) = 1.$

  (b) $\sum_{n=1}^{\infty} \mathbb{E}\xi_n$ *converges and* $\sum_{n=1}^{\infty} \mathbb{E}\left(\xi_n - \mathbb{E}\xi_n\right)^2 < \infty$.

Now we consider an example.

Example 2.2.2. Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \infty$ be independent BERNOULLI random variables, $\alpha_1, \alpha_2, ... \in \mathbb{R}$ and $\beta_1, \beta_2, ... \in \mathbb{R}$. Then

$$\mathbb{P}\left(\sum_{n=1}^{\infty}[\alpha_n + \beta_n\varepsilon_n] \text{ converges}\right) = 1$$

if and only if $\sum_{n=1}^{\infty} \alpha_n$ converges and $\sum_{n=1}^{\infty} \beta_n^2 < \infty$.

Proof. Assuming the conditions on $(\alpha_n)_{n=1}^{\infty}$ and $(\beta_n)_{n=1}^{\infty}$ we get, for $\xi_n := \alpha_n + \beta_n\varepsilon_n$, that

$$\sum_{n=1}^{\infty} \mathbb{E}\xi_n = \sum_{n=1}^{\infty}[\alpha_n + \beta_n \mathbb{E}\varepsilon_n] = \sum_{n=1}^{\infty} \alpha_n$$

and

$$\sum_{n=1}^{\infty} \mathbb{E}[\xi_n - \mathbb{E}\xi_n]^2 = \sum_{n=1}^{\infty} \mathbb{E}[\beta_n\varepsilon_n]^2 = \sum_{n=1}^{\infty} \beta_n^2 < \infty.$$

Hence we can apply the Two-Series-Theorem to obtain the almost sure convergence of $\sum_{n=1}^{\infty}[\alpha_n + \beta_n\varepsilon_n]$. Assume now that

$$\mathbb{P}\left(\sum_{n=1}^{\infty}[\alpha_n + \beta_n\varepsilon_n] \text{ converges}\right) = 1. \tag{2.1}$$

First we want to deduce that $\sup_n |\alpha_n| < \infty$ and $\sup_n |\beta_n| < \infty$. For this purpose we define $\eta_n(\omega) := -\varepsilon_n(\omega)$. It is clear that $\eta_1, \eta_2, ... : \Omega \to \mathbb{R}$ are independent BERNOULLI random variables. Since our assumption (2.1) is an assumption on the *distribution* of $(\varepsilon_1, \varepsilon_2, ...)$ which is the same as the distribution of $(\eta_1, \eta_2, ...)$ we get (2.1) for the sequence $(\eta_n)_{n=1}^\infty$ as well. Hence there are $\Omega_\varepsilon, \Omega_\eta \in \mathcal{F}$ of measure one such that

$$\sum_{n=1}^\infty [\alpha_n + \beta_n \varepsilon_n(\omega)] \quad \text{and} \quad \sum_{n=1}^\infty [\alpha_n + \beta_n \eta_n(\omega')]$$

converge for $\omega \in \Omega_\varepsilon$ and $\omega' \in \Omega_\eta$. Since $\Omega_\varepsilon \cap \Omega_\eta$ is of measure one, there exists at least one $\omega_0 \in \Omega_\varepsilon \cap \Omega_\eta$ such that

$$\sum_{n=1}^\infty [\alpha_n + \beta_n \varepsilon_n(\omega_0)] \quad \text{and} \quad \sum_{n=1}^\infty [\alpha_n + \beta_n \eta_n(\omega_0)]$$

converge. Taking the sum, we get that

$$\sum_{n=1}^\infty [\alpha_n + \beta_n \varepsilon_n(\omega_0)] + \sum_{n=1}^\infty [\alpha_n + \beta_n \eta_n(\omega_0)] = \sum_{n=1}^\infty [\alpha_n + \beta_n \varepsilon_n(\omega_0)] + \sum_{n=1}^\infty [\alpha_n + \beta_n(-\varepsilon_n(\omega_0))] = 2\sum_{n=1}^\infty \alpha_n$$

converges. From that we can deduce in turn that

$$\mathbb{P}\left(\sum_{n=1}^\infty \beta_n \varepsilon_n \text{ converges} \right) = 1.$$

Picking again an $\omega_0' \in \Omega$ such that

$$\sum_{n=1}^\infty \beta_n \varepsilon_n(\omega_0')$$

converges, we get that $\sup_n |\beta_n| = \sup_n |\beta_n||\varepsilon_n(\omega_0)| < \infty$. Hence

$$|\xi_n(\omega)| \le |\alpha_n| + |\beta_n| \le \sup_m |\alpha_m| + \sup_m |\beta_m| < \infty$$

and we can apply Proposition 2.2.1 which gives that

$$\sum_{n=1}^{\infty} \alpha_n = \sum_{n=1}^{\infty} \mathbb{E}\xi_n \quad \text{and} \quad \sum_{n=1}^{\infty} \beta_n^2 = \sum_{n=1}^{\infty} \mathbb{E}|\xi_n - \mathbb{E}\xi_n|^2$$

converge (the convergence of $\sum_{n=1}^{\infty} \alpha_n$ we also derived before). □

As we saw, the Two-Series-Theorem only gives an equivalence in the case that $\sup_{n,\omega} |\xi_n(\omega)| < \infty$. For the general case we have an equivalence as well, the Three-Series-Theorem, which one can quickly deduce from the Two-Series-Theorem. For its formulation we introduce for a random variable $f : \Omega \to \mathbb{R}$ and some constant $c > 0$ the *truncated* random variable

$$f^c(\omega) := \begin{cases} f(\omega) & : \quad |f(\omega)| \leq c \\ c & : \quad f(\omega) > c \\ -c & : \quad f(\omega) < -c \end{cases} .$$

Corollary 2.2.3 (Three-Series-Theorem of KOLMOGOROV)**.** *Assume a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* independent *random variables* $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$. *Then the following conditions are equivalent:*

(i) $\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \textit{ converges }\right) = 1$.

(ii) *For* all *constants* $c > 0$ *the following three conditions are satisfied:*

  (a) $\sum_{n=1}^{\infty} \mathbb{E}\xi_n^c$ *converges,*

  (b) $\sum_{n=1}^{\infty} \mathbb{E}\left(\xi_n^c - \mathbb{E}\xi_n^c\right)^2 < \infty$,

  (c) $\sum_{n=1}^{\infty} \mathbb{P}\left(|\xi_n| \geq c\right) < \infty$.

(iii) *There* exists *one constant* $c > 0$ *such that conditions* (a)*,* (b)*, and* (c) *of item* (ii) *are satisfied.*

Proof. $(iii) \Rightarrow (i)$. The Two-Series-Theorem implies that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n^c \text{ converges}\right) = 1.$$

Consider $\eta_n := \xi_n - \xi_n^c$ and $B_n := \{\omega \in \Omega : \eta_n(\omega) \neq 0\}$. Then

$$\sum_{n=1}^{\infty} \mathbb{P}(B_n) = \sum_{n=1}^{\infty} \mathbb{P}\left(|\xi_n| > c\right) < \infty.$$

The lemma of BOREL-CANTELLI implies that

$$\mathbb{P}\left(\{\omega \in \Omega : \#\{n : \omega \in B_n\} = \infty\}\right) = 0.$$

This implies that $\mathbb{P}\left(\sum_{n=1}^{\infty} \eta_n \text{ converges }\right) = 1$ so that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \text{ converges}\right) = \mathbb{P}\left(\sum_{n=1}^{\infty} [\eta_n + \xi_n^c] \text{ converges}\right) = 1.$$

(ii) $\Longrightarrow$ (iii) is trivial.

(i) $\Longrightarrow$ (ii) The almost sure convergence of $\sum_{n=1}^{\infty} \xi_n(\omega)$ implies that $\lim_n |\xi_n(\omega)| = 0$ a.s. so that

$$\mathbb{P}\left(\limsup_n \{\omega \in \Omega : |\xi_n(\omega)| \geq c\}\right) = 0.$$

The Lemma of BOREL-CANTELLI (note that the random variables ξ_n are independent) gives that

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|\xi_n| \geq c\right) < \infty$$

so that we obtain condition (c). Next, the almost sure convergence of $\sum_{n=1}^{\infty} \xi_n(\omega)$ implies the almost sure convergence of $\sum_{n=1}^{\infty} \xi_n^c(\omega)$ since $\xi_n(\omega) = \xi_n^c(\omega)$ for $n \geq n(\omega)$ for almost all $\omega \in \Omega$. Hence we can apply the Two-Series-Theorem Proposition 2.2.1 to obtain items (a) and (b). □

The following examples demonstrate the basic usage of the Three-Series-Theorem: it shows that not the size of the large values of the random variables is sometimes important, but only their probability.

Example 2.2.4. *Assume independent random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ such that*

$$\mathbb{P}(\xi_n = \alpha_n) = \mathbb{P}(\xi_n = -\alpha_n) = p_n \in \left(0, \frac{1}{2}\right)$$

with $\sum_{n=1}^{\infty} p_n < \infty$, $\mathbb{P}(\xi_n = 0) = 1 - 2p_n$, and $\alpha_n \geq 0$. Then one has that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \xi_n \text{ converges}\right) = 1.$$

Proof. For $c = 1$ we apply the Three-Series-Theorem. By symmetry we have that

$$\sum_{n=1}^{\infty} \mathbb{E}\xi_n^1 = \sum_{n=1}^{\infty} 0 = 0$$

so that item (a) follows. Moreover,

$$\sum_{n=1}^{\infty} \mathbb{E}(\xi_n^1 - \mathbb{E}\xi_n^1)^2 = \sum_{n=1}^{\infty} \mathbb{E}(\xi_n^1)^2 \leq 2 \sum_{n=1}^{\infty} p_n < \infty$$

and

$$\sum_{n=1}^{\infty} \mathbb{P}(|\xi_n| \geq 1) \leq 2 \sum_{n=1}^{\infty} p_n < \infty$$

so that items (b) and (c) follow as well. □

Next we deduce a variant of the STRONG LAW OF LARGE NUMBERS:

Proposition 2.2.5. *Let* $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ *be a sequence of independent random variables and* $\beta_n > 0$ *such that*

(i) $\beta_n \uparrow \infty$,

(ii) $\mathbb{E}\xi_n = 0$,

(iii) $\sum_{n=1}^{\infty} \frac{\mathbb{E}\xi_n^2}{\beta_n^2} < \infty$.

Then one has that

$$\mathbb{P}\left(\frac{1}{\beta_n}(\xi_1 + \cdots + \xi_n) \to_n 0\right) = 1.$$

For the proof we use two analytical lemmata.

Lemma 2.2.6 (TÖPLITZ). [6] *Let* $\alpha_n \geq 0$, $\beta_n := \alpha_1 + \cdots + \alpha_n > 0$, $\beta_n \uparrow \infty$, *and* $(x_n)_{n=1}^{\infty} \subseteq \mathbb{R}$ *with* $\lim_n x_n = x \in \mathbb{R}$. *Then one has that*

$$\frac{1}{\beta_n}(\alpha_1 x_1 + \cdots + \alpha_n x_n) \to_n x.$$

[6]Otto Töplitz 01/08/1881 (Breslau, Germany) - 15/02/1940 (Jerusalem), worked on infinite linear and quadratic forms.

Proof. Let $\varepsilon > 0$ and find some n_0 such that $|x_n - x| < \varepsilon$ for $n \geq n_0$. Then, for $n > n_0$,

$$\begin{aligned}\left|\frac{1}{\beta_n}\sum_{i=1}^{n}\alpha_i x_i - x\right| &= \left|\frac{1}{\beta_n}\sum_{i=1}^{n}\alpha_i x_i - \frac{1}{\beta_n}\sum_{i=1}^{n}\alpha_i x\right| \\ &\leq \frac{1}{\beta_n}\sum_{i=1}^{n}\alpha_i|x_i - x| \\ &= \frac{1}{\beta_n}\sum_{i=1}^{n_0}\alpha_i|x_i - x| + \frac{1}{\beta_n}\sum_{i=n_0+1}^{n}\alpha_i|x_i - x| \\ &\leq \frac{1}{\beta_n}\sum_{i=1}^{n_0}\alpha_i|x_i - x| + \varepsilon.\end{aligned}$$

Since $\beta_n \uparrow \infty$ there is some $n_1 > n_0$ such that for all $n \geq n_1$ one has that

$$\frac{1}{\beta_n}\sum_{i=1}^{n_0}\alpha_i|x_i - x| \leq \varepsilon$$

which implies that

$$\left|\frac{1}{\beta_n}\sum_{i=1}^{n}\alpha_i x_i - x\right| \leq 2\varepsilon$$

for $n \geq n_1$. □

Lemma 2.2.7 (KRONECKER). [7] *Let $\beta_n > 0$, $\beta_n \uparrow \infty$, and $(x_n)_{n=1}^{\infty} \subseteq \mathbb{R}$ such that $\sum_{n=1}^{\infty} x_n$ exists. Then*

$$\frac{1}{\beta_n}\sum_{i=1}^{n}\beta_i x_i \to_n 0.$$

[7] Leopold Kronecker, 7/12/1823 (Liegnitz, Prussia ; now Legnica, Poland)- 29/12/1891 (Berlin, Germany), major contributions in elliptic functions and the theory of algebraic numbers.

Proof. Let $\beta_0 = S_0 = 0$ and $S_n := x_1 + \cdots + x_n$. Then

$$\begin{aligned}\frac{1}{\beta_n}\sum_{i=1}^n \beta_i x_i &= \frac{1}{\beta_n}\sum_{i=1}^n \beta_i (S_i - S_{i-1}) \\ &= \frac{1}{\beta_n}\left(\beta_n S_n - \beta_0 S_0 - \sum_{i=1}^n S_{i-1}(\beta_i - \beta_{i-1})\right) \\ &= S_n - \frac{\beta_0 S_0}{\beta_n} - \frac{1}{\beta_n}\sum_{i=1}^n S_{i-1}(\beta_i - \beta_{i-1}).\end{aligned}$$

Observing that $\lim_n S_n = \sum_{n=1}^\infty x_n \in \mathbb{R}$, $\beta_0 S_0/\beta_n \to 0$ since $\beta_n \uparrow \infty$, and

$$\lim_n \frac{1}{\beta_n}\sum_{i=1}^n S_{i-1}(\beta_i - \beta_{i-1}) = \lim_n S_n = \sum_{n=1}^\infty x_n$$

by Lemma 2.2.6 we finally obtain that

$$\lim_n \frac{1}{\beta_n}\sum_{i=1}^n \beta_i x_i = 0.$$

□

Proof of Proposition 2.2.5. From the Two-Series-Theorem we know that

$$\mathbb{P}\left(\sum_{n=1}^\infty \frac{\xi_n}{\beta_n} \text{ converges}\right) = 1.$$

Hence Lemma 2.2.7 gives that

$$\begin{aligned}1 &= \mathbb{P}\left(\lim_n \frac{1}{\beta_n}\sum_{i=1}^n \beta_i \frac{\xi_i}{\beta_i} = 0\right) \\ &= \mathbb{P}\left(\lim_n \frac{1}{\beta_n}\sum_{i=1}^n \xi_i = 0\right).\end{aligned}$$

□

The STRONG LAW OF LARGE NUMBERS (SLLN) in a more standard form can be deduced from Proposition 2.2.5.

Proposition 2.2.8 (Strong law of large numbers)**.** *Let $\eta_1, \eta_2, ... : \Omega \to \mathbb{R}$ be a sequence of independent random variables such that*

(i) $m = \mathbb{E}\eta_1 = \mathbb{E}\eta_2 = \cdots$,

(ii) $\sum_{n=1}^{\infty} \frac{\mathbb{E}(\eta_n - \mathbb{E}\eta_n)^2}{n^2} < \infty$.

Then one has that

$$\mathbb{P}\left(\frac{1}{n}(\eta_1 + \cdots + \eta_n) \to_n m\right) = 1.$$

Proof. Apply Proposition 2.2.5 to $\beta_n = n$ and $\xi_n := \eta_n - m$ so that

$$\frac{1}{n}(\eta_1(\omega) + \cdots + \eta_n(\omega)) = \frac{1}{\beta_n}(\xi_1(\omega) + \cdots + \xi_n(\omega)) + m$$

and $\sum_{n=1}^{\infty} 1/\beta_n^2 < \infty$. □

An example is the following:

Example 2.2.9. *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be independent random variables such that*

$$\mathbb{P}(\xi_n = 0) = \mathbb{P}(\xi_n = 1) = \frac{1}{2}.$$

Then

$$\lim_n \frac{1}{n} \sum_{i=1}^{n} \xi_i = \mathbb{E}\xi_1 = \frac{1}{2} \; a.s.$$

The proof is an immediate consequence of Proposition 2.2.8. A possible interpretation is the following: We let $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1), \mathcal{B}([0,1)), \lambda)$ where λ is the LEBESGUE [8] -measure. Then writing every number $x \in [0,1)$ as dyadic expansion $0, \xi_1\xi_2\xi_3 \cdots$ where we do not allow an infinite series of digits 1. Then considering $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ we get random variables which satisfy the assumptions of Example 2.2.9. The assertion of this proposition says that zeros and ones are equally likely for the dyadic representation of a number $x \in [0,1)$.

[8]Henri Léon Lebesgue, 28/06/1875 (Beauvais, Oise, Picardie, France) - 26/07/1941 (Paris, France), French mathematician, generalized the Riemann integral to the Lebesgue integral, continuation of work of Emile Borel and Camille Jordan.

Earlier we have shown for independent BERNOULLI random variables $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ that

$$\mathbb{P}\left(\limsup_n \frac{f_n}{\sqrt{n}} = \infty, \quad \liminf_n \frac{f_n}{\sqrt{n}} = -\infty\right) = 1$$

where $f_n := \varepsilon_1 + \cdots + \varepsilon_n$. As an application of Proposition 2.2.5 we show an opposite statement now.

Proposition 2.2.10. *Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ be independent* BERNOULLI *random variables, $f_n := \varepsilon_1 + \cdots + \varepsilon_n$, and $\varepsilon > 0$. Then*

$$\mathbb{P}\left(\lim_n \frac{f_n}{\sqrt{n}(\log(n+1))^{\frac{1}{2}+\varepsilon}} = 0\right) = 1.$$

Proof. Let $\beta_n := \sqrt{n}(\log(n+1))^{\frac{1}{2}+\varepsilon} \uparrow_n \infty$ so that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\varepsilon_n^2}{n(\log(n+1))^{1+2\varepsilon}} = \sum_{n=1}^{\infty} \frac{1}{n(\log(n+1))^{1+2\varepsilon}} < \infty.$$

Applying Proposition 2.2.5 gives that

$$\mathbb{P}\left(\lim_n \frac{f_n}{\sqrt{n}(\log(n+1))^{\frac{1}{2}+\varepsilon}} = 0\right) = \mathbb{P}\left(\lim_n \frac{f_n}{\beta_n} = 0\right) = 1.$$

□

2.2.2 Proof of the Two-Series-Theorem

Now we are going to prove the Two-Series-Theorem. The first inequality, we need, is a deviation inequality due to KOLMOGOROV. Before we state and prove this inequality, let us motivate it. Assume a random variable $f : \Omega \to \mathbb{R}$ with $\mathbb{E}f^2 < \infty$. Then, by CHEBYSHEV [9] 's inequality

$$\varepsilon^2\, \mathbb{P}\left(|f| \geq \varepsilon\right) = \varepsilon^2\, \mathbb{P}\left(|f|^2 \geq \varepsilon^2\right) \leq \int_\Omega |f|^2\, d\,\mathbb{P}$$

[9] Pafnuty Lvovich Chebyshev, 16/05/1821 (Okatovo, Russia) - 08/12/1894 (St Petersburg, Russia), number theory, mechanics, famous for the orthogonal polynomials he investigated.

for $\varepsilon > 0$ so that

$$\mathbb{P}\left(|f| \ \geq \varepsilon\right) \ \leq \frac{\mathbb{E}\,|f|^2}{\varepsilon^2}.$$

Assuming independent random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ and $f_n := \xi_1 + \cdots + \xi_n$ this gives that

$$\mathbb{P}\left(|f_n| \geq \varepsilon\right) \ \leq \frac{\mathbb{E}|f_n|^2}{\varepsilon^2}.$$

Now one can enlarge the left-hand side by replacing $|f_n|$ by $\sup_{k=1,\dots,n} |f_k|$ so that we get a *maximal inequality*.

Lemma 2.2.11 (Inequality of KOLMOGOROV)**.** *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be a sequence of independent random variables such that*

(i) $\mathbb{E}\xi_n = 0$ *for* $n = 1, 2, ...,$

(ii) $\mathbb{E}\xi_n^2 < \infty$ *for* $n = 1, 2, ...$

If $f_n := \xi_1 + \cdots + \xi_n$, $\varepsilon > 0$, and $n = 1, 2, ...,$ then one has that

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |f_k| \geq \ \varepsilon\right) \leq \frac{\mathbb{E}f_n^2}{\varepsilon^2}.$$

Proof. Let $n = 1, 2, \dots$ be fixed, $A := \{\max_{1 \leq k \leq n} |f_k| \geq \varepsilon\}$, and

$$B_k := \{|f_1| < \ \varepsilon, \dots, |f_{k-1}| < \ \varepsilon, |f_k| \geq \varepsilon\}$$

for $k \geq 2$ and $B_1 := \{|f_1| \geq \varepsilon\}$. We get a disjoint union $A = \bigcup_{k=1}^n B_k$. Now

$$\begin{aligned}
\varepsilon^2 \mathbb{P}(A) &= \varepsilon^2 \, \mathbb{P}\left(\bigcup_{k=1}^n B_k\right) \\
&= \sum_{k=1}^n \varepsilon^2 \, \mathbb{P}(B_k) = \sum_{k=1}^n \int_{B_k} \varepsilon^2 d\, \mathbb{P} \\
&\leq \sum_{k=1}^n \int_{B_k} f_k^2 d\, \mathbb{P} \leq \sum_{k=1}^n \int_{B_k} f_n^2 d\, \mathbb{P} \\
&\leq \mathbb{E}\, f_n^2,
\end{aligned}$$

which proves our assertion, where the second to the last inequality can be verified as follows:

$$\begin{aligned}\int_{B_k} f_n^2 d\,\mathbb{P} &= \int_{B_k} (f_k + \xi_{k+1} + \cdots + \xi_n)^2\, d\,\mathbb{P} \\ &= \int_{B_k} f_k^2 d\,\mathbb{P} + 2\int_{B_k} f_k\,(\xi_{k+1} + \cdots + \xi_n)\, d\,\mathbb{P} \\ &\quad + \int_{B_k} (\xi_{k+1} + \cdots + \xi_n)^2\, d\,\mathbb{P} \\ &= \int_{B_k} f_k^2 d\,\mathbb{P} + \int_{B_k} (\xi_{k+1} + \cdots + \xi_n)^2\, d\,\mathbb{P}\end{aligned}$$

since, by independence,

$$\begin{aligned}\int_{B_k} f_k\,(\xi_{k+1} + \cdots + \xi_n)\, d\,\mathbb{P} &= \int_{\Omega} (f_k \chi_{B_k})\,(\xi_{k+1} + \cdots + \xi_n)\, d\mathbb{P} \\ &= \left(\int_{\Omega} f_k \chi_{B_k} d\mathbb{P}\right)\left(\int_{\Omega} (\xi_{k+1} + \cdots + \xi_n)\, d\mathbb{P}\right) \\ &= 0.\end{aligned}$$

□

Remark 2.2.12. There are other inequalities, called LÉVY [10] -OCTAVIANI inequalities: assuming independent random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ and $\varepsilon > 0$ one has

(i) $\mathbb{P}\left(\max_{1 \le k \le n} |f_k| > \varepsilon\right) \le 3 \max_{1 \le k \le n} \mathbb{P}\left(|f_k| > \frac{\varepsilon}{3}\right)$,

(ii) $\mathbb{P}\left(\max_{1 \le k \le n} |f_k| > \varepsilon\right) \le 2\mathbb{P}\left(|f_n| > \varepsilon\right)$ if the sequence $(\xi_n)_{n=1}^{\infty}$ is additionally symmetric that means that for all signs $\theta_1, ..., \theta_n \in \{-1, 1\}$ the distributions of the vectors $(\xi_1, ..., \xi_n)$ and $(\theta_1\xi_1, ..., \theta_n\xi_n)$ are the same.

Example 2.2.13. For BERNOULLI variables $(\varepsilon_n)_{n=1}^{\infty}$ it follows from Lemma 2.2.11 that

$$\mathbb{P}\left(\max_{1 \le k \le n} |\varepsilon_1 + \cdots + \varepsilon_k| \ge \varepsilon\right) \le \frac{\mathbb{E} f_n^2}{\varepsilon^2} = \frac{n}{\varepsilon^2}.$$

[10] Paul Pierre Lévy, 15/09/1886 (Paris, France) - 15/12/1971 (Paris, France), influenced greatly probability theory, also worked in functional analysis and partial differential equations.

Letting $\varepsilon = \theta\sqrt{n}$, $\theta > 0$, this gives

$$\mathbb{P}\left(\max_{1\le k\le n} |\varepsilon_1 + \cdots + \varepsilon_k| \ge \theta\sqrt{n}\right) \le \frac{1}{\theta^2}.$$

The left-hand side describes the probability that the random walk exceeds $-\theta\sqrt{n}$ or $\theta\sqrt{n}$ *up* to step n (not *only* at step n).

Now we need the converse of the above inequality:

Lemma 2.2.14 (Converse inequality of KOLMOGOROV)**.** *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be a sequence of independent random variables such that*

(i) $\mathbb{E}\xi_n = 0$ *for* $n = 1, 2, ...$,

(ii) *there exists a constant $c > 0$ such that $|\xi_n(\omega)| \le c$ for all $\omega \in \Omega$ and* $n = 1, 2, ...$

If $f_n := \xi_1 + \cdots + \xi_n$, $\varepsilon > 0$, $n \in \{1, 2, ...\}$, and $\mathbb{E}f_n^2 > 0$, then

$$\mathbb{P}\left(\max_{1\le k\le n} |f_k| \ge \varepsilon\right) \ge 1 - \frac{(c+\varepsilon)^2}{\mathbb{E}f_n^2}.$$

The next lemma, we need, describes CAUCHY [11] sequences with respect to the convergence in probability.

Lemma 2.2.15. *Let $f_1, f_2, ... : \Omega \to \mathbb{R}$ be a sequence of random variables. Then the following conditions are equivalent:*

(i) $\mathbb{P}\left(\{\omega \in \Omega : (f_n(\omega))_{n=1}^\infty\right.$ *is a* CAUCHY *sequence* $\left.\}\right) = 1$,

(ii) *For all $\varepsilon > 0$ one has that*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{k,l\ge n} |f_k - f_l| \ge \varepsilon\right) = 0.$$

(iii) *For all $\varepsilon > 0$ one has that*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{k\ge n} |f_k - f_n| \ge \varepsilon\right) = 0.$$

[11] Augustin Louis Cauchy, 21/08/1789 (Paris, France)- 23/05/1857 (Sceaux, France), study of real and complex analysis, theory of permutation groups.

Proof. (ii)$\Longleftrightarrow$ (iii) follows from

$$\begin{aligned}\sup_{k\geq n}|f_k - f_n| &\leq \sup_{k,l\geq n}|f_k - f_l| \\ &\leq \sup_{k\geq n}|f_k - f_n| + \sup_{l\geq n}|f_n - f_l| \\ &= 2\sup_{k\geq n}|f_k - f_n|.\end{aligned}$$

(i)$\Longleftrightarrow$ (ii) Let

$$A := \{\omega \in \Omega : (f_n(\omega))_{n=1}^{\infty} \text{ is a Cauchy sequence }\}.$$

Then we get that

$$A = \bigcap_{N=1,2,...}\bigcup_{n=1,2,...}\bigcap_{k>l\geq n}\left\{\omega \in \Omega : |f_k(\omega) - f_l(\omega)| \leq \frac{1}{N}\right\}.$$

Consequently, we have that $\mathbb{P}(A) = 1$ if and only if

$$\mathbb{P}\left(\bigcup_{n=1,2,...}\bigcap_{k>l\geq n}\left\{\omega \in \Omega : |f_k(\omega) - f_l(\omega)| \leq \frac{1}{N}\right\}\right) = 1$$

for all $N = 1, 2, ...$, if and only if

$$\lim_{n\to\infty}\mathbb{P}\left(\bigcap_{k>l\geq n}\left\{\omega \in \Omega : |f_k(\omega) - f_l(\omega)| \leq \frac{1}{N}\right\}\right) = 1$$

for all $N = 1, 2, ...$, if and only if

$$\lim_{n\to\infty}\mathbb{P}\left(\bigcup_{k>l\geq n}\left\{\omega \in \Omega : |f_k(\omega) - f_l(\omega)| > \frac{1}{N}\right\}\right) = 0$$

for all $N = 1, 2, ...$ We can finish the proof by remarking that

$$\bigcup_{k>l\geq n}\left\{\omega \in \Omega : |f_k(\omega) - f_l(\omega)| > \frac{1}{N}\right\} = \left\{\sup_{k>l\geq n}|f_k(\omega) - f_l(\omega)| > \frac{1}{N}\right\}.$$

$\square$

Proof of Proposition 2.2.1. (i) We let $\eta_n := \xi_n - \mathbb{E}\xi_n$ so that $\mathbb{E}\eta_n = 0$ and

$$\sum_{n=1}^{\infty} \mathbb{E}\eta_n^2 = \sum_{n=1}^{\infty} \mathbb{E}(\xi_n - \mathbb{E}\xi_n)^2 < \infty.$$

Since $\sum_{n=1}^{\infty} \mathbb{E}\xi_n$ converges, the convergence of $\sum_{n=1}^{\infty} \eta_n(\omega)$ implies the convergence of $\sum_{n=1}^{\infty} \xi_n(\omega) = \sum_{n=1}^{\infty}(\eta_n(\omega) + \mathbb{E}\xi_n)$. Hence it is sufficient to show that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \eta_n \text{ converges}\right) = 1.$$

Let $g_n := \eta_1 + \cdots + \eta_n$. Applying Lemma 2.2.15 we see that it is enough to prove

$$\lim_n \mathbb{P}\left(\sup_{k \geq n} |g_k - g_n| \geq \varepsilon\right) = 0$$

for all $\varepsilon > 0$. But this follows from

$$\begin{aligned}
\mathbb{P}\left(\sup_{k \geq n} |g_k - g_n| \geq \varepsilon\right) &= \lim_{N \to \infty} \mathbb{P}\left(\sup_{n \leq k \leq n+N} |g_k - g_n| \geq \varepsilon\right) \\
&\leq \lim_{N \to \infty} \frac{\mathbb{E}(g_{n+N} - g_n)^2}{\varepsilon^2} \\
&= \lim_{N \to \infty} \frac{\sum_{k=1}^{N} \mathbb{E}\eta_{n+k}^2}{\varepsilon^2} \\
&= \frac{\sum_{l=n+1}^{\infty} \mathbb{E}\eta_l^2}{\varepsilon^2}
\end{aligned}$$

where we have used the KOLMOGOROV inequality Lemma 2.2.11. Obviously, the last term converges to zero as $n \to \infty$.

(ii) Because of step (i) we only have to prove that $(a) \Rightarrow (b)$. We use again a symmetrization argument and consider a new sequence $\xi_1', \xi_2', ... : \Omega' \to \mathbb{R}$ of independent random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ having the same distribution as the original sequence $\xi_1, \xi_2, ...$, that means

$$\mathbb{P}(\xi_n \leq \lambda) = \mathbb{P}'(\xi_n' \leq \lambda)$$

for all $n = 1, 2, ...$ and all $\lambda \in \mathbb{R}$. We also may assume that $|\xi_n'(\omega')| \leq d$ for all $\omega' \in \Omega'$ and $n = 1, 2, ...$ Taking the product space $(M, \Sigma, \mu) =$

$(\Omega, \mathcal{F}, \mathbb{P}) \times (\Omega', \mathcal{F}', \mathbb{P}')$ we may consider $\xi_n, \xi_n' : M \to \mathbb{R}$ with the convention that $\xi_n(\omega, \omega') = \xi_n(\omega)$ and $\xi_n'(\omega, \omega') = \xi_n'(\omega')$. Now we let

$$\eta_n(\omega, \omega') = \xi_n(\omega, \omega') - \xi_n'(\omega, \omega')$$

and get

$$\begin{aligned} \mathbb{E}_\mu \eta_n &= \mathbb{E}\xi_n - \mathbb{E}\xi_n' = 0, \\ |\eta_n(\omega, \omega')| &\le |\xi_n(\omega, \omega')| + |\xi_n'(\omega, \omega')| \le 2d, \end{aligned}$$

and

$$\begin{aligned} & \mu\left(\left\{(\omega, \omega') \in M : \sum_{n=1}^{\infty} \eta_n(\omega, \omega') \text{ converges}\right\}\right) \\ = \; & \mathbb{P} \times \mathbb{P}'\left(\left\{(\omega, \omega') \in \Omega \times \Omega' : \sum_{n=1}^{\infty} (\xi_n(\omega) - \xi_n'(\omega')) \text{ converges}\right\}\right) \\ = \; & 1. \end{aligned}$$

Letting $g_n := \eta_1 + \cdots + \eta_n$ and $\varepsilon > 0$, Lemma 2.2.15 implies that there is an $n \in \{1, 2, ...\}$ such that

$$\mu\left(\sup_{k \ge n} |g_k - g_n| \ge \varepsilon\right) < \frac{1}{2}.$$

Exploiting Lemma 2.2.14 gives, for $N \ge 1$, that

$$1 - \frac{(2d + \varepsilon)^2}{\sum_{k=n+1}^{n+N} \mathbb{E}\eta_k^2} = 1 - \frac{(2d + \varepsilon)^2}{\mathbb{E}(g_{n+N} - g_n)^2} \le \mu\left(\sup_{k=n,\dots,n+N} |g_k - g_n| \ge \varepsilon\right) < \frac{1}{2}.$$

But from this it follows that

$$\sum_{k=n+1}^{n+N} \mathbb{E}\eta_k^2 < 2(2d + \varepsilon)^2 < \infty$$

for all N, so that

$$\sum_{k=1}^{\infty} \mathbb{E}\eta_k^2 < \infty.$$

This gives that

$$\sum_{n=1}^{\infty} \mathbb{E}(\xi_n - \mathbb{E}\xi_n)^2 = \frac{1}{2}\sum_{n=1}^{\infty} \mathbb{E}([\xi_n - \mathbb{E}\xi_n] - [\xi_n' - \mathbb{E}\xi_n'])^2 = \sum_{n=1}^{\infty} \mathbb{E}\eta_n^2 < \infty.$$

It remains to show that $\sum_{n=1}^{\infty} \mathbb{E}\xi_n$ exists. Since $\sum_{n=1}^{\infty} \xi_n$ converges almost surely and $\sum_{n=1}^{\infty}(\xi_n - \mathbb{E}\xi_n)$ converges almost surely because of step (i) and $\sum_{n=1}^{\infty} \mathbb{E}(\xi_n - \mathbb{E}\xi_n)^2 < \infty$ proved right now, we have to have that $\sum_{n=1}^{\infty} \mathbb{E}\xi_n$ converges as well. □

2.3 The law of iterated logarithm

From the previous sections we got two bounds for the random walk

$$f_n(\omega) := \varepsilon_1(\omega) + \cdots + \varepsilon_n(\omega)$$

where $\varepsilon_1, ..., \varepsilon_n$ are independent random variables such that $\mathbb{P}(\varepsilon_n = 1) = \mathbb{P}(\varepsilon_n = -1) = 1/2$, namely

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{f_n}{\sqrt{n}} = \infty, \quad \liminf_{n\to\infty} \frac{f_n}{\sqrt{n}} = -\infty\right) = 1$$

from Corollary 2.1.10 and

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{f_n}{\sqrt{n}(\log(n+1))^{\frac{1}{2}+\varepsilon}} = 0\right) = 1$$

from Proposition 2.2.10. But what is the right scaling factor between $\sqrt{n}$ and $\sqrt{n}\log(n+1)$? The answer is given by the famous LAW OF ITERATED LOGARITHM:

Proposition 2.3.1 (Law of iterated logarithm)**.** *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be a sequence of independent and identically distributed random variables such that $\mathbb{E}\xi_n = 0$ and $\mathbb{E}\xi_n^2 = \sigma^2 \in (0, \infty)$. Then, for $f_n := \xi_1 + \cdots + \xi_n$, one has that*

$$\mathbb{P}\left(\limsup_{\binom{n\to\infty}{n\geq 3}} \frac{f_n}{\psi(n)} = 1, \quad \liminf_{\binom{n\to\infty}{n\geq 3}} \frac{f_n}{\psi(n)} = -1\right) = 1$$

with $\psi(n) := \sqrt{2\sigma^2 n \log\log n}$.

Remark 2.3.2. (i) The conditions

$$\limsup_{\substack{n\to\infty\\ n\geq 3}} \frac{f_n}{\psi(n)} = 1 \text{ a.s.} \quad \text{and} \quad \liminf_{\substack{n\to\infty\\ n\geq 3}} \frac{f_n}{\psi(n)} = -1 \text{ a.s.}$$

are equivalent since one may consider the random variables $(-\xi_n)_{n=1}^\infty$ which satisfy the assumptions of the (LIL) as well.

(ii) The statement can be reformulated in terms of the two conditions

$$\begin{aligned}\mathbb{P}\left(\#\left\{n\geq 3 : f_n \geq (1-\varepsilon)\psi(n)\right\} = \infty\right) &= 1,\\ \mathbb{P}\left(\#\left\{n\geq 3 : f_n \geq (1+\varepsilon)\psi(n)\right\} = \infty\right) &= 0\end{aligned}$$

for all $\varepsilon \in (0,1)$.

(iii) KHINTCHINE proved the (LIL) in 1924 in the case that $|\xi_n(\omega)| \leq c$. Later on, KOLMOGOROV extended the law to other random variables in 1929. Finally, the above version was proved by WIENER [12] and HARTMAN in 1941.

We will give the idea of the proof of the (LIL) in the case that $\xi_n = g_n \sim N(0,1)$. Let us start with an estimate for the distribution of g_n.

Lemma 2.3.3. *For $g \sim N(0,\sigma^2)$ with $\sigma > 0$ one has that*

$$\lim_{\substack{\lambda\to\infty\\ \lambda>0}} \frac{\mathbb{P}(g>\lambda)}{\frac{\sigma}{\sqrt{2\pi}\lambda} e^{-\frac{\lambda^2}{2\sigma^2}}} = 1.$$

Proof. We get that

$$\begin{aligned}\lim_{\substack{\lambda\to\infty\\ \lambda>0}} \frac{\mathbb{P}(g>\lambda)}{\frac{\sigma}{\sqrt{2\pi}\lambda} e^{-\frac{\lambda^2}{2\sigma^2}}} &= \lim_{\substack{\lambda\to\infty\\ \lambda>0}} \frac{\int_\lambda^\infty e^{-\frac{y^2}{2\sigma^2}} \frac{dy}{\sqrt{2\pi}\sigma}}{\frac{\sigma}{\sqrt{2\pi}\lambda} e^{-\frac{\lambda^2}{2\sigma^2}}}\\ &= \lim_{\substack{\lambda\to\infty\\ \lambda>0}} \frac{e^{-\frac{\lambda^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma}}{e^{-\frac{\lambda^2}{2\sigma^2}} \left(\frac{\sigma}{\sqrt{2\pi}\lambda} + \frac{\sigma\lambda}{\sqrt{2\pi}\lambda\sigma^2}\right)}\\ &= 1\end{aligned}$$

[12]Norbert Wiener, 26/11/1894 (Columbia, USA) - 18/03/1964 (Stockholm, Sweden), worked on Brownian motion, from where he progressed to harmonic analysis, and won the Bôcher prize from his studies on Tauberian theorems.

where we apply the rule of L'HOSPITAL [13]. ☐

Next we need again a maximal inequality similar to one of that already mentioned in Remark 2.2.12.

Lemma 2.3.4. *Let $\xi_1, ..., \xi_n : \Omega \to \mathbb{R}$ be independent random variables which are symmetric, that means*

$$\mathbb{P}(\xi_k \leq \lambda) = \mathbb{P}(-\xi_k \leq \lambda)$$

for all $k = 1, ..., n$ and $\lambda \in \mathbb{R}$. Then one has that

$$\mathbb{P}\left(\max_{k=1,\ldots,n} f_k > \varepsilon\right) \leq 2\mathbb{P}(f_n > \varepsilon)$$

where $f_k := \xi_1 + \cdots + \xi_k$ and $\varepsilon > 0$.

Proof. Let $1 \leq k \leq n$. Because of $\mathbb{P}(f_n - f_k > 0) = \mathbb{P}(f_n - f_k < 0)$ and

$$1 = \mathbb{P}(f_n - f_k > 0) + \mathbb{P}(f_n - f_k < 0) + \mathbb{P}(f_n - f_k = 0)$$

we get that $\mathbb{P}(f_n - f_k \geq 0) \geq 1/2$. Now, again as in the proof of KOLMOGOROV's maximal inequality, let $B_1 := \{f_1 > \varepsilon\}$ and

$$B_k := \{f_1 \leq \varepsilon, ..., f_{k-1} \leq \varepsilon, f_k > \varepsilon\}$$

for $k = 2, ..., n$. Then we get that, where $0/0 := 1$,

$$\begin{aligned} \mathbb{P}\left(\max_{k=1,\ldots,n} f_k > \varepsilon\right) &= \sum_{k=1}^{n} \mathbb{P}(B_k) \\ &= \sum_{k=1}^{n} \mathbb{P}(B_k) \frac{\mathbb{P}(f_n \geq f_k)}{\mathbb{P}(f_n \geq f_k)} \\ &= \sum_{k=1}^{n} \frac{\mathbb{P}(B_k \cap \{f_n \geq f_k\})}{\mathbb{P}(f_n \geq f_k)} \end{aligned}$$

[13] Guillaume Francois Antoine Marquis de L'Hôpital, 1661 (Paris, France)- 2/2/1704 (Paris, France), French mathematician who wrote the first textbook on calculus, which consisted of the lectures of his teacher Johann Bernoulli.

where we used the independence of B_k and $f_n - f_k = \xi_{k+1} + \cdots + \xi_n$ if $k < n$. Using that

$$\frac{1}{\mathbb{P}(f_n \geq f_k)} \leq 2$$

we end up with

$$\begin{aligned}\mathbb{P}\left(\max_{k=1,\ldots,n} f_k > \varepsilon\right) &\leq 2\sum_{k=1}^{n} \mathbb{P}(B_k \cap \{f_n \geq f_k\}) \\ &\leq 2\mathbb{P}(f_n > \varepsilon).\end{aligned}$$

□

Proof of Proposition 2.3.1 for $\xi_n = g_n$. By symmetry we only need to show that

$$\mathbb{P}\left(\limsup_{\substack{n\to\infty \\ n\geq 3}} \frac{f_n}{\psi(n)} \leq 1\right) = 1 \quad \text{and} \quad \mathbb{P}\left(\limsup_{\substack{n\to\infty \\ n\geq 3}} \frac{f_n}{\psi(n)} \geq 1\right) = 1.$$

This is equivalent that for all $\varepsilon \in (0,1)$ one has that

$$\mathbb{P}\Big(\big\{\omega \in \Omega : \exists n_0 \geq 3 \forall\ n \geq n_0\ \ f_n(\omega) \leq (1+\varepsilon)\psi(n)\big\}\Big) = 1 \qquad (2.2)$$

and

$$\mathbb{P}\Big(\big\{\omega \in \Omega : \#\left\{n \geq 3\ \ f_n(\omega) \geq (1-\varepsilon)\psi(n)\right\}\big\} = \infty\Big) = 1. \qquad (2.3)$$

First we turn to (2.2): let $\lambda := 1 + \varepsilon$ and $n_k := \lambda^k$ for $k \geq k_0$ such that $n_{k_0} \geq 3$. Define

$$A_k := \{\omega \in \Omega : \exists n \in (n_k, n_{k+1}]\ \ f_n(\omega) > \lambda\psi(n)\}$$

for $k \geq k_0$. Then $\mathbb{P}(\limsup_k A_k) = 0$ would imply (2.2). According to the Lemma of BOREL-CANTELLI it is sufficient to show that $\sum_{k=k_0}^{\infty} \mathbb{P}(A_k) < \infty$. This follows from

$$\begin{aligned}\mathbb{P}(A_k) &\leq \mathbb{P}\left(\exists n \in [1, n_{k+1}] : f_n > \lambda\psi(n_k)\right) \\ &\leq 2\mathbb{P}\left(f_{[n_{k+1}]} > \lambda\psi(n_k)\right) \\ &\leq 2c\frac{\sigma(f_{[n_{k+1}]})}{\sqrt{2\pi}(\lambda\psi(n_k))} e^{-\frac{(\lambda\psi(n_k))^2}{2\sigma(f_{[n_{k+1}]})^2}}\end{aligned}$$

where we have used the maximal inequality from Lemma 2.3.4 and the estimate from Lemma 2.3.3 and where $\sigma(f_n)^2$ is the variance of f_n. Finally, by $\sigma(f_n)^2 = n$ we get (after some computation) that

$$\sum_{k=k_0}^{\infty} \mathbb{P}(A_k) \leq \sum_{k=k_0}^{\infty} c_1 k^{-\lambda} < \infty.$$

Now we give the idea for the second equation (2.3). Applying (2.2) to $(-f_n)_{n=1}^{\infty}$ and $\varepsilon = 1$ gives that

$$\mathbb{P}\Big(\big\{\omega \in \Omega : \exists n_0 \geq 3 \forall \;\; n \geq n_0 \quad - f_n(\omega) \leq 2\psi(n)\big\}\Big) = 1. \tag{2.4}$$

We set $n_k := N^k$ for some $N > 1$ and put $Y_k := f_{[n_k]} - f_{[n_{k-1}]}$ for $k > k_0$ and $\lambda := 1 - \varepsilon$ with $\varepsilon \in (0,1)$. Assume that we can show that

$$Y_k > \lambda\psi(n_k) + 2\psi(n_{k-1}) \tag{2.5}$$

happens infinitely often with probability one, which means by definition that

$$f_{[n_k]} - f_{[n_{k-1}]} > \lambda\psi(n_k) + 2\psi(n_{k-1})$$

happens infinitely often with probability one. Together with (2.4) this would imply that

$$f_{[n_k]} > \lambda\psi(n_k)$$

happens infinitely often with probability one. Hence we have to show (2.5). For this choose a $\lambda' \in (\lambda, 1)$ and notice that there exists an $N > 1$ such that for all k it holds

$$\begin{aligned} \lambda'[2(N^k - N^{k-1}) \log\log N^k]^{\frac{1}{2}} &> \lambda[2(N^k \log\log N^k]^{\frac{1}{2}} + 2[2N^{k-1} \log\log N^{k-1}]^{\frac{1}{2}} \\ &= \lambda\psi(N^k) + 2\psi(N^{k-1}) \end{aligned}$$

Hence, by Lemma 2.3.3 we get that

$$\begin{aligned} \mathbb{P}\left(Y_k > \lambda\psi(n_k) + 2\psi(n_{k-1})\right) &\geq \mathbb{P}\left(Y_k > \lambda'[2(N^k - N^{k-1}) \log\log N^k]^{\frac{1}{2}}\right) \\ &\geq \frac{c_2}{k \log k} \end{aligned}$$

so that

$$\sum_{k=k_0+1}^{\infty} \mathbb{P}\left(Y_k > \lambda\psi(n_k) + 2\psi(n_{k-1})\right) = \infty.$$

An application of the Lemma of BOREL-CANTELLI implies (2.5). □

Now we have also answered questions (Q5) and (Q7) from Section 1.1.

Chapter 3

Martingales in discrete time

First let us raise to natural questions related to martingales which we are able to answer later:

- **First run of three sixes:** A fair die is thrown independently each time. The gambler wins if 3 sixes consecutively appear. What is the mean number of throws until this happens?

- **A strategy to win without risk?** We interpret the simple symmetric random walk as a game. Let $(Y_n)_{n\in\mathbb{N}}$ be independent random variables with

  $$\mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = -1) = \frac{1}{2}.$$

  Then $X_0 := 0$ and $X_n := \sum_{k=1}^n Y_k$ describes the amount of money the gambler has at time n. Assume he can quit the game at any time he wishes and therefore he decides to quit the first time $X_n = 1$, which means he quits at the random time

  $$T(\omega) = \inf\{n \geq 0 : X_n(\omega) = 1\}.$$

  Since stopping at T leaves the gambler always with 1 unit while he started with 0 so it seems to be that this game is not a fair one. Perhaps from other courses ((X_n) is a recurrent Markov chain) it is known to the reader that $\mathbb{P}(T < \infty) = 1$. Can one compute the expected time $\mathbb{E}T$ to reach the state 1?

Originally martingales were intended to model *fair games*. Having this in mind, it is not surprising that martingale theory plays an important role in areas connected to stochastic modeling such as Stochastic Finance, Biology, and Physics. On the other hand side, martingale theory enters many branches of pure mathematics as well: for example it is a powerful tool in harmonic and functional analysis and gives new insight into certain phenomena.

In this lecture notes we want to develop some basic parts of the martingale theory. One can distinguish between two cases: discrete time and continuous time. The first case deals with sequences of random variables and is sometimes easier to treat. In the second case we have to work with families of random variables having uncountably many elements, which is technically partially more advanced, is however based on the discrete time case in many cases.

3.1 Some prerequisites

First we summarize some notation and basic facts used from now on.

The Lebesgue spaces. The Lebesgue spaces are named after HENRI LEBESGUE, 1875-1941.

Definition 3.1.1. (i) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $1 \leq p < \infty$. Then a random variable $f : \Omega \to \mathbb{R}$ belongs to $\mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ if and only if

$$\|f\|_{\mathcal{L}_p} = \|f\|_p := \left(\int_\Omega |f(\omega)|^p d\mathbb{P}(\omega) \right)^{\frac{1}{p}} < \infty.$$

(ii) The space of all equivalence classes of $\mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ if we identify f and g whenever $\mathbb{P}(f = g) = 1$ is denoted by $L_p(\Omega, \mathcal{F}, \mathbb{P})$ and equipped with $\|\hat{f}\|_{L_p} := \|f\|_{\mathcal{L}_p}$ where $f \in \hat{f}$.

Let us give some examples.

Example 3.1.2.

(a) Let $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1))$ and $\mathbb{P} = \lambda$ be the Lebesgue-measure, that means

$$\lambda([a, b)) = b - a$$

for $0 \leq a < b \leq 1$. Assume a continuous $f : [0,1) \to \mathbb{R}$. Then $f \in \mathcal{L}_p([0,1))$ if and only if

$$\|f\|_p = \left(\int_0^1 |f(x)|^p dx \right)^{\frac{1}{p}} < \infty$$

where we can take the RIEMANN [1] -integral. For example, for $\theta > 0$,

$$f(x) := \frac{1}{x^\theta} \in \mathcal{L}_p([0,1))$$

if and only if $p\theta < 1$.

(b) Let $\Omega = \{1, 2, 3, \ldots\}$, $\mathcal{F} := 2^\Omega$ (i.e. $\mathcal{F}$ is the set of all subsets of Ω), and $\mathbb{P}(\{k\}) = q_k$ with $\sum_{k=1}^\infty q_k = 1$ and $0 \leq q_k \leq 1$. Then $\mathbb{P}(A) = \sum_{k \in A} q_k$ and $f : \Omega \to \mathbb{R} \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ if and only if

$$\|f\|_p = \left(\int_\Omega |f(\omega)|^p d\mathbb{P}(\omega) \right)^{\frac{1}{p}} = \left(\sum_{k=1}^\infty |f(k)|^p q_k \right)^{\frac{1}{p}} < \infty.$$

Let us give some basic properties of the LEBESGUE spaces. Part of the proofs can be found in the script [1].

Proposition 3.1.3 (Basic properties of $\mathcal{L}_p$)**.**

(i) *If* $1 \leq p_1 \leq p_2 < \infty$, *then* $\mathcal{L}_{p_2}(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \mathcal{L}_{p_1}(\Omega, \mathcal{F}, \mathbb{P})$ *with* $\|f\|_{p_1} \leq \|f\|_{p_2}$ *for* $f \in \mathcal{L}_{p_1}(\Omega, \mathcal{F}, \mathbb{P})$.

(ii) MINKOWSKI [2] *'s inequality: if* $1 \leq p < \infty$, $\lambda \in \mathbb{R}$ *and* $f, g \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, *then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad \text{and} \quad \|\lambda f\|_p = |\lambda| \|f\|_p.$$

(iii) *Completeness: if* $1 \leq p < \infty$ *and* $(f_n)_{n=1}^\infty \subseteq \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ *such that for all* $\varepsilon > 0$ *there exists an* $n_0 \geq 1$ *such that for all* $m \geq n \geq n_0$ *one has*

$$\|f_m - f_n\|_p < \varepsilon,$$

[1] Georg Friedrich Bernhard Riemann, 17/09/1826 (Breselenz, Hanover)- 20/07/1866 (Selasca, Italy), clarified the notion of an integral (Riemann integral).

[2] Hermann Minkowski, 22/06/1864 (Alexotas, Russia)- 12/01/1909 (Göttingen, Germany), developed a new view of space and time, foundation of the theory of relativity.

then there exists an $f \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ *such that*

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

(iv) *Equivalence-classes: for* $f \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ *one has that* $\|f\|_p = 0$ *if and only if* $\mathbb{P}(f \neq 0) = 0$. *The relation* $f \sim g$ *if* $\mathbb{P}(f = g) = 1$ *defines an equivalence class relation.*

Remark 3.1.4.

(i) Because of assertion (ii) of Proposition 3.1.3 the expression $\|\cdot\|_p$ is a *semi-norm* on $\mathcal{L}_p$. Moreover, (ii) and (iv) imply that $\|\cdot\|_{L_p}$ is a norm on L_p, that means $[L_p, \|\cdot\|_{L_p}]$ is a normed space.

(ii) Items (iii) and (iv) say that every *Cauchy-sequence* in L_p converges to a limit in L_p. Hence $[L_p, \|\cdot\|_p]$ is a *complete* normed space. A complete normed space is called *Banach*[3] *space*.

(iii) If $p = 2$, then $[L_2, \|\cdot\|_2]$ is a *Hilbert* [4] *space* where the inner product is given by

$$\langle \hat{f}, \hat{g} \rangle := \int_\Omega fg d\mathbb{P} \quad \text{for} \quad f \in \hat{f} \text{ and } g \in \hat{g}.$$

Conditional expectation. There are different ways to approach conditional expectations. Let us explain the approach of the best approximation. Assume $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}([0,1))$ and $\mathbb{P} = \lambda$ to be the Lebesgue measure. Define the sub-σ-algebras

$$\mathcal{F}_n^{\text{dyad}} = \sigma\left(\left[0, \tfrac{1}{2^n}\right), \left[\tfrac{1}{2^n}, \tfrac{2}{2^n}\right), \dots, \left[\tfrac{2^n-1}{2^n}, 1\right)\right), \qquad n = 0, 1, 2, ...,$$

which is the system of all possible unions of sets of type $\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)$ for $k = 1, 2, 3, \dots, 2^n$ where $n \geq 1$ is fixed. Then $\mathcal{F}_n^{\text{dyad}} \subseteq \mathcal{F}_N^{\text{dyad}} \subseteq \mathcal{F}$ for $1 \leq n \leq N < \infty$. We know that $g : [0,1) \to \mathbb{R}$ is $\mathcal{F}_n^{\text{dyad}}$-measurable if and only if g is constant on all $\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)$.

[3] Stefan Banach, 20/3/1892 (Krakow, Austria-Hungary)- 31/8/1945 (Lvov), founded modern functional analysis and made major contributions to the theory of topological vector spaces.

[4] David Hilbert, 23/1/1862 (Königsberg, Prussia)- 14/2/1943 (Göttingen, Germany), Hilbert's work in geometry had the greatest influence in that area after Euclid.

Let $f : [0,1) \to \mathbb{R}$ be a continuous function with $f \in \mathcal{L}_2([0,1)) = \mathcal{L}_2([0,1), \mathcal{B}([0,1)), \lambda)$. What is the best L_2-approximation of f by a function $g : [0,1) \to \mathbb{R}$ measurable with respect to $\mathcal{F}_n^{\text{dyad}}$? In other words, we would like to minimize

$$\inf_g \|f-g\|_{L_2} = \inf_g \left(\int_0^1 |f(x)-g(x)|^2 dx \right)^{\frac{1}{2}},$$

where $g : [0,1) \to \mathbb{R}$ is $\mathcal{F}_n^{\text{dyad}}$-measurable.

Proposition 3.1.5. (i) *The best approximation g is of form*

$$g(x) = 2^n \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(y)dy \textit{ for } x \in \left[\tfrac{k-1}{2^n}, \tfrac{k}{2^n}\right).$$

(ii) *For all $A \in \mathcal{F}_n^{dyad}$ one has that*

$$\int_A f(x)dx = \int_A g(x)dx.$$

Proof. (i) Take another $\mathcal{F}_n^{\text{dyad}}$-measurable function $h : [0,1) \to \mathbb{R}$. Then

$$\int_0^1 |f(x)-h(x)|^2 dx = \sum_{k=1}^{2^n} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} |f(x)-h(x)|^2 dx.$$

We will prove that

$$\begin{aligned} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} |f(x)-h(x)|^2 dx &\geq \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} |f(x)-g(x)|^2 dx \\ &= \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} \left| f(x) - 2^n \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(y)dy \right|^2 dx \end{aligned}$$

Consider

$$\begin{aligned} \varphi(y) :&= \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} |f(x)-y|^2 dx \\ &= \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} |f(x)|^2 dx - 2y \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(x)dx + \frac{1}{2^n} y^2. \end{aligned}$$

Then the minimum is attained for $y = 2^n \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(x)dx$ and we are done.

(ii) $A \in \mathcal{F}_n^{\text{dyad}}$ means $A = \bigcup_{k\in I} \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)$. Then

$$\begin{aligned}\int_A f(x)dx &= \sum_{k\in I} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(x)dx \\ &= \sum_{k\in I} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} g(x)dx \\ &= \int_A g(x)dx.\end{aligned}$$

□

We add another example that motivates the conditional expectation as the best approximation in L_2.

Example 3.1.6. Consider the following situation: Given the space $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$ choose a set $A \in \mathcal{F}$ with $\mathbb{P}(A) \in (0, 1)$ and let $\mathcal{G} = \{\Omega, \emptyset, A, A^c\}$. Then

$$\mathcal{L}_2(\Omega, \mathcal{G}, \mathbb{P}) \subseteq \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P}).$$

For a given $f \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$, does there exist a $g_0 \in \mathcal{L}_2(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\|f - g_0\|_2 = \inf_{g\in\mathcal{L}_2(\Omega,\mathcal{G},\mathbb{P})} \|f - g\|_2?$$

Since any $g \in \mathcal{L}_2(\Omega, \mathcal{G}, \mathbb{P})$ is of the type $g = a\chi_A + b\chi_{A^c}$ for some real a and b we have

$$\begin{aligned}\inf_{g\in\mathcal{L}_2(\Omega,\mathcal{G},\mathbb{P})} \|f - g\|_2^2 &= \inf_{a,b\in\mathbb{R}} \|f - a\chi_A + b\chi_{A^c}\|_2^2 \\ &= \inf_{a,b\in\mathbb{R}} \mathbb{E}(f - a\chi_A + b\chi_{A^c})^2 \\ &= \inf_{a,b\in\mathbb{R}} (\mathbb{E}f^2 - 2a\mathbb{E}(f\chi_A) - 2b\mathbb{E}(f\chi_{A^c}) + a^2\mathbb{P}(A) + b^2\mathbb{P}(A^c))\end{aligned}$$

So we want to compute the infimum of the smooth real function

$$(a, b) \mapsto F(a, b) = \mathbb{E}f^2 - 2a\mathbb{E}(f\chi_A) - 2b\mathbb{E}(f\chi_{A^c}) + a^2\mathbb{P}(A) + b^2\mathbb{P}(A^c).$$

From $\frac{\partial F}{\partial a} = 0$ and $\frac{\partial F}{\partial b} = 0$ we get $a = \frac{\mathbb{E}(f\chi_A)}{\mathbb{P}(A)}$ and $b = \frac{\mathbb{E}(f\chi_{A^c})}{\mathbb{P}(A^c)}$ for which we can easily verify that for these we have the minimum, so that

$$g_0 = \frac{\mathbb{E}(f\chi_A)}{\mathbb{P}(A)}\chi_A + \frac{\mathbb{E}(f\chi_{A^c})}{\mathbb{P}(A^c)}\chi_{A^c}.$$

From definition 3.1.8 below we will see that g_0 is the conditional expectation of f given $\mathcal{G}$:

$$g_0 = \mathbb{E}(f \mid \mathcal{G}).$$

Behind the above example there is the general notion of the expected value we introduce now.

Proposition 3.1.7 (Conditional expectation)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-algebra. Let* $f \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$.

(i) *There exists a* $g \in \mathcal{L}_1(\Omega, \mathcal{G}, \mathbb{P})$ *such that*

$$\int_B f d\mathbb{P} = \int_B g d\mathbb{P} \text{ for all } B \in \mathcal{G}.$$

(ii) *If* g *and* g' *are like in* (i)*, then* $\mathbb{P}(g \neq g') = 0$.

The main point of the theorem above is that g is $\mathcal{G}$-measurable. This leads to the following definition:

Definition 3.1.8 (Conditional expectation)**.** The $\mathcal{G}$ measurable and integrable random variable g from Proposition 3.1.7 is called *conditional expectation* of f with respect to $\mathcal{G}$ and is denoted by

$$g = \mathbb{E}(f \mid \mathcal{G}).$$

One has to keep in mind that the conditional expectation is only unique up to null sets from $\mathcal{G}$.

We will prove Proposition 3.1.7 later. First we continue with some basic properties of conditional expectations.

Proposition 3.1.9. *Let* $f, g, f_1, f_2 \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ *and let* $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ *be sub* σ*-algebras of* $\mathcal{F}$*. Then the following holds true:*

(i) *Linearity: if* $\mu, \lambda \in \mathbb{R}$*, then*

$$\mathbb{E}(\lambda f + \mu g \mid \mathcal{G}) = \lambda \mathbb{E}(f \mid \mathcal{G}) + \mu \mathbb{E}(g \mid \mathcal{G}) \text{ a.s.}$$

(ii) *Monotonicity: if* $f_1 \leqq f_2$ *a.s., then* $\mathbb{E}(f_1 \mid \mathcal{G}) \leqq \mathbb{E}(f_2 \mid \mathcal{G})$ *a.s.*

(iii) *Positivity: If* $f \geqq 0$ *a.s., then* $\mathbb{E}(f \mid \mathcal{G}) \geqq 0$ *a.s.*

(iv) *Convexity: one has that* $|\mathbb{E}(f \mid \mathcal{G})| \leqq \mathbb{E}(|f| \mid \mathcal{G})$ *a.s.*

(v) *Projection property: if* f *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}(f \mid \mathcal{G}) = f$ *a.s.*

(vi) *Advanced projection property (or tower property):*

$$\mathbb{E}(\mathbb{E}(f \mid \mathcal{G}) \mid \mathcal{H}) = \mathbb{E}(\mathbb{E}(f \mid \mathcal{H}) \mid \mathcal{G}) = \mathbb{E}(f \mid \mathcal{H}) \text{ a.s.}$$

(vii) *'Take out what is known': If* $h : \Omega \to \mathbb{R}$ *is* $\mathcal{G}$*-measurable and* $fh \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$*, then*

$$\mathbb{E}(hf \mid \mathcal{G}) = h\mathbb{E}(f \mid \mathcal{G}) \text{ a.s.}$$

(viii) *If* $\mathcal{G} = \{\emptyset, \Omega\}$*, then* $\mathbb{E}(f|\mathcal{G}) = \mathbb{E}f$*.*

(ix) *The case that* f *is independent from* $\mathcal{G}$*: if for all* $B \in \mathcal{B}(\mathbb{R})$ *and all* $A \in \mathcal{G}$ *one has that*

$$\mathbb{P}(\{f \in B\} \cap A) = \mathbb{P}(f \in B)\mathbb{P}(A),$$

then $\mathbb{E}(f|\mathcal{G}) = \mathbb{E}f$ *a.s.*

(x) *Monotone convergence: assume* $f \geq 0$ *a.s. and random variables* $0 \leq h_n \uparrow f$ *a.s. Then*

$$\lim_n \mathbb{E}(h_n|\mathcal{G}) = \mathbb{E}(f|\mathcal{G}) \text{ a.s.}$$

Proof. (i) is an exercise.

(ii) Assume that (ii) does not hold. We find an $A \in \mathcal{G}$, $\mathbb{P}(A) > 0$ and $\alpha < \beta$ such that

$$\mathbb{E}(f_2 \mid \mathcal{G}) \leq \alpha < \beta \leq \mathbb{E}(f_1 \mid \mathcal{G}) \text{ on } A.$$

Hence

$$\int_A f_2 d\mathbb{P} \leq \alpha\mathbb{P}(A) < \beta\mathbb{P}(A) \leq \int_A f_1 d\mathbb{P}$$

which is a contradiction.

(iii) Apply (ii) to $0 = f_1 \leq f_2 = f$.

(iv) The inequality $f \leq |f|$ gives $\mathbb{E}(f|\mathcal{G}) \leqq \mathbb{E}(|f||\mathcal{G})$ a.s. and $-f \leq |f|$ gives $-\mathbb{E}(f|\mathcal{G}) \leqq \mathbb{E}(|f||\mathcal{G})$ a.s., so that we are done.

(v) follows directly from the definition.

(vi) Since $\mathbb{E}(f|\mathcal{H})$ is $\mathcal{H}$-measurable and hence $\mathcal{G}$-measurable, item (v) implies that $\mathbb{E}(\mathbb{E}(f|\mathcal{H})|\mathcal{G}) = \mathbb{E}(f|\mathcal{H})$ a.s. so that one equality is shown. For the other equality we have to show that

$$\int_A \mathbb{E}(\mathbb{E}(f|\mathcal{G})|\mathcal{H})d\mathbb{P} = \int_A f d\mathbb{P}$$

for $A \in \mathcal{H}$. Letting $h := \mathbb{E}(f|\mathcal{G})$ this follows from

$$\begin{aligned}
\int_A \mathbb{E}(\mathbb{E}(f|\mathcal{G})|\mathcal{H})d\mathbb{P} &= \int_A \mathbb{E}(h|\mathcal{H})d\mathbb{P} \\
&= \int_A h d\mathbb{P} \\
&= \int_A \mathbb{E}(f|\mathcal{G})d\mathbb{P} \\
&= \int_A f d\mathbb{P}
\end{aligned}$$

since $A \in \mathcal{H} \subseteq \mathcal{G}$.

(vii) Assume first that $h = \sum_{n=1}^N \alpha_n \chi_{A_n}$, where $\bigcup_{n=1}^N A_n = \Omega$ is a disjoint partition with $A_n \in \mathcal{G}$. For $A \in \mathcal{G}$ we get, a.s., that

$$\begin{aligned}
\int_A h f d\mathbb{P} &= \sum_{n=1}^N \alpha_n \int_A \chi_{A_n} f d\mathbb{P} \\
&= \sum_{n=1}^N \alpha_n \int_{A \cap A_n} f d\mathbb{P} \\
&= \sum_{n=1}^N \alpha_n \int_{A \cap A_n} \mathbb{E}(f|\mathcal{G})d\mathbb{P} \\
&= \int_A \left(\sum_{n=1}^N \alpha_n \chi_{A_n} \right) \mathbb{E}(f|\mathcal{G})d\mathbb{P} \\
&= \int_A h \mathbb{E}(f|\mathcal{G})d\mathbb{P}.
\end{aligned}$$

Hence $\mathbb{E}(hf|\mathcal{G}) = h\mathbb{E}(f|\mathcal{G})$ a.s. For the general case we can assume that $f, h \geq 0$ since we can decompose $f = f^+ - f^-$ and $h = h^+ - h^-$ with $f^+ := \max\{f, 0\}$ and $f^- := \max\{-f, 0\}$ (and in the same way we proceed with h). We find step-functions $0 \leq h_n \leq h$ such that $h_n(\omega) \uparrow h(\omega)$. Then, by our fist step, we get that

$$h_n\mathbb{E}(f|\mathcal{G}) = \mathbb{E}(h_n f|\mathcal{G}) \text{ a.s.}$$

By $n \to \infty$ the left-hand side follows. The right-hand side is a consequence of the monotone convergence given in (x).

(viii) Clearly, we have that

$$\int_\emptyset f d\mathbb{P} = 0 = \int_\emptyset (\mathbb{E}f) d\mathbb{P} \quad \text{and} \quad \int_\Omega f d\mathbb{P} = \mathbb{E}f = \int_\Omega (\mathbb{E}f) d\mathbb{P}$$

so that (viii) follows.

(ix) is an exercise.

(x) is an exercise. □

Proof of Proposition 3.1.7. For this purpose we need the theorem of RADON [5] -NIKODYM[6] .

Definition 3.1.10 (Signed measures)**.** Let $(\Omega, \mathcal{F})$ be a measurable space.

(i) A map $\mu : \mathcal{F} \to \mathbb{R}$ is called (finite) *signed measure* if and only if

$$\mu = \alpha\mu^+ - \beta\mu^-,$$

where $\alpha, \beta \geq 0$ and μ^+ and μ^- are probability measures on $\mathcal{F}$.

(ii) Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and that μ is a signed measure on $(\Omega, \mathcal{F})$. Then $\mu \ll \mathbb{P}$ (μ is *absolutely continuous* with respect to $\mathbb{P}$) if and only if

$$\mathbb{P}(A) = 0 \quad \text{implies} \quad \mu(A) = 0.$$

[5]Johann Radon, 16/12/1887 (Tetschen, Bohemia)- 25/05/1956 (Vienna, Austria), worked on the calculus of variations, differential geometry and measure theory.

[6]Otton Marcin Nikodym, 13/8/1887 (Zablotow, Austria-Hungary ; now Ukraine)- 4/05/1974 (Utica, USA), worked in measure theory, functional analysis, projections onto convex sets with applications to Dirichlet problem, generalized solutions of differential equations, descriptive set theory and the foundations of quantum mechanics.

Example 3.1.11. Let $L \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ and

$$\mu(A) := \int_A L d\mathbb{P}.$$

Then μ is a signed measure and $\mu \ll \mathbb{P}$.

Proof. We let $L^+ := \max\{L, 0\}$ and $L^- := \max\{-L, 0\}$ so that $L = L^+ - L^-$. Assume that $\int_\Omega L^\pm d\mathbb{P} > 0$ and define

$$\mu^\pm(A) = \frac{\int_\Omega \chi_A L^\pm d\mathbb{P}}{\int_\Omega L^\pm d\mathbb{P}}.$$

Now we check that $\mu^\pm$ are probability measures. First we have that

$$\mu^\pm(\Omega) = \frac{\int_\Omega \chi_\Omega L^\pm d\mathbb{P}}{\int_\Omega L^\pm d\mathbb{P}} = 1.$$

Then assume $A_n \in \mathcal{F}$ to be disjoint sets such that $A = \bigcup\limits_{n=1}^{\infty} A_n$. Set $\alpha := \int_\Omega L^+ d\mathbb{P}$. Then

$$\begin{aligned}
\mu^+\left(\bigcup_{n=1}^{\infty} A_n\right) &= \frac{1}{\alpha}\int_\Omega \chi_{\bigcup_{n=1}^{\infty} A_n} L^+ d\mathbb{P} \\
&= \frac{1}{\alpha}\int_\Omega \left(\sum_{n=1}^{\infty} \chi_{A_n}(\omega)\right) L^+ d\mathbb{P} \\
&= \frac{1}{\alpha}\int_\Omega \lim_{N\to\infty} \left(\sum_{n=1}^{N} \chi_{A_n}\right) L^+ d\mathbb{P} \\
&= \frac{1}{\alpha}\lim_{N\to\infty}\int_\Omega \left(\sum_{n=1}^{N} \chi_{A_n}\right) L^+ d\mathbb{P} \\
&= \sum_{n=1}^{\infty} \mu^+(A_n)
\end{aligned}$$

where we have used LEBESGUE's dominated convergence theorem. The same can be done for L^-. □

Theorem 3.1.12 (Radon-Nikodym)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and μ a signed measure with $\mu \ll \mathbb{P}$. Then there exists an $L \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$\mu(A) = \int_A L(\omega) d\mathbb{P}(\omega), \ A \in \mathcal{F}. \tag{3.1}$$

The random variable L is unique in the following sense. If L and L' are random variables satisfying (3.1), then

$$\mathbb{P}(L \neq L') = 0.$$

Definition 3.1.13. L is called RADON-NIKODYM derivative. We shall write

$$L = \frac{d\mu}{d\mathbb{P}}.$$

We should keep in mind the rule

$$\mu(A) = \int_\Omega \chi_A d\mu = \int_\Omega \chi_A L d\mathbb{P},$$

so that '$d\mu = L d\mathbb{P}$'.

Proof of Proposition 3.1.7. Define

$$\mu(A) := \int_A f d\mathbb{P} \text{ for } A \in \mathcal{G}$$

so that μ is a signed measure on $\mathcal{G}$. Applying the Theorem of RADON-NIKODYM gives an $g \in \mathcal{L}_1(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\mu(A) = \int_A g d\mathbb{P}$$

so that

$$\int_A g d\mathbb{P} = \mu(A) = \int_A f d\mathbb{P}.$$

Assume now another $g' \in \mathcal{L}_1(\Omega, \mathcal{G}, \mathbb{P})$ with

$$\int_A g d\mathbb{P} = \int_A g' d\mathbb{P}$$

for all $A \in \mathcal{G}$ and assume that $\mathbb{P}(g \neq g') > 0$. Hence we find a set $A \in \mathcal{G}$ with $\mathbb{P}(A) > 0$ and real numbers $\alpha < \beta$ such that

$$g(\omega) \leq \alpha < \beta \leq g'(\omega) \quad \text{for} \quad \omega \in A$$

or

$$g'(\omega) \leq \alpha < \beta \leq g(\omega) \quad \text{for} \quad \omega \in A.$$

Consequently (for example in the first case)

$$\int_A g d\mathbb{P} \leq \alpha \mathbb{P}(A) < \beta \mathbb{P}(A) \leq \int_A g' d\mathbb{P}$$

which is a contradiction. □

3.2 Definition and examples of martingales

Looking up the word *martingale* from an encyclopedia (for example in www.dict.org/bin/Dict) gives the following:

(i) A strap fastened to a horse's girth, passing between his fore legs, and fastened to the bit, or now more commonly ending in two rings, through which the reins pass. It is intended to hold down the head of the horse, and prevent him from rearing.

(ii) (Naut.) A lower stay of rope or chain for the jib boom or flying jib boom, fastened to, or reeved through, the dolphin striker. Also, the dolphin striker itself.

(iii) (Gambling) The act of doubling, at each stake, that which has been lost on the preceding stake; also, the sum so risked; – metaphorically derived from the bifurcation of the martingale of a harness. [Cant] –Thackeray.

We start with the notation of a filtration which describes the information we have at a certain time-point.

Definition 3.2.1 (Filtration and adapted process)**.** (i) Let $(\Omega, \mathcal{F})$ be a measurable space. An increasing sequence $(\mathcal{F}_n)_{n=0}^\infty$ of σ-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_n \subseteq \ldots \subseteq \mathcal{F}$ is called *filtration*. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, then $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$ is called *filtered* probability space.

(ii) A stochastic process $(X_n)_{n=0}^\infty$, where $X_n : \Omega \to \mathbb{R}$ are random variables, is called *adapted* with respect to the filtration $(\mathcal{F}_n)_{n=0}^\infty$ (or $(\mathcal{F}_n)_{n=0}^\infty$-adapted) provided that X_n is $\mathcal{F}_n$-measurable for all $n = 0, 1, ...$

(iii) Given a stochastic process $X = (X_n)_{n=0}^\infty$ of random variables $X_n : \Omega \to \mathbb{R}$, then the filtration $(\mathcal{F}_n^X)_{n=0}^\infty$ given by

$$\mathcal{F}_n^X := \sigma(X_0, ..., X_n)$$

is called *natural filtration* of the process X.

It is clear that a process X is adapted with respect to $(\mathcal{F}_n^X)_{n=0}^\infty$.

Definition 3.2.2 (Martingale)**.** Let $M = (M_n)_{n=0}^\infty$ with $M_n \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ for $n = 0, 1, ...$ and let $(\mathcal{F}_n)_{n=0}^\infty$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$.

(i) The process M is called *martingale* if M is $(\mathcal{F}_n)_{n=0}^\infty$-adapted and

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = M_n \text{ a.s. for } n = 0, 1, ...$$

(ii) The process M is called *sub-martingale* or *super-martingale*, if M is $(\mathcal{F}_n)_{n=0}^\infty$-adapted and

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) \geqq M_n \text{ a.s.} \quad \text{or} \quad \mathbb{E}(M_{n+1} \mid \mathcal{F}_n) \leqq M_n \text{ a.s.},$$

for $n = 0, 1, ...$, respectively.

The definition of a (sub-, super-) martingale depends on the filtration and the measure. To emphasize this, one sometimes uses the phrases $(\mathcal{F}_n)_{n=0}^\infty$-*martingale* or $\mathbb{P}$-*martingale* (and the same for the sub- and super-martingales). Now we consider some examples.

Random walk. Let $0 < p, q < 1$, $p + q = 1$ and $\varepsilon_1^{(p)}, \varepsilon_2^{(p)}, \ldots : \Omega \to \mathbb{R}$ be independent random variables such that

$$\mathbb{P}(\varepsilon_n^{(p)} = -1) = p \quad \text{and} \quad \mathbb{P}(\varepsilon_n^{(p)} = 1) = q$$

(see Definition 2.1.1). As filtration we use $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and

$$\mathcal{F}_n := \sigma(\varepsilon_1^{(p)}, \ldots, \varepsilon_n^{(p)})$$

for $n \geq 1$ so that $\mathcal{F}_n$ consists of all possible unions of sets of type

$$\{\varepsilon_1^{(p)} = \theta_1, \dots, \varepsilon_n^{(p)} = \theta_n\}$$

with $\theta_1, \dots, \theta_n \in \{1, -1\}$. In dependence on p we get that our random walk considered in Chapter 2 is a (sub-, super-) martingale.

Proposition 3.2.3. *Let $M = (M_n)_{n=0}^{\infty}$ be given by $M_n := \varepsilon_1^{(p)} + \dots + \varepsilon_n^{(p)}$ for $n \geq 1$ and $M_0 := 0$. Then one has the following:*

(i) *If $q = p = \frac{1}{2}$, then M is a martingale.*

(ii) *If $q > \frac{1}{2}$, then M is a sub-martingale.*

(iii) *If $q < \frac{1}{2}$, then M is a super-martingale.*

Proof. Using the rules for conditional expectations we get that

$$\begin{aligned} \mathbb{E}(M_{n+1} \mid \mathcal{F}_n) &= \mathbb{E}(M_n + \varepsilon_{n+1} \mid \mathcal{F}_n) \\ &= \mathbb{E}(M_n \mid \mathcal{F}_n) + \mathbb{E}(\varepsilon_{n+1} \mid \mathcal{F}_n) \\ &= M_n + \mathbb{E}\varepsilon_{n+1}. \end{aligned}$$

It remains to observe that $\mathbb{E}\varepsilon_{n+1} = q - p$ is negative, zero, or positive in dependence on p. □

The random walk from Proposition 3.2.3 has an *additive* structure since, for $0 \leq k \leq n$,

- $M_n - M_k$ is independent from $\mathcal{F}_k$,
- $M_n - M_k$ has the same distribution as M_{n-k}.

Exponential random walk. In applications, for example in Stochastic Finance, one needs instead of the additive structure a multiplicative structure. For instance, changes in share price processes are always relative to the current price. Hence we are looking for a process $M_n > 0$ such that, for $0 \leq k \leq n$, M_n/M_k is independent from $\mathcal{F}_k$ and has the same distribution as M_{n-k}. One process of this type is given by

Proposition 3.2.4. *Let* $p \in (0,1)$, $\sigma > 0$, $c \in \mathbb{R}$, $M_0 := 1$,

$$M_n := e^{\sigma \sum_{i=1}^{n} \varepsilon_i^{(p)} + cn}$$

for $n = 1, 2, ...$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, *and* $\mathcal{F}_n = \sigma(\varepsilon_1^{(p)}, \ldots, \varepsilon_n^{(p)})$. *Then one has the following:*

(i) *If* $p(e^{-\sigma} - e^{\sigma}) = e^{-c} - e^{\sigma}$, *then* $M = (M_n)_{n=0}^{\infty}$ *is a martingale.*

(ii) *If* $p(e^{-\sigma} - e^{\sigma}) \geq e^{-c} - e^{\sigma}$, *then* $M = (M_n)_{n=0}^{\infty}$ *is a sub-martingale.*

(iii) *If* $p(e^{-\sigma} - e^{\sigma}) \leq e^{-c} - e^{\sigma}$, *then* $M = (M_n)_{n=0}^{\infty}$ *is a super-martingale.*

Proof. We have that

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(e^{\sigma \varepsilon_{n+1}^{(p)} + c} M_n \mid \mathcal{F}_n) = M_n \mathbb{E} e^{\sigma \varepsilon_{n+1}^{(p)} + c} = M_n (p e^{c-\sigma} + q e^{c+\sigma})$$

and obtain a martingale if and only if $pe^{-\sigma} + (1-p)e^{\sigma} = e^{-c}$ or

$$p(e^{-\sigma} - e^{\sigma}) = e^{-c} - e^{\sigma}.$$

If we have an inequality, then we get a sub- or super-martingale. □

Dyadic martingales. Let $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}([0,1))$, $f \in \mathcal{L}_1([0,1))$, and $\mathbb{P} = \lambda$ be the LEBESGUE measure. As filtration we define

$$\mathcal{F}_n^{\text{dyadic}} = \sigma\left(\left[0, \tfrac{1}{2^n}\right), \left[\tfrac{1}{2^n}, \tfrac{2}{2^n}\right), \ldots, \left[\tfrac{2^n-1}{2^n}, 1\right)\right)$$

for $n = 0, 1, \ldots$ Finally, we let

$$M_n(t) := \mathbb{E}(f | \mathcal{F}_n^{\text{dyad}})(t).$$

The process $(M_n)_{n=0}^{\infty}$ is called *dyadic* martingale. This process is a martingale as shown by the more general

Proposition 3.2.5. *For* $f \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ *and a filtration* $(\mathcal{F}_n)_{n=0}^{\infty}$ *the sequence* $(M_n)_{n=0}^{\infty}$ *with* $M_n := \mathbb{E}(f | \mathcal{F}_n)$ *is a martingale.*

Proof. One has that, a.s.,

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(f \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n) = \mathbb{E}(f \mid \mathcal{F}_n) = M_n.$$

□

Remark 3.2.6. The function f in Proposition 3.2.5 is called the *closure* of the martingale $(\mathbb{E}(f|\mathcal{F}_n))_{n=0}^{\infty}$.

Remark 3.2.7. Assume a continuous $f : [0,1] \to \mathbb{R}$ and $M_n(t) := \mathbb{E}(f|\mathcal{F}_n^{\text{dyad}})$ as above. Then one can easily show that

$$\lim_n M_n(t) = f(t)$$

for all $t \in [0,1)$ and obtain our first limit theorem. To check this, fix $t_0 \in [0,1)$, find $k_0, k_1, ...$ such that

$$\frac{k_n - 1}{2^n} \leq t_0 < \frac{k_n}{2^n},$$

and let $I_n := \{t \in [0,1) : |t_0 - t| \leq 2^{-n}\}$. Then

$$\inf_{t \in I_n} f(t) \leq M_n(t_0) = 2^n \int_{\frac{k_n-1}{2^n}}^{\frac{k_n}{2^n}} f(t)dt \leq \sup_{t \in I_n} f(t).$$

Since f is continuous, we finish by

$$\lim_n \left(\inf_{t \in I_n} f(t) \right) = \lim_n \left(\sup_{t \in I_n} f(t) \right) = f(t_0).$$

Branching processes. Here we come back to the branching process example considered in Section 1.2. Now we can show that we get a martingale if we rescale the process, which models the number of persons having a given name, properly.

Proposition 3.2.8. *Let* $\{X_i^{(n)} : i, n = 1, 2, \ldots\}$ *be independent random variables,* $q_0, \ldots, q_N \in [0,1]$, $q_0 + \cdots + q_N = 1$, *and* $\mathbb{P}(X_i^{(n)} = k) = q_k$ *for* $k = 0, \ldots, N$ *and let* $X_0^{(n)} := 0$. *Assume that* $\mu := \sum_{k=0}^{N} k q_k > 0$, $f_0 := 1$, *and*

$$f_{n+1}(\omega) := X_0^{(n+1)}(\omega) + \cdots + X_{f_n(\omega)}^{(n+1)}(\omega) \quad \textit{and} \quad M_n := \frac{f_n}{\mu^n}.$$

Then $M = (M_n)_{n=0}^{\infty}$ *is a martingale with respect to the filtration* $(\mathcal{F}_n)_{n=0}^{\infty}$ *given by* $\mathcal{F}_n := \sigma(M_0, \ldots, M_n)$ *for* $n \geq 0$.

Proof. Clearly, M_n is $\mathcal{F}_n$-measurable by definition of $\mathcal{F}_n$. Next we show $M_n \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. We proceed by induction and assume that $\mathbb{E}M_{n-1} < \infty$ or equivalently, that $\mathbb{E}f_{n-1} < \infty$. We get

$$\begin{aligned}
\mathbb{E}M_n &= \frac{1}{\mu^n}\mathbb{E}(X_0^{(n)} + \ldots + X_{f_{n-1}}^{(n)}) \\
&= \sum_{k=1}^{\infty} \frac{1}{\mu^n}\mathbb{P}(f_{n-1} = k)\mathbb{E}(X_0^{(n)} + \ldots + X_k^{(n)}) \\
&= \frac{\mathbb{E}X_1^{(n)}}{\mu^n}\sum_{k=1}^{\infty}\mathbb{P}(f_{n-1} = k)k \\
&= \frac{\mathbb{E}X_1^{(n)}}{\mu^n}\mathbb{E}f_{n-1} \\
&< \infty.
\end{aligned}$$

The martingale property is obtained by

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = \frac{1}{\mu^{n+1}}\mathbb{E}(X_0^{(n+1)} + \ldots + X_{f_n}^{(n+1)} \mid \mathcal{F}_n) = \frac{1}{\mu^{n+1}} f_n \mathbb{E}X_1^{(n+1)} \text{ a.s.}$$

where the last equality is justified in the same way we used to get $M \subset \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ in our first step: for $B_n \in \mathcal{F}_n$, one has that

$$\begin{aligned}
&\int_{B_n} (X_0^{(n+1)} + X_1^{(n+1)} + \cdots + X_{f_n}^{(n+1)})d\mathbb{P} \\
&= \sum_{k=1}^{\infty}\int_{B_n \cap \{f_n = k\}} (X_0^{(n+1)} + X_1^{(n+1)} + \cdots + X_k^{(n+1)})d\mathbb{P} \\
&= \sum_{k=1}^{\infty}\int_{B_n \cap \{f_n = k\}} d\mathbb{P} \int_{\Omega} (X_0^{(n+1)} + X_1^{(n+1)} + \cdots + X_k^{(n+1)})d\mathbb{P} \\
&= \sum_{k=1}^{\infty}\mathbb{E}X_1^{(n+1)}\int_{B_n \cap \{f_n = k\}} k\,d\mathbb{P} \\
&= \mathbb{E}X_1^{(n+1)}\int_{B_n} f_n d\mathbb{P} \\
&= \int_{B_n} [f_n \mathbb{E}X_1^{(n+1)}]d\mathbb{P}.
\end{aligned}$$

Finally, we finish with

$$\frac{1}{\mu^{n+1}} f_n \mathbb{E}X_1^{(n+1)} = \frac{1}{\mu} M_n \mathbb{E}X_1^{(n+1)} = M_n.$$

□

3.3 Some elementary properties of martingales

Here we collect some very basic properties of martingales. It will be useful to introduce for a martingale $M = (M_n)_{n=0}^\infty$ the sequence of *martingale differences* $dM_0 := M_0$, and

$$dM_n := M_n - M_{n-1} \quad \text{for} \quad n = 1, 2, ...$$

We start with

Proposition 3.3.1. *Let $M = (M_n)_{n=0}^\infty$ be a martingale on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$. Then one has*

(i) $\mathbb{E}(M_m \mid \mathcal{F}_n) = M_n$ *a.s. for* $0 \le n < m < \infty$,

(ii) $\mathbb{E}M_0 = \mathbb{E}M_n$ *for all* $n = 1, 2, ...$.

(iii) *If* $\mathbb{E}|M_n|^2 < \infty$ *for all* $n = 0, 1, \ldots$, *then*

$$\int_\Omega dM_n dM_m d\mathbb{P} = 0$$

for $m, n \ge 0$ *with* $m \ne n$.

Item (iii) means the martingale differences $(dM_k)_{k=0}^\infty$ are orthogonal if the martingale is quadratic integrable.

Proof of Proposition 3.3.1. (i) Successively we obtain, a.s.,

$$\begin{aligned}
\mathbb{E}(M_m|\mathcal{F}_n) &= \mathbb{E}(\mathbb{E}(M_m|\mathcal{F}_{m-1})|\mathcal{F}_n) \\
&= \mathbb{E}(M_{m-1}|\mathcal{F}_n) \\
&\ \ \vdots \\
&= \mathbb{E}(M_{n+1}|\mathcal{F}_n) \\
&= M_n.
\end{aligned}$$

(ii) Applying item (i) to $n = 0$ gives $\mathbb{E}(M_m|\mathcal{F}_0) = M_0$ a.s., so that

$$\mathbb{E}M_m = \int_\Omega M_m d\mathbb{P} = \int_\Omega \mathbb{E}(M_m|\mathcal{F}_0) d\mathbb{P} = \int_\Omega M_0 d\mathbb{P} = \mathbb{E}M_0$$

since $\Omega \in \mathcal{F}_0$.

(iii) Assume that $0 \leq m < n$. Then

$$\begin{aligned}
& \int_\Omega dM_n dM_m d\mathbb{P} \\
= & \int_\Omega (M_n - M_{n-1})(M_m - M_{m-1}) d\mathbb{P} \\
= & \int_\Omega M_n M_m d\mathbb{P} - \int_\Omega M_n M_{m-1} d\mathbb{P} - \int_\Omega M_{n-1} M_m d\mathbb{P} + \int_\Omega M_{n-1} M_{m-1} d\mathbb{P} \\
= & \int_\Omega \mathbb{E}(M_n M_m | \mathcal{F}_m) d\mathbb{P} - \int_\Omega \mathbb{E}(M_n M_{m-1} \mid \mathcal{F}_{m-1}) d\mathbb{P} \\
& - \int_\Omega \mathbb{E}(M_{n-1} M_m | \mathcal{F}_m) d\mathbb{P} + \int_\Omega \mathbb{E}(M_{n-1} M_{m-1} | \mathcal{F}_{m-1}) d\mathbb{P} \\
= & \int_\Omega M_m \mathbb{E}(M_n \mid \mathcal{F}_m) d\mathbb{P} - \int_\Omega M_{m-1} \mathbb{E}(M_n \mid \mathcal{F}_{m-1}) d\mathbb{P} \\
& - \int_\Omega M_m \mathbb{E}(M_{n-1} | \mathcal{F}_m) d\mathbb{P} + \int_\Omega M_{m-1} \mathbb{E}(M_{n-1} | \mathcal{F}_{m-1}) d\mathbb{P} \\
= & \int_\Omega M_m^2 d\mathbb{P} - \int_\Omega M_{m-1}^2 d\mathbb{P} - \int_\Omega M_m^2 d\mathbb{P} + \int_\Omega M_{m-1}^2 d\mathbb{P} \\
= & \ 0
\end{aligned}$$

where we have used Proposition 3.1.9. □

Lemma 3.3.2 (JENSEN's inequality)**.** [7] *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function, $f \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ such that $\varphi(f) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-algebra. Then*

$$\varphi(\mathbb{E}(f|\mathcal{G})) \leq (\mathbb{E}(\varphi(f)|\mathcal{G})) \ \textit{a.s.}$$

Proof. Any convex function $\varphi : \mathbb{R} \to \mathbb{R}$ is continuous, hence measurable. To check the continuity we fix $x \in \mathbb{R}$ and observe first that for $x_n \to x$ we have that

$$\limsup_n \varphi(x_n) \leq \varphi(x)$$

(if this is not true we would get, by drawing a picture, a contradiction to the convexity of φ). Assume now that $a_n \uparrow x$ and $b_n \downarrow x$ and that either $\liminf_n \varphi(a_n)$

[7] Johan Ludwig William Valdemar Jensen, 08/05/1859 (Nakskov, Denmark)- 05/ 03/1925 (Copenhagen, Denmark), studied infinite series, the gamma function and inequalities for convex functions. Only did mathematics in his spare time, his actual job was in a telephone company.

or $\liminf_n \varphi(b_n)$ (or both) is (are) strictly less than $\varphi(x)$. Picking $\theta_n \in (0,1)$ such that $x = (1-\theta_n)a_n + \theta_n b_n$ we would get that

$$(1-\theta_n)\varphi(a_n) + \theta_n \varphi(b_n) < \varphi(x)$$

for $n \geq n_0$ which is a contradiction to the convexity of φ. Next we observe that there exists a countable set D of linear functions $L : \mathbb{R} \to \mathbb{R}$ such that, for all $x \in \mathbb{R}$, one has that

$$\varphi(x) = \sup_{L \in D} L(x).$$

To get that D we proceed as follows. For $q \in \mathbb{Q}$ we let

$$a_q := \lim_{x \uparrow q} \frac{\varphi(q) - \varphi(x)}{q - x} \quad \text{and} \quad b_q := \lim_{x \downarrow q} \frac{\varphi(q) - \varphi(x)}{q - x}.$$

By convexity of φ we have that $-\infty < a_q \leq b_q < \infty$. Now we collect in D all linear functions $L(x) := a_q(x-q) + \varphi(q)$ and $L(x) := b_q(x-q) + \varphi(q)$, $q \in \mathbb{Q}$. By construction we have that

$$\sup_D L(x) \leq \varphi(x)$$

for all $x \in \mathbb{R}$ and

$$\sup_D L(q) = \varphi(q)$$

for $q \in \mathbb{Q}$. Since $\sup_D L(x)$ is convex as well and hence continuous, we have to have that $\sup_D L(x) = \varphi(x)$. Now we come back to probability: given $L \in D$, we get that, a.s.,

$$L(\mathbb{E}(f|\mathcal{G})) = \mathbb{E}(L(f)|\mathcal{G}) \leq \mathbb{E}(\varphi(f)|\mathcal{G}).$$

Taking the supremum on the left-hand side over all $L \in D$ we get the desired inequality (in order to handle the exceptional sets, which are null-sets, we use that D is countable). □

Proposition 3.3.3. *Let $M = (M_n)_{n=0}^\infty$ be a martingale with respect to $(\mathcal{F}_n)_{n=0}^\infty$ and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $(\varphi(M_n))_{n=0}^\infty \subset \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. Then $(\varphi(M_n))_{n=0}^\infty$ is a sub-martingale.*

Proof. We simply apply the JENSEN inequality for the conditional expectations and get, a.s.,

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

□

An important application is the case $\varphi(x) := |x|^p$ with $1 \leq p < \infty$:

Proposition 3.3.4. *Let $p \in [1,\infty)$ and $M = (M_n)_{n=0}^\infty \subset \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ be a martingale. Then $(|M_n|^p)_{n=0}^\infty$ is a sub-martingale.*

Now we describe an easy version of *martingale transforms*. These transforms are, for example, a main ingredient in Stochastic Finance. Let us start with the definition:

Definition 3.3.5. Assume a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$.

(i) A sequence of random variables $v = (v_n)_{n=1}^\infty$, $v_n : \Omega \to \mathbb{R}$, is called *predictable* if v_n is $\mathcal{F}_{n-1}$-measurable for $n = 1, 2, ...$

(ii) For a predictable sequence v and a martingale $M = (M_n)_{n=0}^\infty$ we define its martingale transform $M^v = (M_n^v)_{n=0}^\infty$ by $M_0^v := M_0$ and

$$M_n^v := M_0 + \sum_{k=1}^n v_k (M_k - M_{k-1}) \quad \text{for} \quad n = 1, 2, ...$$

Proposition 3.3.6. *If $M = (M_n)_{n=0}^\infty$ is a martingale and $v = (v_n)_{n=1}^\infty$ is predictable such that $\mathbb{E}|v_n dM_n| < \infty$ for $n = 1, 2, ...$ Then the martingale transform $(M_n^v)_{n=0}^\infty$ is a martingale as well.*

Proof. Because of $\mathbb{E}|v_n dM_n| < \infty$ for $n = 1, 2, ...$ we get that $(M_n^v)_{n=0}^\infty \subset \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, M_n^v is $\mathcal{F}_n$-measurable since all terms of the defining sum are $\mathcal{F}_n$-measurable. The martingale property follows since one has, a.s.,

$$\begin{aligned}
\mathbb{E}(M_{n+1}^v \mid \mathcal{F}_n) &= \mathbb{E}(M_0 \mid \mathcal{F}_n) + \sum_{k=1}^{n+1} \mathbb{E}(v_k(M_k - M_{k-1}) \mid \mathcal{F}_n) \\
&= M_0 + \sum_{k=1}^{n+1} v_k \mathbb{E}(M_k - M_{k-1} \mid \mathcal{F}_n) \\
&= M_0 + \sum_{k=1}^{n} v_k \mathbb{E}(M_k - M_{k-1} \mid \mathcal{F}_n) + \\
&\qquad + v_{n+1} \mathbb{E}(M_{n+1} - M_n \mid \mathcal{F}_n) \\
&= M_0 + \sum_{k=1}^{n} v_k (M_k - M_{k-1}) = M_n^v.
\end{aligned}$$

□

3.4 Stopping times

An important tool in the martingale theory and their applications are *stopping times*. For example, we would like to stop a martingale before it is getting too large values. However the stopping should be done such that the stopped object is again a martingale.

Definition 3.4.1. Assume a measurable space $(\Omega, \mathcal{F})$ equipped with a filtration $(\mathcal{F}_n)_{n=0}^{\infty}$. A random time $\tau : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ is called *stopping time* provided that

$$\{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n \text{ for } n = 0, 1, 2, \ldots.$$

Let us start with some examples. The most easiest one is

Example 3.4.2. Consider $\tau : \Omega \to \{0, 1, ...\}$ with $\tau(\omega) \equiv n_0$. Then τ is a stopping time, since

$$\{\tau = n\} = \begin{cases} \Omega : & n = n_0 \\ \emptyset : & n \neq n_0. \end{cases}$$

One typical and often used class of stopping times is the class of *hitting times*.

Example 3.4.3 (Hitting time)**.** Let $B \in \mathcal{B}(\mathbb{R})$, the stochastic process $X = (X_n)_{n=0}^{\infty}$ be adapted, and

$$\sigma_B(\omega) := \inf\{n \geq 0 : X_n(\omega) \in B\}$$

with $\inf \emptyset := \infty$. Then σ_B is a stopping time and called *hitting time* of B.

Proof. That σ_B is a stopping time follows from

$$\{\sigma_B = n\} = \{X_0 \notin B\} \cap \ldots \cap \{X_{n-1} \notin B\} \cap \{X_n \in B\} \in \mathcal{F}_n$$

for $n \geq 1$ and $\{\sigma_B = 0\} = \{X_0 \in B\} \in \mathcal{F}_0$. □

Example 3.4.4. Let $X = (X_n)_{n=0}^{\infty}$ and $Y = (Y_n)_{n=0}^{\infty}$ be adapted processes and

$$\sigma(\omega) := \inf\{n \geq 0 : \ X_n(\omega) = Y_n(\omega)\}$$

with $\inf \emptyset =: \infty$. Then σ is a stopping time. In fact, if we let $Z_n := X_n - Y_n$ and $B := \{0\}$, then σ is the hitting time of B with respect to the adapted process Z.

Example 3.4.5. Let $X = (X_n)_{n=0}^{\infty}$ be adapted. Then

$$\sigma(\omega) := \inf\{n \geq 0 : \ X_{n+1}(\omega) > 1\}$$

with $\inf \emptyset := \infty$ is, in general, *not* a stopping time.

Now we continue with some general properties of stopping times.

Proposition 3.4.6. *Let $\sigma, \tau : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ be stopping times. Then one has the following:*

(i) *The map $\sigma : \Omega \to \mathbb{R} \cup \{\infty\}$ is a random variable, that means $\sigma^{-1}(\infty) \in \mathcal{F}$ and $\sigma^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}(\mathbb{R})$.*

(ii) *The time $\max\{\sigma, \tau\}$ is a stopping time.*

(iii) *The time $\min\{\sigma, \tau\}$ is a stopping time.*

(iv) *The time $\sigma + \tau$ is a stopping time.*

Proof. Item (i) follows from

$$\sigma^{-1}(B) = \bigcup_{n\in B, n\geq 0} \sigma^{-1}(\{n\}) \in \bigcup_{n\geq 0} \mathcal{F}_n \subseteq \mathcal{F}$$

and

$$\sigma^{-1}(\infty) = \Omega \setminus \bigcup_{n=0}^{\infty} \sigma^{-1}(\{n\}) \in \mathcal{F}.$$

Items (ii) and (iii) are consequences of

$$\begin{aligned}\{\max\{\sigma,\tau\} = n\} &= \{\sigma = n, \tau \leq n\} \cup \{\sigma \leq n, \tau = n\} \\ &= (\{\sigma = n\} \cap \{\tau \leq n\}) \cup (\{\sigma \leq n\} \cap \{\tau = n\}) \\ &\in \mathcal{F}_n\end{aligned}$$

and

$$\begin{aligned}\{\min\{\sigma,\tau\} = n\} &= (\{\sigma = n\} \cap \{\tau \geq n\}) \cup (\{\sigma \geq n\} \cap \{\tau = n\}) \\ &= [\{\sigma = n\} \cap (\Omega \setminus \{\tau < n\})] \cup [\{\tau = n\} \cap (\Omega \setminus \{\sigma < n\})] \\ &\in \mathcal{F}_n.\end{aligned}$$

(iv) is an exercise. □

Now we introduce the σ-algebra $\mathcal{F}_\tau$ for a stopping time τ. The σ-algebra will contain all events one can decide until the stopping time τ occurs. For example one can decide

$$A = \{\tau = n_0\}$$

until the stopping time τ: We take $\omega \in \Omega$ and wait until τ is going to happen, that means on $\{\tau = 0\}$ the stopping time occurs at time 0, on $\{\tau = 1\}$ at time 1, and so on. If $\tau(\omega)$ is going to happen, and $\tau(\omega) \neq n_0$, then $\omega \notin A$, if $\tau(\omega) = n_0$, then $\omega \in A$. The abstract definition is as follows:

Definition 3.4.7. Let $\tau : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ be a stopping time, then a set $A \in \mathcal{F}$ belongs to $\mathcal{F}_\tau$ if and only if

$$A \cap \{\tau = n\} \in \mathcal{F}_n$$

for all $n = 0, 1, 2, \ldots$.

Proposition 3.4.8. *$\mathcal{F}_\tau$ is a σ-algebra.*

Proof. We have that $\emptyset \cap \{\tau = n\} = \emptyset \in \mathcal{F}_n$ and $\Omega \cap \{\tau = n\} = \{\tau = n\} \in \mathcal{F}_n$ for all $n \in \{0, 1, 2, ...\}$ so that $\emptyset, \Omega \in \mathcal{F}_\tau$. Now let $A \in \mathcal{F}_\tau$. Then $A^c \cap \{\tau = n\} = \{\tau = n\} \setminus (A \cap \{\tau = n\}) \in \mathcal{F}_n$ so that $A^c \in \mathcal{F}_\tau$. Finally, let $A_1, A_2, \ldots \in \mathcal{F}_\tau$. Then $A_k \cap \{\tau = n\} \in \mathcal{F}_n$ and $(\bigcup_{k=1}^\infty A_k) \cap \{\tau = n\} \in \mathcal{F}_n$ and $\bigcup_{k=1}^\infty A_k \in \mathcal{F}_\tau$. □

Example 3.4.9. If $\tau \equiv n_0$, then $\mathcal{F}_\tau = \mathcal{F}_{n_0}$. In fact, by definition $A \in \mathcal{F}_\tau$ if and only if $A \cap \{\tau = n\} \in \mathcal{F}_n$ for $n = 0, 1, 2, ...$ Since $n \neq n_0$ immediately gives that $A \cap \{\tau = n\} = \emptyset \in \mathcal{F}_n$ the condition reduces to $n = n_0$ which means that $A \in \mathcal{F}_\tau$ if and only if $A \cap \{\tau = n_0\} \in \mathcal{F}_{n_0}$ which simply means that $A \in \mathcal{F}_{n_0}$.

Proposition 3.4.10. *Let τ be a stopping time. Then $A \in \mathcal{F}_\tau$ if and only if $A = A_\infty \cup \bigcup_{n=0}^\infty A_n$ with $A_n \in \mathcal{F}_n$ and $A_n \subseteq \{\tau = n\}$, and $A_\infty \in \mathcal{F}$ and $A_\infty \subseteq \{\tau = \infty\}$.*

Let us illustrate this by an example.

Example 3.4.11. Let $X_n = \varepsilon_1 + \ldots + \varepsilon_n$, $X_0 = 0$, and

$$\tau(\omega) = \inf\{n \geq 0 \,:\, X_n(\omega) \geq 2\}.$$

Then one has

$$\begin{aligned} A &= A \cap \bigcup_{k \in \{\infty, 0, 1, ...\}} \{\tau = k\} = \bigcup_{k \in \{\infty, 0, 1, ...\}} [A \cap \{\tau = k\}] \\ &=: \bigcup_{k \in \{\infty, 0, 1, ...\}} A_k \end{aligned}$$

with $\{\tau = 0\} = \{\tau = 1\} = \emptyset$ (after 0 or 1 steps one cannot have the value 2) and

$$\begin{aligned} \{\tau = 2\} &= \{\varepsilon_1 = \varepsilon_2 = 1\}, \\ \{\tau = 3\} &= \emptyset, \\ \{\tau = 4\} &= \{\varepsilon_1 = -1, \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = 1\} \cup \{\varepsilon_2 = -1, \varepsilon_1 = \varepsilon_3 = \varepsilon_4 = 1\} \cup \\ &\quad\; \{\varepsilon_3 = -1, \varepsilon_1 = \varepsilon_2 = \varepsilon_4 = 1\} \cup \{\varepsilon_4 = -1, \varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1\}, \\ &\quad\; \cdots \end{aligned}$$

We continue with some general properties.

Proposition 3.4.12. *Let* $X = (X_n)_{n=0}^{\infty}$ *be an adapted process and* $\tau : \Omega \to \{0, 1, \ldots\}$ *a stopping time. Then*

$$Z(\omega) := X_{\tau(\omega)}(\omega) : \Omega \to \mathbb{R}$$

is $\mathcal{F}_\tau$*-measurable.*

Proof. We have to show that

$$\{\omega \in \Omega : \ Z(\omega) \in B\} \in \mathcal{F}_\tau \text{ for all } B \in \mathcal{B}(\mathbb{R}).$$

By definition this is equivalent to

$$\{\omega \in \Omega : \ Z(\omega) \in B\} \cap \{\omega \in \Omega : \ \tau(\omega) = n\} \in \mathcal{F}_n,$$

or

$$\{\omega \in \Omega : \ X_{\tau(\omega)}(\omega) \in B, \, \tau(\omega) = n\} \in \mathcal{F}_n,$$

or

$$\{\omega \in \Omega : \ X_n(\omega) \in B, \, \tau(\omega) = n\} \in \mathcal{F}_n,$$

or

$$\{X_n \in B\} \cap \{\tau = n\} \in \mathcal{F}_n,$$

which is true. □

Proposition 3.4.13. *Let* $0 \le S \leqq T \le \infty$ *be stopping times. Then*

$$\mathcal{F}_S \subseteq \mathcal{F}_T.$$

Proof. We have that $A \in \mathcal{F}_S$ if and only if $A \cap \{S = n\} \in \mathcal{F}_n$ for $n = 0, 1, 2, \ldots$ Hence

$$A \cap \{T \leqq n\} = A \cap \{S \leqq n\} \cap \{T \leqq n\} \in \mathcal{F}_n$$

since $A \cap \{S \leqq n\} \in \mathcal{F}_n$ and $\{T \leqq n\} \in \mathcal{F}_n$. But this implies that $A \cap \{T = n\} \in \mathcal{F}_n$.

□

3.5 DOOB-decomposition

The DOOB[8] -decomposition of a sub-martingale consists in a martingale and in a predictable increasing process. In this way, we get two components which can be treated separately.

Theorem 3.5.1. *A process $Y = (Y_n)_{n=0}^{\infty}$ is a sub-martingale if and only if for $n = 0, 1, ...$ it holds*

$$Y_n = M_n + A_n$$

where $A = (A_n)_{n=0}^{\infty}$ and $M = (M_n)_{n=0}^{\infty}$ are processes such that

(i) $0 = A_0(\omega) \le A_1(\omega) \le \ldots$ *a.s.,*

(ii) A *is integrable and* A_n *is* $\mathcal{F}_{n-1}$*-measurable for* $n = 1, 2, \ldots$*,*

(iii) M *is a martingale.*

The process $A = (A_n)_{n=0}^{\infty}$ is unique in the following sense: if there are two such processes A and A', then $\mathbb{P}(A_n = A'_n) = 1$ for $n = 0, 1, 2, \ldots$

Proof. In order to deduce the formula for A_n let us assume that we already have the decomposition $Y_n = M_n + A_n$. We get that, a.s.,

$$\mathbb{E}(Y_{n+1} - A_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = M_n = Y_n - A_n$$

and

$$A_{n+1} = \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n + A_n.$$

Now let us go the other way round. Define

$$A_0 := 0 \quad \text{and} \quad A_{n+1} := \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n + A_n.$$

The properties of $A = (A_n)_{n=0}^{\infty}$ are that

- A_{n+1} is $\mathcal{F}_n$-measurable for $n \ge 0$,
- that $A_{n+1} - A_n = \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n \ge 0$ a.s. since $Y = (Y_n)_{n=0}^{\infty}$ is a sub-martingale, and that

[8]Joseph Leo Doob, 27/02/1910 (Cincinnati, USA)- 7/06/2004 (Clark-Lindsey Village, USA), DOOB's work was in probability and measure theory. Made major contributions in stochastic processes, martingales, optimal stopping, potential theory.

- the process M is a martingale because, a.s.,

$$\begin{aligned}\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) &= \mathbb{E}(Y_{n+1} - A_{n+1} \mid \mathcal{F}_n) \\ &= \mathbb{E}(Y_{n+1} - \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) + Y_n - A_n \mid \mathcal{F}_n) \\ &= Y_n - A_n \\ &= M_n.\end{aligned}$$

The uniqueness can be seen as follows: if $(A_n)_{n=0}^\infty$ and $(A'_n)_{n=0}^\infty$ are such processes, then, a.s.,

$$\begin{aligned}A_{n+1} &= \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n + A_n \\ A'_{n+1} &= \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n + A'_n\end{aligned}$$

with $A_0 = A'_0 \equiv 0$, which implies by induction that $A_n = A'_n$. □

Definition 3.5.2. The decomposition

$$Y_n = M_n + A_n$$

from Theorem 3.5.1 is called DOOB-decomposition.

One might think that the process $(A_n)_{n=0}^\infty$ measures how far a sub-martingale is from a martingale. Let us consider now some examples.

Example 3.5.3. Let $0 = Y_0 \le Y_1 \le Y_2 \le \ldots$ be real numbers. For any filtration we get a sub-martingale since $\mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) = Y_{n+1} \ge Y_n$ a.s. The DOOB-decomposition is then $M_n = 0$ a.s. and $A_n = Y_n$ a.s.

Example 3.5.4. If $Y = (Y_n)_{n=0}^\infty$ is a martingale, then $A_n = 0$ a.s.

Example 3.5.5. We consider independent random variables $\varepsilon_1^{(p)}, \varepsilon_2^{(p)}, ... : \Omega \to \mathbb{R}$ such that $\mathbb{P}(\varepsilon_n^{(p)} = -1) = p$ and $\mathbb{P}(\varepsilon_n^{(p)} = 1) = q$, where $0 < p, q < 1$ with $p + q = 1$. If $q > \frac{1}{2}$, then the process $Y = (Y_n)_{n=0}^\infty$ with $Y_0 = 0$ and

$$Y_n = \varepsilon_1^{(p)} + \ldots + \varepsilon_n^{(p)}$$

is a sub-martingale with respect to $\mathcal{F}_n := \sigma(\varepsilon_1^{(p)}, ..., \varepsilon_n^{(p)})$ for $n \ge 1$ and $\mathcal{F}_0 := \{\emptyset, \Omega\}$, and

$$A_{n+1} = \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) - Y_n + A_n = \mathbb{E}\varepsilon_{n+1} + A_n = (n+1)\mathbb{E}\varepsilon_1 = (n+1)(q-p).$$

Example 3.5.6. Using the same filtration and random variables as in the previous example we consider the exponential random walk $S_n = e^{\sigma \sum_{i=1}^n \varepsilon_i^{(p)} + cn}$ with $S_0 := 1$. If $p(e^{-\sigma} - e^{\sigma}) \geq e^{-c} - e^{\sigma}$ we get a sub-martingale and

$$\begin{aligned} A_{n+1} &= \mathbb{E}(S_{n+1} \mid \mathcal{F}_n) - S_n + A_n \\ &= S_n \mathbb{E}\left[e^{\sigma \varepsilon_{n+1}^{(p)} + c}\right] - S_n + A_n \\ &= S_n[\alpha - 1] + A_n \\ &= S_n[\alpha - 1] + S_{n-1}[\alpha - 1] + A_{n-1} \\ &= (\alpha - 1)(S_n + S_{n-1} + \cdots + S_0). \end{aligned}$$

Moreover, we have that $\alpha \geq 1$ if and only if $p(e^{-\sigma} - e^{\sigma}) \geq e^{-c} - e^{\sigma}$.

3.6 Optional stopping theorem

If $M = (M_n)_{n=0}^{\infty}$ is a martingale and $0 \leq S \leq T < \infty$ are integers, then

$$\mathbb{E}(M_T \mid \mathcal{F}_S) = M_S.$$

What is going to happen if $S \leq T$ are replaced by stopping times $\sigma, \tau : \Omega \to \{0, 1, 2, \ldots\}$? To get an idea, let us consider the following example.

Example 3.6.1. Our starting capital is 1 Euro and we assume an amount of c Euro, $c > 1$, we would like to win by the following game: if you have at time n an amount of N_n, we toss a fair coin for $2^n + |N_n|$ Euro, that means that

$$N_0 \equiv 1 \quad \text{and} \quad N_{n+1} := N_n + \varepsilon_{n+1}(2^n + |N_n|)$$
$$= \begin{cases} -2|N_n| - 2^n & \text{for } N_n < 0, \varepsilon_{n+1} = -1 \\ -2^n & \text{for } N_n > 0, \varepsilon_{n+1} = -1 \\ 2^n & \text{for } N_n < 0, \varepsilon_{n+1} = 1 \\ 2|N_n| + 2^n & \text{for } N_n > 0, \varepsilon_{n+1} = 1 \end{cases}$$

where $\varepsilon_1, \varepsilon_2, ...$ are independent symmetric BERNOULLI random variables. Letting also $\varepsilon_0 = 1$, we take $\mathcal{F}_n := \sigma(\varepsilon_0, ..., \varepsilon_n)$. We obtain

$$N_{n+1} \geqq 2^n \text{ if and only if } \varepsilon_{n+1} = 1,$$

this is what we want, and that the process $(N_n)_{n=0}^{\infty}$ is a martingale transform of $(\varepsilon_0 + \cdots + \varepsilon_n)_{n=0}^{\infty}$ since

$$N_{n+1} - N_n = (2^n + |N_n|)\varepsilon_{n+1}.$$

Now we introduce our stopping strategy $\tau : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ by

$$\tau(\omega) := \inf\{n \geq 0 \ : \ N_n(\omega) \geq c\}.$$

The properties of τ are summarized in

Proposition 3.6.2. *The random time τ is a stopping time with*

(i) $\mathbb{P}(\{\omega \in \Omega : \tau(\omega) < \infty\}) = 1$,

(ii) $N_{\tau(\omega)}(\omega) \geq c$ *on* $\{\tau < \infty\}$,

(iii) *for all* $T > 0$ *one has that* $\mathbb{P}(\{\omega \in \Omega : \tau(\omega) > T\}) > 0$.

Proof. The process $(N_n)_{n=0}^{\infty}$ is adapted and τ can be considered as the hitting time of the set $B = [c, \infty)$. Hence we have a stopping time.

(i) We observe that one has, for $n \geq 0$,

$$\mathbb{P}(N_{n+1} \geqq 2^n) = \mathbb{P}(\varepsilon_{n+1} = 1) = \frac{1}{2}.$$

Assuming now an $n_0 \geq 0$ with $2^{n_0} \geqq c$. It holds

$$\begin{array}{rcl} \{\varepsilon_{n_0+1} = 1\} & = & \{N_{n_0+1} \geq 2^{n_0}\} \subseteq \{N_{n_0+1} \geq c\} \\ \{\varepsilon_{n_0+1} = -1, ..., \varepsilon_{n_0+k} = -1, \varepsilon_{n_0+k+1} = 1\} & = & \{N_{n_0+1} < 0, ..., N_{n_0+k} < 0 N_{n_0+k+1} \geq 2^{n_0+k}\} \\ & \subseteq & \{N_{n_0+k+1} \geq c\}, \end{array}$$

and since

$$\begin{array}{rcl} \{\tau < \infty\} & = & \{\omega \in \Omega : \exists k \geq 0 : N_k(\omega) \geq c\} \\ & \supseteq & \bigcup_{k=n_0+1}^{\infty} \{N_k \geq c\} \end{array}$$

this gives that

$$\begin{array}{rcl} \mathbb{P}(\tau < \infty) & \geqq & \mathbb{P}(\varepsilon_{n_0+1} = 1) + \mathbb{P}(\varepsilon_{n_0+1} = -1, \varepsilon_{n_0+2} = 1) \\ & & + \mathbb{P}(\varepsilon_{n_0+1} = -1, \varepsilon_{n_0+2} = -1, \varepsilon_{n_0+3} = 1) + \cdots \\ & = & \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \cdots \\ & = & 1. \end{array}$$

(ii) follows by the very definition of τ.

(iii) We can take $\varepsilon_1 = \ldots = \varepsilon_T = -1$ and have for this path that

$$N_0(\omega) < c, \ldots, N_T(\omega) < c$$

so that $\mathbb{P}(\tau > T) \geq 2^{-T}$. □

From that we get

$$N_{\tau(\omega)}(\omega) \geq c \text{ a.s.} \quad \text{and} \quad \mathbb{E}(N_\tau \chi_{\{\tau < \infty\}}) \geq c$$

Hence, given two stopping times with $0 \leq \sigma \leq \tau < \infty$ a.s. and a martingale $M = (M_n)_{n=0}^\infty$ we cannot expect that

$$\mathbb{E}(M_\tau \mid \mathcal{F}_\sigma) = M_\sigma \text{ a.s..}$$

Where is the problem? The problem is that there is no $T > 0$ such that

$$\mathbb{P}(\{\omega \in \Omega : \tau(\omega) \leq T\}) = 1.$$

Proposition 3.6.3 (Optional stopping theorem)**.** *Let $Y = (Y_n)_{n=0}^\infty$ be a sub-martingale with respect to $(\mathcal{F}_n)_{n=0}^\infty$ and let $S, T : \Omega \to \{0, 1, 2, \ldots\}$ be stopping times such that*

$$S(\omega) \leqq T(\omega) \leqq T_0 < \infty$$

for all $\omega \in \Omega$ and some $T_0 > 0$. Then

$$\mathbb{E}(Y_T \mid \mathcal{F}_S) \geq Y_S \textit{ a.s.}$$

One has equality in the case that $Y = (Y_n)_{n=0}^\infty$ is a martingale.

Proof. (a) First we assume that Y is a martingale. We have to show that

$$\mathbb{E}(Y_T \mid F_S) = Y_S \text{ a.s.}$$

For $n \geqq 0$ we let

$$H_n(\omega) := \chi_{\{n \leqq T(\omega)\}} - \chi_{\{n \leqq S(\omega)\}} = \begin{cases} 0 & n \leqq S(\omega) \leqq T(\omega) \\ 0 & S(\omega) \leqq T(\omega) < n \\ 1 & S(\omega) < n \leqq T(\omega) \end{cases}$$

and get that $H_0(\omega) \equiv 0$ and that H_n is $\mathcal{F}_{n-1}$-measurable for $n \geqq 1$ since

$$\{H_n = 1\} = \{S < n\} \cap \{n \leq T\} = \{S \leq n-1\} \cap (\Omega \setminus \{T < n\}) \in \mathcal{F}_{n-1}$$

and

$$\{H_n = 0\} = \Omega \setminus \{H_n = 1\} \in \mathcal{F}_{n-1}.$$

Next we observe $Y_T - Y_S = \sum_{n=1}^m H_n(Y_n - Y_{n-1})$ for any fixed $m > T_0$ so that

$$\begin{aligned} \mathbb{E}(Y_T - Y_S) &= \sum_{n=1}^m \mathbb{E}(H_n(Y_n - Y_{n-1})) \\ &= \sum_{n=1}^m \mathbb{E}(\mathbb{E}(H_n(Y_n - Y_{n-1})|\mathcal{F}_{n-1})) \\ &= \sum_{n=1}^m \mathbb{E}(H_n \mathbb{E}((Y_n - Y_{n-1})|\mathcal{F}_{n-1})) \\ &= 0 \end{aligned}$$

because $\mathbb{E}((Y_n - Y_{n-1})|\mathcal{F}_{n-1}) = 0$ a.s. so that $\mathbb{E}Y_T = \mathbb{E}Y_S$. Let now $B \in \mathcal{F}_S$ and define the new random times

$$S_B := S\chi_B + m\chi_{B^c},$$

$$T_B := T\chi_B + m\chi_{B^c}.$$

The random times S_B and T_B are stopping times since, for example for S_B, one has

$$\{S_B = n\} = (B \cap \{S = n\}) \cup (B^c \cap \{m = n\})$$

where the first term belongs to $\mathcal{F}_n$ by the definition of the σ-algebra $\mathcal{F}_S$ and the second term belongs to $\mathcal{F}_n$ since it is equal the empty set if $m \neq n$ and $B^c \in \mathcal{F}_S \subseteq \mathcal{F}_m = \mathcal{F}_n$ if $m = n$. Hence we have stopping times with

$$0 \leqq S_B \leqq T_B \leqq m$$

so that

$$\mathbb{E}(Y_T\chi_B + Y_m\chi_{B^c}) = \mathbb{E}Y_{T_B} = \mathbb{E}Y_{S_B} = \mathbb{E}(Y_S\chi_B + Y_m\chi_{B^c})$$

and

$$\int_B Y_T d\mathbb{P} = \int_B Y_S d\mathbb{P}.$$

Together with Proposition 3.4.12 this implies that $\mathbb{E}(Y_T \mid \mathcal{F}_S) = Y_S$ a.s.

(b) Now we assume that we have a sub-martingale: we first apply the DOOB-decomposition and get that

$$Y_n = M_n + A_n.$$

By step (a) we have

$$\mathbb{E}(M_T \mid \mathcal{F}_S) = M_S \text{ a.s.},$$

so that, a.s.,

$$\begin{aligned} \mathbb{E}(Y_T \mid \mathcal{F}_S) &= \mathbb{E}(M_T + A_T \mid \mathcal{F}_S) \\ &= M_S + \mathbb{E}(A_T \mid \mathcal{F}_S) \\ &\geqq M_S + \mathbb{E}(A_S \mid \mathcal{F}_S) \\ &= Y_S. \end{aligned}$$

□

Example 3.6.1 (continued) Assume now that we have to stop our game at some time $T_0 > 0$. Our new strategy is then

$$T(\omega) := \min\{\tau(\omega), T_0\}$$

which is as the minimum of two stopping times again a stopping time. Moreover $T \leqq T_0$. Letting $S \equiv 0$, we have

$$0 \leqq S \leqq T \leqq T_0$$

and, by the optional stopping theorem, $\mathbb{E}(M_T|\mathcal{F}_S) = M_S$ a.s. so that

$$\mathbb{E}M_T = \mathbb{E}(\mathbb{E}(M_T|\mathcal{F}_S)) = \mathbb{E}M_S = 1.$$

Hence our unfair doubling strategy does not work anymore as before, we cannot become rich, if we only have finite time at our disposal.

3.7 Doob's maximal inequalities

DOOB's maximal inequality is a replacement of KOLMOGOROV's inequality in Lemma 2.2.11. In the following we also work with (sub-, super-) martingales $X =$

$(X_n)_{n=0}^N$ over the finite time scale $\{0,...,N\}$. Moreover, given $X = (X_n)_{n=0}^N$ or $X = (X_n)_{n=0}^\infty$, we need the *maximal functions*

$$X_k^*(\omega) := \sup_{0\le n\le k} |X_n(\omega)| \quad \text{and} \quad X^*(\omega) := \sup_{n=0,1,...} |X_n(\omega)|,$$

where the latter is used for $X = (X_n)_{n=0}^\infty$ only.

Proposition 3.7.1 (DOOB's maximal inequalities). *Let $X = (X_n)_{n=0}^N$ be a martingale or non-negative sub-martingale. Then the following holds true:*

(i) *For $\lambda \ge 0$ one has*

$$\lambda \mathbb{P}\left(X_N^* \ge \lambda\right) \le \mathbb{E}\chi_{\{X_N^*\ge\lambda\}}|X_N|.$$

(ii) *For $p \in (1,\infty)$ one has*

$$\|X_N^*\|_{\mathcal{L}_p} \le \frac{p}{p-1}\|X_N\|_{\mathcal{L}_p}.$$

Proof. If $X = (X_n)_{n=0}^N$ is a martingale, then $(|X_n|)_{n=0}^N$ is a non-negative sub-martingale, so that it remains to consider this case only.

(i) For $\lambda \ge 0$ we define the stopping times

$$\sigma(\omega) := \inf\{n : X_n \ge \lambda\} \wedge N \quad \text{and} \quad \tau(\omega) := N$$

so that $0 \le \sigma \le \tau \le N$. By the Optional Stopping Theorem we may deduce

$$X_\sigma \le \mathbb{E}(X_\tau \mid \mathcal{F}_\sigma) = \mathbb{E}(X_N \mid \mathcal{F}_\sigma) \text{ a.s.}$$

and

$$\begin{aligned}
\mathbb{E}X_N &\ge \mathbb{E}X_\sigma = \mathbb{E}(X_\sigma \chi_{\{X_N^*\ge\lambda\}}) + \mathbb{E}(X_\sigma \chi_{\{X_N^*<\lambda\}}) \\
&\ge \mathbb{E}(\lambda \chi_{\{X_N^*\ge\lambda\}}) + \mathbb{E}(X_N \chi_{\{X_N^*<\lambda\}})
\end{aligned}$$

so that $\lambda\mathbb{P}(X_N^* \ge \lambda) \le \mathbb{E}X_N\chi_{\{X_N^*\ge\lambda\}}$.

(ii) Let $c > 0$ be fixed. Then

$$\begin{aligned}
\mathbb{E}((X_N^* \wedge c)^p) &= \mathbb{E}\int_0^{(X_N^*\wedge c)} p\lambda^{p-1} d\lambda \\
&= \mathbb{E}\int_0^c p\lambda^{p-1}\chi_{\{X_N^*\ge\lambda\}} d\lambda \\
&= \int_0^c p\lambda^{p-1}\mathbb{P}(X_N^* \ge \lambda) d\lambda.
\end{aligned}$$

By step (i) we continue with

$$\begin{aligned}
\mathbb{E}((X_N^* \wedge c)^p) &\leq \int_0^c p\lambda^{p-2}\mathbb{E}(|X_N|\chi_{\{X_N^* \geq \lambda\}})d\lambda \\
&= p\mathbb{E}[|X_N| \int_0^{X_N^* \wedge c} \lambda^{p-2} d\lambda] \\
&= \frac{p}{p-1}\mathbb{E}[|X_N|(X_N^* \wedge c)^{p-1}] \\
&\leq \frac{p}{p-1}(\mathbb{E}|X_N|^p)^{\frac{1}{p}}(\mathbb{E}((X_N^* \wedge c)^{p^*(p-1)})^{\frac{1}{p^*}}
\end{aligned}$$

with $1 = \frac{1}{p} + \frac{1}{p^*}$. Because of $p^*(p-1) = p$ this implies that

$$\mathbb{E}((X_N^* \wedge c)^p)^{1-\frac{1}{p^*}} \leq \frac{p}{p-1}(\mathbb{E}|X_N|^p)^{\frac{1}{p}}.$$

By $c \uparrow \infty$ we conclude the proof. □

Now we reprove Lemma 2.2.11.

Corollary 3.7.2 (Inequality of KOLMOGOROV)**.** *Let $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ be a sequence of independent mean zero and quadratic integrable random variables. If $M_n := \xi_1 + \cdots + \xi_n$ for $n \geq 1$ and $M_0 := 0$, then one has that*

$$\mathbb{P}\left(M_n^* \geq \lambda\right) \leq \frac{\mathbb{E}M_n^2}{\lambda^2}$$

for all $\lambda > 0$.

Proof. Applying Proposition 3.3.3 we get that $(M_n^2)_{n=0}^\infty$ is a sub-martingale. Hence we may apply Proposition 3.7.1 and get that

$$\lambda\mathbb{P}\left((M_n^*)^2 \geq \lambda\right) \leq \mathbb{E}M_n^2$$

or

$$\lambda^2\mathbb{P}\left(M_n^* \geq \lambda\right) \leq \mathbb{E}M_n^2.$$

□

Corollary 3.7.3 (Exit probability of a random walk)**.** *Let* $\varepsilon_1, \varepsilon_2, \ldots : \Omega \to \mathbb{R}$ *be independent* BERNOULLI *random variables,* $M_0 :\equiv 0$*, and* $M_n = \varepsilon_1 + \ldots + \varepsilon_n$ *for* $n = 1, 2, ...$ *Then, for* $\lambda \geq 0$ *and* $n = 1, 2, ...$*,*

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} M_n^* \geq \lambda\right) \leq 2e^{-\frac{\lambda^2}{2}}.$$

Proof. We can assume $\lambda > 0$. We define the stochastic exponential

$$Z_n := e^{\alpha M_n - \frac{\beta^2}{2} n}$$

where $\alpha, \beta > 0$ are fixed later. The process $Z = (Z_n)_{n=0}^\infty$ is a positive martingale if $\frac{1}{2}(e^{-\alpha} + e^{\alpha}) = e^{\frac{\beta^2}{2}}$ under the natural filtration $(\mathcal{F}_n)_{n=0}^\infty$ given by $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(\varepsilon_1, \ldots, \varepsilon_n)$ for $n = 1, 2, ...$ In fact, we have

$$\mathbb{E}(Z_{n+1} \mid \mathcal{F}_n) = Z_n$$

if and only if

$$\mathbb{E}(e^{\alpha \varepsilon_{n+1} - \frac{\beta^2}{2}}) = 1,$$

or

$$\tfrac{1}{2}(e^{-\alpha} + e^{\alpha}) = e^{\frac{\beta^2}{2}}.$$

Hence we are in a position to apply DOOB's maximal inequality and obtain

$$\begin{aligned}
\mathbb{P}\left(\sup_{0 \leq k \leq n} M_k \geq \lambda\right) &= \mathbb{P}\left(e^{\alpha \sup_{0 \leq k \leq n} M_k - \frac{\beta^2}{2} n} \geq e^{\alpha \lambda - \frac{\beta^2}{2} n}\right) \\
&\leq \mathbb{P}\left(e^{\sup_{0 \leq k \leq n}\left(\alpha M_k - \frac{\beta^2}{2} k\right)} \geq e^{\alpha \lambda - \frac{\beta^2}{2} n}\right) \\
&= \mathbb{P}\left(Z_n^* \geq e^{\alpha \lambda - \frac{\beta^2}{2} n}\right) \\
&\leq e^{-\alpha \lambda + \frac{\beta^2}{2} n} \mathbb{E} Z_n \\
&= e^{-\alpha \lambda + \frac{\beta^2}{2} n}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}\left(\tfrac{1}{\sqrt{n}} \sup_{0 \leq k \leq n} M_k \geq \lambda\right) &\leq e^{-\alpha \sqrt{n} \lambda + \frac{\beta^2}{2} n} = e^{-\alpha \sqrt{n} \lambda} \left(\frac{e^{-\alpha} + e^{\alpha}}{2}\right)^n \\
&\leq e^{-\alpha \sqrt{n} \lambda} e^{\frac{\alpha^2}{2} n} = e^{\frac{\alpha^2}{2} n - \alpha \sqrt{n} \lambda},
\end{aligned}$$

where the inequality $e^{-\alpha} + e^{\alpha} \le 2e^{\frac{\alpha^2}{2}}$ can be verified by a TAYLOR[9] -expansion of both sides and the comparison of the expansions. Finally, we have to find the minimum of

$$f(\alpha) := \tfrac{\alpha^2}{2}n - \alpha\sqrt{n}\lambda.$$

Since $f'(\alpha) = \alpha n - \sqrt{n}\lambda$, we get $\alpha_0 = \frac{\lambda}{\sqrt{n}}$ and

$$\mathbb{P}\left(\tfrac{1}{\sqrt{n}} \sup_{0\le k\le n} M_k \ge \lambda\right) \le e^{\frac{\lambda^2}{2}-\lambda^2} = e^{-\frac{\lambda^2}{2}}.$$

We finish the proof by

$$\mathbb{P}\left(\tfrac{M_n^*}{\sqrt{n}} \ge \lambda\right) \le \mathbb{P}\left(\tfrac{1}{\sqrt{n}} \sup_{0\le k\le n} M_k \ge \lambda\right) + \mathbb{P}\left(\tfrac{1}{\sqrt{n}} \sup_{0\le k\le n} (-M_k) \ge \lambda\right) \le 2e^{\frac{-\lambda^2}{2}}.$$

□

3.8 Uniformly integrable martingales

In this section we consider uniformly integrable martingales. Let us recall the definition of *uniformly integrable*.

Definition 3.8.1. A family $(f_i)_{i\in I}$ of random variables $f_i : \Omega \to \mathbb{R}$ is called *uniformly integrable* provided that for all $\varepsilon > 0$ there exists a $c > 0$ such that

$$\sup_{i\in I} \int_{|f_i|\ge c} |f_i| d\mathbb{P} \le \varepsilon.$$

An equivalent condition to the above one is

$$\lim_{c\to\infty} \sup_{i\in I} \int_{|f_i|\ge c} |f_i| d\mathbb{P} = 0.$$

Moreover, we need the following criteria for uniform integrability:

[9]Brook Taylor, 18/08/1685 (Edmonton, England) - 29/12/1731 (Somerset House, England), 'calculus of finite differences', integration by parts, discovered the formula known as Taylor's expansion.

Lemma 3.8.2. *Let $G : [0,\infty) \to [0,\infty)$ be non-negative and increasing such that*

$$\lim_{t\to\infty} \frac{G(t)}{t} = \infty$$

and $(f_i)_{i\in I}$ be a family of random variables $f_i : \Omega \to \mathbb{R}$ such that

$$\sup_{i\in I} \mathbb{E}G(|f_i|) < \infty.$$

Then $(f_i)_{i\in I}$ is uniformly integrable.

Proof. We let $\varepsilon > 0$ and $M := \sup_{i\in I} \mathbb{E}G(|f_i|)$ and find a $c > 0$ such that

$$\frac{M}{\varepsilon} \le \frac{G(t)}{t}$$

for $t \ge c$. Then

$$\int_{|f_i|\ge c} |f_i| d\mathbb{P} \le \frac{\varepsilon}{M} \int_{|f_i|\ge c} G(|f_i|) d\mathbb{P} \le \varepsilon.$$

□

Corollary 3.8.3. *Let $p \in (1,\infty)$ and $X = (X_n)_{n=0}^\infty$ be a martingale or a non-negative sub-martingale with $\sup_{n=0,1,\ldots} \|X_n\|_{L_p} < \infty$. Then $(X_n)_{n=0}^\infty$ is uniformly integrable and $\mathbb{E}\sup_n |X_n|^p < \infty$.*

Proof. We apply Lemma 3.8.2 with $G(t) = t^p$ and DOOB's maximal inequality. □

What can we deduce from that? For example, in applications it is often important to know whether a martingale $M = (M_t)_{t\ge 0}$ has a limit as $t \to \infty$ and in what sense the convergence takes place.

Definition 3.8.4. A martingale $(M_n)_{n=0}^\infty$ is *closable* if there exists an $M_\infty : \Omega \to \mathbb{R}$ with $\mathbb{E}|M_\infty| < \infty$ such that, for $n = 0, 1, 2, \ldots$,

$$M_n = \mathbb{E}(M_\infty \mid \mathcal{F}_n) \text{ a.s.}$$

Does every martingale have a closure? The answer is *NO* as shown by

Example 3.8.5. Let $\Omega = [0,1)$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \sigma([0,\frac{1}{2}))$, $\mathcal{F}_2 = \sigma([0,\frac{1}{4}),[\frac{1}{4},\frac{1}{2}))$, $\mathcal{F}_3 = \sigma([0,\frac{1}{8}),[\frac{1}{8},\frac{1}{4}),[\frac{1}{4},\frac{1}{2}))$, $\ldots$. $M_0 \equiv 1$, $M_1 = 2\chi_{[0,\frac{1}{2})}$, $M_2 = 2^2\chi_{[0,\frac{1}{2^2})}$, $\ldots$, $M_n = 2^n\chi_{[0,\frac{1}{2^n})}$,... The process $(M_n)_{n=0}^\infty$ is a martingale, which is not closable. Assume it would be closable. Then

$$\int_{\frac{1}{2^n}}^1 M_\infty(t)dt = \int_{\frac{1}{2^n}}^1 M_n(t)dt = 0,\ n = 1,2,\ldots$$

so that $\int_{\frac{1}{2^n}}^1 M_\infty(t)dt = 0$ for $n = 1,2,...$ and $\int_\Omega M_\infty(t)dt = \int_{(0,1)} M_\infty(t)dt = 0$ which is a contradiction as $\int_{(0,1)} M_0(t)dt = 1$.

Proposition 3.8.6. *Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$ be a stochastic basis with*

$$\mathcal{F} = \sigma\left(\bigcup_{n=0}^\infty \mathcal{F}_n\right).$$

and $M = (M_n)_{n=0}^\infty$ be a martingale. Then the following assertions are equivalent:

(i) *The sequence $(M_n)_{n=0}^\infty$ is a Cauchy-sequence in L_1, that means for all $\varepsilon > 0$ there exists an $n_0 \geq 0$ such that for all $m, n \geq n_0$ one has that*

$$\|M_m - M_n\|_{L_1} < \varepsilon.$$

(ii) *There exists an $Z \in L_1$ such that $\lim_{n\to\infty} M_n = Z$ in L_1 that means $\lim_{n\to\infty} \|M_n - Z\|_{L_1} = 0$.*

(iii) *There exists an $M_\infty \in L_1$ such that $M_n = \mathbb{E}(M_\infty \mid \mathcal{F}_n)$ a.s. for $n \geq 0$.*

(iv) *The family $M = (M_n)_{n\geq 0}$ is uniformly integrable.*

If these equivalent conditions hold, then

(a) $Z = M_\infty$ *a.s.,*

(b) $M_\infty = \lim_{n\to\infty} M_n$ *a.s.*

(c) *If $\sup_n \mathbb{E}|M_n|^p < \infty$ for some $1 < p < \infty$, then*

$$\lim_{n\to\infty} \|M_n - M_\infty\|_{L_p} = 0.$$

Remark 3.8.7. *If* $\sup_n \mathbb{E}|M_n|^p < \infty$ *for some* $1 < p < \infty$*, then the assertions (i) - (iv) hold true, and we also have* $\lim_{n\to\infty} \|M_n - M_\infty\|_{L_p} = 0$.

Part I of the proof of Proposition 3.8.6:

(i)$\Longleftrightarrow$ (ii) follows from the completeness of L_1 which means that a sequence of functions $(f_n)_{n=0}^\infty \in \mathcal{L}_1$ is a Cauchy sequence if and only if there exists $f_\infty \in \mathcal{L}_1$ with $\lim_{n\to\infty} f_n = f_\infty$ in $\mathcal{L}_1$.

(ii)$\Longrightarrow$ (iii) We show that $M_n = \mathbb{E}(Z \mid \mathcal{F}_n)$ a.s. In fact, for $m \geq n$ we get that

$$\begin{aligned}
\|M_n - \mathbb{E}(Z \mid \mathcal{F}_n)\|_{\mathcal{L}_1} &= \|M_n - \mathbb{E}(Z - M_m + M_m \mid \mathcal{F}_n)\|_{\mathcal{L}_1} \\
&= \|M_n - \mathbb{E}(Z - M_m \mid \mathcal{F}_n) - \mathbb{E}(M_m \mid \mathcal{F}_n)\|_{\mathcal{L}_1} \\
&= \|\mathbb{E}(Z - M_m \mid \mathcal{F}_n)\|_{\mathcal{L}_1} \\
&= \int_\Omega |\mathbb{E}(Z - M_m \mid \mathcal{F}_n)| d\mathbb{P} \\
&\leq \int_\Omega \mathbb{E}(|Z - M_m| \mid \mathcal{F}_n) d\mathbb{P} \\
&= \int_\Omega |Z - M_m| d\mathbb{P}.
\end{aligned}$$

Since $\lim_m \int_\Omega |Z - M_m| d\mathbb{P} = 0$ we conclude that $\mathbb{E}|M_n - \mathbb{E}(Z \mid \mathcal{F}_n)| = 0$ and $M_n = \mathbb{E}(Z \mid \mathcal{F}_n)$ a.s.

(iii)$\Longrightarrow$ (iv) For $c > 0$ we get that

$$\begin{aligned}
\int_{\{|M_n|>c\}} |M_n| d\mathbb{P} &= \int_{\{|M_n|>c\}} |\mathbb{E}(M_\infty \mid \mathcal{F}_n)| d\mathbb{P} \\
&\leq \int_{\{|M_n|>c\}} \mathbb{E}(|M_\infty| \mid \mathcal{F}_n) d\mathbb{P} \\
&= \int_{\{|M_n|>c\}} |M_\infty| d\mathbb{P} \\
&\leq \int_{\{\sup_m |M_m|>c\}} |M_\infty| d\mathbb{P}.
\end{aligned}$$

Since for $c > 0$ DOOB's inequality gives

$$\begin{aligned}
\mathbb{P}\left(\sup_{m\geq 0}|M_m| > c\right) &= \sup_{N\geq 1}\mathbb{P}\left(\sup_{0\leq m\leq N}|M_m| > c\right) \\
&\leq \sup_{N\geq 1}\frac{1}{c}\mathbb{E}|M_N| \\
&= \sup_{N\geq 1}\frac{1}{c}\mathbb{E}|\mathbb{E}(M_\infty|\mathcal{F}_N)| \\
&\leq \sup_{N\geq 1}\frac{1}{c}\mathbb{E}\mathbb{E}(|M_\infty||\mathcal{F}_N) \\
&= \frac{1}{c}\mathbb{E}|M_\infty| \to 0
\end{aligned}$$

as $c \to \infty$, we are done. □

Now we interrupt the proof to prepare the implication $(iv) \Rightarrow (ii)$. We need the notion of *down-crossings of a function*: let $I \subseteq \mathbb{R}$, $f : I \to \mathbb{R} \cup \{-\infty, \infty\}$, $-\infty < a < b < \infty$, and $F \subseteq I$ with $F = \{t_1, \ldots, t_d\}$ and $t_1 < t_2 < \ldots < t_d$. Then we define

$$\begin{aligned}
s_1 &:= \inf\{t_i \,:\, f(t_i) > b\}, \\
s_2 &:= \inf\{t_i > s_1 \,:\, f(t_i) < a\}, \\
&\vdots \\
s_{2n+1} &:= \inf\{t_i > s_{2n} \,:\, f(t_i) > b\}, \\
s_{2n+2} &:= \inf\{t_i > s_{2n+1} \,:\, f(t_i) < a\},
\end{aligned}$$

where we use $\inf \emptyset := t_d$, and

$$D_0(f, F, [a, b]) := \#\{n \,:\, s_{2n} < t_d\}.$$

Definition 3.8.8.

$$D(f, I, [a, b]) := \sup\{D_0(f, F, [a, b]) \,:\, F \subseteq I \text{ finite}\}$$

is the *number of down-crossings* of f over $[a, b]$.

Proposition 3.8.9. *Let $X = (X_n)_{n=0}^\infty$ be a sub-martingale and let $I \subseteq \{0, 1, 2, ...\}$ be non-empty. Then one has for all $-\infty < a < b < \infty$ that*

$$(b - a)\mathbb{E}(D(X, I, [a, b])) \leqq \sup_{n\in I}\mathbb{E}(X_n - b)^+.$$

Proof. Applying monotone convergence we see that it is enough to prove that

$$(b-a)\mathbb{E}(D_0(X,F,[a,b])) \leqq \sup_{n\in F} \mathbb{E}(X_n-b)^+$$

for a non-empty finite subset $F \subseteq I$. By changing the notation we can, without loss of generality, assume that $F = \{0,\dots,N\}$. Then the sequence $s_i : \Omega \to T$, $i = 1,2,...$, is a sequence of stopping times

$$s_1(\omega) \le s_2(\omega) \le s_3(\omega) \le \cdots \le N$$

with respect to the filtration $(\mathcal{F}_n)_{n=0}^N$ for some $N \ge 1$ (for $N = 1$ the number of down-crossings is zero). We let $s_i(\omega)$ be the previously defined sequence, now for the random function $n \to f_\omega(n) := X_n(\omega)$. According to the optional stopping theorem, the Proposition 3.6.3, the process $(X_{s_k})_{k=1}^\infty$ is now a sub-martingale with respect to the filtration $(\mathcal{F}_{s_k})_{k=1}^\infty$. We define

$$A_k := \{s_k < N\} \in \mathcal{F}_{s_k}$$

and get a non-increasing sequence

$$A_1 \supseteq A_2 \supseteq \cdots \supseteq A_{2n-1} \supseteq A_{2n} \supseteq \cdots$$

By the sub-martingale property of $(X_{s_k})_{k=1}^\infty$ we obtain that

$$\begin{aligned} 0 &\le \int_{A_{2n-1}} (X_{s_{2n-1}} - b)d\mathbb{P} \\ &\le \int_{A_{2n-1}} (X_{s_{2n}} - b)d\mathbb{P} \\ &= \int_{A_{2n}} (X_{s_{2n}} - b)d\mathbb{P} + \int_{A_{2n-1}\setminus A_{2n}} (X_{s_{2n}} - b)d\mathbb{P} \\ &\le (a-b)\mathbb{P}(A_{2n}) + \int_{A_{2n-1}\setminus A_{2n}} (X_{s_{2n}} - b)d\mathbb{P} \end{aligned}$$

because on A_{2n-1} we have $X_{s_{2n-1}} > b$ and on A_{2n} we have that $X_{s_{2n}} < a$. Hence

$$(b-a)\mathbb{P}(A_{2n}) \le \int_{A_{2n-1}\setminus A_{2n}} (X_{s_{2n}} - b)d\mathbb{P}.$$

Since $s_{2n} = N$ on A_{2n}^c this implies that

$$(b-a)\mathbb{P}(A_{2n}) \le \int_{A_{2n-1}\setminus A_{2n}} (X_N - b)d\mathbb{P}.$$

Summing over n one gets

$$(b-a)\sum_{n=1}^{\infty}\mathbb{P}(D_0(X,F,[a,b]) \geqq n) = (b-a)\sum_{n=1}^{\infty}\mathbb{P}(A_{2n}) \leq \int_{\Omega}(X_N-b)^+ d\mathbb{P}.$$

□

Corollary 3.8.10. *Let $(X_n)_{n=0}^{\infty}$ be a sub-martingale such that*

$$\sup_n \mathbb{E}X_n^+ < \infty.$$

Then there exists a random variable $X_\infty : \Omega \to \mathbb{R}$ such that

$$\lim_n X_n = X_\infty \text{ a.s.}$$

In particular, every non-negative martingale converges almost surely.

Proof. FATOU's lemma implies that

$$\mathbb{E}\liminf_n X_n^+ \leq \liminf_n \mathbb{E}X_n^+ \leq \sup_n \mathbb{E}X_n^+ < \infty$$

so that

$$\liminf_n X_n^+ < \infty \text{ a.s.} \quad \text{and} \quad \liminf_n X_n < \infty \text{ a.s.}$$

Assuming

$$\mathbb{P}(\liminf_n X_n < \limsup_n X_n) > 0$$

implies the existence of $-\infty < a < b < \infty$ such that

$$\mathbb{P}(\liminf_n X_n < a < b < \limsup_n X_n) > 0$$

and that

$$\mathbb{P}(D((X_n)_{n=0}^{\infty}, \{0,1,2,\ldots\}, [a,b]) = \infty) > 0.$$

But this is a contradiction to Proposition 3.8.9 so that

$$\mathbb{P}(\liminf_n X_n = \limsup_n X_n) = 1$$

and $Z := \lim_{n\to\infty} X_n$ exists almost surely as an extended random variable with values in $\mathbb{R} \cup \{-\infty, \infty\}$. Since $\liminf_n X_n < \infty$ a.s. we get that $Z < \infty$ a.s.

It remains to show that $Z > -\infty$ a.s. First we observe that

$$\mathbb{E}X_n^+ - \mathbb{E}X_n^- = \mathbb{E}X_n \geqq \mathbb{E}X_1 = \mathbb{E}X_1^+ - \mathbb{E}X_1^-$$

so that

$$\mathbb{E}X_n^- \leq \mathbb{E}X_1^- + \mathbb{E}X_n^+ - \mathbb{E}X_1^+ \leqq \mathbb{E}X_1^- + \sup_{m\geq 1} \mathbb{E}X_m^+ < \infty.$$

This implies as above that $\liminf_n(-X_n) < \infty$ a.s. and $\limsup_n X_n > -\infty$ a.s. so that $Z > -\infty$ a.s. □

Proposition 3.8.11. *Let $(f_n)_{n=0}^\infty \subseteq \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ be uniformly integrable and $f_\infty : \Omega \to \mathbb{R}$ be a random variable such that $f_\infty = \lim_{n\to\infty} f_n$ a.s. Then one has that*

(i) $f_\infty \in L_1(\Omega, \mathcal{F}, \mathbb{P})$,

(ii) $\lim_{n\to\infty} \mathbb{E}|f_\infty - f_n| = 0$.

Proof. (i) Let $\varepsilon > 0$ and find $c > 0$ such that

$$\int_{|f_n|\geq c} |f_n| d\mathbb{P} \leq \varepsilon$$

for $n = 0, 1, 2, ...$ Then $\mathbb{E}|f_n| \leq c + \varepsilon$ and $\sup_n \mathbb{E}|f_n| < \infty$. The lemma of FATOU gives that

$$\mathbb{E}|f_\infty| = \mathbb{E}\lim_n |f_n| \leq \liminf_n \mathbb{E}|f_n| \leq c + \varepsilon.$$

(ii) Again we use $\int_{|f_n|\geq c} |f_n| d\mathbb{P} \leq \varepsilon$ for $n \geq 0$ and define

$$g^c := \begin{cases} c & g \geq c \\ g & |g| < c \\ -c & g \leq -c \end{cases}.$$

We get

$$\begin{aligned} \mathbb{E}|f_n - f_\infty| &\leq \mathbb{E}|f_n - f_n^c| + \mathbb{E}|f_n^c - f_\infty^c| + \mathbb{E}|f_\infty^c - f_\infty| \\ &\leq \varepsilon + \mathbb{E}|f_n^c - f_\infty^c| + \mathbb{E}|f_\infty^c - f_\infty| \end{aligned}$$

so that

$$\limsup_n \mathbb{E}|f_n - f_\infty| \le \varepsilon + \mathbb{E}|f_\infty^c - f_\infty|.$$

By $\varepsilon \downarrow 0$ this improves to

$$\limsup_n \mathbb{E}|f_n - f_\infty| \le \mathbb{E}|f_\infty^c - f_\infty|.$$

We conclude by $c \uparrow \infty$ which gives $\limsup_n \mathbb{E}|f_n - f_\infty| = 0$. □

Part II of the proof of Proposition 3.8.6:
$(iv) \Rightarrow (ii)$ follows from Corollary 3.8.10 and Proposition 3.8.11.
(a) Let $\mathcal{A}$ be the system of all $A \in \mathcal{F}$ such that

$$\int_A (M_\infty - Z)d\mathbb{P} = 0.$$

The system $\mathcal{A}$ is a λ-system, that means

- $\Omega \in \mathcal{A}$,
- if $A, B \in \mathcal{A}$ with $A \subseteq B$, then $B \setminus A \in \mathcal{A}$,
- if $A_1 \subseteq A_2 \subseteq \cdots$ and $A_1, A_2, ... \in \mathcal{A}$, then $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$.

From the implication $(ii) \Rightarrow (iii)$ we get that

$$\mathbb{E}(M_\infty|\mathcal{F}_n) = M_n = \mathbb{E}(Z|\mathcal{F}_n) \text{ a.s.}$$

so that $\mathcal{B} := \bigcup_{n=1}^\infty \mathcal{F}_n \subseteq \mathcal{A}$. Since $\mathcal{B}$ is a π-system (closed under intersections) which generates $\mathcal{F}$, the π-λ-Theorem implies $\sigma(\mathcal{B}) \subseteq \mathcal{A}$. Hence $\mathcal{A} = \mathcal{F}$ which gives $Z = M_\infty$ a.s.

(b) From Corollary 3.8.10 we know that $\lim_n M_n$ exists almost surely. At same time we know that there is a subsequence such that $\lim_k M_{n_k} = Z$ a.s. Hence $\lim_n M_n = Z$ a.s. Applying (a) gives the assertion of (b).

(c) By DOOB's maximal inequality we have

$$\mathbb{E}(M^*)^p \le \left(\frac{p}{p-1}\right)^p \sup_n \mathbb{E}|M_n|^p < \infty.$$

Hence

$$\lim_n \mathbb{E}|M_n - M_\infty|^p = \mathbb{E} \lim_n |M_n - M_\infty|^p = 0$$

by LEBESGUE's Dominated Convergence Theorem. □

3.9 Applications

3.9.1 A martingale proof for KOLMOGOROV's zero-one law

Alternative proof of Proposition 2.1.6: Let $\xi_k : \Omega \to \mathbb{R}$ be independent random variables, $\mathcal{F}_n^\infty := \sigma(\xi_n, \xi_{n+1}, \ldots)$, $\mathcal{F}^\infty = \bigcap_{n=1}^\infty \mathcal{F}_n^\infty$, $A \in \mathcal{F}^\infty$, $M_\infty := \chi_A$, and $M_n := \mathbb{E}(M_\infty \mid \mathcal{F}_n)$ with $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(\xi_1, ..., \xi_n)$. By Proposition 3.8.6 we have that $(M_n)_{n=0}^\infty$ is uniformly integrable and that

$$\lim_{n\to\infty} M_n = M_\infty \text{ a.s. and in } L_1.$$

Since

$$M_n = \mathbb{E}(M_\infty \mid \mathcal{F}_n) = \mathbb{E}M_\infty \text{ a.s.}$$

we get that

$$M_\infty = \mathbb{E}M_\infty \text{ a.s.}$$

Since M_∞ takes the values 0 or 1 only we may deduce that

$$\mathbb{P}(A) = \mathbb{E}M_\infty \in \{0, 1\}.$$

□

3.9.2 The branching process example

The key-point is that we use Corollary 3.8.10 to get a limit of our re-normalized branching process.

Proof of Proposition 1.2.1: Corollary 3.8.10 says that there exists a limit $M_\infty = \lim_{n\to\infty} M_n$ a.s. Let

$$\pi_n := \mathbb{P}(M_n = 0)$$

for $n = 0, 1, ...$ Since by construction $M_n(\omega) = 0$ implies that $M_{n+1}(\omega) = 0$, we get a bounded increasing sequence

$$0 = \pi_0 < q_0 = \pi_1 \le \pi_2 \le \pi_3 \le \cdots \le 1$$

with the limit $\alpha := \lim_n \pi_n$. We shall investigate the value of α by a functional method. For that purpose we let the generating function of f_n, $n = 0, 1, ...$, be

$$\Phi_n(\theta) := \sum_{k=0}^{\infty} \theta^k \mathbb{P}(f_n = k) = \mathbb{E}\theta^{f_n}, \theta \in [0,1],$$

and, in particular,

$$\varphi(\theta) := \Phi_1(\theta) = \sum_{k=0}^{N} q_k \theta^k$$

the generating function of the off-spring distribution. Then

$$\Phi_{n+1} = \Phi_n \cdot \Phi_1 = \Phi_1 \cdots \Phi_1 \cdot \Phi_1$$

since

$$\begin{aligned}
\Phi_{n+1}(\theta) &= \mathbb{E}\theta^{f_{n+1}} \\
&= \sum_{k=0}^{\infty} \mathbb{P}(f_n = k) \mathbb{E}\theta^{X_1^{(n+1)} + \cdots + X_k^{(n+1)}} \\
&= \sum_{k=0}^{\infty} \mathbb{P}(f_n = k) \left(\mathbb{E}\theta^{X_1^{(n+1)}} \right)^k \\
&= \Phi_n \left(\mathbb{E}\theta^{X_1^{(n+1)}} \right) \\
&= \Phi_n \cdot \Phi_1(\theta)
\end{aligned}$$

with the convention that $0^0 := 1$ and 'empty' sums are treated as zero. We get that

(i) $0 < q_0 = \varphi(0) < \varphi(1) = 1$,

(ii) $\pi_{n+1} = \varphi(\pi_n)$,

(iii) $\mu = \varphi'(1)$,

(iv) $\varphi(\alpha) = \alpha$,

Assertion (i) is clear by our assumptions. Part (ii) follows from $\pi_n = \Phi_n(0)$ and

$$\pi_{n+1} = \Phi_{n+1}(0) = \varphi\left(\Phi_n(0)\right) = \varphi(\pi_n).$$

To check (iii) we see that

$$\varphi'(\theta) = \sum_{k=1}^{N} q_k k \theta^{k-1} \quad \text{and} \quad \mu = \sum_{k=1}^{N} q_k k = \varphi'(1).$$

Finally, (iv) follows from (ii), $\lim_n \pi_n = \alpha$, and the continuity of φ. Now we consider the different cases for the expected number of sons μ.

(a) $0 < \mu \leq 1$: The only fix-point of φ is $\theta = 1$ so that $\alpha = 1$. Hence $\lim_{n\to\infty} \mathbb{P}(f_n = 0) = 1$ and

$$\mathbb{P}(M_\infty \in (-\varepsilon, \varepsilon)) = \lim_n \mathbb{P}(M_n \in (-\varepsilon, \varepsilon)) \geq \lim_n \mathbb{P}(f_n = 0) = 1$$

as $M_n = \frac{f_n}{\mu^n}$. Since this is true for all $\varepsilon > 0$ we conclude that $\mathbb{P}(M_\infty = 0) = 1$.

(b) $\mu > 1$: Here we have exactly one fix-point θ_0 in $(0, 1)$. Moreover, by drawing a picture we check that

$$\alpha = \lim_n \pi_n = \theta_0 \in (0, 1).$$

The equality $\mathbb{E}M_\infty = 1$ we do not prove here. □

To understand the different cases better, here are two figures:

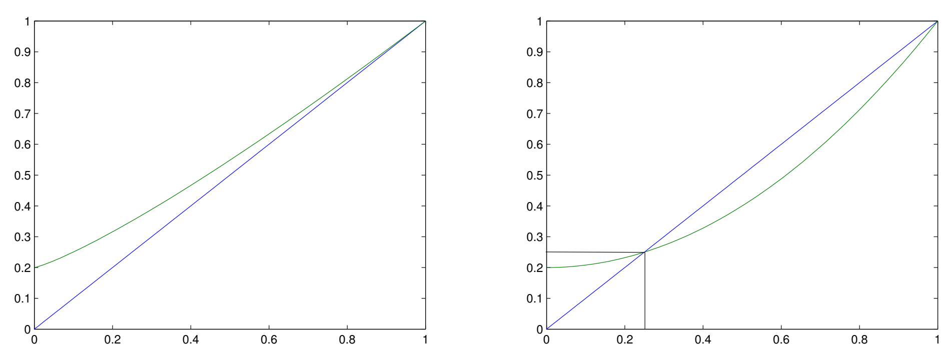


Figure 3.1: On the left $0 < \mu \leq 1$, on the right $\mu > 1$ and $\theta_0 = 0.25$.

Remark 3.9.1. In dependence on the parameter μ one distinguishes the following cases for our branching process:

$0 < \mu < 1$	:	sub-critical case
$\mu = 1$	:	critical case
$\mu > 1$	:	supercritical case

3.9.3 The Theorem of RADON-NIKODYM.

The following proof is according to P.-A. Meyer [4] (see also D. Williams [7]). Assume a measurable space $(\Omega, \mathcal{F})$ and two finite measures $\mathbb{P}$ and ν on $(\Omega, \mathcal{F})$. We recall that μ is absolutely continuous with respect to $\mathbb{P}$ provided that $\mathbb{P}(A) = 0$ implies that $\mu(A) = 0$. The aim of this section is to prove the following fundamental theorem by martingale methods.

Proposition 3.9.2 (Theorem of RADON-NIKODYM). *Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a finite measure μ on $(\Omega, \mathcal{F})$. If μ is absolutely continuous with respect to $\mathbb{P}$, then there exists a non-negative random variable $L \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$\mu(A) = \int_A L d\mathbb{P} \quad \textit{for all} \quad A \in \mathcal{F}.$$

First we show that our assumption implies some quantitative uniformity.

Lemma 3.9.3. *Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a finite measure μ on $(\Omega, \mathcal{F})$ such that μ is absolutely continuous with respect to $\mathbb{P}$. Then, given $\varepsilon > 0$ there is some $\delta \in (0,1)$ such that $\mathbb{P}(A) \le \delta$ implies that $\mu(A) \le \varepsilon$.*

Proof. Assume the contrary, that means that there is an $\varepsilon_0 > 0$ such that for all $\delta \in (0,1)$ there is a set $A(\delta)$ such that

$$\mathbb{P}(A(\delta)) \le \delta \quad \text{but} \quad \mu(A(\delta)) > \varepsilon_0.$$

In particular, we get a sequence $B_n := A(2^{-n})$ for $n = 1, 2, 3, ...$ From the Lemma of FATOU we know that

$$\varepsilon_0 \le \limsup_n \mu(B_n) \le \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} B_k\right).$$

On the other hand, for all $n_0 \ge 1$,

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} B_k\right) \le \mathbb{P}\left(\bigcup_{k=n_0}^{\infty} B_k\right) \le \sum_{k=n_0}^{\infty} \mathbb{P}(B_k) \le \frac{2}{2^{n_0}}$$

so that

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} B_k\right) = 0 \quad \text{but} \quad \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} B_k\right) \ge \varepsilon_0$$

which is a contradiction to the absolute continuity of μ with respect to $\mathbb{P}$. □

Proof of Proposition 3.9.2. (a) First we assume a filtration $(\mathcal{F}_n)_{n=0}^\infty$ such that

- $\mathcal{F}_n = \sigma\left(A_1^{(n)}, ..., A_{L_n}^{(n)}\right)$,
- the $A_1^{(n)}, ..., A_{L_n}^{(n)}$ are pair-wise disjoint and $\bigcup_{l=1}^{L_n} A_l^{(n)} = \Omega$,
- every $A_l^{(n)}$ is a union of elements from $\left\{A_1^{(n+1)}, ..., A_{L_{n+1}}^{(n+1)}\right\}$,
- $\mathcal{F} = \sigma\left(A_l^{(n)} : n = 0, 1, ... \text{ and } l = 1, ..., L_n\right)$.

We define the random variables $M_n : \Omega \to \mathbb{R}$ by

$$M_n(\omega) := \begin{cases} \frac{\mu(A_l^{(n)})}{\mathbb{P}(A_l^{(n)})} & : \quad \mathbb{P}(A_l^{(n)}) > 0 \\ 1 & : \quad \mathbb{P}(A_l^{(n)}) = 0 \end{cases}$$

whenever $\omega \in A_l^{(n)}$. The process $M = (M_n)_{n=0}^\infty$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n=0}^\infty$. Clearly $\mathbb{E}|M_n| < \infty$ since it is a finite step-function. Moreover, M_n is $\mathcal{F}_n$-measurable since M_n is constant over the sets $A_l^{(n)} \in \mathcal{F}_n$. To check the martingale property it is enough to check that

$$\int_{A_l^{(n)}} M_n d\mathbb{P} = \int_{A_l^{(n)}} M_{n+1} d\mathbb{P}$$

for all $l = 1, ..., L_n$ such that $\mathbb{P}(A_l^{(n)}) > 0$. By assumption we can express $A_l^{(n)}$ as disjoint union

$$A_l^{(n)} = \bigcup_{m \in I} A_m^{(n+1)}$$

for some index set $I \subseteq \{1, ..., L_{n+1}\}$. Hence

$$\begin{aligned}
\int_{A_l^{(n)}} M_{n+1} d\mathbb{P} &= \sum_{m\in I} \int_{A_m^{(n+1)}} M_{n+1} d\mathbb{P} \\
&= \sum_{m\in I} \mathbb{P}(A_m^{(n+1)}) \frac{\mu(A_m^{(n+1)})}{\mathbb{P}(A_m^{(n+1)})} \\
&= \sum_{m\in I} \mu(A_m^{(n+1)}) \\
&= \mu(A_l^{(n)}) \\
&= \int_{A_l^{(n)}} M_n d\mathbb{P}
\end{aligned}$$

with the convention that $0/0 := 1$ (note that our assumption of absolute continuity is used for the equality $\mathbb{P}(A_m^{(n+1)})(\mu(A_m^{(n+1)})/\mathbb{P}(A_m^{(n+1)})) = \mu(A_m^{(n+1)})$). In particular, the above computation shows that M_n is the density of μ with respect to $\mathbb{P}$ on $\mathcal{F}_n$, i.e.

$$\mu(A) = \int_A M_n d\mathbb{P} \quad \text{for all} \quad A \in \mathcal{F}_n.$$

Next we show that the martingale $M = (M_n)_{n=0}^{\infty}$ is uniformly integrable. We know that, for $c > 0$, one has that

$$\mathbb{P}(M_n \geq c) \leq \frac{\mathbb{E}M_n}{c} = \frac{\mu(\Omega)}{c}.$$

Applying Lemma 3.9.3, for any $\varepsilon > 0$ we find a $c > 0$ such that $\mathbb{P}(A) \leq \mu(\Omega)/c$ implies that $\mu(A) \leq \varepsilon$. Hence

$$\int_{M_n \geq c} M_n d\mathbb{P} = \mu(M_n \geq c) \leq \varepsilon$$

and the uniform integrability is proved. Applying Proposition 3.8.6 we have a limit $L \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$M_n = \mathbb{E}\left(L | \mathcal{F}_n\right) \text{ a.s.}$$

The random variable L is the density we were looking for: defining a measure ν on $\mathcal{F}$ by

$$\nu(A) := \int_A L d\mathbb{P}$$

we get $\mu = \nu$ on $\bigcup_{n=0}^{\infty} \mathcal{F}_n$ which is a generating algebra of $\mathcal{F}$. Hence $\mu = \nu$ on $\mathcal{F}$ according to CARATHÉODORY[10]'s extension theorem.

(b) Next we assume the general case. Let $\mathcal{A}$ be the collection of all sub-σ-algebras $\mathcal{G}$ of $\mathcal{F}$ such that

$$\mathcal{G} = \sigma(B_1, ..., B_L)$$

for $L = 1, 2, ...$ and $B_l \in \mathcal{F}$. As we saw in step (a) for every $\mathcal{G} \in \mathcal{A}$ there is a non-negative $L_{\mathcal{G}} \in \mathcal{L}_1(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\frac{d\mu}{d\mathbb{P}}|_{\mathcal{G}} = L_{\mathcal{G}}.$$

Fact 3.9.4. *For all $\varepsilon > 0$ there is an $\mathcal{G}_\varepsilon \in \mathcal{A}$ such that*

$$\mathbb{E}|L_{\mathcal{G}_1} - L_{\mathcal{G}_2}| < \varepsilon$$

for all $\mathcal{G}_1 \supseteq \mathcal{G}_\varepsilon$ and $\mathcal{G}_2 \supseteq \mathcal{G}_\varepsilon$.

Proof. Assuming the contrary gives an $\varepsilon_0 > 0$ and a sequence $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{G}_3 \subseteq \cdots$ from $\mathcal{A}$ such that

$$\mathbb{E}|L_{\mathcal{G}_{n+1}} - L_{\mathcal{G}_n}| \geq \varepsilon_0$$

for $n = 1, 2, ...$ However, using the argument from step (a) says that $(L_{\mathcal{G}_n})_{n=1}^{\infty}$ should be a uniformly integrable martingale with respect to the filtration $(\mathcal{G}_n)_{n=1}^{\infty}$ which yields to a contradiction so that the fact is proved. □

Now we define our candidate for the density L: We take $\mathcal{G}^{(n)} \in \mathcal{A}$ such that

$$\mathbb{E}|L_{\mathcal{G}_1} - L_{\mathcal{G}_2}| < \frac{1}{2^{n+1}}$$

for all $\mathcal{G}_1 \supseteq \mathcal{G}^{(n)}$ and $\mathcal{G}_2 \supseteq \mathcal{G}^{(n)}$. Letting $\mathcal{H}_n := \sigma(\mathcal{G}^{(1)}, ..., \mathcal{G}^{(n)})$ we obtain a martingale $(L_{\mathcal{H}_n})_{n=1}^{\infty}$ with respect to the filtration $(\mathcal{H}_n)_{n=1}^{\infty}$ which converges in $\mathcal{L}_1$ and almost surely to a limit L. Now, for $B \in \mathcal{F}$ we get

$$\begin{aligned}\left|\mu(B) - \int_B L d\mathbb{P}\right| &= \left|\int_B L_{\sigma(\mathcal{H}_n, B)} d\mathbb{P} - \int_B L d\mathbb{P}\right| \leq \mathbb{E}\left|L_{\sigma(\mathcal{H}_n,B)} - L\right| \\ &\leq \|L_{\sigma(\mathcal{H}_n,B)} - L_{\mathcal{H}_n}\|_{L_1} + \|L_{\mathcal{H}_n} - L\|_{L_1} \leq \frac{1}{2^{n+1}} + \|L_{\mathcal{H}_n} - L\|_{L_1} \to_{n\to\infty} 0\end{aligned}$$

so that $\mu(B) = \int_B L d\mathbb{P}$. □

[10] Constantin Carathéodory, 13/09/1873 (Berlin, Germany)- 2/02/1950 (Münich, Germany), significant contributions to the calculus of variations, the theory of measure, and the theory of functions of a real variable.

3.9.4 On a theorem of KAKUTANI

Martingales that are generated by products of independent random variables play an important role in stochastic modeling and in understanding equivalent probabilities on a given measurable space. Let us start with the KAKUTANI's theorem

Proposition 3.9.5. *Let $X_0, X_1, X_2, ... : \Omega \to \mathbb{R}$ be independent and non-negative random variables with $X_0 \equiv 1$ and $\mathbb{E}X_n = 1$. Define*

$$M_n := X_0 X_1 X_2 \cdots X_n$$

and $\mathcal{F}_n := \sigma(X_0, ..., X_n)$ for $n = 0, 1, 2, ...$

(i) *Then $M = (M_n)_{n=0}^{\infty}$ is a non-negative martingale with respect to $(\mathcal{F}_n)_{n=0}^{\infty}$ so that $M_\infty := \lim M_n$ exists almost surely.*

(ii) *The following assertions are equivalent:*

   (a) $\sum_{n=0}^{\infty}(1 - \mathbb{E}\sqrt{X_n}) < \infty$.

   (b) $\prod_{n=0}^{\infty} \mathbb{E}\sqrt{X_n} > 0$.

   (c) $\mathbb{E}M^* < \infty$.

   (d) M *is uniformly integrable.*

   (e) $\mathbb{E}M_\infty = 1$.

(iii) *If M is not uniformly integrable, then $\mathbb{P}(M_\infty = 0) = 1$.*

Before we prove the proposition we consider a simple example.

Example 3.9.6. Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ be independent BERNOULLI random variables, i.e. $\mathbb{P}(\varepsilon_n = -1) = \mathbb{P}(\varepsilon_n = 1) = 1/2$. Define $M_0 := 1$ and

$$M_n := e^{\varepsilon_1 + \cdots + \varepsilon_n - cn} \quad \text{such that} \quad \mathbb{E}e^{\varepsilon_1 - c} = 1 \quad \text{with} \quad c = \log\left(\frac{e + \frac{1}{e}}{2}\right).$$

Then $X_n = e^{\varepsilon_n - c}$ for $n \geq 1$ and

$$\sum_{n=1}^{\infty}(1 - \mathbb{E}e^{\frac{\varepsilon_n - c}{2}}) = \infty$$

because we have that $\mathbb{E}e^{\frac{\varepsilon_n - c}{2}} = \mathbb{E}e^{\frac{\varepsilon_1 - c}{2}} < 1$ by a computation.

Proof of Proposition 3.9.5. (i) The martingale property follows simply from, a.s.,

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

Corollary 3.8.10 gives the almost sure existence of $M_\infty := \lim M_n$.

(ii) Let $\alpha_n := \mathbb{E}\sqrt{X_n}$ for $n = 0, 1, ...$ Then

$$\alpha_n \le (\mathbb{E}(\sqrt{X_n})^2)^{\frac{1}{2}} = 1$$

and $\alpha_n > 0$ (note that $\alpha_n = 0$ would imply that $X_n = 0$ a.s. so that $\mathbb{E}X_n = 1$ would not be possible).

(a)$\iff$ (b) is a simple fact from analysis.

(b)$\Longrightarrow$ (c) Defining

$$N_n := \frac{\sqrt{X_0}}{\alpha_0} \cdots \frac{\sqrt{X_n}}{\alpha_n} = \frac{\sqrt{M_n}}{\alpha_0 \cdots \alpha_n}$$

we get again a martingale $N = (N_n)_{n=0}^\infty$. Since

$$\mathbb{E}N_n^2 = \frac{1}{\alpha_0^2 \cdots \alpha_n^2} \le \prod_{k=0}^\infty \frac{1}{\alpha_k^2} < \infty$$

the martingale N is bounded in L_2 and it follows from DOOB's maximal inequality that

$$\mathbb{E}M^* \le \mathbb{E}(N^*)^2 \le 4 \sup_{n \ge 0} \mathbb{E}N_n^2 < \infty.$$

(c)$\Longrightarrow$ (d) is clear.

(d)$\Longrightarrow$ (e) follows from $\mathbb{E}M_n = 1$ and Proposition 3.8.6 (see also Proposition 3.8.11).

(e)$\Longrightarrow$ (b) Suppose that (b) is not satisfied. Since $\lim_n N_n$ exists almost surely we have to have that

$$\lim_n \sqrt{X_0} \cdots \sqrt{X_n} = 0$$

almost surely, so that $\mathbb{P}(M_\infty = 0) = 1$. This proves $(e) \Rightarrow (b)$ and at same time item (iii). □

Now let $\Omega = \mathbb{R}^{\mathbb{N}}$ the space of all real valued sequences $\omega = (\omega_n)_{n=1}^\infty \subseteq \mathbb{R}$ and

$$Y_n : \Omega \to \mathbb{R} \quad \text{given by} \quad Y_n(\omega) := \omega_n.$$

As filtration we use $\mathcal{F}_0 := \{\Omega, \emptyset\}$ and

$$\mathcal{F}_n := \sigma(Y_1, ..., Y_n).$$

Finally, we let $\mathcal{F} := \vee_{n=0}^{\infty} \mathcal{F}_n$. Assume BOREL-measurable $f_n, g_n : \mathbb{R} \to \mathbb{R}$ such that $f_n(x) > 0$ and $g_n(x) > 0$ for all $x \in \mathbb{R}$ and such that

$$\int_{\mathbb{R}} f_n(x)dx = \int_{\mathbb{R}} g_n(x)dx = 1$$

for all $n = 1, 2, ...$ which means that f_n and g_n are densities of probability measures μ_n and ν_n, respectively, on $\mathbb{R}$. Define the measures

$$\mathbb{P} := \times_{n=1}^{\infty} \mu_n \quad \text{and} \quad Q := \times_{n=1}^{\infty} \nu_n$$

on the measurable space $(\Omega, \mathcal{F})$ and

$$X_n(\omega) := \frac{g_n(Y_n(\omega))}{f_n(Y_n(\omega))} = \frac{g_n(\omega_n)}{f_n(\omega_n)}.$$

The random variables X_n form a *Likelihood-Ratio Test*. We get the following

Proposition 3.9.7. *The following assertions are equivalent:*

(i) *The measure Q is absolutely continuous with respect to $\mathbb{P}$, that means that $\mathbb{P}(A) = 0$ implies that $Q(A) = 0$.*

(ii) $\prod_{n=1}^{\infty} \mathbb{E}_{\mathbb{P}} \sqrt{X_n} = \prod_{n=1}^{\infty} \int_{\mathbb{R}} \sqrt{f_n(x) g_n(x)} dx > 0.$

(iii) $\sum_{n=1}^{\infty} \int_{\mathbb{R}} (\sqrt{f_n(x)} - \sqrt{g_n(x)})^2 dx < \infty.$

The measure Q is not absolutely continuous with respect to $\mathbb{P}$ if and only if $X_1 X_2 \cdots X_n \to 0$ $\mathbb{P}$-a.s.

Proof. Let $X_0 \equiv 1$. By construction it is clear that $X_0, X_1, X_2, ...$ are independent and non-negative random variables with respect to $(\Omega, \mathcal{F}, \mathbb{P})$ and that

$$\mathbb{E}_{\mathbb{P}} X_n = \mathbb{E}_{\mathbb{P}} \frac{g_n(Y_n)}{f_n(Y_n)} = \int_{\mathbb{R}} \frac{g_n(x)}{f_n(x)} f_n(x) dx = \int_{\mathbb{R}} g_n(x) dx = 1$$

for $n \geq 1$. Moreover,

$$\mathbb{E}_{\mathbb{P}} \sqrt{X_n} = \int_{\mathbb{R}} \sqrt{\frac{g_n(x)}{f_n(x)}} f_n(x) dx = \int_{\mathbb{R}} \sqrt{g_n(x) f_n(x)} dx$$

and

$$\begin{aligned}\sum_{n=0}^{\infty}(1-\mathbb{E}_{\mathbb{P}}\sqrt{X_n}) &= \sum_{n=1}^{\infty}\left(1-\int_{\mathbb{R}}\sqrt{\frac{g_n(x)}{f_n(x)}}f_n(x)dx\right) \\ &= \sum_{n=1}^{\infty}\left(1-\int_{\mathbb{R}}\sqrt{f_n(x)g_n(x)}dx\right) \\ &= \frac{1}{2}\sum_{n=1}^{\infty}\int_{\mathbb{R}}(\sqrt{f_n(x)}-\sqrt{g_n(x)})^2dx.\end{aligned}$$

Applying KAKUTANI[11] 's theorem Proposition 3.9.5 gives that conditions (ii) and (iii) are equivalent to the uniform integrability of the martingale $M=(M_n)_{n=0}^{\infty}$ given by

$$M_n := X_0 \cdots X_n.$$

Moreover, we have that

$$M_n = \frac{dQ}{d\mathbb{P}}|_{\mathcal{F}_n}$$

since

$$\begin{aligned}\times_{k=1}^{n}\nu_k(B) &= \int_B g_1(x_1)\cdots g_n(x_n)dx_1...dx_n \\ &= \int_B \frac{g_1(x_1)\cdots g_n(x_n)}{f_1(x_1)\cdots f_n(x_n)}d(\times_{k=1}^{n}\mu_k)(x).\end{aligned}$$

(ii) $\Longrightarrow$ (i) Here we use the proof of the RADON-NIKODYM theorem where we exploit that M is uniformly integrable.

(i)$\Longrightarrow$ (ii) Here we get a density $L=\frac{dQ}{d\mathbb{P}}$ and

$$\mathbb{E}(L|\mathcal{F}_n) = M_n \text{ a.s.}$$

But then Proposition 3.8.6 gives the uniform integrability of M.

Finally, if Q is not absolutely continuous with respect to $\mathbb{P}$, then M is not uniformly integrable and KAKUTANI's theorem says that in this case $M_n \to 0$ a.s. The other way round: if $M_n \to 0$ a.s. but, $\mathbb{E}M_n = 1$ for all $n=1,2,...$, then

[11]Shizuo Kakutani, 28/08/1911 (Osaka, Japan)- 17/08/2004 (New Haven, USA), contributed to several areas of mathematics: complex analysis, topological groups, fixed point theorems, Banach spaces and Hilbert spaces, Markov processes, measure theory, flows, Brownian motion, and ergodic theory.

$(M_n)_{n=0}^{\infty}$ cannot be uniformly integrable and Q cannot be absolutely continuous with respect to $\mathbb{P}$. $\square$

Looking at the symmetry in conditions (ii) and (iii) of Proposition 3.9.7 with respect to the densities f_n and g_n one gets immediately the following

Corollary 3.9.8. *One can add to the equivalent conditions of Proposition 3.9.7 the following ones:*

(iv) *The measure $\mathbb{P}$ is absolutely continuous with respect to Q, that means that $Q(A) = 0$ implies that $\mathbb{P}(A) = 0$.*

(v) *The measures $\mathbb{P}$ and Q are equivalent, that means that $Q(A) = 0$ if and only if $\mathbb{P}(A) = 0$.*

But we have still more, an alternative. To formulate this we introduce

Definition 3.9.9. *Let $(\Omega, \mathcal{F})$ be a measurable space with probability measures $\mathbb{P}$ and Q. The measures $\mathbb{P}$ and Q are said to be* singular *or* orthogonal *($\mathbb{P} \perp Q$) if there is a $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 1$ and $Q(A^c) = 1$.*

Now we can close the chapter with

Proposition 3.9.10 (KAKUTANI's alternative)**.** *The measures $\mathbb{P}$ and Q defined above are either singular or equivalent.*

Proof. Assume that Q is not absolutely continuous with respect to $\mathbb{P}$. According to Proposition 3.9.7 we have that

$$\lim_{n\to\infty} X_1 \cdots X_n = 0 \ \mathbb{P}\text{-a.s.}$$

For all $\varepsilon \in (0,1)$ this gives an $n(\varepsilon) \geq 1$ such that

$$\mathbb{P}(X_1 \cdots X_n \geq \varepsilon) \leq \varepsilon \text{ for } n \geq n(\varepsilon).$$

Taking $A := \left\{X_1 \cdots X_{n(\varepsilon)} < \varepsilon\right\}$ yields

$$\left(\times_{n=1}^{n(\varepsilon)} \mu_n\right)(A) \geq 1 - \varepsilon \text{ and } \left(\times_{n=1}^{n(\varepsilon)} \nu_n\right)(A) \leq \varepsilon.$$

Letting $\widetilde{A} := A \times \mathbb{R} \times \mathbb{R} \cdots$, we continue with

$$\mathbb{P}(A) \geq 1 - \varepsilon \text{ and } Q(A) \leq \varepsilon.$$

Now we apply this to $\varepsilon = 2^{-n}$ and find A_n with

$$\mathbb{P}(A_n) \geq 1 - 2^{-n} \text{ and } Q(A_n) \leq 2^{-n}.$$

Setting $B := \limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ we end up with

$$\mathbb{P}(B) = \mathbb{P}(\limsup_n A_n) \geq \limsup_n \mathbb{P}(A_n) = 1$$

and

$$Q(B) \leq Q\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \sum_{k=n}^{\infty} Q(A_k) \leq \sum_{k=n}^{\infty} 2^{-k} \to 0 \text{ as } n \to \infty,$$

so that $\mathbb{P}(B) = 1$ but $Q(B) = 0$. □

3.10 Backward martingales

Definition 3.10.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_n)_{n=-1}^{-\infty}$ be a sequence of σ-algebras such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all $n \leq -2$. A process $M = (M_n)_{n=-1}^{-\infty}$ is called *backward martingale* if and only if

(i) M_n is $\mathcal{F}_n$-measurable,

(ii) $\mathbb{E}|M_n| < \infty$,

(iii) $\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = M_n$ a.s. for $n \leq -2$.

Consequently, M is a martingale, however, the time scale is $(-\infty, -1]$.

Example 3.10.2. *Let* $\xi_1, \xi_2, \cdots : \Omega \to \mathbb{R}$ *be such that*

(i) $\xi_1, \xi_2, \cdots : \Omega \to \mathbb{R}$ *are independent,*

(ii) $\xi_1, \xi_2, \ldots$ *have the same distribution,*

(iii) $\mathbb{E}|\xi_1| < \infty$.

Then $M_n := \frac{\xi_1+\cdots+\xi_{-n}}{-n}$ *is a backward martingale with respect to* $\mathcal{F}_n := \sigma(M_k : k \le n)$.

As $n \to -\infty$, backward martingales get simpler and they have always a closure, namely M_{-1}. Hence the following theorem is not surprising:

Theorem 3.10.3. *Let* $M = (M_n)_{n=-1}^{-\infty}$ *be a backward martingale with respect to* $(\mathcal{F}_n)_{n=-1}^{-\infty}$*, and let* $\mathcal{F}_{-\infty} := \bigcap_{n=-1}^{-\infty} \mathcal{F}_n$. *Then*

$$\lim_{n\to-\infty} M_n = M_{-\infty} := \mathbb{E}\left(M_{-1} \,\middle|\, \mathcal{F}_{-\infty}\right) \text{ a.s. and in } L_1.$$

We start by

Lemma 3.10.4. *Let* $B \in \mathcal{F}$ *such that for all* $n = -1, -2, \ldots$ *there exists* $A_n \in \mathcal{F}_n$ *with*

$$\mathbb{P}(A_n \Delta B) = 0.$$

Then there exists $A \in \mathcal{F}_{-\infty}$ *with*

$$\mathbb{P}(A \Delta B) = 0.$$

Proof. We let

$$A := \limsup_{n\to-\infty} A_n = \bigcap_{m=-1}^{-\infty} \bigcup_{n=m}^{-\infty} A_n.$$

Then

$$A = \bigcap_{m=N}^{-\infty} \bigcup_{n=m}^{-\infty} A_n \in \mathcal{F}_N \quad \text{for all} \quad N \le -1$$

so that

$$A \in \bigcap_{N=-1}^{-\infty} \mathcal{F}_N = \mathcal{F}_{-\infty}.$$

Moreover

$$\mathbb{P}\left(\left(\bigcup_{n=m}^{-\infty} A_n\right) \setminus B\right) = \mathbb{P}\left(\bigcup_{n=m}^{-\infty} (A_n \setminus B)\right) \le \sum_{n=m}^{-\infty} \mathbb{P}(A_n \setminus B) = 0$$

and

$$\mathbb{P}(A \setminus B) \le \mathbb{P}\left(\left(\bigcup_{n=m}^{-\infty} A_n\right) \setminus B\right) = 0.$$

On the other hand

$$\begin{aligned}\mathbb{P}(B\setminus A) = \mathbb{P}\left(B\setminus \bigcap_{m=-1}^{-\infty}\bigcup_{n=m}^{-\infty} A_n\right) &= \lim_{m\to-\infty}\mathbb{P}\left(B\setminus\bigcup_{n=m}^{-\infty} A_n\right)\\ &\leq \lim_{m\to-\infty}\mathbb{P}(B\setminus A_m) = 0.\end{aligned}$$

Combining the above yields $\mathbb{P}(A\Delta B) = 0$. □

Lemma 3.10.5. *Let*

$$\mathscr{H}_n := L_2(\Omega,\mathcal{F}_n,\mathbb{P}) \subseteq \mathscr{H} := L_2(\Omega,\mathcal{F},\mathbb{P}),\ \textit{i.e.}$$

$f_n \in \mathscr{H}_n$ *if and only if there exists an* $\mathcal{F}_n$ *-measurable* $g_n : \Omega \to \mathbb{R}$ *with* $\mathbb{E}g_n^2 < \infty$ *and* $f_n = g_n$ *a.s. Then*

$$\mathscr{H}_{-\infty} := \bigcap_{n=-1}^{-\infty} \mathscr{H}_n = L_2(\Omega,\mathcal{F}_{-\infty},\mathbb{P}),\ \textit{i.e.}$$

$f \in \mathscr{H}_{-\infty}$ *if and only if there exists an* $\mathcal{F}_{-\infty}$*-measurable* $g : \Omega \to \mathbb{R}$ *with* $\mathbb{E}g^2 < \infty$ *and* $f = g$ *a.s.*

Proof. (a) If $f \in L_2(\Omega,\mathcal{F}_{-\infty},\mathbb{P})$, then we find a $g : \Omega \to \mathbb{R}$ which is $\mathcal{F}_{-\infty}$-measurable with $\mathbb{E}g^2 < \infty$ and $f = g$ a.s. But this g is also $\mathcal{F}_n$-measurable for all $n \leq -1$ so that

$$g \in L_2(\Omega,\mathcal{F}_n,\mathbb{P}) = \mathscr{H}_n \quad \text{and} \quad f \in \bigcap_{n=-1}^{-\infty}\mathscr{H}_n.$$

(b) Let us assume that $f \in \bigcap_{n=-1}^{-\infty}\mathscr{H}_n$. We find $\mathcal{F}_n$-measurable $g_n : \Omega \to \mathbb{R}$ with $f = g_n$ a.s. Hence, for all $n \leq -1$,

$$\mathbb{P}(\{f>\lambda\}\,\Delta\,\{g_n>\lambda\}) = 0 \quad \text{and} \quad \mathbb{P}(\{f>\lambda\}\,\Delta A_\lambda) = 0$$

for some $A_\lambda \in \mathcal{F}_{-\infty}$ by Lemma 3.10.4. Let

$$h(\omega) := \sup\{\lambda \in Q : \omega \in A_\lambda\} \in [-\infty,\infty].$$

Then, for $\mu \in \mathbb{R}$, we have that $h(\omega) > \mu$ if and only if there exists an $\lambda \in Q$ with $\lambda > \mu$ and $\omega \in A_\lambda$, i.e.

$$\omega \in \bigcup_{\lambda > \mu, \lambda \in Q} A_\lambda$$

so that

$$\{h > \mu\} = \bigcup_{\lambda > \mu, \lambda \in Q} A_\lambda \in \mathcal{F}_{-\infty}.$$

Hence $h : \Omega \to [-\infty, \infty]$ is an extended $\mathcal{F}_{-\infty}$-measurable random variable. Finally $h = f$ on

$$\Omega_0 := \left(\bigcup_{\lambda \in Q} (\{f > \lambda\} \,\Delta A_\lambda) \right)^c = \bigcap_{\lambda \in Q} (\{f > \lambda\} \,\Delta A_\lambda)^c$$

where

$$\mathbb{P}(\Omega_0^c) = \mathbb{P} \left(\bigcup_{\lambda \in Q} (\{f > \lambda\} \,\Delta A_\lambda) \right) = 0.$$

In fact, for $\lambda \in \mathbb{R}$ we have that

$$\Omega_0 \cap \{f > \lambda\} = \Omega_0 \cap A_\lambda$$

so that

$$\Omega_0 \cap \{h > \mu\} = \Omega_0 \cap \bigcup_{\lambda > \mu, \lambda \in Q} A_\lambda = \Omega_0 \cap \bigcup_{\lambda > \mu, \lambda \in Q} \{f > \lambda\} = \Omega_0 \cap \{f > \mu\}.$$

□

Lemma 3.10.6. *If $\varepsilon > 0$ and $M = (M_n)_{n=-1}^{-\infty}$ is a backward martingale, then*

$$\mathbb{P}\left(\sup_{k \le n} |M_k| > \varepsilon \right) \le \frac{\mathbb{E}|M_n|}{\varepsilon} \quad \textit{for all} \quad n \le -1.$$

Proof. By the monotonicity of the measure it is enough to check that

$$\mathbb{P}\left(\sup_{N \le k \le n} |M_k| > \varepsilon \right) \le \frac{\mathbb{E}|M_n|}{\varepsilon}.$$

But this follows from DOOB's maximal inequality. □

Proof of Theorem 3.10.3 (a) First we assume that $\mathbb{E}M_{-1}^2 < \infty$ and denote

$$dM_n := M_n - M_{n-1} \quad \text{for} \quad n \le -1.$$

By orthogonality of the martingale differences it follows that

$$\|M_{-1}\|_2^2 = \|dM_{-1}\|_2^2 + \cdots + \|dM_{n+1}\|_2^2 + \|M_n\|_2^2.$$

Consequently

$$\sum_{n=-1}^{-\infty} \|dM_n\|_2^2 \le \|M_{-1}\|_2^2 < \infty$$

and there exists (by HILBERT space theory) some

$$M_{-\infty} := \left(\sum_{n=-1}^{-\infty} dM_n \right) - M_{-1} \text{ in } L_2(\Omega, \mathcal{F}, \mathbb{P}).$$

Since

$$\begin{aligned} M_{-\infty} &= -M_{-1} + dM_{-1} + \cdots + dM_{n+1} + \cdots \\ &= -M_n + dM_n + \cdots \in \mathscr{H}_n \subseteq \mathscr{H}_{n+1} \subseteq \cdots \subseteq \mathscr{H}_{-1} \end{aligned}$$

we get that

$$M_{-\infty} \in \bigcap_{n=-1}^{-\infty} \mathscr{H}_n = \mathscr{H}_{-\infty}. \text{ By lemma 3.10.5, } M_{-\infty} \in L_2\left(\Omega, \mathcal{F}_{-\infty}, \mathbb{P}\right).$$

(b) We show that

$$\mathbb{E}(M_{-1} \,|\, \mathcal{F}_{-\infty}) = M_{-\infty} \text{ a.s.}$$

For all $B \in \mathcal{F}_{-\infty}$ we have

$$\int_B M_n d\mathbb{P} = \int_B M_{-1} d\mathbb{P}$$

by $n \to -\infty$ this implies

$$\int_B M_{-\infty} d\mathbb{P} = \int_B M_{-1} d\mathbb{P}$$

and by the uniqueness of the conditional expectation

$$\mathbb{E}\left(M_{-1} \middle| \mathcal{F}_{-\infty}\right) = M_{-\infty} \text{ a.s.}$$

by construction of L_2, and by HÖLDER[12] 's inequality, we have L_1- convergence.

(c) We show that $M_n \to M_{-\infty}$ a.s. Applying Lemma 3.10.6 to the backward martingale $(M_n - M_{-\infty})_{n \le -1}$ gives that

$$\mathbb{P}(\sup_{k \le m} |M_k - M_{-\infty}| > \varepsilon) \le \tfrac{1}{\varepsilon}\mathbb{E}|M_m - M_{-\infty}|$$

for all $\varepsilon > 0$ and

$$\lim_{m \to -\infty} \mathbb{P}\left(\sup_{k \le m} |M_k - M_{-\infty}| > \varepsilon\right) = 0.$$

But this implies $M_n \to M_{-\infty}$ a.s.

(d) Finally, we weaken the condition $M_{-1} \in L_2$ to $M_{-1} \in L_1$. Given $\varepsilon > 0$ we find an $Y_{-1} \in L_2$ such that

$$\mathbb{E}|M_{-1} - Y_{-1}| < \varepsilon^2.$$

Let $Y_n := \mathbb{E}\left(Y_{-1} \middle| \mathcal{F}_n\right)$. By DOOB's maximal inequality,

$$\mathbb{P}(\sup_n |M_n - Y_n| > \varepsilon) \le \tfrac{1}{\varepsilon}\mathbb{E}|M_{-1} - Y_{-1}| \le \varepsilon$$

and therefore

$$\mathbb{P}(\limsup_n M_n - \liminf_n M_n > 2\varepsilon) \le \varepsilon.$$

By $\varepsilon \downarrow 0$,

$$\mathbb{P}(\limsup_n M_n > \liminf_n M_n) = 0$$

and $M_{-\infty} = \lim_{n \to -\infty} M_n$ exists almost surely. Since $(M_n)_{n=-1}^{-\infty}$ is closable, it is also uniformly integrable, so that we have

$$M_{-\infty} = \lim_{n \to -\infty} M_n \text{ in } L_1.$$

As before, we get

$$\mathbb{E}\left(M_{-1} \middle| \mathcal{F}_{-\infty}\right) = M_{-\infty} \text{ a.s.}$$

□

[12] Otto Ludwig Hölder, 22/12/1859 (Stuttgart, Germany) - 29/08/1937 (Leipzig, Germany), worked on the convergence of Fourier series, in 1884 he discovered the inequality named after him. He became interested in group theory through KRONECKER and KLEIN and proved the uniqueness of the factor groups in a composition series.

Chapter 4

Exercises

4.1 Introduction

1. Proposition 1.1.1.

   (a) Prove Proposition 1.1.1(i) for $p = q = \frac{1}{2}$ directly.

   (b) Deduce Proposition 1.1.1(i) for $p = q = \frac{1}{2}$ from Proposition 1.1.1(i) for $p \neq q$ by L'Hospital's rule.

   (c) Prove Proposition 1.1.1(ii).

4.2 Sums of independent random variables

1. Prove Lemma 2.1.7 from the course using the following monotone class theorem:

   Let $(\Omega, \mathcal{F})$ be a measurable space. If a system of subsets $\mathcal{G} \subseteq \mathcal{F}$ is a monotone class, that means $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ with $A_k \in \mathcal{G}$ implies that $\bigcup_k A_k \in \mathcal{G}$, and if $\mathcal{G}$ is an algebra such that $\sigma(\mathcal{G}) = \mathcal{F}$, then $\mathcal{G} = \mathcal{F}$.

2. Let $\varepsilon_1, \varepsilon_2, \varepsilon_3, ... : \Omega \to \mathbb{R}$ independent random variables such that $\mathbb{P}(\varepsilon_k = -1) = p$ and $\mathbb{P}(\varepsilon_k = 1) = q = 1-p$ with $p \in (0,1)$. Let $f_n := \varepsilon_1 + \cdots + \varepsilon_n$. Given an integer $k \geq 0$, compute

$$\mathbb{P}(f_n = k).$$

3. Given independent random variables $f_1, f_2, f_3, ... : \Omega \to \mathbb{R}$ and $S_n := f_1 + \cdots + f_n$. What is the probability of

$$\left\{\omega \in \Omega : \limsup_{n\to\infty} \frac{S_n(\omega)}{n^2} \geq 1\right\}?$$

4. Let $\xi_1, \xi_2, ... : \Omega \to [0,1]$ be a sequence of independent random variables and $c \in [0,1]$.

   (a) Using Kolmogorow's 0-1-law, one shows that

   $$p_c = \mathbb{P}\left(\limsup_n \xi_n = c\right) \in \{0,1\}.$$

   (b) Can one choose $c \in [0,1]$ such that $p_c = 1$?

5. Let $-\infty < A < B < \infty$ and $f_n := \varepsilon_1 + \cdots + \varepsilon_n$, where $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ are independent such that $\mathbb{P}(\varepsilon_k = -1) = \mathbb{P}(\varepsilon_k = 1) = 1/2$. We say that a trajectory $(f_n(\omega))_{n=1}^\infty$ has infinitely many up-crossings provided that there exists a sequence $n_1(\omega) < N_1(\omega) < n_2(\omega) < N_2(\omega) < \cdots$ such that

   $$f_{n_k(\omega)}(\omega) < A < B < f_{N_k(\omega)}(\omega)$$

   for $k = 1, 2, ...$ Let C be the event that $(f_n)_{n=1}^\infty$ has infinitely many up-crossings.

   (a) Prove by Hewitt-Savage that $\mathbb{P}(C) \in \{0,1\}$.

   (b) Use another statement from the course to prove that $\mathbb{P}(C) = 1$.

6. Assume independent random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ such that $\mathbb{P}(\xi_k = -1) = 1/4$, $\mathbb{P}(\xi_k = 0) = 1/2$, and $\mathbb{P}(\xi_k = 1) = 1/4$. Let $f_n := \xi_1 + \cdots + \xi_n$. Does the random walk leave the strip $[-10, 10]$ with probability one?

7. Let $(\xi_n)_{n=1}^\infty$ be a sequence of independent random variables with $\xi_n(\omega) \geq 0$ for all $\omega \in \Omega$. One shows that

   $$\sum_{n=1}^\infty \mathbb{E}\frac{\xi_n}{1+\xi_n} < \infty \Longrightarrow \mathbb{P}\left(\sum_{n=1}^\infty \xi_n \quad \text{converges}\right) = 1.$$

8. From
$$\lim_{x\to\infty} \frac{\int_x^\infty e^{-\frac{y^2}{2}} dy}{\frac{1}{x} e^{-\frac{x^2}{2}}} = 1$$
deduce the following: If $(g_n)_{n=1}^\infty$ are independent random variables such that $g_n \sim N(0,1)$, then
$$\sum_{n=1}^\infty \mathbb{P}\left(g_n > (1-\varepsilon)\sqrt{2\log n}\right) = \infty \quad \text{for all} \quad 0 \le \varepsilon < 1$$
and
$$\sum_{n=1}^\infty \mathbb{P}\left(g_n > (1+\varepsilon)\sqrt{2\log n}\right) < \infty \quad \text{for all} \quad \varepsilon > 0.$$

9. Show using exercise 8 and the Lemma of BOREL-CANTELLI that
$$\mathbb{P}\left(\limsup_{n\to\infty (n\ge 2)} \frac{g_n}{\sqrt{2\log n}} = 1 \right) = 1.$$

4.3 Martingales in discrete time

1. Show Proposition 3.1.9 (ix) and (x).

2. For $f \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ prove that
$$\left(\mathbb{E}\left(|f| \,\middle|\, \mathcal{G} \right) \right)^2 \le \mathbb{E}\left(|f|^2 \,\middle|\, \mathcal{G} \right) \; a.s.$$

3. Let $0 < p < \infty$, $\Omega = [0,1)$, and $\mathcal{F}_n := \sigma\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) : k = 1, ..., 2^n\right)$ and λ be the Lebesgue measure. Define $M_n(t) := 2^{\frac{n}{p}}$ for $t \in [0, 2^{-n})$ and $M_n(t) := 0$ for $t \in [2^{-n}, 1)$ for $n = 0, 1, ...$

   (a) Classify $(M_n)_{n=0}^\infty$ to be a martingale, super- or sub-martingale in dependence on p.

   (b) For what p is the sequence $(M_n)_{n=0}^\infty$ a Cauchy-sequence in L_1?

   (c) Is $(M_n)_{n=0}^\infty$ uniformly integrable if $p = 1$?

4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ independent random variables such that $\mathbb{P}(\varepsilon_n = 1) = \mathbb{P}(\varepsilon_n = -1) = 1/2$. Let $v_0 \in \mathbb{R}$ and $v_n : \mathbb{R}^n \to \mathbb{R}$ be functions. Define $M_0 := 0$, $M_1(\omega) := \varepsilon_1(\omega) v_0$,
$$M_n(\omega) := \varepsilon_1(\omega) v_0 + \sum_{k=2}^{n} \varepsilon_k(\omega) v_{k-1}\Big(\varepsilon_1(\omega), ..., \varepsilon_{k-1}(\omega)\Big),$$
$\mathcal{F}_0 := \{\emptyset, \Omega\}$, and $\mathcal{F}_n := \sigma(\varepsilon_1, ..., \varepsilon_n)$ (check that $(M_n)_{n=0}^{\infty}$ is a martingale). Let $Z_0 := 1$ and
$$Z_n(\omega) := e^{M_n(\omega) - \frac{1}{2}\sum_{k=1}^{n} |v_{k-1}(\varepsilon_1(\omega), ..., \varepsilon_{k-1}(\omega))|^2}.$$
Show that $(Z_n)_{n=0}^{\infty}$ is a super-martingale.

5. Assume that $\varepsilon_1, ..., \varepsilon_n : \Omega \to \mathbb{R}$ are independent random variables such that $\mathbb{P}(\varepsilon_i = 1) = p$ and $\mathbb{P}(\varepsilon_i = -1) = q$ for some $p, q \in (0,1)$ with $p + q = 1$. Define the stochastic process $X_k := e^{a(\varepsilon_1 + ... + \varepsilon_k) + bk}$ for $k = 1, ..., n$ and $X_0 := 1$ with $a > 0$ and $b \in \mathbb{R}$ and the filtration $(\mathcal{F}_k)_{k=0}^{n}$ with $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_k := \sigma(\varepsilon_1, ..., \varepsilon_k)$.

   (a) Assume that $-a + b < 0 < a + b$. Find $p, q \in (0, 1)$ such that the process $(X_k)_{k=0}^{n}$ is a martingale.

   (b) Assume that $-a + b > 0$. Why there cannot exist random variables $\varepsilon_1, ..., \varepsilon_n : \Omega \to \{-1, 1\}$ such that $(X_k)_{k=0}^{n}$ is a martingale?

6. Let $M_0 := 0$ and $M_n := \varepsilon_1 + \cdots + \varepsilon_n$, $n \geq 1$, where $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ are independent random variables such that $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2}$. Let $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(\varepsilon_1, ..., \varepsilon_n)$. Are the following functions stopping times ($\inf \emptyset := \infty$)?

   (a) $\sigma(\omega) := \inf \{n \geq 0 : M_n \in (10, 12)\}$
   (b) $\sigma(\omega) := \inf \{n \geq 0 : M_n \in (10, 12)\} - 1$
   (c) $\sigma(\omega) := \inf \{n \geq 0 : M_n \in (10, 12)\} + 1$
   (d) $\sigma(\omega) := \inf \{n \geq 0 : M_{n+1} \in (10, 12)\}$
   (e) $\sigma(\omega) := \inf \{n \geq 0 : M_{n+1} \in (10, 11)\}$
   (f) $\sigma(\omega) := \inf \{n \geq 1 : M_{n-1} = 10\}$
   (g) $\sigma(\omega) := \inf \{n \geq 1 : M_{n-1} = 10\} - 1$

7. Assume that $\varepsilon_1, \varepsilon_2, \varepsilon_3 : \Omega \to \{-1, 1\}$ are independent random variables such that $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = 1/2$ and that $f : \mathbb{R}^3 \to \mathbb{R}$ is a function. Check that
$$\mathbb{E}(f(\varepsilon_1, \varepsilon_2, \varepsilon_3)|\sigma(\varepsilon_1, \varepsilon_2)) = g(\varepsilon_1, \varepsilon_2)$$
with $g(\varepsilon_1, \varepsilon_2) := (1/2)(f(\varepsilon_1, \varepsilon_2, -1) + f(\varepsilon_1, \varepsilon_2, 1))$.

8. Let $(\mathcal{F}_n)_{n=0}^\infty$ be a filtration and $\sigma, \tau : \Omega \to \{0, 1, 2, ...\} \cup \{\infty\}$ stopping times.
   (a) Show that $\sigma + \tau$ is a stopping time.
   (b) Show that $\mathcal{F}_{\sigma \wedge \tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

9. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$ be a stochastic basis and $M = (M_n)_{n=0}^\infty$ be a martingale such that $M_0 = 0$ a.s. and $\mathbb{E}M_n^2 < \infty$ for all $n = 0, 1, 2, ...$ Define $\langle M \rangle_0 := 0$,
$$\langle M \rangle_n := \sum_{k=1}^n \mathbb{E}((M_k - M_{k-1})^2 | \mathcal{F}_{k-1}) \in [0, \infty).$$
The process $\langle M \rangle := (\langle M \rangle_n)_{n=0}^\infty$ is called *(predictable) bracket process*. Show that
   (a) $(M_n^2 - \langle M \rangle_n)_{n=0}^\infty$ is a martingale,
   (b) $\mathbb{E}M_n^2 = \mathbb{E}\langle M \rangle_n$ for all $n = 0, 1, ...$

10. Assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_k)_{k=0}^n)$ with $\Omega = \{\omega_1, ..., \omega_N\}$, $\mathbb{P}(\{\omega_i\}) > 0$, and a process $(Z_k)_{k=0}^n$ such that Z_k is $\mathcal{F}_k$-measurable. Define
$$U_n := Z_n$$
and, backwards,
$$U_k := \max\{Z_k, \mathbb{E}(U_{k+1}|\mathcal{F}_k)\}$$
for $k = 0, ..., n-1$.
    - Show that $(U_k)_{k=0}^n$ is a super-martingale.
    - Show that $(U_k)_{k=0}^n$ is the smallest super-martingale which dominates $(Z_k)_{k=0}^n$: if $(V_k)_{k=0}^n$ is a super-martingale with $Z_k \le V_k$, then $U_k \le V_k$ a.s.

- Show that $\tau(\omega) := \inf\{k = 0, ..., n : Z_k(\omega) = U_k(\omega)\}$ ($\inf \emptyset := n$) is a stopping time.

The process $(U_k)_{k=0}^n$ is called SNELL-envelop of $(Z_k)_{k=0}^n$.

11. Assume i.i.d. random variables $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ such that $\mathbb{P}(\varepsilon_k = 1) = \mathbb{P}(\varepsilon_k = -1) = 1/2$ and define $M_0 := a$ and $M_n := a + \sum_{k=1}^n \varepsilon_k$. Assume that $-\infty < b < a < c < \infty$, where a, b, c are integers. Let

$$\tau(\omega) := \inf\{n \geq 0 : M_n \in \{b, c\}\}$$

with $\inf \emptyset := \infty$.

- Using statements from the course prove that $\mathbb{P}(\tau < \infty) = 1$.
- Using statements from the course prove that

$$\mathbb{E}M_{\tau \wedge N} = a$$

where $N \geq 1$ is a fixed integer.
- Deduce that $\mathbb{E}M_\tau = a$ and compute the probability that the process M hits first b.

12. Prove for the process M from Example 3.8.5 that M is a martingale, which fails to be uniformly integrable, and that

$$\lim_N \int_0^1 \sup_{n=1,...,N} M_n(t)dt = \infty \quad \text{but} \quad \int_0^1 M_N(t)dt = 1,$$

i.e. DOOB's maximal inequality does not hold for $p = 1$.

13. Let $M_n := \varepsilon_1 + \cdots + \varepsilon_n$ and $M_0 := 0$, where $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ are i.i.d. with $\mathbb{P}(\varepsilon_k = \pm 1) = 1/2$ and let

$$\tau(\omega) := \inf\{n \geq 0 : M_n(\omega) = -10\}.$$

- Prove that $\mathbb{P}(\tau < \infty) = 1$.
- Is there a constant $c > 0$ such that $\tau \leq c$ a.s.?
- Is $\{\omega \in \Omega : \inf_n M_n(\omega) \leq -10\} \in \mathcal{F}_\tau$?
- Is $\{\omega \in \Omega : \inf_n M_n(\omega) \geq 0\} \in \mathcal{F}_\tau$?

- Is $\{\omega \in \Omega : \sup_n M_n(\omega) \geq 2\} \in \mathcal{F}_\tau$?

14.

Theorem 4.3.1. *Let $M = (M_n)_{n=0}^\infty$ be a martingale with respect to the filtration $(\mathcal{F}_n)_{n=0}^\infty$, where $\mathcal{F}_n := \sigma(M_0, ..., M_n)$. Assume a stopping time $\tau \geq 1$ with $\mathbb{E}\tau < \infty$ and that*

$$\chi_{\{\tau \geq n\}} \mathbb{E}(|M_{n+1} - M_n| | \mathcal{F}_n) \leq c \; a.s.$$

for some $c > 0$ and all $n = 0, 1, 2, ...$ Then

$$\mathbb{E}|M_\tau| < \infty \quad \textit{and} \quad \mathbb{E}M_\tau = \mathbb{E}M_0.$$

- WALD's Identity: Assume i.i.d. random variables $\xi_1, \xi_2, ... : \Omega \to \mathbb{R}$ with $\mathbb{E}|\xi_1| < \infty$, $\mathcal{F}_0 := \{\emptyset, \Omega\}$, $\mathcal{F}_n := \sigma(\xi_1, ..., \xi_n)$ for $n \geq 1$, and let $\tau \geq 1$ be a stopping time with $\mathbb{E}\tau < \infty$. Prove that

$$\mathbb{E}(\xi_1 + \cdots + \xi_\tau) = \mathbb{E}\xi_1 \mathbb{E}\tau. \tag{4.1}$$

  Hint: Use $M_n := \xi_1 + \cdots + \xi_n - n\mathbb{E}\xi_1$ and Theorem 4.3.1.

- Prove the WALD identity (4.1) directly if τ and $\xi_1, \xi_2, ...$ are independent.

  Hint: See the proof of Proposition 3.2.8.

- Under the assumptions of Exercise 13 : Prove that $\mathbb{E}\tau = \infty$.

  Hint: Assume that $\mathbb{E}\tau < \infty$ and use WALD's identity.

15. Let $\Omega := [0, 1)$ and $M_n(t) := h_0(t) + \cdots + h_n(t)$, where

$$h_n(t) := 2^n \chi_{[0,1/2^{n+1})} - 2^n \chi_{[1/2^{n+1},1/2^n)},$$

and let $\mathcal{F}_n := \sigma(h_0, ..., h_n)$. (Check that $M = (M_n)_{n \geq 0}$ is a martingale.)

- Is there a constant $c > 0$ such that for all $N = 1, 2, ...$ one has

$$\int_0^1 \sup_{n=1,...,N} |M_n(t)| dt \leq c \int_0^1 |M_N(t)| dt?$$

- Is there a random variable $M_\infty : [0, 1) \to \mathbb{R} \in L_1$ such that

$$M_n = \mathbb{E}(M_\infty | \mathcal{F}_n) \quad a.s.?$$

16. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$ be a stochastic basis, $\mathcal{F}_\infty := \sigma(\bigcup_{n=0}^\infty \mathcal{F}_n)$, and $Z \in L_1$. Regarding the almost sure and L_1-convergence, what is the potential difference of the behavior of $\mathbb{E}(Z|\mathcal{F}_n) \to_n \mathbb{E}(Z|\mathcal{F}_\infty)$ and $\mathbb{E}(Z|\mathcal{F}_n) \to_n Z$? Prove Your statements by propositions of the course or give counterexamples.

17. Let $f : [0,1] \to \mathbb{R}$ be a Lipschitz function, i.e. $|f(x) - f(y)| \le L|x-y|$. Let

$$\xi_n(t) := \sum_{k=1}^{2^n} \frac{k-1}{2^n} \chi_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(t),$$

$\Omega := [0,1)$, $\mathcal{F}_n := \sigma(\xi_n)$, and

$$M_n(t) := \frac{f(\xi_n(t) + 2^{-n}) - f(\xi_n(t))}{2^{-n}}.$$

- Prove that $(\mathcal{F}_n)_{n=0}^\infty$ is a filtration and that $\mathcal{B}([0,1)) = \sigma(\bigcup_{n=0}^\infty \mathcal{F}_n)$.
- Prove that $(M_n)_{n=0}^\infty$ is a martingale with $|M_n(t)| \le L$.
- Prove that there is an integrable function $g : [0,1) \to \mathbb{R}$ such that $M_n = \mathbb{E}(g|\mathcal{F}_n)$ a.s.
- Prove that $f(\frac{k}{2^n}) = f(0) + \int_0^{\frac{k}{2^n}} g(t)dt$ for $k = 0, ..., .2^n - 1$.
- Prove that $f(x) = f(0) + \int_0^x g(t)dt$ for $x \in [0,1]$, i.e. g is the generalized derivative of f.

18.

Theorem 4.3.2 (Backward martingales)**.** *Let $(\mathcal{F}_n)_{n=1}^\infty$ be a decreasing sequence of σ-algebras, $\mathcal{F}_\infty := \bigcap_n \mathcal{F}_n$, and let $Z \in L_1$. Then one has that*

$$\lim_n \mathbb{E}(Z|\mathcal{F}_n) = \mathbb{E}(Z|\mathcal{F}_\infty) \textit{ in } L_1 \textit{ and a.s.}$$

Let $\varepsilon_1, \varepsilon_2, ... : \Omega \to \mathbb{R}$ be i.i.d. Bernoulli random variables, i.e. $\mathbb{P}(\varepsilon_k = \pm 1) = 1/2$, and let $S_n := \varepsilon_1 + \cdots + \varepsilon_n$.

- Prove that

$$\mathbb{E}(\varepsilon_1 | \sigma(S_n, S_{n+1}, S_{n+2}, ...)) = \frac{S_n}{n} \quad a.s.$$

- Using Theorem 4.3.2 prove that

$$\lim_n \frac{S_n}{n} = m \text{ in } L_1 \text{ and a.s.}$$

where $m = \mathbb{E}\varepsilon_1 = 0$.

19. Let $\mathbb{P}_k$ and μ_k be measures on $\Omega_k := \{-1, 1\}$ with $\mathcal{F}_k := 2^{\Omega_k}$ such that $\mathbb{P}_k(\{-1\}) = \mathbb{P}_k(\{1\}) = 1/2$, and $\mu_k(\{-1\}) = p_k$ and $\mu_k(\{1\}) = q_k$ with $p_k + q_k = 1$, where

$$p_k := \frac{1}{2} + \sqrt{\frac{1}{k} - \frac{1}{k^2}} \quad \text{for} \quad k \geq 4$$

and $p_k, q_k \in (0, 1)$ for $k = 1, 2, 3, 4$. Decide whether $\times_{n=1}^{\infty} \mu_k$ is absolutely continuous with respect to $\times_{k=1}^{\infty} \mathbb{P}_k$.

Hint: Use a discrete version of Proposition 3.9.7.

Index

Bibliography

[1] C. Geiss and S. Geiss. *An introduction to probability theory*. Lecture Notes 60, Department of Mathematics and Statistics, University of Jyväskylä, 2009.

[2] S. Geiss. *An introduction to probability theory II*. Lecture Notes 61, Department of Mathematics and Statistics, University of Jyväskylä, 2009.

[3] P. Hitczenko. *Probabilistic Analysis of Sorting Algorithms*. Report 96, Department of Mathematics and Statistics, University of Jyväskylä, 2004.

[4] P.-A. Meyer. *Probability and Potential*. Blaisdell, Waltham, Mass., 1966.

[5] J. Neveu. *Discrete parameter martingales*. North–Holland, Amsterdam, 1975.

[6] A.N. Širjaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, 1984.

[7] D. Williams. Probability with martingales. 1991.