A few exercises in analysis, probability, and combinatorics

Benjamin Dadoun

v20200202

Contents

CHAPTER 1

ANALYSIS

1.1 Measure theory

Exercise 1.1.1. Give an example of a set E, a σ-algebra $\mathscr{A}$ on E and an application $f\colon E \to F$ such that
$$\left\{ f(A)\colon A \in \mathscr{A} \right\}$$
is *not* a σ-algebra on $f(E)$.

Solution of Exercise 1.1.1. Take for instance $E := \{1,2,3\}$, $\mathscr{A} := \sigma(\{\{3\}\}) = \{\emptyset, \{1,2\}, \{3\}, E\}$, and f defined on E by $f(i) := (-1)^i$. Then $f(\emptyset) = \emptyset$, $f(\{1,2\}) = f(E) = \{-1,1\}$, and $f(\{3\}) = \{-1\}$, but
$$\left\{ f(A)\colon A \in \mathscr{A} \right\} = \left\{ \emptyset, \{-1\}, \{-1,1\} \right\}$$
is not a σ-algebra on $f(E)$ (it is not stable by complement, as it does not contain $\{1\} = f(E) \setminus \{-1\}$). (This is due to the lack of injectivity of f.) ∎

Exercise 1.1.2. Let
$$\mathscr{C} := \left\{ [a,b)\colon a,b \in \mathbb{Q},\ a < b \right\}.$$
Prove that the σ-algebra $\sigma(\mathscr{C})$ generated by $\mathscr{C}$ is the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ of $\mathbb{R}$.
Hint. Recall that $\mathbb{Q}$ is dense in $\mathbb{R}$.

Solution of Exercise 1.1.2. By definition $\mathscr{B}(\mathbb{R})$ is a σ-algebra containing all open subsets of $\mathbb{R}$. As $[a,b) = (a-1,b) \setminus (a-1,a)$ is the difference of two such subsets, it is clear that $\mathscr{C} \subseteq \mathscr{B}(\mathbb{R})$. Since $\sigma(\mathscr{C})$ is the smallest σ-algebra containing $\mathscr{C}$ we have $\sigma(\mathscr{C}) \subseteq \mathscr{B}(\mathbb{R})$.

Let $b \in \mathbb{R}$. For all $n \geqslant 1$ there exists a rational number b_n in the interval $(b, b+1/n)$. Then $b < b_n < b + 1/n$ for all $n \geqslant 1$ and b_n tends to b as $n \to \infty$. We observe that
$$(-\infty, b] = \bigcup_{m\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} [-m, b_n).$$

Indeed, if $x \in \mathbb{R}$ with $x \leqslant b$ then we can find $m \in \mathbb{N}$ such that $-m \leqslant x$, and because $b < b_n$ we have $x \in [-m, b_n)$ for all $n \in \mathbb{N}$. Conversely, if $x \in \mathbb{R}$ is such that $-m \leqslant x < b_n$ for some $m \in \mathbb{N}$ and for all $n \in \mathbb{N}$, then taking the limit $n \to \infty$ gives $-m \leqslant x \leqslant b$ and in particular $x \in (-\infty, b]$. Since $\sigma(\mathscr{C})$ is stable by countable unions and intersections, we deduce that $\sigma(\mathscr{C})$ contains the family of intervals $(-\infty, b]$ with $b \in \mathbb{R}$. But the smallest σ-algebra containing this family is precisely $\mathscr{B}(\mathbb{R})$, so $\sigma(\mathscr{C}) \supseteq \mathscr{B}(\mathbb{R})$. In conclusion, $\sigma(\mathscr{C}) = \mathscr{B}(\mathbb{R})$. ∎

Exercise 1.1.3. Let E, F be two sets, $\mathscr{A}$ and $\mathscr{B}$ two σ-algebras on E and F respectively and $f\colon E \to F$ an application. Recall the notion of *inverse image*

$$f^{-1}\langle B\rangle := \left\{x \in E\colon f(x) \in B\right\}$$

of a subset $B \subseteq F$ by f.

1. Prove that the set of subsets of E defined by

$$f^{-1}\langle \mathscr{B}\rangle := \left\{f^{-1}\langle B\rangle\colon B \in \mathscr{B}\right\}$$

is a σ-algebra on E.

2. Prove that the set of subsets of F defined by

$$f[\mathscr{A}] := \left\{B \subseteq F\colon f^{-1}\langle B\rangle \in \mathscr{A}\right\}$$

is a σ-algebra on F.

Solution of Exercise 1.1.3. In both questions we need to check that the three axioms defining a σ-algebra are fulfilled.

1. First, $f^{-1}\langle F\rangle = E$ since $f(x) \in F$ for all $x \in E$, so $E \in f^{-1}\langle \mathscr{B}\rangle$. Second, if $A \in f^{-1}\langle \mathscr{B}\rangle$ then $A = f^{-1}\langle B\rangle$ for some $B \in \mathscr{B}$, so $F \setminus B$ is also in $\mathscr{B}$ because $\mathscr{B}$ is a σ-algebra on F, and since

$$\begin{aligned} f(x) \in F \setminus B &\iff f(x) \notin B \\ &\iff x \notin f^{-1}\langle B\rangle \\ &\iff x \in E \setminus f^{-1}\langle B\rangle, \end{aligned}$$

which means that $f^{-1}\langle F \setminus B\rangle = E \setminus f^{-1}\langle B\rangle = E \setminus A$, the set $f^{-1}\langle \mathscr{B}\rangle$ is stable by complement. Third, if $A_1, A_2, \ldots \in f^{-1}\langle \mathscr{B}\rangle$, then there exist $B_1, B_2, \ldots \in \mathscr{B}$ such that $A_n = f^{-1}\langle B_n\rangle$ for all $n \geqslant 1$. Since $\bigcup_{n\geqslant 1} B_n$ is again in $\mathscr{B}$ (because $\mathscr{B}$ is a σ-algebra), and

$$\begin{aligned} f(x) \in \bigcup_{n\geqslant 1} B_n &\iff \exists n \geqslant 1,\ f(x) \in B_n \\ &\iff \exists n \geqslant 1,\ x \in f^{-1}\langle B_n\rangle \\ &\iff x \in \bigcup_{n\geqslant 1} f^{-1}\langle B_n\rangle, \end{aligned}$$

we have that

$$\bigcup_{n\geqslant 1} A_n = f^{-1}\left\langle \bigcup_{n\geqslant 1} B_n \right\rangle \in f^{-1}\langle \mathscr{B}\rangle,$$

and $f^{-1}\langle \mathscr{B}\rangle$ is stable by countable union. Hence $f^{-1}\langle \mathscr{B}\rangle$ is a σ-algebra on E.

2. First $f^{-1}\langle F\rangle = E \in \mathscr{A}$ since $\mathscr{A}$ is a σ-algebra, so $F \in f[\mathscr{A}]$. Second, if $B \in f[\mathscr{A}]$ then $f^{-1}\langle B\rangle \in \mathscr{A}$, and because again $\mathscr{A}$ is a σ-algebra on E, $E \setminus f^{-1}\langle B\rangle = f^{-1}\langle F \setminus B\rangle \in \mathscr{A}$, proving that $F \setminus B \in f[\mathscr{A}]$ and so $f[\mathscr{A}]$ is stable by complement. Third, if $B_1, B_2, \ldots \in \mathscr{B}$ then $f^{-1}\langle B_1\rangle, f^{-1}\langle B_2\rangle, \ldots \in \mathscr{A}$ and thus
$$f^{-1}\left\langle \bigcup_{n\geqslant 1} B_n \right\rangle = \bigcup_{n\geqslant 1} f^{-1}\langle B_n\rangle \in \mathscr{A}$$
(because $\mathscr{A}$ is a σ-algebra!), so $f[\mathscr{A}]$ is stable by countable union. Hence $f[\mathscr{A}]$ is a σ-algebra on F. ∎

Exercise 1.1.4. Let $E := [0,1)$, $n \in \{1,2,\ldots\}$ and $0 =: a_0 < a_1 < \cdots < a_n := 1$. Give $\sigma(\mathscr{C})$, the smallest σ-algebra on E which contains all elements of
$$\mathscr{C} := \Big\{[a_{i-1}, a_i) \colon 1 \leqslant i \leqslant n\Big\}.$$

Solution of Exercise 1.1.4. Write $[n] := \{1,2,\ldots,n\}$ and $\mathscr{P}_n$ for the powerset of $[n]$. Let
$$\mathscr{A} := \left\{\bigcup_{i\in I}[a_{i-1}, a_i) \colon I \in \mathscr{P}_n\right\}.$$
Then $\sigma(\mathscr{C}) = \mathscr{A}$. Indeed:

(i) $\mathscr{A}$ is a σ-algebra: it contains E (given by $I = [n]$) and is stable by complement (if $A \in \mathscr{A}$ is given by $I \in \mathscr{P}_n$, then $E \setminus A$ is given by $[n] \setminus I \in \mathscr{P}_n$) and by (countable) union (take $I = \cup_j I_j$);

(ii) $\mathscr{A}$ contains the elements of $\mathscr{C}$ (given by $I = \{1\}, \{2\}, \ldots, \{n\}$);

(iii) $\mathscr{A}$ is contained in $\sigma(\mathscr{C})$, since the σ-algebra $\sigma(\mathscr{C})$ must contain (countable) unions of elements from $\mathscr{C}$. ∎

Exercise 1.1.5. Let E be a set. Show that
$$\mathscr{A} := \Big\{A \subseteq E \colon A \text{ or } E \setminus A \text{ is finite or countably infinite}\Big\}$$
is a σ-algebra on E.

Solution of Exercise 1.1.5. First, $\emptyset = E \setminus E$ is finite so $E \in \mathscr{A}$. Second, if $A \in \mathscr{A}$, then either $E \setminus A$ or $A = E \setminus (E \setminus A)$ is finite or countably infinite, so $E \setminus A \in \mathscr{A}$ and $\mathscr{A}$ is thus stable by set difference. Third, let $A_1, A_2, \ldots \in \mathscr{A}$ and $A := \bigcup_{n\geqslant 1} A_n$. If for each $n \geqslant 1$, A_n is finite or countably infinite then so is A. If otherwise there exists $n \geqslant 1$ such that $E \setminus A_n$ is finite or countably infinite, then $E \setminus A$ must be also finite or countably infinte, since $E \setminus A \subseteq E \setminus A_n$. In any case, either A or $E \setminus A$ is finite or countably infinite, so $\mathscr{A}$ is stable by union. It follows by definition that $\mathscr{A}$ is a σ-algebra on E. ∎

Exercise 1.1.6. Let E be a set, $\mathscr{A}$ a σ-algebra on E and μ, ν two measures on $(E, \mathscr{A})$ such that $\mu(E) = \nu(E) = 1$. Prove that the set
$$\mathscr{D} := \Big\{A \in \mathscr{A} \colon \mu(A) = \nu(A)\Big\}$$
is a Dynkin system.

Solution of Exercise 1.1.6. First, $E \in \mathscr{D}$ since $\mu(E) = 1 = \nu(E)$. Second, if $A \in E$ then $E \setminus A$ is in $\mathscr{A}$ (because $\mathscr{A}$ is a σ-algebra) and

$$\mu(E \setminus A) = \mu(E) - \mu(A) = \nu(E) - \nu(A) = \nu(E \setminus A),$$

hence $E \setminus A \in \mathscr{D}$ (note that the equalities above hold because all quantities are finite). Third, if $A_1, A_2, \ldots$ are pairwise disjoints elements of $\mathscr{D}$ then $\bigsqcup_{n\geqslant 1} A_n \in \mathscr{A}$ (because $A_1, A_2, \ldots \in \mathscr{A}$ and $\mathscr{A}$ is a σ-algebra) and

$$\mu\left(\bigsqcup_{n\geqslant 1} A_n\right) = \sum_{n\geqslant 1} \mu(A_n) = \sum_{n\geqslant 1} \nu(A_n) = \nu\left(\bigsqcup_{n\geqslant 1} A_n\right)$$

by the σ-additivity of measures. It follows by definition that $\mathscr{D}$ is a Dynkin system. ∎

Exercise 1.1.7. Let E be a set and $\mathscr{A}$ a σ-algebra on E. We suppose that $\mathscr{A}$ is finite or countably infinite. For x in E we define

$$A(x) := \bigcap_{\substack{A\in\mathscr{A} \\ \text{s.t. } x\in A}} A.$$

1. Show that $A(x) \in \mathscr{A}$ and that $A(x)$ is the smallest element of $\mathscr{A}$ which contains x. (Prove that for all $A \in \mathscr{A}, \quad x \in A \implies A(x) \subseteq A$.)

2. Show that for all $x, y \in E, \quad y \in A(x) \implies A(x) = A(y)$.

   Hint. Use that $E \setminus A(y) \in \mathscr{A}$.

3. Let $\mathscr{E} := \{B \subseteq E : \exists x \in E,\ B = A(x)\}$. Prove that $\mathscr{A} = \sigma(\mathscr{E})$.

4. Let $\mathscr{P}(\mathscr{E})$ denote the powerset of $\mathscr{E}$. Show that the application

$$\begin{aligned} \Phi \colon \mathscr{P}(\mathscr{E}) &\longrightarrow \mathscr{A} \\ \mathscr{B} &\longmapsto \bigcup_{B\in\mathscr{B}} B \end{aligned}$$

   is injective.

Remark. This exercise proves that there is no countably infinite σ-algebra (as the powerset of any set cannot be countably infinite).

Solution of Exercise 1.1.7.

1. By assumption, $\mathscr{A}$ is either finite or countably infinite, so $A(x)$ is an intersection over a finite or countably infinite family of elements of $\mathscr{A}$. Because σ-algebras are stable by countable intersection, we have $A(x) \in \mathscr{A}$. The second assertion is immediate.

2. Assume $y \in A(x)$. The result of Question 1 with $A = A(x)$ gives the inclusion $A(y) \subseteq A(x)$. To show $A(y) \subseteq A(x)$ it is thus enough to prove that $x \in A(y)$. But if $x \in E \setminus A(y)$, applying the result of Question 1 with $A = E \setminus A(y)$ gives $A(x) \subseteq E \setminus A(y)$, and therefore $A(y) \subseteq A(x) \subseteq E \setminus A(y)$ which yields the contradiction $y \in A(y) \cap (E \setminus A(y))$. Hence $A(x) = A(y)$.

3. Because $A(x) \in \mathscr{A}$ for all $x \in E$, it is immediate that $\mathscr{E} \subseteq \mathscr{A}$ and thus $\sigma(\mathscr{E}) \subseteq \mathscr{A}$ because $\mathscr{A}$ is a σ-algebra. Since for all $A \in \mathscr{A}$,
$$A = \bigcup_{\substack{B \in \mathscr{A} \text{ s.t.} \\ \exists x \in A,\ B = A(x)}} B,$$
where the union is over a family which is at most countably infinite, the reverse inclusion holds.

4. First, as $\mathscr{E} \subseteq \mathscr{A}$ the set $\mathscr{E}$ and therefore any of its subsets $\mathscr{B} \in \mathscr{P}(\mathscr{E})$ is finite or countably infinite (so that $\Phi(\mathscr{B}) \in \mathscr{A}$); Φ is thus well defined from $\mathscr{P}(\mathscr{E})$ to $\mathscr{A}$. Second, let $\mathscr{B}, \mathscr{B}' \in \mathscr{P}(\mathscr{E})$ be different. By eventually exchanging $\mathscr{B}$ and $\mathscr{B}'$ we can assume that there exists $x \in E$ such that $A(x) \in \mathscr{B}$ and $A(x) \notin \mathscr{B}'$. Then $\Phi(\mathscr{B}) \neq \Phi(\mathscr{B}')$ since $x \in \Phi(\mathscr{B})$ (because $A(x) \in \mathscr{B}$ and $x \in A(x) \subseteq \Phi(\mathscr{B})$) but $x \notin \Phi(\mathscr{B}')$ (because $x \in B'$ for some $B' \in \mathscr{B}'$ would imply $B' = A(x)$ by Question 2, but this is impossible since $A(x) \notin \mathscr{B}'$ by assumption). ■

Exercise 1.1.8. Let $(\Omega, \mathscr{A}, \mu)$ be a measured space. We write
$$\mathscr{N}_\mu := \Big\{ N \subseteq \Omega \colon\ \exists B \in \mathscr{A},\ N \subseteq B \text{ and } \mu(B) = 0 \Big\}$$
for the set of μ-negligible subsets of Ω. Recall also the completion of $\mathscr{A}$ w.r.t. μ:
$$\mathscr{A}_\mu := \Big\{ A \subseteq \Omega \colon\ \exists (E,F) \in \mathscr{A}^2,\ E \subseteq A \subseteq F \text{ and } \mu(F \setminus E) = 0 \Big\}.$$
It is known that $\mathscr{A}_\mu \supseteq \mathscr{A}$ is still a σ-algebra on Ω.
Show that $\mathscr{A}_\mu = \{A \subseteq \Omega \colon\ \exists (E,N) \in \mathscr{A} \times \mathscr{N}_\mu,\ A = E \cup N\}$.

Solution of Exercise 1.1.8. We proceed by double inclusion. Let $A := E \cup N$ with $E \in \mathscr{A}$ and $N \in \mathscr{N}_\mu$. There exists $B \in \mathscr{A}$ such that $N \subseteq B$ and $\mu(B) = 0$. Then $E \subseteq E \cup N \subseteq E \cup B =: F \in \mathscr{A}$ with $\mu(F \setminus E) \leqslant \mu(B) = 0$, so $A \in \mathscr{A}_\mu$. Conversely, let $A \in \mathscr{A}_\mu$. Then there exist $E, F \in \mathscr{A}$ such that $E \subseteq A \subseteq F$ and $\mu(B) = 0$, where $B := F \setminus E$. But $N := A \setminus E \subseteq B \in \mathscr{A}$, so we have $A = E \cup N$ with $E \in \mathscr{A}$ and $N \in \mathscr{N}_\mu$. The equality is thus established. ■

Exercise 1.1.9. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space and $f_n \colon \Omega \to [-\infty, \infty]$, $n \in \mathbb{N}$, be a sequence of measurable functions such that
$$f(\omega) := \lim_{n \to \infty} f_n(\omega)$$
exists for μ-almost every $\omega \in \Omega$. We denote by D the domain of the function f.

1. Recall briefly why $D \in \mathscr{A}$ and $f \colon D \to [-\infty, \infty]$ is measurable.

2. Recall what "μ-almost every" means in general, and here in terms of $\Omega \setminus D$.

3. We suppose that $f_n \geqslant 0$ for all $n \in \mathbb{N}$, and that the limit
$$L := \lim_{n \to \infty} \int f_n \, \mathrm{d}\mu$$
exists in $[0, \infty)$.

a) What can you say about $\displaystyle\int f\,\mathrm{d}\mu$? Does it exist, is it finite? What if $L = 0$?

b) Show with the help of a counterexample that in general, $\displaystyle\int f_n\,\mathrm{d}\mu \not\longrightarrow \int f\,\mathrm{d}\mu$.

c) What additional *sufficient* condition on (f_n) would imply $\displaystyle\int f_n\,\mathrm{d}\mu \longrightarrow \int f\,\mathrm{d}\mu$?

4. We no longer make the assumptions of Question 3, and suppose instead that f_n is integrable for every $n \in \mathbb{N}$.

a) Show with the help of a counterexample that f is not necessarily integrable.

b) What additional *sufficient* condition on the sequence (f_n) would guarantee both the integrability of f and the convergence $\displaystyle\int f_n\,\mathrm{d}\mu \longrightarrow \int f\,\mathrm{d}\mu$?

Solution of Exercise 1.1.9.

1. The domain of f can be written as the inverse image $D = \Phi^{-1}\langle\Delta\rangle$, where

$$\Delta := \Big\{(x, y) \in [-\infty,\infty]^2 : x = y\Big\}$$

is a closed (thus measurable) set of $[-\infty,\infty]^2$, and $\Phi := (\limsup f_n, \liminf f_n)$ is a measurable function from Ω to $[-\infty,\infty]^2$ since its components are two measurable functions, *e.g.*

$$\{\omega \in \Omega : \limsup f_n(\omega) \geqslant t\} = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geqslant n} \Big\{\omega \in \Omega : f_k(\omega) \geqslant t - \frac{1}{n}\Big\} \in \mathscr{A}, \qquad t \in \mathbb{R}.$$

Then $f : D \to [-\infty,\infty]$ is measurable because it is the restriction to $D \in \mathscr{A}$ of the measurable function $\limsup f_n$.

2. We say that a (not necessarily measurable) set $A \subseteq \Omega$ occurs μ-*almost everywhere* if its complement $\Omega \setminus A$ is μ-*negligible*, *i.e*, there is $N \in \mathscr{A}$ with $N \supset \Omega \setminus A$ and $\mu(N) = 0$. Here, "$\lim f_n(\omega)$ exists for μ-a.e. $\omega \in \Omega$" means that D occurs μ-almost everywhere, which is also equivalent to $\mu(\Omega \setminus D) = 0$ because $D \in \mathscr{A}$.

3. a) As measurable nonnegative functions, the integrals (with respect to μ) of f and f_n, $n \in \mathbb{N}$, are well defined (and nonnegative). Moreover

$$\int f\,\mathrm{d}\mu = \int \liminf f_n\,\mathrm{d}\mu \leqslant \liminf \int f_n\,\mathrm{d}\mu = L < \infty$$

by Fatou's lemma. If $L = 0$, then the integral of f is 0 and consequently $f = 0$ μ-almost everywhere (that is, $f = 0$ up to a μ-negligible set of points in Ω): indeed, if it were not the case we would have, by monotonicity of the measure,

$$\lim_{n\to\infty}\uparrow \mu\Big(\Big\{\omega \in D : f(\omega) \geqslant \frac{1}{n}\Big\}\Big) = \mu\big(\{\omega \in D : f(\omega) > 0\}\big) > 0,$$

hence

$$\int f\,\mathrm{d}\mu \geqslant \frac{1}{n}\,\mu\left(\left\{\omega\in D\colon f(\omega)\geqslant\frac{1}{n}\right\}\right)$$

(this is Markov's inequality[1]) would be positive for some $n\in\mathbb{N}$.

b) Take $(\Omega,\mathscr{A},\mu) := ((0,1),\mathscr{B}((0,1)),\mathrm{d}x)$, and $f_n := n\mathbb{1}_{(0,\frac{1}{n})}$ for $n\in\mathbb{N}$. Then as $n\to\infty$, $f_n(x)\to 0 =: f(x)$ for every $x\in(0,1)$. But Fatou's inequality is strict here: for all $n\in\mathbb{N}$,

$$\int f\,\mathrm{d}\mu = 0 < 1 = \int f_n\,\mathrm{d}\mu.$$

c) The integrals converge if (f_n) is a non-decreasing sequence of nonnegative funtions, that is $0\leqslant f_n\leqslant f_{n+1}$ for all $n\in\mathbb{N}$ (monotone convergence theorem).

4. a) Take $(\Omega,\mathscr{A},\mu) := ([-1,1],\mathscr{B}([-1,1]),\mathrm{d}x)$, and $f_n\colon x\mapsto nx/(1+nx^2)$, $n\in\mathbb{N}$. These are odd continuous functions on $[-1,1]$, there are thus integrable with

$$\int_{[-1,1]} f_n(x)\,\mathrm{d}x = 0 \qquad (\star)$$

for all $n\in\mathbb{N}$; however the pointwise limit $\lim_{n\to\infty} f_n(x) = \mathbb{1}_{\{x\neq 0\}}/x$ is not integrable on $[-1,1]$, even though here the integrals $(\star)$ obviously converge.

b) An additional sufficient condition would be $|f_n|\leqslant g$ for every $n\in\mathbb{N}$, where $g\in\mathrm{L}^1(\Omega,\mathscr{A},\mu)$ is some integrable function (Lebesgue's dominated convergence theorem). ■

Exercise 1.1.10. Let λ_2 denote the Lebesgue measure on $(\mathbb{R}^2,\mathscr{B}(\mathbb{R}^2))$, and

$$D := \{(s,s)\colon s\in(0,1)\},\quad E := \{(s,s+t)\colon s,t\in(0,1)\}.$$

Justify that $D,E\in\mathscr{B}(\mathbb{R}^2)$ and use the translation invariance of λ_2 to show that

1. $\lambda_2(D)=0$,
2. $\lambda_2(E)=1$.

Solution of Exercise 1.1.10. Let $f\colon (s,t)\mapsto t-s$. It is a continuous and therefore measurable map.

1. We see that $D=\{(s,t)\in(0,1)^2\colon f(s,t)\geqslant 0\}\cap\{(s,t)\in(0,1)^2\colon f(s,t)\leqslant 0\}$, so $D\in\mathscr{B}(\mathbb{R}^2)$. Let $A_i := (\frac{i}{n},\frac{i}{n})\in\mathbb{R}^2$. Then for all $n\in\mathbb{N}$,

$$D\subseteq\bigcup_{i=0}^{n-1}\left(A_i+\left[0,\frac{1}{n}\right]^2\right),$$

so $\lambda_2(D)\leqslant n\lambda_2([0,\frac{1}{n}]^2)\leqslant\frac{1}{n}$ by translation invariance. Hence $\lambda_2(D)=0$.

[1] a.k.a. Chebyshev's (first) inequality.

2. Here $E = \{(s,t) \in (0,1) \times \mathbb{R} \colon f(s,t) < 1\} \cap \{(s,t) \in (0,1) \times \mathbb{R} \colon f(s,t) > 0\}$, so $E \in \mathscr{B}(\mathbb{R}^2)$. Moreover, E is the disjoint union $E = T_1 \sqcup S \sqcup ((0,1) + T_2)$, where $S := (0,1) \times \{1\}$ is a segment and

$$T_1 := \{(s,t) \in (0,1)^2 \colon s < t\}, \quad T_2 := \{(s,t) \in (0,1)^2 \colon t < s\}$$

are triangles (draw a picture). First we have $\lambda_2(S) = \lambda_1((0,1))\,\lambda_1(\{1\}) = 0$, and second $\lambda_2((0,1) + T_2) = \lambda_2(T_2)$ by translation invariance. Consequently $\lambda_2(E) = \lambda_2(T_1 \sqcup T_2)$. As $(0,1)^2 = T_1 \sqcup T_2 \sqcup D$, we deduce that $\lambda_2(E) = 1$. ■

Exercise 1.1.11 (True or false?). Let λ denote the Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Prove or disprove (with a counterexample) the following statements:

1. Let $A \in \mathscr{B}(\mathbb{R})$.
   a) If $B \subseteq A$ then $B \in \mathscr{B}(\mathbb{R})$.
   b) If $\lambda(A) = \infty$ then A is an unbounded set.
   c) If $\lambda(A) < \infty$ then A is a bounded set.
   d) If $\lambda(A) = 0$ then A is a bounded set.
   e) If A is an open set then $\lambda(A) > 0$.
   f) If $\lambda(A \cap (0,1)) = 1$ then $A \cap (0,1)$ is dense in $(0,1)$.
   g) If $A \cap (0,1)$ is dense in $(0,1)$ then $\lambda(A \cap (0,1)) > 0$.
   h) If $\lambda(A) > 0$ then A has a non-empty interior.
2. (In the following statements, measurability is meant w.r.t. the Borel σ-field.)
   a) If $f \colon \mathbb{R} \to \mathbb{R}$ is differentiable, then f' is measurable.
   b) If $f_1, f_2, \ldots \colon \mathbb{R} \to \mathbb{R}$ are measurable functions, then the set $B := \{x \in \mathbb{R} \colon \lim_{n\to\infty} f_n(x) \text{ exists}\}$ is measurable.
   c) If $f \colon [0,1] \to \mathbb{R}$ is such that $\{x \in [0,1] \colon f(x) = c\}$ is measurable for all $c \in \mathbb{R}$, then f is measurable.

Solution of Exercise 1.1.11.

1. a) False, since there exists $B \subset \mathbb{R}$ with $B \notin \mathscr{B}(\mathbb{R})$.
   b) True. If A is a bounded set then there exists $r > 0$ such that $A \subseteq [-r, r]$, and therefore $\lambda(A) \leqslant 2r < \infty$.
   c) False, since for instance the unbounded Borel set

   $$\bigsqcup_{n=0}^{\infty} \left[n, n + \frac{1}{n!} \right)$$

   has Lebesgue measure $e < \infty$.
   d) False, since for instance $\mathbb{Q}$ is an unbounded Borel set with $\lambda(\mathbb{Q}) = 0$.

e) False, since $\emptyset$ is an open set with $\lambda(\emptyset) = 0$. However, any Borel set $A \in \mathscr{B}(\mathbb{R})$ with *non-empty* interior (in particular, any non-empty open set A) must contain some non-empty interval (a, b), so $\lambda(A) \geqslant b - a > 0$.

f) True. Indeed $\lambda((0,1) \setminus A) = 1 - \lambda(A \cap (0,1)) = 0$ so by the previous answer $(0,1) \setminus A$ cannot have a non-empty interior, which precisely rephrases that $A \cap (0,1)$ is dense in $(0,1)$.

g) False, since $\mathbb{Q} \cap (0,1)$ is dense in $(0,1)$ and $\lambda(\mathbb{Q} \cap (0,1)) \leqslant \lambda(\mathbb{Q}) = 0$.

h) False, since $(0,1) \setminus \mathbb{Q}$ has an empty interior and $\lambda((0,1) \setminus \mathbb{Q}) = 1$. See also the fat Cantor set for an example of a *nowhere dense* set having yet a positive Lebesgue measure.

2. a) True, since f' is a limit of measurable functions:

$$f'(x) = \lim_{n \to \infty} \frac{f(x + \frac{1}{n}) - f(x)}{\frac{1}{n}}.$$

■

b) True, since $B = \{x \in \mathbb{R}\colon \liminf f_n(x) = \limsup f_n(x)\}$.

c) False. Let $A \subset [0,1]$ be some non-measurable set. Define $f\colon [0,1] \to \mathbb{R}$ by $f(x) = x$ if $x \in A$ and $f(x) = -x$ else. Then $\{x \in \mathbb{R}\colon f(x) = c\}$ is a subset of $\{\pm c\}$, so measurable (for any c), but $f^{-1}\langle [0,1] \rangle = A \notin \mathscr{B}(\mathbb{R})$.

Exercise 1.1.12. Let λ_n denote the Lebesgue measure on $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. Show that for any hyperplane $H \subset \mathbb{R}^n$, $\lambda_n(H) = 0$.

Hint. Show first $\lambda_n(H_0) = 0$ for the hyperplane $H_0 := \{(x_1, \dots, x_n) \in \mathbb{R}^n\colon x_n = 0\}$.

Solution of Exercise 1.1.12. As $H_0 = \mathbb{R}^{n-1} \times \{0\}$, we have $\lambda_n(H_0) = 0$ (*could you explain why?*). Now if $H \subset \mathbb{R}^n$ is any hyperplane, there exists a non-zero vector $e_n \in \mathbb{R}^n$ such that $H \oplus \mathbb{R}e_n = \mathbb{R}^n$. Let $(e_1, \dots, e_{n-1})$ be any basis of H and $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be the linear isomorphism which maps the standard basis of $\mathbb{R}^n$ onto $(e_1, \dots, e_n)$. Then

$$\lambda_n(H) = \lambda_n(T(H_0)) = |\det(T)|\, \lambda_n(H_0) = 0.$$

■

Exercise 1.1.13. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space and $f\colon \Omega \to [-\infty, \infty]$.

1. We suppose that $f \in \mathrm{L}^1(\Omega, \mathscr{A}, \mu)$. Show that $|f(\omega)| < \infty$ for μ-a.e. $\omega \in \Omega$.

2. We suppose that there is a sequence f_n, $n \in \mathbb{N}$, converging to f in $\mathrm{L}^1(\Omega, \mathscr{A}, \mu)$. Show that there is a subsequence (f_{n_k}) of (f_n) converging to f μ-a.e., that is

$$\lim_{k \to \infty} f_{n_k}(\omega) = f(\omega)$$

for μ-a.e. $\omega \in \Omega$.

Solution of Exercise 1.1.13.

1. By monotonicity of the measure, we have
$$\forall n \in \mathbb{N}, \qquad \mu(\{\omega \in \Omega \colon |f(\omega)| = \infty\}) \leqslant \mu(\{\omega \in \Omega \colon |f(\omega)| > n\}),$$
where, applying Markov's inequality,
$$\mu(\{\omega \in \Omega \colon |f(\omega)| > n\}) \leqslant \frac{\|f\|_1}{n} \xrightarrow[n\to\infty]{} 0.$$
Hence $\mu(\{\omega \in \Omega \colon |f(\omega)| = \infty\}) = 0$, that is $|f(\omega)| < \infty$ for μ-a.e. $\omega \in \Omega$.

2. Since $f_n \to f$ in $\mathrm{L}^1(\Omega, \mathscr{A}, \mu)$, we can build (by induction on $k \in \mathbb{N}$) an increasing sequence $(n_k)_{k\in\mathbb{N}}$ of integers such that $\|f_{n_k} - f\|_1 < 2^{-k}$ for all $k \in \mathbb{N}$. But then by monotone convergence theorem we have
$$\int \left(\sum_{k=1}^{\infty} |f_{n_k} - f| \right) \mathrm{d}\mu = \sum_{k=1}^{\infty} \|f_{n_k} - f\|_1 < \infty,$$
which using the result of Question 1 implies that the series
$$\sum_{k=1}^{\infty} |f_{n_k}(\omega) - f(\omega)|$$
converges for μ-a.e. $\omega \in \Omega$; in particular $|f_{n_k}(\omega) - f(\omega)| \to 0$ at least for such ω. ∎

Exercise 1.1.14. Let $a \in \mathbb{C}$ with $|a| < 1$. Show that the two sums
$$\sum_{n=1}^{\infty} \frac{a^n}{1 - a^{2n}} \qquad \text{and} \qquad \sum_{m=1}^{\infty} \frac{a^{2m-1}}{1 - a^{2m-1}}$$
are well defined and equal.
Hint. Introduce $f_{n,m} := a^{n(2m-1)}$ for $m, n \in \mathbb{N}$ and apply Fubini's theorem.

Solution of Exercise 1.1.14. We see $f \colon (n,m) \in \mathbb{N}^2 \mapsto f_{n,m} \in \mathbb{C}$ as a measurable function on the product of the σ-finite measure space $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \#)$ with itself, where $\#$ is the counting measure. Then
$$\int \#(\mathrm{d}m) \int \#(\mathrm{d}n)\, |f_{n,m}| \overset{\text{def}}{=} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (|a|^{2m-1})^n = \sum_{m=1}^{\infty} \frac{|a|^{2m-1}}{1 - |a|^{2m-1}},$$
which is a converging series; indeed
$$\sum_{m=1}^{\infty} \frac{|a|^{2m-1}}{1 - |a|^{2m-1}} \leqslant \frac{1}{1 - |a|} \sum_{m=1}^{\infty} |a|^{2m-1} = \frac{|a|}{(1 - |a|)(1 - |a|^2)} < \infty$$
— we could also check the convergence by the ratio test. Now, theorems of Fubini-Tonelli and Fubini-Lebesgue say that f is integrable and that we may interchange the orders of summation. The two sums
$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f_{n,m} = \sum_{m=1}^{\infty} \frac{a^{2m-1}}{1 - a^{2m-1}}$$
and
$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f_{n,m} = \sum_{n=1}^{\infty} a^{-n} \sum_{m=1}^{\infty} (a^{2n})^m = \sum_{n=1}^{\infty} \frac{a^n}{1 - a^{2n}}$$
are thus well defined and equal. ∎

Exercise 1.1.15. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space.

1. Let $n \in \mathbb{N}$ and $f_1, \dots, f_n \in \mathrm{L}^n(\Omega, \mathscr{A}, \mu)$. Show that
$$\|f_1 \cdots f_n\|_1 \leqslant \prod_{i=1}^{n} \|f_i\|_n.$$
Hint. Proceed by induction and recall Hölder's inequality.

2. We suppose here that $\mu = \mathbb{P}$ is a probability measure (*i.e*, $\mathbb{P}(\Omega) = 1$). Show that for every finite family $\{A_1, \dots, A_n\} \subseteq \mathscr{A}$ of events on Ω,
$$\mathbb{P}(A_1 \cap \cdots \cap A_n) \leqslant \left[\mathbb{P}(A_1) \cdots \mathbb{P}(A_n)\right]^{1/n}.$$
Remark. By comparison between arithmetic and geometric means, this inequality is sharper than the (trivial) inequality
$$\mathbb{P}(A_1 \cap \cdots \cap A_n) \leqslant \frac{\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)}{n}.$$

Solution of Exercise 1.1.15.

1. We show by induction on $1 \leqslant k \leqslant n$ that $f_1 \cdots f_k \in \mathrm{L}^{n/k}(\Omega, \mathscr{A}, \mu)$ holds together with the inequality
$$\|f_1 \cdots f_k\|_{n/k} \leqslant \prod_{i=1}^{k} \|f_i\|_n. \tag{$\star$}$$
Case $k = 1$ is trivial. Suppose $2 \leqslant k \leqslant n$. By induction we know that $f_1 \cdots f_{k-1} \in \mathrm{L}^{n/(k-1)}(\Omega, \mathscr{A}, \mu)$. Now if we apply Hölder's inequality to the product of $(f_1 \cdots f_{k-1})^{n/k} \in \mathrm{L}^{k/(k-1)}(\Omega, \mathscr{A}, \mu)$ by $f_k^{n/k} \in \mathrm{L}^k(\Omega, \mathscr{A}, \mu)$, with the conjugate exponents $k/(k-1)$ and k respectively (as $\frac{1}{k/(k-1)} + \frac{1}{k} = 1$), then we precisely get that $f_1 \cdots f_k \in \mathrm{L}^{n/k}(\Omega, \mathscr{A}, \mu)$ and
$$\|f_1 \cdots f_k\|_{n/k} \leqslant \|f_1 \cdots f_{k-1}\|_{n/(k-1)} \|f_k\|_n,$$
which, using the induction hypothesis, gives ($\star$) as desired. This inequality is therefore true for $k = n$, what we wanted to show.

2. This follows directly from Question 1 with $f_i := \mathbb{1}_{A_i}$, that is $f_i(\omega) = 1$ if $\omega \in A_i$ and 0 otherwise. Indeed, $\|f_1 \cdots f_n\|_1 = \mathbb{P}(A_1 \cap \cdots \cap A_n)$ and $\|f_i\|_n = \mathbb{P}(A_i)^{1/n}$. ■

Exercise 1.1.16. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. We suppose that there exists a measurable function $f : \Omega \to (0, \infty)$ such that f and $1/f$ are integrable (w.r.t. μ). Prove that μ is finite.

Solution of Exercise 1.1.16. We have $1 \leqslant \frac{1}{2}(x + \frac{1}{x})$ for all $x > 0$. Therefore, by linearity of $\int$,
$$\mu(\Omega) = \int_\Omega 1 \,\mathrm{d}\mu \leqslant \frac{1}{2}\left(\int_\Omega f \,\mathrm{d}\mu + \int_\Omega \frac{1}{f} \,\mathrm{d}\mu\right) < \infty.$$
■

Exercise 1.1.17. Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and $f\colon \Omega \to [0,\infty]$ be a measurable function. Let $E_t := \{\omega \in \Omega\colon f(\omega) > t\}$ for each $t > 0$. Prove that

$$\int_\Omega f \,\mathrm{d}\mu = \int_{(0,\infty)} \mu(E_t)\,\mathrm{d}\lambda_1(t).$$

Solution of Exercise 1.1.17. Note that f is a nonnegative measurable function and for all $t > 0$, $\mu(E_t) = \int_\Omega \mathbb{1}_{\{f(\omega)>t\}}\,\mu(\mathrm{d}\omega)$. Therefore, by Fubini–Tonelli's theorem,

$$\begin{aligned}\int_{(0,\infty)} \mu(E_t)\,\mathrm{d}\lambda_1(t) &= \int_\Omega \left(\int_{(0,\infty)} \mathbb{1}_{\{f(\omega)>t\}}\,\mathrm{d}\lambda_1(t)\right)\mu(\mathrm{d}\omega)\\ &= \int_\Omega \left(\int_{(0,f(\omega))} \mathrm{d}\lambda_1\right)\mu(\mathrm{d}\omega)\\ &= \int_\Omega f\,\mathrm{d}\mu.\end{aligned}$$

■

Exercise 1.1.18.

1. Compute the double integral
$$\iint_{(0,\infty)^2} \frac{\mathrm{d}\lambda_2(x,y)}{(1+y)(1+yx^2)}.$$

2. Deduce that
$$\int_{(0,\infty)} \frac{\log x}{x^2-1}\,\mathrm{d}\lambda_1(x) = \frac{\pi^2}{4}.$$
Hint. Observe that $\dfrac{1}{(1+y)(1+x^2y)} = \dfrac{1}{x^2-1}\left(\dfrac{x^2}{1+yx^2} - \dfrac{1}{1+y}\right)$.

3. Show that
$$\int_{(0,1)} \frac{\log x}{x^2-1}\,\mathrm{d}\lambda_1(x) = \frac{\pi^2}{8}.$$

Solution of Exercise 1.1.18.

1. The integrand is positive and continuous as a function of $(x,y) \in (0,\infty)^2$, so the integral is well defined. Applying Fubini–Tonelli's theorem gives
$$\begin{aligned}\iint_{(0,\infty)^2} \frac{\mathrm{d}\lambda_2(x,y)}{(1+y)(1+yx^2)} &= \int_{(0,\infty)} \frac{1}{1+y}\left(\int_{(0,\infty)} \frac{\mathrm{d}\lambda_1(x)}{1+yx^2}\right)\mathrm{d}\lambda_1(y)\\ &= \int_{(0,\infty)} \frac{\mathrm{d}\lambda_1(y)}{1+y}\left[\frac{\arctan(x\sqrt{y})}{\sqrt{y}}\right]_{x=0}^{x\to\infty}\\ &= \pi\int_{(0,\infty)} \frac{\mathrm{d}\lambda_1(y)}{2\sqrt{y}(1+y)}\\ &= \frac{\pi^2}{2},\end{aligned}$$
where the last integral is easily computed by substituting u to $\sqrt{y}$.

2. Tonelli's theorem gives also

$$\iint_{(0,\infty)^2} \frac{\mathrm{d}\lambda_2(x,y)}{(1+y)(1+yx^2)} = \int_{(0,\infty)} \frac{\mathrm{d}\lambda_1(x)}{x^2-1} \int_{(0,\infty)} \left(\frac{x^2}{1+yx^2} - \frac{1}{1+y}\right) \mathrm{d}\lambda_1(y)$$
$$= \int_{(0,\infty)} \frac{\mathrm{d}\lambda_1(x)}{x^2-1} \left[\log\left(\frac{1+yx^2}{1+y}\right)\right]_{y=0}^{y\to\infty}$$
$$= 2\int_{(0,\infty)} \frac{\log x}{x^2-1}\,\mathrm{d}\lambda_1(x),$$

and we conclude by the result of Question 1.

3. The change of variable $y = 1/x$ from $(0,1)$ to $(1,\infty)$ yields

$$\int_{(0,1)} \frac{\log x}{x^2-1}\,\mathrm{d}\lambda_1(x) = \int_{(1,\infty)} \frac{\log y}{y^2-1}\,\mathrm{d}\lambda_1(y),$$

and we conclude using Question 2 and the linearity of $\int$. ■

Exercise 1.1.19. Let $f\colon \mathbb{R}^2 \to [0,\infty)$ be a measurable function, and

$$I := \iint_{(0,1)^2} f\left(\sqrt{-2\log u}\cos(2\pi v), \sqrt{-2\log u}\sin(2\pi v)\right)\mathrm{d}\lambda_2(u,v).$$

Show that

$$I = \iint_{\mathbb{R}^2} f(x,y)\,\frac{e^{-\frac{x^2+y^2}{2}}}{(\sqrt{2\pi})^2}\,\mathrm{d}\lambda_2(x,y).$$

Solution of Exercise 1.1.19. Clearly, the integrand in I is a nonnegative measurable function of (u,v). We make first the change of variable $(r,\theta) = (\sqrt{-2\log u}, 2\pi v) =: \varphi(u,v)$ which is a $\mathscr{C}^1$-diffeomorphism from $(0,1)^2$ onto $(0,\infty)\times(0,2\pi)$ (the inverse map is given by $\varphi^{-1}(r,\theta) = (e^{-r^2/2}, \theta/2\pi)$). We have

$$\det(D\varphi(u,v)) = \begin{vmatrix} -\dfrac{1}{ur} & 0 \\ 0 & 2\pi \end{vmatrix} = -\left(\frac{e^{-r^2/2}}{2\pi}\,r\right)^{-1},$$

so (by the change of variable formula)

$$I = \iint_{(0,\infty)\times(0,2\pi)} f(r\cos(\theta), r\sin(\theta))\,\frac{e^{-r^2/2}}{2\pi}\,r\,\mathrm{d}\lambda_2(r,\theta).$$

This is an integral in polar form which we finally rewrite into cartesian coordinates with the classical transformation $(x,y) = (r\cos\theta, r\sin\theta)$ (polar coordinates). Then

$$I = \iint_{\mathbb{R}^2} f(x,y)\,\frac{e^{-\frac{x^2+y^2}{2}}}{(\sqrt{2\pi})^2}\,\mathrm{d}\lambda_2(x,y)$$

as stated. ■

Exercise 1.1.20.

1. Let $t > 0$. Show that
$$\int_{(0,t)} \frac{\sin x}{x} \, \mathrm{d}\lambda_1(x) = \int_{(0,\infty)} \left(\int_{(0,t)} e^{-xy} \sin x \, \mathrm{d}\lambda_1(x) \right) \mathrm{d}\lambda_1(y).$$

2. Deduce that
$$\int_{(0,t)} \frac{\sin x}{x} \, \mathrm{d}\lambda_1(x) = \int_{(0,\infty)} \frac{1 - e^{-ty}(y \sin t + \cos t)}{1 + y^2} \, \mathrm{d}\lambda_1(y)$$
for all $t > 0$, and conclude that
$$\lim_{t\to\infty} \int_{(0,t)} \frac{\sin x}{x} \, \mathrm{d}\lambda_1(x) = \frac{\pi}{2}.$$
Hint. Apply (*properly!*) the dominated convergence theorem.

3. Is the function $x \mapsto \dfrac{\sin x}{x}$ Lebesgue-integrable on $(0, \infty)$?

Solution of Exercise 1.1.20.

1. The map $f \colon (x, y) \mapsto e^{-xy} \sin x$ is (jointly) measurable since it is continuous in (x, y). Fubini–Tonelli's theorem shows that
$$\begin{aligned}\iint\limits_{(0,t)\times(0,\infty)} |f(x,y)| \, \mathrm{d}\lambda_2(x,y) &= \int_{(0,t)} \left(\int_{(0,\infty)} e^{-xy} \, \mathrm{d}\lambda_1(y) \right) |\sin x| \, \mathrm{d}\lambda_1(x) \\ &= \int_{(0,t)} \frac{|\sin x|}{x} \, \mathrm{d}\lambda_1(x) \\ &< \infty\end{aligned}$$
(there is no divergence at 0 because we may extend $x \mapsto \frac{|\sin x|}{x}$ continuously with the value 1). Therefore f is integrable on $(0, t) \times (0, \infty)$ and by Fubini–Lebesgue's theorem we can compute the integral in any order:
$$\int_{(0,t)} \left(\int_{(0,\infty)} f(x,y) \, \mathrm{d}\lambda_1(y) \right) \mathrm{d}\lambda_1(x) = \int_{(0,\infty)} \left(\int_{(0,t)} f(x,y) \, \mathrm{d}\lambda_1(x) \right) \mathrm{d}\lambda_1(y).$$
But this is exactly
$$\int_{(0,t)} \frac{\sin x}{x} \, \mathrm{d}\lambda_1(x) = \int_{(0,\infty)} \left(\int_{(0,t)} e^{-xy} \sin x \, \mathrm{d}\lambda_1(x) \right) \mathrm{d}\lambda_1(y).$$

2. The first part is easy: writing $\Im z$ for the imaginary part of $z \in \mathbb{C}$, we have
$$\begin{aligned}\int_{(0,t)} e^{-xy} \sin x \, \mathrm{d}\lambda_1(x) &= \Im\left(\int_{(0,t)} e^{(-y+i)x} \, \mathrm{d}\lambda_1(x)\right) \\ &= -\Im\left(\frac{y+i}{1+y^2}\left[e^{(-y+i)x}\right]_{x=0}^{x=t}\right) \\ &= \frac{1-e^{-ty}(y\sin t + \cos t)}{1+y^2}\end{aligned}$$
for all (real) $y > 0$. Now, for any sequence of positive reals $(t_n)_{n\in\mathbb{N}}$ going to infinity, the sequence of measurable maps on $(0,\infty)$
$$f_n\colon y \mapsto \frac{1-e^{-t_n y}(y\sin t_n + \cos t_n)}{1+y^2}, \qquad n \in \mathbb{N}$$
i) converges pointwise to the function $g\colon y \mapsto \frac{1}{1+y^2}$, and *ii)* is dominated by the integrable function $2g$, so we may apply the dominated convergence theorem. We get
$$\begin{aligned}\int_{(0,t_n)} \frac{\sin x}{x} \, \mathrm{d}\lambda_1(x) &= \int_{(0,\infty)} f_n(y) \, \mathrm{d}\lambda_1(y) \\ &\xrightarrow[n\to\infty]{} \int_{(0,\infty)} g(y) \, \mathrm{d}\lambda_1(y) = \Big[\arctan y\Big]_{y=0}^{y\to\infty} = \frac{\pi}{2},\end{aligned}$$
and as this is true for any sequence $(t_n)_{n\in\mathbb{N}}$ going to ∞, the result follows.

3. No, it is not! Indeed, since $|\sin(x+k\pi)| = |\sin x|$ for any integer k, we have
$$\begin{aligned}\int_{(0,\infty)} \frac{|\sin x|}{x} \, \mathrm{d}\lambda_1(x) &\geqslant \sum_{k=1}^{\infty} \int_{((k-1)\pi,k\pi)} \frac{|\sin x|}{x} \, \mathrm{d}\lambda_1(x) \\ &\geqslant \sum_{k=1}^{\infty} \frac{1}{k\pi} \int_{((k-1)\pi,k\pi)} |\sin x| \, \mathrm{d}\lambda_1(x) \\ &= \sum_{k=1}^{\infty} \frac{1}{k\pi} \int_{(0,\pi)} \sin x \, \mathrm{d}\lambda_1(x) \\ &= \sum_{k=1}^{\infty} \frac{2}{k\pi} \\ &= \infty.\end{aligned}$$
■

Exercise 1.1.21. Let $(\Omega, \mathscr{A}, \mu)$ be a measure space and $f\colon \Omega \to \mathbb{R}$ be a measurable function. For $p \in [1,\infty]$, we set $\|f\|_p := \infty$ if $f \notin \mathrm{L}^p(\Omega, \mathscr{A}, \mu)$.

1. Let $1 \leqslant p < q \leqslant \infty$ and suppose *for this question only* that $\mu(\Omega) < \infty$. Show that
$$\|f\|_p \leqslant \mu(\Omega)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q$$
(with the convention $1/\infty = 0$).

   Remark. We get $\mathrm{L}^q(\Omega, \mathscr{A}, \mu) \subset \mathrm{L}^p(\Omega, \mathscr{A}, \mu)$ (under the above conditions).

2. Suppose that $f \in \mathrm{L}^r(\Omega, \mathscr{A}, \mu)$ for some $1 \leqslant r < \infty$. Prove that

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty.$$

Hint. Show $\liminf_{p\to\infty} \|f\|_p \geqslant \|f\|_\infty$ using Chebyshev's inequality.

Solution of Exercise 1.1.21.

1. We assume $\|f\|_q < \infty$, otherwise the inequality holds trivially. If $q = \infty$, then the inequality is clear since $|f| \leqslant \|f\|_\infty$ μ-a.e. and

$$\|f\|_p^p = \int_\Omega |f|^p \,\mathrm{d}\mu \leqslant \int_\Omega \|f\|_\infty^p \,\mathrm{d}\mu = \mu(\Omega)\,\|f\|_\infty^p.$$

Suppose now $q \neq \infty$. We apply Hölder's inequality to the function $|f|^p \in \mathrm{L}^{q/p}(\Omega, \mathscr{A}, \mu)$ and the constant function $g :\equiv 1 \in \mathrm{L}^r(\Omega, \mathscr{A}, \mu)$ (since $\mu(\Omega) < \infty$), where $1 \leqslant r < \infty$ is such that $p/q + 1/r = 1$:

$$\|f\|_p^p = \||f|^p \cdot g\|_1 \leqslant \||f|^p\|_{q/p}\,\|g\|_r = \mu(\Omega)^{1-\frac{p}{q}}\,\|f\|_q^p.$$

Raising this to the power $1/p$ yields the desired inequality.

2. Let $0 < \varepsilon < \|f\|_\infty$. The set $A := \{\omega \in \Omega : |f(\omega)| > \|f\|_\infty - \varepsilon\} \in \mathscr{A}$ has positive measure by definition of $\|f\|_\infty$, and $\|f\|_p \geqslant (\|f\|_\infty - \varepsilon)\,\mu(A)^{1/p}$ by Chebyshev's inequality, so

$$\liminf_{p\to\infty} \|f\|_p \geqslant \|f\|_\infty - \varepsilon.$$

(If $\mu(A) = \infty$, then $\liminf \|f\|_p = \infty$ and the above inequality clearly holds.) As ε is arbitrary we get $\liminf \|f\|_p \geqslant \|f\|_\infty$. To conclude, it remains to show that $\limsup \|f\|_p \leqslant \|f\|_\infty$. By assumption there exists $1 \leqslant r < \infty$ such that $\|f\|_r < \infty$. For $r < p < \infty$ we observe that $|f|^p \leqslant \|f\|_\infty^{p-r}\,|f|^r$ μ-a.e., so $\|f\|_p \leqslant \|f\|_\infty^{1-r/p}\,\|f\|_r^{r/p}$ and thus

$$\limsup_{p\to\infty} \|f\|_p \leqslant \|f\|_\infty.$$

■

Exercise 1.1.22. Suppose that $(\Omega, \mathscr{A}) := (\mathbb{Z}, \mathscr{P}(\mathbb{Z}))$ and μ is the counting measure on $\mathbb{Z}$ and consider the sequence space $\ell^p := \mathrm{L}^p(\Omega, \mathscr{A}, \mu)$ for $p \in [1, \infty]$. As above, we set $\|f\|_p := \infty$ if $f \notin \ell^p$. Show that

$$\|f\|_q \leqslant \|f\|_p$$

whenever $1 \leqslant p < q \leqslant \infty$. In particular there is the inclusion $\ell^p \subset \ell^q$.

Solution of Exercise 1.1.22. If $q = \infty$, then clearly $\|f\|_\infty := \sup_{n\in\mathbb{Z}} |f(n)| \leqslant \|f\|_p$. Suppose $q \neq \infty$ and $f \not\equiv 0$, otherwise there is nothing more to prove. Dividing both sides of the inequality by $\|f\|_p$ we may consider $f/\|f\|_p$ instead of f and it is therefore equivalent to show that

$$\sum_{n\in\mathbb{Z}} |f(n)|^q \leqslant 1$$

under the assumption $\|f\|_p = 1$. But this is immediate since then $|f(n)| \leqslant 1$ and, because $p < q$, $|f(n)|^q \leqslant |f(n)|^p$ for all $n \in \mathbb{Z}$. ■

Exercise 1.1.23. Let $p \neq q$ in $[1,\infty]$. Prove that $\mathrm{L}^p(\mathbb{R}) \setminus \mathrm{L}^q(\mathbb{R}) \neq \emptyset$.

Solution of Exercise 1.1.23. Actually, for each $p \in [1,\infty]$, we can give an example of measurable function $f\colon \mathbb{R} \to \mathbb{R}$ such that for all $q \in [1,\infty]$, $f \in \mathrm{L}^q(\mathbb{R})$ *if and only if* $q = p$. If $p = \infty$, then the constant function $f :\equiv 1$ is in $\mathrm{L}^\infty(\mathbb{R})$ but not in $\mathrm{L}^q(\mathbb{R})$ for any $q \in [1,\infty)$. If $p \neq \infty$, let us introduce the measurable function defined by

$$f(x) := \begin{cases} 0, & \text{if } x \leqslant 0, \\ x^{-1/p}(1+\log^2 x)^{-1}, & \text{if } x > 0. \end{cases}$$

We have $|f(x)|^p \leqslant x^{-1}(1+\log^2 x)^{-1}$ for all $x > 0$, so

$$\int_{\mathbb{R}} |f(x)|^p \,\lambda_1(\mathrm{d}x) \leqslant \int_{(0,\infty)} \frac{\lambda_1(\mathrm{d}x)}{x(1+\log^2 x)} = \Big[\arctan(\log x)\Big]_{x\to 0}^{x\to\infty} = \pi < \infty,$$

hence $f \in \mathrm{L}^p(\mathbb{R})$. We now show that however $f \notin \mathrm{L}^q(\mathbb{R})$ for any $q \in [1,\infty] \setminus \{p\}$. Since $f(x) \to \infty$ as $x \to 0$, we have $f \notin \mathrm{L}^\infty(\mathbb{R})$. Let $q \in (p,\infty)$. Then as $x \to 0$, $|f(x)|^q = x^{-q/p}(1+\log^2 x)^{-q}$ dominates $2x^{-r/p}$ for $p < r < q$, so there exists $\alpha > 0$ small enough such that

$$\int_{\mathbb{R}} |f(x)|^q \,\lambda_1(\mathrm{d}x) \geqslant \int_{(0,\alpha]} |f(x)|^q \,\lambda_1(\mathrm{d}x) \geqslant \int_{(0,\alpha]} x^{-r/p} \,\lambda_1(\mathrm{d}x) = \infty.$$

Suppose finally $q \in [1,p)$. Then as $x \to \infty$,

$$\int_{\mathbb{R}} |f(x)|^q \,\lambda_1(\mathrm{d}x) \geqslant \int_{[\beta,\infty)} |f(x)|^q \,\lambda_1(\mathrm{d}x) \geqslant \int_{[\beta,\infty)} x^{-s/p} \,\lambda_1(\mathrm{d}x) = \infty.$$

■

Exercise 1.1.24 (Riesz–Scheffé's lemma). Let $(\Omega,\mathscr{A},\mu)$ be a measure space, and $f, f_1, f_2, \ldots \in \mathrm{L}^p(\Omega)$ with $p \in [1,\infty)$. We suppose that, as $n \to \infty$, $f_n(\omega) \to f(\omega)$ for μ-a.e. $\omega \in \Omega$ and that $\|f_n\|_p \to \|f\|_p$. Let $\operatorname{sign}\colon \mathbb{R} \to \{-1,1\}$ denote a function such that $|x| = (\operatorname{sign} x)x$ for all $x \in \mathbb{R}$, and write

$$f_n^* := f_n \mathbb{1}_{\{|f_n| \leqslant |f|\}} + (\operatorname{sign} f_n)|f|\mathbb{1}_{\{|f_n| > |f|\}}$$

for every $n \in \mathbb{N}$.

1. Show that $\|f_n^* - f\|_p \to 0$ as $n \to \infty$.

2. Show that $\|f_n - f_n^*\|_p \to 0$ as $n \to \infty$. Conclude that $f_n \to f$ in $\mathrm{L}^p(\Omega,\mathscr{A},\mu)$.

   Hint. Use the convexity inequality $(y-x)^p \leqslant y^p - x^p$ for $0 \leqslant x \leqslant y$.

Solution of Exercise 1.1.24.

1. First, f_n^* is obviously measurable, and $|f_n^*| \leqslant |f|$ with $f \in \mathrm{L}^p(\Omega,\mathscr{A},\mu)$, so the sequence of measurable functions $|f_n^* - f|^p$, $n \in \mathbb{N}$, is dominated by the integrable function $2^p|f|^p$. Second, it follows from the first assumption that, as $n \to \infty$, $|f_n^*(\omega) - f(\omega)| \to 0$ for μ-a.e. $\omega \in \Omega$. Dominated convergence theorem then entails $\|f_n^* - f\|_p \to 0$ as $n \to \infty$.

2. Clearly,
$$|f_n - f_n^*| = \mathbb{1}_{\{|f|<|f_n|\}}(|f_n| - |f|) \leqslant \mathbb{1}_{\{|f|<|f_n|\}}(|f_n| - |f_n^*|) \leqslant |f_n| - |f_n^*|,$$
so $|f_n - f_n^*|^p \leqslant |f_n|^p - |f_n^*|^p$ using the indication. Hence, with the result of Question 1 and the second assumption,
$$\|f_n - f_n^*\|_p^p = \int_\Omega |f_n - f_n^*|^p \,\mathrm{d}\mu \leqslant \|f_n\|_p^p - \|f_n^*\|_p^p \xrightarrow[n\to\infty]{} \|f\|_p^p - \|f\|_p^p = 0.$$
We conclude from Minkowski's inequality that $f_n \to f$ in $\mathrm{L}^p(\Omega, \mathscr{A}, \mu)$. ■

Remark. One-line proof:
$$\limsup_{n\to\infty} \|f_n - f\|_p^p \;=\; 2^p\|f\|_p^p - \liminf_{n\to\infty} \int_\Omega \underbrace{\left[2^{p-1}(|f_n|^p + |f|^p) - |f_n - f|^p\right]}_{\geqslant 0 \text{ (by convexity)}} \mathrm{d}\mu \overset{\text{(Fatou)}}{\leqslant} 0.$$

Exercise 1.1.25. If $f\colon \mathbb{R} \to \mathbb{R}$ is measurable and $h \in \mathbb{R}$, we define $\tau_h f\colon x \mapsto f(x+h)$ "the translation of f by h" which is obviously also measurable. Let $1 \leqslant p < q \leqslant \infty$ such that $1/p + 1/q = 1$, $f \in \mathrm{L}^p(\mathbb{R})$ and $g \in \mathrm{L}^q(\mathbb{R})$. Recall that the convolution $f \star g$ of f and g is given by
$$f \star g(x) := \int_{\mathbb{R}} f(y) g(x-y)\,\lambda_1(\mathrm{d}y), \qquad \lambda_1\text{-a.e.} \tag{$\star$}$$

1. Show that $\tau_h f \to f$ in $\mathrm{L}^p(\mathbb{R})$ as $h \to 0$.

   Hint. Approximate f smoothly; note that $p < \infty$.

2. In the special case $p = 1$ (so $q = \infty$), show that the definition in ($\star$) is actually valid *everywhere* and makes $f \star g$ be a bounded and uniformly continuous function.

Solution of Exercise 1.1.25.

1. The change of variable $y \leftarrow y + h$ easily shows that for any $h \in \mathbb{R}$, $\tau_h f \in \mathrm{L}^p(\mathbb{R})$ with $\|\tau_h f\|_p = \|f\|_p$. Let $g \in \mathscr{C}_c(\mathbb{R})$ be a continuous function with compact support. Then clearly $\tau_h g \to g$ pointwise as $h \to 0$ (and even uniformly since g is uniformly continuous). *A fortiori* $\tau_h g \to g$ in $\mathrm{L}^p(\mathbb{R})$. (This last assertion follows for instance from the dominated convergence theorem, or from either Exercise 1.1.21 (Question 1) or Exercise 1.1.24...) Let $\varepsilon > 0$ and recall that $p < \infty$. By the density theorem in L^p, there exists $g \in \mathscr{C}_c(\mathbb{R})$ such that $\|f - g\|_p \leqslant \varepsilon$. Minkowski's inequality then gives
$$\begin{aligned}\|\tau_h f - f\|_p &\leqslant \|\tau_h f - \tau_h g\|_p + \|\tau_h g - g\|_p + \|g - f\|_p \\ &= 2\|f - g\|_p + \|\tau_h g - g\|_p \\ &\leqslant 2\varepsilon + \|\tau_h g - g\|_p\end{aligned}$$
for any $h \in \mathbb{R}$. Using what precedes we get
$$\limsup_{h\to 0} \|\tau_h f - f\|_p \leqslant 2\varepsilon,$$
and since ε is arbitrary, we conclude that $\tau_h f \to f$ in $\mathrm{L}^p(\mathbb{R})$.

2. For all $x \in \mathbb{R}$,

$$\int_{\mathbb{R}} |f(y)g(x-y)|\,\lambda_1(\mathrm{d}y) \leqslant \|g\|_\infty \int_{\mathbb{R}} |f(y)|\,\lambda_1(\mathrm{d}y) = \|f\|_1\,\|g\|_\infty < \infty,$$

so the definition of $f \star g$ by ($\star$) is valid everywhere. Furthermore the triangle inequality gives $|f \star g(x)| \leqslant \|f\|_1\,\|g\|_\infty$, so $f \star g \colon \mathbb{R} \to \mathbb{R}$ is a bounded function. Finally, simple changes of variables successively lead to

$$\begin{aligned} |f\star g(x) - f\star g(y)| &= \left|\int_{\mathbb{R}} [f(t+x) - f(t+y)]g(-t)\,\lambda_1(\mathrm{d}t)\right| \\ &\leqslant \|g\|_\infty \int_{\mathbb{R}} |f(t+x-y) - f(t)|\,\lambda_1(\mathrm{d}t) \\ &= \|g\|_\infty\,\|\tau_{x-y}f - f\|_1 \end{aligned}$$

for all $x, y \in \mathbb{R}$. Using Question 1, we get that $|f\star g(x) - f\star g(y)|$ vanishes as $|x-y| \to 0$, which means that $f \star g$ is also uniformly continuous. ∎

Exercise 1.1.26. Let λ denote the Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Recall that, by translation invariance of λ, for any $E \in \mathscr{B}(\mathbb{R})$ with $\lambda(E) > 0$ the set

$$E - E := \{x - y\colon\ x, y \in E\}$$

contains some open interval centered at 0: $\exists \varepsilon > 0,\ (-\varepsilon, \varepsilon) \subset E - E$. In this exercise we suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is a measurable function such that

$$\forall (x,y) \in \mathbb{R}^2, \quad f(x+y) = f(x) + f(y). \tag{$\star$}$$

1. For $k \in \mathbb{N}$, justify that the set $E_k := \{x \in \mathbb{R}\colon |f(x)| < k\}$ is in $\mathscr{B}(\mathbb{R})$ and, by observing the identity

$$\bigcup_{k\in\mathbb{N}}\!\uparrow E_k = \mathbb{R},$$

show that there exist $k \in \mathbb{N}$ and $\varepsilon > 0$ such that: $\quad |x| < \varepsilon \implies |f(x)| < 2k$.

2. Deduce that $f(x) \to 0$ as $x \to 0$.

3. Conclude that $f(x) = f(1)x$ for all $x \in \mathbb{R}$.

   Hint. Use the density of $\mathbb{Q}$ in $\mathbb{R}$.

Solution of Exercise 1.1.26.

1. We have $E_k = \{x \in \mathbb{R}\colon f(x) < k\} \cap \{x \in \mathbb{R}\colon f(x) > -k\} \in \mathscr{B}(\mathbb{R})$ by the measurability of f. Next, the observation is clear since $E_k \subseteq E_{k+1}$ and any $x \in \mathbb{R}$ lies in E_k for $k \in \mathbb{N}$ with $k > |f(x)|$. By monotonicity of the measure, we have $\infty = \lambda(\mathbb{R}) = \lim_{k\to\infty}\lambda(E_k)$ so in particular there exists $k \in \mathbb{N}$ such that $\lambda(E_k) > 0$. Thanks to the recalled result there exists also some $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \subseteq E_k - E_k$. This means that when $|x| < \varepsilon$ we can find $u, v \in E_k$ with $x = u - v$, and thus

$$|f(x)| = |f(u-v)| = |f(u) - f(v)| \leqslant |f(u)| + |f(v)| < 2k$$

(where the second equality follows easily from ($\star$)).

2. Property ($\star$) also entails $f(rx) = rf(x)$ if $r \in \mathbb{Z}$, and even if $r \in \mathbb{Q}$ since $qf(p/q \cdot x) = f(px) = pf(x)$ for all $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. Then with the previous notations we have $|f(x)| = |f(qx)|/q \leqslant 2k/q$ when $q|x| < \varepsilon$, so $\limsup |f(x)| \leqslant 2k/q$ as $x \to 0$. But the latter holds for any $q \in \mathbb{N}$; hence $f(x) \to 0$ as $x \to 0$.

3. Let $x \in \mathbb{R}$. By the density of $\mathbb{Q}$ in $\mathbb{R}$ there exist rational numbers $r_n \in \mathbb{Q}$ such that $r_n \to x$ as $n \to \infty$. Now using the preceding observations and the triangle inequality,

$$|f(x) - xf(1)| \overset{(\star)}{=} |f(x) - f(r_n) + r_n f(1) - xf(1)| \leqslant |f(x - r_n)| + |r_n - x||f(1)|$$

where the right-hand side tends to 0 as $n \to \infty$, giving $f(x) = xf(1)$. ■

1.2 Linear differential equations

Exercise 1.2.1. Solve (over $\mathbb{R}$) the following systems of linear differential equations:

1. $\begin{cases} x' = x + z \\ y' = -y - z \\ z' = 2y + z \end{cases}$

2. $\begin{cases} x' = 2x - y + 4t \\ y' = x + e^{-t} \end{cases}$

3. $\begin{cases} x' = \cos(t)x - \sin(t)y \\ y' = \sin(t)x + \cos(t)y \end{cases}$

   Hint. Rewrite the system as a first order differential equation in $z := x + iy$.

Solution of Exercise 1.2.1.

1. The system can be rewritten as $X'(t) = AX(t)$ where A is the (constant) matrix

$$A := \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 1 \end{pmatrix}.$$

   The characteristic polynomial of A is $P(X) := (X-1)(X+i)(X-i)$. Hence A is diagonalizable over $\mathbb{C}$, with simple eigenvalues 1, i, $-i$. Eigenspaces are:

$$\ker(A - I_3) = \mathbb{C}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \ker(A - iI_3) = \mathbb{C}\begin{pmatrix} 1+i \\ 1-i \\ -2 \end{pmatrix}, \ \ker(A + iI_3) = \mathbb{C}\begin{pmatrix} 1-i \\ 1+i \\ -2 \end{pmatrix},$$

from which we deduce a basis of (complex) solutions for the system. The general (complex) solution has then the form

$$t \mapsto \alpha e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta e^{it} \begin{pmatrix} 1+i \\ 1-i \\ -2 \end{pmatrix} + \gamma e^{-it} \begin{pmatrix} 1-i \\ 1+i \\ -2 \end{pmatrix}, \qquad \alpha, \beta, \gamma \in \mathbb{C}.$$

The general real solution to the system is obtained by taking the real and imaginary parts:

$$t \mapsto \lambda e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} \cos t - \sin t \\ \cos t + \sin t \\ -2\cos t \end{pmatrix} + \nu \begin{pmatrix} \sin t + \cos t \\ \sin t - \cos t \\ -2\sin t \end{pmatrix}, \qquad \lambda, \mu, \nu \in \mathbb{R}.$$

2. Suppose that (x, y) is a solution. As then $x' = 2x - y + 4t$ where the right-hand side is $\mathscr{C}^1$, expression of y' we have $x'' = 2x' - y' + 4 = 2x' - x + 4 - e^{-t}$. Two independent solutions to the associated homogeneous, linear, second order differential equation with constant coefficients $x'' - 2x' + x = 0$ are $t \mapsto e^t$ and $t \mapsto te^t$. Hence the general solution to this homogeneous equation is of the form
$$x_0\colon t \mapsto (a + bt)e^t,$$
where $a, b \in \mathbb{R}$. To get a solution to the inhomogeneous differential equation $x'' - 2x' + x = 4 - e^{-t}$, there are several methods:

First method. We rewrite the equation and the previous solution into the "first order style":

$$\begin{pmatrix} x \\ x' \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix} + \begin{pmatrix} 0 \\ 4 - e^{-t} \end{pmatrix}, \tag{1.1}$$

$$X\colon t \mapsto a \begin{pmatrix} e^t \\ e^t \end{pmatrix} + b \begin{pmatrix} te^t \\ (1+t)e^t \end{pmatrix}. \tag{1.2}$$

Then we let the constants a and b vary with t in X. We find that X is a solution to (1.1) if and only if:
$$\begin{cases} a'(t)e^t + b'(t)te^t &= 0, \\ a'(t)e^t + b'(t)(1+t)e^t &= 4 - e^{-t}, \end{cases}$$
giving (recall Cramer's rule):
$$\begin{cases} a'(t) = te^{-2t} - 4te^{-t}, \\ b'(t) = 4e^{-t} - e^{-2t}. \end{cases}$$
Integrating with respect to t yields:
$$\begin{cases} a(t) = \frac{1}{4}e^{-2t}(16(1+t)e^t - 2t - 1) + A, \\ b(t) = \frac{1}{2}e^{-2t}(1 - 8e^t) + B, \end{cases}$$
where $A, B \in \mathbb{R}$. We now plug this into the first line of (1.2). We deduce that if (x, y) is a solution to the system, then *necessarily* x is of the form
$$x(t) = 4 - \frac{1}{4}e^{-t} + (A + Bt)e^t.$$

Now $y = 2x - x' + 4t$ gives:

$$y(t) = 4t - \frac{3}{4}e^{-t} + (A - B + Bt)e^t + 8.$$

Since the space of solutions to the system has dimension 2, we have here all the solutions.

Second method. With some intuition we can guess that there should exist a solution to $x'' - 2x' + x = 4 - e^t$ of the form $x(t) = \lambda + \mu e^{-t}$. Plugging in we find that $\lambda := 4$ and $\mu := -\frac{1}{4}$ give indeed a solution. We then add the general solution to the homogeneous equation for x, and deduce y as in the first method.

Third method. We exploit the matrix notation $U'(t) = MU(t) + V(t)$, where

$$U(t) := \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad M := \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad V(t) := \begin{pmatrix} 4t \\ e^{-t} \end{pmatrix}.$$

Unfortunately M is not diagonalizable over $\mathbb{C}$. But we may notice (and check by induction) that for $n = 0, 1, 2, \dots$

$$M^n = \begin{pmatrix} n+1 & -n \\ n & 1-n \end{pmatrix}$$

(or, observe that $M = I_2 + N$ with $N^2 = 0$, and use the binomial theorem). This makes the exponentiation of M easy. For each $t \in \mathbb{R}$,

$$\exp(tM) = \sum_{n=0}^{\infty} \frac{t^n}{n!} M^n = \begin{pmatrix} (1+t)e^t & -te^t \\ te^t & (1-t)e^t \end{pmatrix}.$$

We then multiply the matrix equation by $\exp(-tM) \in \mathsf{GL}_n(\mathbb{R})$ to get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\exp(-tM)U(t)\Big] = \exp(-tM)\Big[U'(t) - MU(t)\Big] = \exp(-tM)V(t)$$

(beware that M does *not* depend on t). Therefore

$$\exp(-tM)U(t) = \int \exp(-tM)V(t)\,\mathrm{d}t + \begin{pmatrix} A \\ B \end{pmatrix}$$

with $A, B \in \mathbb{R}$ (we integrate component-wise), and finally

$$U(t) = \exp(tM)\left[\int \exp(-tM)V(t)\,\mathrm{d}t + \begin{pmatrix} A \\ B \end{pmatrix}\right].$$

3. We follow the indication. The system is equivalent to $z'(t) - e^{it}z(t) = 0$, or (multiplying by $\exp(ie^{it}) \neq 0$)

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\exp(ie^{it})z(t)\Big] = 0.$$

Thus the solution is $z(t) = C\exp(-ie^{it}) = C\exp(\sin(t) - i\cos(t))$ with $C := A + iB \in \mathbb{C}$. We go back to x and y by taking respectively the real and imaginary parts of z. In conclusion, the general solution to the system is

$$\begin{cases} x(t) = \exp(\sin(t))\,[A\cos(\cos(t)) + B\sin(\cos(t))], \\ y(t) = \exp(\sin(t))\,[B\cos(\cos(t)) - A\sin(\cos(t))], \end{cases}$$

with $A, B \in \mathbb{R}$. ■

Exercise 1.2.2. We consider the following Cauchy problem.

$$\begin{aligned}(1+t^2)x'' - t(1-t^2)x' + (1-t^2)x = 0,\\ x(0)=1, \quad x'(0)=1.\end{aligned} \tag{E}$$

1. Show that the functions $t \mapsto At$, $A \in \mathbb{R}$, are solutions to (E) but that none of them is a solution to the Cauchy problem.

2. Find all solutions to (E) by letting the constant A vary with t.

   Hint:
   $$-\frac{2+t^2+t^4}{t(1+t^2)} = \frac{2t}{1+t^2} - t - \frac{2}{t}.$$

   Now, solve the Cauchy problem.

Solution of Exercise 1.2.2.

1. The verification is immediate.

2. We look for solutions of the form $x(t) = A(t)\cdot t$. We find that x is a solution to (E) if and only if

   $$(2+t^2+t^4)A'(t) + t(1+t^2)A''(t) = 0.$$

   This is a homogeneous, linear, first order differential equation in A' where the variables can be separated:

   $$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\log|A'(t)|\Big] = -\frac{2+t^2+t^4}{t(1+t^2)} = \frac{2t}{1+t^2} - t - \frac{2}{t}.$$

   Hence
   $$A'(t) = \frac{\lambda(1+t^2)}{t^2}e^{-t^2/2},$$

   and finally
   $$A(t) = \frac{\lambda}{t}e^{-t^2/2} + \mu$$

   with $\lambda, \mu \in \mathbb{R}$. The general solution to (E) is then of the form

   $$x(t) = \lambda e^{-t^2/2} + \mu t, \qquad \lambda, \mu \in \mathbb{R}.$$

   We have $x(0) = x'(0) = 1$ if and only if $\lambda = \mu = 1$. In conclusion, the solution to the Cauchy problem is

   $$t \mapsto e^{-t^2/2} + t.$$

■

Exercise 1.2.3. The goal is to find all twice differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 1$ and

$$\forall (s,t) \in \mathbb{R}^2, \quad f(s+t) + f(s-t) = 2f(s)f(t). \tag{$\star$}$$

Let f be such a function.

1. Show that f is an even function.

2. Show that f is a solution to $x'' = \lambda x$ for some constant $\lambda \in \mathbb{R}$.

3. Conclude.

Solution of Exercise 1.2.3.

1. Substitute 0 for s into the relation ($\star$).

2. On the one hand, differentiating the relation twice with respect to s:

$$f''(s+t) + f''(s-t) = 2f(t)f''(s).$$

On the other hand, differentiating twice with respect to t:

$$f''(s+t) + f''(s-t) = 2f(s)f''(t).$$

As a consequence: $$f(t)f''(s) = f(s)f''(t)$$

for all $s, t \in \mathbb{R}$. Fixing $s = 0$ shows that f is a solution to $x'' = \lambda x$ with $\lambda := f''(0)$.

3. We now discuss regarding the sign of λ:

   (i) Case $\lambda = 0$: so $f'' = 0$, implying f is a linear function $t \mapsto At + B$. Since $f(0) = 1$ and f is an even function, necessarily $f \equiv 1$.

   (ii) Case $\lambda < 0$: here $f'' + \omega^2 f = 0$ with $\omega := \sqrt{-\lambda}$. The general solution to this equation is $t \mapsto A\cos(\omega t) + B\sin(\omega t)$. Since $f(0) = 1$ and f is an even function, we deduce that $f\colon t \mapsto \cos(\sqrt{-\lambda}\, t)$.

   (iii) Case $\lambda > 0$: similar. We find that f must be equal to $t \mapsto \cosh(\sqrt{\lambda}\, t)$.

   To summarize, any solution to the problem is *necessarily* either

$$\begin{array}{lll} & t \mapsto 1, & \\ \text{or} & t \mapsto \cos(at), & a \in \mathbb{R}, \\ \text{or} & t \mapsto \cosh(bt), & b \in \mathbb{R}. \end{array}$$

   We easily (nevertheless have to) check that these functions are indeed solutions. ■

Exercise 1.2.4. Let $A(t) := (a_{i,j}(t)) \in \mathbb{R}^{n\times n}$ be a matrix, and $X_1(t), \ldots, X_n(t) \in \mathbb{R}^n$ be n solutions to the linear differential equation

$$X'(t) = A(t)X(t). \tag{F}$$

We define

$$W(t) := \Big[X_1(t) \,\Big|\, X_2(t) \,\Big|\, \cdots \,\Big|\, X_n(t) \Big] \in \mathbb{R}^{n\times n}$$

and

$$w(t) := \det(W(t)) = \det(W_1(t), \ldots, W_n(t)),$$

where $W_1(t), \ldots, W_n(t)$ are the rows of the matrix $W(t)$.

1. Recalling that the determinant is a multilinear form, prove that
$$w'(t) = \sum_{i=1}^{n} \det\Big(W_1(t), \dots, W_{i-1}(t), W_i'(t), W_{i+1}(t), \dots, W_n(t)\Big).$$
Check also that $W_i'(t) = \sum_{j=1}^{n} a_{i,j}(t) W_j(t)$ for every $i = 1, \dots, n$.
2. Recalling that the determinant is an alternating form, deduce that w is a solution to the homogeneous, first order, linear differential equation
$$y' = \operatorname{tr}(A(t))\, y.$$
3. Prove that either $(\forall t \in \mathbb{R},\ w(t) = 0)$ or $(\forall t \in \mathbb{R},\ w(t) \neq 0)$, and that the latter happens if and only if $(X_1, X_2, \dots, X_n)$ is a basis of solutions to (F).

   Hint. Recall the isomorphism $X \mapsto X(0)$ from the solutions to (F) onto $\mathbb{R}^n$.

Solution of Exercise 1.2.4.

1. To find $w'(t)$ use the expression of the determinant:
$$w(t) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) W_{1,\sigma(1)}(t) W_{2,\sigma(2)}(t) \cdots W_{n,\sigma(n)}(t),$$
or more generally, use the n-linearity to establish the following Taylor expansion of $w(t+h)$ as $h \to 0$:
$$\begin{aligned} w(t+h) &= \det\Big(W_1(t) + hW_1'(t) + o(h), \dots, W_n(t) + hW_n'(t) + o(h)\Big) \\ &= w(t) + h \sum^{n} \det\Big(W_1(t), \dots, W_{i-1}(t), W_i'(t), W_{i+1}(t) \dots, W_n(t)\Big) + o(h). \end{aligned}$$
The expression stated for $W_i'(t)$ is a direct consequence to the matrix identity $W'(t) = A(t)W(t)$.
2. The determinant is zero whenever two rows are equal. Therefore, using the n-linearity and the results of Question 1,
$$\begin{aligned} w'(t) &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}(t) \det\Big(W_1(t), \dots, W_{i-1}(t), W_j(t), W_{i+1}(t), \dots, W_n(t)\Big) \\ &= \sum_{i=1}^{n} a_{i,i}(t) \det\Big(W_1(t), \dots, W_{i-1}(t), W_i(t), W_{i+1}(t), \dots, W_n(t)\Big) \\ &= \operatorname{tr}(A(t))\, w(t). \end{aligned}$$
3. It therefore follows that $w(t) = w(0)\exp(\int_0^t \operatorname{tr}(A(s))\,\mathrm{d}s)$, so either $w \equiv 0$ ($w(0) = 0$), or w has no zeros ($w(0) \neq 0$). Moreover, the existence and uniqueness of a solution to the Cauchy problem $\{X'(t) = A(t)X(t),\ X(0) = X_0\}$ for any $X_0 \in \mathbb{R}^n$ shows that $\Phi\colon X \mapsto X(0)$ is an isomorphism from the space of solutions to (F) onto $\mathbb{R}^n$. Clearly,
$$\begin{aligned} (X_1, \dots, X_n) \text{ basis of solutions to (F)} &\iff (\Phi(X_1), \dots, \Phi(X_n)) \text{ basis of } \mathbb{R}^n \\ &\iff (X_1(0), \dots, X_n(0)) \text{ basis of } \mathbb{R}^n \\ &\iff w(0) \neq 0 \\ &\iff \forall t \in \mathbb{R},\ w(t) \neq 0. \end{aligned}$$
■

CHAPTER 2

PROBABILITY

2.1 Combinatorial probability

Exercise 2.1.1. In an urn, there are 17 green, 5 blue, and 11 red, indistinguishable balls. Answer the following questions (specify in each case the probability space):

1. We pick two balls simultaneously (without replacement). What is the probability that none of these balls is red?
2. We pick three balls one after the other, with replacement. What is the probability that at most two of these balls are green?

Solution of Exercise 2.1.1. We label the green balls $\mathsf{G} := \{1,\dots,17\}$, the blue balls $\mathsf{B} := \{18,\dots,22\}$ and the red balls $\mathsf{R} := \{23,\dots,33\}$.

1. We can take
$$\Omega := \{\text{subsets of } \mathsf{G} \cup \mathsf{B} \cup \mathsf{R} \text{ having two elements}\}.$$
which has $|\Omega| = \binom{33}{2}$ elements. The subset $\Omega_1 := \{\omega \in \Omega : \omega \cap \mathsf{R} = \emptyset\}$ of outcomes where no red ball is picked has $|\Omega_1| = \binom{22}{2}$ elements (where $22 = |\mathsf{G} \cup \mathsf{B}|$), hence the probability $P(\Omega_1) = |\Omega_1|/|\Omega| = 7/16$.

2. Since balls are replaced in the urn, we now model the experiment by
$$\Omega := (\mathsf{G} \cup \mathsf{B} \cup \mathsf{R})^3,$$
which has 33^3 elements. The subset $\Omega_2 := \Omega \setminus \mathsf{G}^3$ of outcomes where at most two green balls are picked has $|\Omega_2| = 33^3 - 17^3$ elements, hence the probability $P(\Omega_2) = |\Omega_2|/|\Omega| = 1 - (17/33)^3$. ■

Exercise 2.1.2. We consider a 5-card hand from a traditional deck of 52 cards. Specify the probability space and find the probability that the hand contains...

1. five cards of the same suit;
2. four cards of the same rank;
3. five cards of sequential rank (the aces having both the lowest and highest ranks);
4. three cards of the same rank and two other cards of another rank.

Solution of Exercise 2.1.2. Setting $\mathsf{R} := \{1,\dots,13\}$ for the ranks and $\mathsf{S} := \{A,B,C,D\}$ for the suits, we can take

$$\Omega := \{\text{subsets of } \mathsf{R}\times\mathsf{S} \text{ with 5 elements}\},$$

which contains $|\Omega| = \binom{52}{5}$ outcomes (all possible hands).

1. Let s denote the projection $\mathsf{R}\times\mathsf{S}\to\mathsf{S}$. The subset of outcomes where all cards have the same suit is $\Omega_1 := \{\omega\in\Omega\colon s(\omega)\in\{\{A\},\{B\},\{C\},\{D\}\}$, which contains $|\Omega_1| = 4\cdot\binom{13}{5} = 5148$ elements. Hence the probability
$$P(\Omega_1) = \frac{|\Omega_1|}{|\Omega|} = \frac{33}{16660}.$$
2. Let r denote the projection $\mathsf{R}\times\mathsf{S}\to\mathsf{R}$. The subset of outcomes with four cards of the same rank is $\Omega_2 := \{\omega\in\Omega\colon \exists h\in\omega,\ r(\omega\setminus\{h\})\in\{\{1\},\dots,\{13\}\}\}$, which contains $|\Omega_2| = 13\cdot 48 = 624$ elements (fixing the rank, there are 48 possibilities for the fifth card). Hence the probability
$$P(\Omega_2) = \frac{|\Omega_2|}{|\Omega|} = \frac{1}{4165}.$$
3. The possible sequences of ranks are
$$\mathscr{S} := \{\{1,2,3,4,5\},\{2,3,4,5,6\},\dots\{9,10,11,12,13\},\{10,11,12,13,1\}\}.$$
Fixing the first component $i\in\{1,\dots,10\}$ of such sequences and choosing the suit of each card, there are 4^5 ways to obtain $\{i,i+1,\dots,i+4\}$ (identifying the ranks 1 and 14 of the aces). We deduce that the subset $\Omega_3 := \{\omega\in\Omega\colon r(\omega)\in\mathscr{S}\}$ of outcomes with five cards of sequential rank contains $|\Omega_3| = 10\cdot 4^5 = 10240$ elements. Hence the probability
$$P(\Omega_3) = \frac{|\Omega_3|}{|\Omega|} = \frac{128}{32487}.$$
4. There are $\binom{4}{3}$ ways of combining three cards of a given rank, and $12\cdot\binom{4}{2}$ pairs of some other rank, so the subset $\Omega_4\subseteq\Omega$ of outcomes having three cards of the same rank and two other cards of another rank has $|\Omega_4| = 13\cdot\binom{4}{3}\cdot 12\cdot\binom{4}{2} = 3744$ elements. Hence the probability
$$P(\Omega_4) = \frac{|\Omega_4|}{|\Omega|} = \frac{6}{4165}.$$
∎

Exercise 2.1.3. Let X be a Poisson random variable with parameter $\lambda > 0$, that is

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad \text{for } k = 0,1,2,\dots$$

1. Show that $\mathbb{E}[X] = \lambda$.
2. Show that $\mathrm{Var}(X) = \lambda$.

Solution of Exercise 2.1.3.

1. Clearly,

$$\mathbb{E}[X] \overset{\text{def}}{=} \sum_{k=1}^{\infty} k\,\mathbb{P}(X = k) = \lambda\, e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

2. Using the equality above,

$$\begin{aligned}\mathrm{Var}(X) &\overset{\text{def}}{=} \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \sum_{k=1}^{\infty} k^2\, e^{-\lambda} \frac{\lambda^k}{k!} - \lambda^2 \\ &= \sum_{k=0}^{\infty} (k+1)\, e^{-\lambda} \frac{\lambda^{k+1}}{k!} - \lambda^2 \\ &= \lambda \mathbb{E}[X] + \lambda - \lambda^2 \\ &= \lambda.\end{aligned}$$

■

Exercise 2.1.4. Let $n \in \mathbb{N}$, $x \in [0,1]$ and X_n be a random variable having the binomial distribution with parameter (n, p), that is

$$\mathbb{P}(X_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, 2, \dots, n.$$

1. Show that $\mathbb{E}[X_n] = np$.
2. Show that $\mathrm{Var}(X_n) = np(1-p)$.

Solution of Exercise 2.1.4.

1. Recall that $k\binom{n}{k} = n\binom{n-1}{k-1}$. Hence

$$\mathbb{E}[X_n] \overset{\text{def}}{=} \sum_{k=1}^{n} k\,\mathbb{P}(X = k) = n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-1-k} = np(p + 1 - p)^{n-1} = np.$$

2. Using what precedes,

$$\begin{aligned}\mathrm{Var}(X_n) &\overset{\text{def}}{=} \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \sum_{k=1}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} - n^2 p^2 \\ &= n \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} p^{k+1} (1-p)^{n-1-k} - n^2 p^2 \\ &= np(1 + \mathbb{E}[X_{n-1}]) - n^2 p^2 \\ &= np(1-p).\end{aligned}$$

Remark. Another way to prove the above statements is to say that X_n just counts the number of successes in a sequence of n independent Bernoulli trials, *i.e*, $X_n = Y_1 + \cdots + Y_n$ with $Y_1, \dots, Y_n$ i.i.d. Bernoulli(p) random variables. Then by linearity of $\mathbb{E}$ we have $\mathbb{E}[X_n] = n\mathbb{E}[Y_1] = n(1\cdot p + 0\cdot(1-p)) = np$, and further, *because of independence*, $\mathrm{Var}(X_n) = n\mathrm{Var}(Y_1) = np(1-p)$ (indeed $\mathrm{Var}(Y_1) = \mathbb{E}[Y_1] - \mathbb{E}[Y_1]^2$ since $Y_1^2 = Y_1$). ■

Exercise 2.1.5. Show that $\mathscr{C} := \{[a,b)\colon a,b \in \mathbb{Q}\}$ generates the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ of $\mathbb{R}$.

Solution of Exercise 2.1.5. Recall that $\mathscr{B}(\mathbb{R})$ is the σ-algebra generated by the topology of $\mathbb{R}$. Clearly, $\sigma(\mathscr{C}) \subseteq \mathscr{B}(\mathbb{R})$. Conversely, since open sets are countable union of open intervals and σ-algebras are stable by countable unions, we just need to show that $\sigma(\mathscr{C}) \ni (a,b)$ for every $-\infty \leqslant a < b \leqslant \infty$. By density of $\mathbb{Q}$ in $\mathbb{R}$, there exist a sequence $(a_n)_{n\in\mathbb{N}}$ of rational numbers decreasing to a, and a sequence $(b_n)_{n\in\mathbb{N}}$ of rational numbers increasing to b. Then

$$(a,b) = \bigcup_{n\in\mathbb{N}} \underbrace{[a_n, b_n)}_{\in\mathscr{C}} \in \sigma(\mathscr{C}).$$

■

Exercise 2.1.6. Let $\mathscr{C} := \{C_i\}_{1\leqslant i\leqslant n}$ be a finite partition of Ω, *i.e*, $\Omega = \bigcup_{i=1}^n C_i$ with $C_1, \dots, C_n$ all non-empty and pairwise disjoint. Describe $\sigma(\mathscr{C})$, the smallest σ-algebra containing $\mathscr{C}$.

Solution of Exercise 2.1.6. Since $\sigma(\mathscr{C})$ is a σ-algebra containing $C_1, \dots, C_n$, it must also contain all possible unions of these, namely the sets

$$C(I) := \bigcup_{i\in I} C_i, \qquad I \subseteq \{1,2,\dots,n\} =: [n].$$

Let $\mathscr{A} := \{C(I)\colon I \subseteq [n]\}$. We show that conversely, $\mathscr{A} \supseteq \sigma(\mathscr{C})$. By minimality, it suffices to show that $\mathscr{A} \supseteq \mathscr{C}$ and $\mathscr{A}$ is a σ-algebra. The first point is clear because $C_i = C(\{i\}) \in \mathscr{A}$ for every $i \in [n]$. For the second point, we check the three requirements of the definition:

(i) $\mathscr{A}$ contains $C([n]) = \Omega$;

(ii) $\mathscr{A}$ is stable by complement since $\Omega \setminus C(I) = C([n] \setminus I)$ for every $I \subseteq [n]$;

(iii) $\mathscr{A}$ is stable by countable union since $\bigcup_{k\in\mathbb{N}} C(I_k) = C(\bigcup_{k\in\mathbb{N}} I_k)$ for any $I_k \subseteq [n]$.

Hence $\sigma(\mathscr{C}) = \mathscr{A}$. (This σ-algebra has 2^n elements.) ■

Exercise 2.1.7. Let $(\Omega, \mathscr{A}, P)$ be a probability space.

1. Let A, B be two events, and its *symmetric difference* $A\Delta B := (A\cup B) \setminus (A \cap B)$. Prove using the axioms of probability that
$$|P(A) - P(B)| \leqslant P(A\Delta B).$$

2. Let $A_n, n \geqslant 1$, be a sequence of events with $P(A_n) = 1$ for every n. Prove that
$$P\left(\bigcap_{n\geqslant 1} A_n\right) = 1.$$

Solution of Exercise 2.1.7.

1. We observe that $A \subseteq (A\Delta B) \cup B$. We deduce that $P(A) \leqslant P(A\Delta B) + P(B)$, *i.e*, $P(A) - P(B) \leqslant P(A\Delta B)$. We conclude by exchanging the roles of A and B.

2. First note that $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) \geqslant 1 + 1 - 1 = 1$. Iterating, we deduce that $P(B_n) = 1$ for $B_n := \bigcap_{1\leqslant k\leqslant n} A_k$, $n \geqslant 1$. Then, by a consequence of the σ-additivity property,

$$P\left(\bigcap_{n\geqslant 1} A_n\right) = P\left(\bigcap_{n\geqslant 1}\downarrow B_n\right) = \lim_{n\to\infty}\downarrow P(B_n) = 1.$$

   Remark. Alternatively we could have applied the so called "union bound" (which also follows from the axiom of σ-additivity):

$$1 - P\left(\bigcap_{n\geqslant 1} A_n\right) = P\left(\bigcup_{n\geqslant 1} A_n^{\complement}\right) \leqslant \sum_{n\geqslant 1} P(A_n^{\complement}) = 0.$$

■

Exercise 2.1.8. Let X be a random variable with values in $\mathbb{N}$. Prove that

$$\mathbb{E}[X] = \sum_{\ell=1}^{\infty} \mathbb{P}(X \geqslant \ell)$$

(with the convention that $\mathbb{E}[X] = \infty$ in case the first moment of X does not exist).

Solution of Exercise 2.1.8. The summands are nonnegative, so

$$\begin{aligned}
\sum_{\ell=1}^{\infty} \mathbb{P}(X \geqslant \ell) &= \sum_{\ell=1}^{\infty}\sum_{k=\ell}^{\infty} \mathbb{P}(X = k) \\
&= \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty} \mathbb{1}_{\{k\geqslant\ell\}}\mathbb{P}(X = k) \\
\text{(Fubini–Tonelli)} \qquad &= \sum_{k=1}^{\infty} \mathbb{P}(X = k) \sum_{\ell=1}^{\infty} \mathbb{1}_{\{\ell\leqslant k\}} \\
&= \sum_{k=1}^{\infty} k\,\mathbb{P}(X = k) \\
&= \mathbb{E}[X].
\end{aligned}$$

Remark. This is a particular case of Exercise 1.1.17. ■

2.2 Distributions, independence

Exercise 2.2.1. Suppose a distribution function F is given by

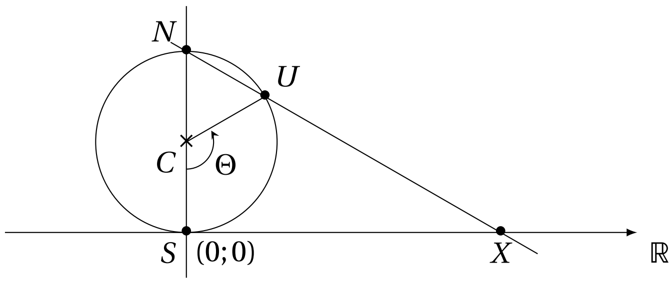
$$F(x) = \frac{1}{4}\mathbb{1}_{[0,\infty)}(x) + \frac{1}{2}\mathbb{1}_{[1,\infty)}(x) + \frac{1}{4}\mathbb{1}_{[2,\infty)}(x).$$

Let P be the probability measure, $P((-\infty, x]) := F(x)$, $x \in \mathbb{R}$. Find the probability of:

$$A = (-12, 12), \quad B = (-12, 32), \quad C = (23, 52), \quad D = [0, 2), \quad E = (3, \infty).$$

Solution of Exercise 2.2.1. Using that $P((a,b)) = F(b-) - F(a)$ and $P([a,b)) = F(b-) - F(a-)$, we find $P(A) = 1$, $P(B) = 1$, $P(C) = 0$, $P(D) = 3/4$ and $P(E) = 0$. ■

Exercise 2.2.2. For each point $U \neq N$ on the circle with center $C\,(0;1/2)$ and diameter 1 below, the line (NU) intersects the real axis at a unique point — we call X its abscissa:



We suppose that U has a uniform distribution, namely we consider that the measure Θ of the oriented angle $(\overrightarrow{CS};\overrightarrow{CU})$ is uniformly distributed on $(-\pi,\pi)$.
Show that X has the standard Cauchy distribution.

Solution of Exercise 2.2.2. By the inscribed angle theorem, a measure of $(\overrightarrow{NS},\overrightarrow{NX})$ is $\Theta/2$. Since $NS = 1$, the abscissa X is then given by $X = \tan(\Theta/2)$, where $\Theta/2$ is uniformly distributed on the interval $(-\frac{\pi}{2},\frac{\pi}{2})$. As a result, X has the standard Cauchy distribution — indeed, for every bounded continuous function $g\colon \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}[g(X)] = \mathbb{E}[g(\tan(\Theta/2))] = \int_{-\pi}^{\pi} g(\tan(\theta/2))\,\frac{\mathrm{d}\theta}{2\pi} = \int_{\mathbb{R}} g(x)\,\frac{\mathrm{d}x}{\pi(1+x^2)}.$$ ■

Exercise 2.2.3. Let X, Y be two independent Bernoulli(1/2) r.v. and $Z := \frac{1}{2}\big(1 + (-1)^{X+Y}\big)$.

1. Show that Z is a Bernoulli(1/2) r.v. which is independent of X and of Y.
2. Check that Z is *not* independent of (X,Y).

Solution of Exercise 2.2.3.

1. Clearly, $\mathbb{P}(Z=1) = \mathbb{P}(X=Y) = 1/2$ so Z is a Bernoulli(1/2) r.v. Observing that $\{Z=1, X=1\} = \{X=1\} \cap \{Y=1\}$, it is also clear that Z is independent of X. Symmetrically, Z is independent of Y as well.
2. However $\mathbb{P}(Z=1,(X,Y)=(0,1)) = 0 \neq \frac{1}{8} = \mathbb{P}(Z=1)\,\mathbb{P}((X,Y)=(0,1))$, so Z is not independent of (X,Y). ■

Exercise 2.2.4.

1. Let $X_1, X_2, \ldots$ be identically distributed real r.v. and N be a $\mathbb{N}_0$-valued r.v. We suppose N and X_ℓ independent for each $\ell \in \mathbb{N}$, and that $\mathbb{E}[|X_1|] < \infty$, $\mathbb{E}[N] < \infty$. Let the random sum

$$S(\omega) := \sum_{\ell=1}^{N(\omega)} X_\ell(\omega), \qquad \omega \in \Omega.$$

Show that S is integrable and $\mathbb{E}[S] = \mathbb{E}[N]\,\mathbb{E}[X_1]$.

Hint. Recall Exercise 2.1.8.

2. With one initial bet of 50 CHF, you are allowed to roll two fair traditional dice. Each time the sum of the two faces up is greater than or equal to 7, you win either 30 CHF or 40 CHF depending on the result of a fair coin toss, and moreover you can roll the dice again. If however the sum is less than 7, then the game is over. Is this game favorable to you?

Solution of Exercise 2.2.4.

1. We observe that $S = \lim_{n\to\infty} \sum_{\ell=1}^{n} \mathbb{1}_{\{\ell \leqslant N\}} X_\ell$, where

$$\left|\sum_{\ell=1}^{n} \mathbb{1}_{\{\ell \leqslant N\}} X_\ell\right| \leqslant \sum_{\ell=1}^{\infty} \mathbb{1}_{\{\ell \leqslant N\}} |X_\ell| \qquad (\triangle\text{-inequality + monotonicity})$$

for every $n \in \mathbb{N}$, with further

$$\begin{aligned}
\mathbb{E}\left[\sum_{\ell=1}^{\infty} \mathbb{1}_{\{\ell \leqslant N\}} |X_\ell|\right] &= \sum_{\ell=1}^{\infty} \mathbb{E}\big[\mathbb{1}_{\{\ell \leqslant N\}} |X_\ell|\big] && \text{(monotone convergence theorem)}\\
&= \sum_{\ell=1}^{\infty} \mathbb{P}(N \geqslant \ell)\,\mathbb{E}[|X_\ell|] && (N \text{ and } X_\ell \text{ independent})\\
&= \mathbb{E}[|X_1|] \sum_{\ell=1}^{\infty} \mathbb{P}(N \geqslant \ell) && (X_1, X_2, \ldots \text{ equally distributed})\\
&= \mathbb{E}[N]\,\mathbb{E}[|X_1|] && (\text{Exercise 2.1.8})\\
&< \infty.
\end{aligned}$$

It thus follows from the dominated convergence theorem that S is integrable and

$$\mathbb{E}[S] = \lim_{n\to\infty} \sum_{\ell=1}^{n} \mathbb{E}\big[\mathbb{1}_{\{\ell \leqslant N\}} X_\ell\big].$$

Repeating the last three equalities above (without the $|\cdot|$), we obtain $\mathbb{E}[S] = \mathbb{E}[N]\,\mathbb{E}[X_1]$.

Remark. This formula is known as *Wald's identity.*

2. Let S be the amount you would win in total. The sum of two dice is greater than or equal to 7 with probability $21/36 = 7/12$, so the number $N \in \{0,1,2,\ldots\}$ of times you would win either 30 CHF or 40 CHF has the geometric distribution with (success) parameter $5/12$; thus

$$\mathbb{E}[N] = \sum_{k=0}^{\infty} k \left(\frac{7}{12}\right)^k \cdot \frac{5}{12} = \frac{7}{5}.$$

The part X_k you would get at time $k = 1, 2, \dots$ is uniformly distributed on $\{30, 40\}$ and has therefore mean 35. Further, $S = X_1 + \cdots + X_N$ where N is independent of $X_1, X_2, \dots$ The result of Question 1 then applies and

$$\mathbb{E}[S] = \mathbb{E}[N]\,\mathbb{E}[X_1] = \frac{7}{5} \cdot 35 = 49 < 50.$$

On average, the game is not favorable to the player. ■

Exercise 2.2.5. For any real r.v. X, let F_X denote its cumulative distribution function.

1. Check that $\lim_{t\to-\infty} F_X(t) = 0$ and $\lim_{t\to\infty} F_X(t) = 1$.
2. Let X and Y be two *independent* r.v. having the exponential distribution with rates $\lambda > 0$ and $\mu > 0$ respectively, *e.g.*

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{if } t \geqslant 0, \\ 0, & \text{otherwise.} \end{cases}$$

   a) Let $\theta > 0$. Show that θX has the exponential distribution with rate λ/θ.
   b) Show that $Z := \min(X, Y)$ has the exponential distribution with rate $\lambda + \mu$.

3. Let $X, X_1, X_2, \dots$ be i.i.d. real r.v. We suppose that for every $n \in \mathbb{N}$, the r.v. $Z_n := n \min(X_1, \dots, X_n)$ has the same law as X and we note $S := 1 - F_X$.

   a) Show that $S(nt) = S(t)^n$ for every $n \in \mathbb{N}$ and every $t \in \mathbb{R}$.
   b) Deduce that $\mathbb{P}(X < 0) = 0$, and $S(r) = S(1)^r$ for every rational $r > 0$.
   c) Show that if $S(1) = 0$, then $\mathbb{P}(X = 0) = 1$.
   d) Assume now $S(1) \neq 0$. Show then that $0 < S(1) < 1$, and conclude that X has the exponential distribution with rate $\log(1/S(1))$.

Solution of Exercise 2.2.5.

1. Using that X has values in $\mathbb{R}$, and monotonicity (of $\mathbb{P}$, and F), we have

$$0 = \mathbb{P}(X = -\infty) = \mathbb{P}\Big(\bigcap_{n\in\mathbb{N}} \{X \leqslant -n\}\Big) = \lim_{n\to\infty} F_X(-n) = \lim_{t\to-\infty} F_X(t).$$

   Similarly, $\lim_{t\to\infty} F_X(t) = 1$.

2. a) This fact is immediately derived from the identity $F_{\theta X}(\cdot) = F_X(\cdot/\theta)$.
   b) For every $t \in \mathbb{R}$,

$$1 - F_Z(t) = \mathbb{P}\big(\min(X, Y) > t\big) = \mathbb{P}(X > t, Y > t) = \big(1 - F_X(t)\big)\big(1 - F_Y(t)\big)$$

   (the last equality holds by independence), so $1 - F_Z(t) = \min(1, e^{-(\lambda+\mu)t})$. Hence Z has the exponential distribution with rate $\lambda + \mu$.

3. a) Similarly to Question 2, for every $t \in \mathbb{R}$,

$$S(nt) = \mathbb{P}\big(\cancel{n}\min(X_1,\dots,X_n) > \cancel{n}t\big) = \mathbb{P}\big(X_1 > t,\dots,X_n > t\big) = S(t)^n,$$

where the last equality holds because $X_1,\dots,X_n$ are i.i.d. r.v. with common distribution function $F_X = 1 - S$.

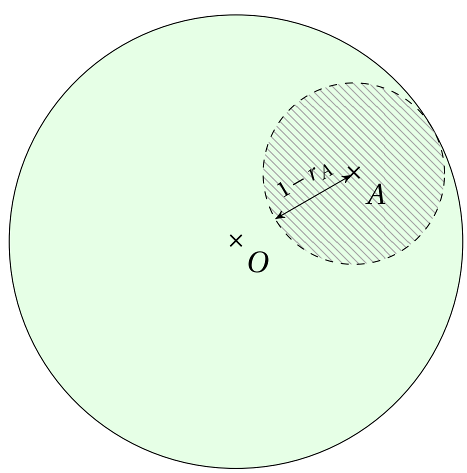
b) For $t < 0$, Question 1 entails that $S(nt) = S(t)^n$ must tend to 1 as $n \to \infty$, and since $S(t) \in [0,1]$, this forces $S(t) = 1$, *i.e*, $F_X(t) = 0$. Taking the limit as $t \to 0-$ yields $\mathbb{P}(X < 0) = F_X(0-) = 0$. Now with $t = 1/n$, we have first $S(1) = S(1/n)^n$, and second $S(m/n) = S(m \cdot 1/n) = S(1/n)^m = S(1)^{m/n}$ for arbitrary $m, n \in \mathbb{N}$.

c) If $S(1) = 0$, then $S(1/n) = S(1)^{1/n} = 0$ for each $n \in \mathbb{N}$ gives $S(0) = 0$ by right-continuity, *i.e*, $\mathbb{P}(X \leqslant 0) = 1$. Because $\mathbb{P}(X < 0) = 0$ by Question 3.b), we deduce that $\mathbb{P}(X = 0) = 1$.

d) Assume $S(1) \neq 0$, so $S(1) > 0$. The fact that $S(n) = S(1)^n$ must tend to 0 as $n \to \infty$ (Question 1) forces $S(1) < 1$. Hence $\lambda := \log(1/S(1)) \in (0,\infty)$. Now for each $t \geqslant 0$, there is a sequence (r_n) of positive rational numbers decreasing toward t as $n \to \infty$. The identity $S(r_n) = S(1)^{r_n} = \exp(-\lambda r_n)$ for each $n \in \mathbb{N}$ implies $S(t) = \exp(-\lambda t)$ by right-continuity. We conclude that X has the exponential distribution with rate $\log(1/S(1))$. (Conversely, by Question 2, if $X_1, X_2, \dots$ are i.i.d. $\mathrm{Exp}(\lambda)$-distributed r.v. for some $\lambda > 0$, then indeed $n\min(X_1,\dots,X_n)$ is also $\mathrm{Exp}(\lambda)$-distributed for every $n \in \mathbb{N}$.) ■

Exercise 2.2.6. Let A and B be two points picked independently and uniformly inside the unit disk $D := D(0;1)$. Write $Z := |AB|$ for the distance between A and B. Find the probability that the disk $D(A,Z)$ with center A and radius Z lies inside D.

Solution of Exercise 2.2.6. We begin with a picture:



We can see that "the disk of center A and radius $Z = |AB|$ lies inside D" means exactly "B belongs to the disk of center A and radius $1 - r_A$". As B is uniformly distributed, for each fixed value a of the random variable A, this has probability

$$\frac{\text{dashed area}}{\text{green area}} = \frac{\pi(1 - r_a)^2}{\pi} = (1 - r_a)^2.$$

By independence between A and B, we just need to integrate this with respect to the law of r_A. Let us find it: for every bounded continuous function $g\colon \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}[g(r_A)] &= \int_{\mathbb{R}^2} g\left(\sqrt{x^2+y^2}\right) \underbrace{\frac{1}{\pi}\mathbb{1}_{\{x^2+y^2\leqslant 1\}}\,\mathrm{d}x\mathrm{d}y}_{\text{density function of } A} \\
&= \int_0^\infty \int_0^{2\pi} g(r)\frac{1}{\pi}\mathbb{1}_{\{r\leqslant 1\}}\, r\,\mathrm{d}\theta\,\mathrm{d}r && \text{(polar coordinates)} \\
&= \int_0^\infty g(r)\left(2r\mathbb{1}_{\{r\leqslant 1\}}\right)\mathrm{d}r.
\end{aligned}$$

Hence the real random variable r_A admits the density $2r\mathbb{1}_{\{r\leqslant 1\}}$, $r>0$. The desired probability is therefore

$$\int_{\mathbb{R}} (1-r_a)^2\, P_{r_A}(\mathrm{d}r_a) = \int_0^\infty (1-r)^2\, 2r\mathbb{1}_{\{r\leqslant 1\}}\,\mathrm{d}r = \int_0^1 2r(1-r)^2\,\mathrm{d}r = \frac{1}{6}.$$

■

Exercise 2.2.7. Let X be a geometric random variable with parameter $p \in [0,1]$, that is

$$\mathbb{P}(X=k) = (1-p)^{k-1}p, \qquad k = 1,2,\ldots$$

1. Compute the c.d.f. of X.
2. Let $q \in [0,1]$ and Y be a Geometric(q) random variable independent of X. Show that $Z := \min(X,Y)$ has the geometric distribution with parameter $1-(1-p)(1-q)$.

Solution of Exercise 2.2.7.

1. Because X is integer-valued, $\mathbb{P}(X\leqslant t) = \mathbb{P}(X \leqslant \lfloor t\rfloor) = 1-(1-p)^{\lfloor t\rfloor}$ (geometric sum), for all $t\geqslant 0$.
2. By independence and Question 1, $\mathbb{P}(Z\leqslant t) = 1-\mathbb{P}(X>t)\mathbb{P}(Y>t) = 1-\big((1-p)(1-q)\big)^{\lfloor t\rfloor}$ for all $t\geqslant 0$, that is the c.d.f. of a geometric distribution with parameter $1-(1-p)(1-q)$. ■

Exercise 2.2.8.

1. Give an example of c.d.f. having an infinite number of discontinuities.
2. Show that every c.d.f. has at most countably many discontinuities.
3. Let X, Y be random variables with c.d.f. F, G respectively, and B be a Bernoulli(1/2) r.v. independent of X and of Y. Compute the c.d.f. of $Z := BX + (1-B)Y$.

Solution of Exercise 2.2.8.

1. For instance, the c.d.f. of a geometric distribution has a discontinuity at any $t \in \mathbb{N}$.

2. We first show a lemma: *Let μ be a probability measure and $(A_t)_{t\in I}$ be a sequence of pairwise disjoint events with $\mu(A_t) > 0$ for all $t \in I$. Then I is at most countable.*

   Indeed, the set
   $$I_k := \left\{ t \in I \colon \mu(A_t) \in \left(\frac{1}{k+1}, \frac{1}{k}\right]\right\},$$
   cannot have more than k elements (otherwise, if $t_1, \dots, t_{k+1} \in I_k$ are distinct then $1 \geqslant \mu(\bigcup_{i=1}^{k+1} A_{t_i}) = \sum_{i=1}^{k+1} \mu(A_{t_i}) > 1$ ↯). Therefore $I = \bigcup_{k\in\mathbb{N}} I_k$ is at most countable.

   Let now F be any c.d.f. and P the associated probability measure on $\mathbb{R}$. Applying the lemma to $\mu := P$, $I := \{t \in \mathbb{R} \colon F(t) - F(t-) > 0\}$, and $A_t := \{t\}$, $t \in I$, shows that the set I of discontinuities of F, is at most countable.

3. Partitioning w.r.t. the values of B and using independence,
   $$\begin{aligned}
   \mathbb{P}(Z \leqslant t) &= \mathbb{P}(B = 1, Z \leqslant t) + \mathbb{P}(B = 0, Z \leqslant t) \\
   &= \mathbb{P}(B = 1, X \leqslant t) + \mathbb{P}(B = 0, Y \leqslant t) \\
   &\overset{\perp\!\!\!\perp}{=} \mathbb{P}(B = 1)\,\mathbb{P}(X \leqslant t) + \mathbb{P}(B = 0)\,\mathbb{P}(Y \leqslant t) \\
   &= \frac{F(t) + G(t)}{2}.
   \end{aligned}$$
   ■

Exercise 2.2.9 (True or false?). Prove, or disprove (by giving a counterexample), briefly the following statements. We consider real r.v. on some general probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

1. *About the laws of random variables.*

   a) For every measurable function $f \colon \mathbb{R} \to \mathbb{R}$,
   b) If $\mathbb{P}(X = t) = \mathbb{P}(Y = t)$ for all $t \in \mathbb{R}$, then $\mathbb{P}(X = Y) = 1$.
   c) If $\mathbb{P}(X \leqslant t) = \mathbb{P}(Y \leqslant t)$ for all $t \in \mathbb{R}$, then $\mathbb{P}(X = Y) = 1$.
   d) If $\mathbb{P}(X = t) = \mathbb{P}(Y = t)$ for all $t \in \mathbb{R}$, then X and Y have the same law.
   e) If X and Y have same law and $X \geqslant 0$ a.s., then $Y \geqslant 0$ a.s.
   f) If X and Y have same law, then $\mathbb{P}(X < Y) = \mathbb{P}(X > Y)$.
   g) If X and Y have same law and $X \in \mathrm{L}^1(\mathbb{P})$, then $Y \in \mathrm{L}^1(\mathbb{P})$ and $\mathbb{E}[X] = \mathbb{E}[Y]$.
   h) If X and Y have same law, then $X + Z$ and $Y + Z$ also have same law.

2. *About independence.*

   a) If $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$, then X and Y are independent.
   b) If X and Y are independent, then $\mathbb{P}(X = Y) = 0$.
   c) If X and Y are independent, then $\mathbb{P}(X = Y) < 1$.
   d) If Z is independent of both X and Y, then Z is independent of (X, Y).
   e) If X, Y are independent, then so are $f(X), g(Y)$ for $f, g \colon \mathbb{R} \to \mathbb{R}$ measurable.

Solution of Exercise 2.2.9.

1. a) False: $\mathbb{E}[f(X)]$ is generally defined in $\mathbb{R}$ (resp. in $[0,\infty]$) for $f(X) \in \mathrm{L}^1(\mathbb{P})$ (resp. $f \geqslant 0$). For instance $\mathbb{E}[X]$ does not exist if X has a Cauchy distribution.

   b) False, even for discrete laws: take *e.g.* X a Bernoulli$(1/2)$ r.v. and $Y := 1-X$.

   c) False: same counterexample as in 1.b).

   d) False: consider X and Y having distinct continuous distributions. (X has a continuous distribution means that $\mathbb{P}(X=t)=0$ for all $t \in \mathbb{R}$...)

   e) True, because then $1 = \mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$ for $A := [0,\infty) \in \mathscr{B}(\mathbb{R})$.

   f) False[1]: take *e.g.* X uniformly distributed on $\{0,1,2\}$, and $Y := X+1 \mod 3$. Then $\mathbb{P}(X < Y) = \mathbb{P}(X \neq 2) = 2/3 \neq 1/3 = \mathbb{P}(X=2) = \mathbb{P}(X > Y)$. One other example involving continuous distributions could be $Y := 2\min(X,X')$ with X, X' independent standard exponential r.v.: then $X \sim Y$ (exercise), but

$$\mathbb{P}(X<Y) = \mathbb{P}(2X' > X) \overset{\perp\!\!\!\perp}{=} \int_0^\infty \mathbb{P}(X' > x/2)\, e^{-x}\,\mathrm{d}x = \int_0^\infty e^{-3x/2}\,\mathrm{d}x = \frac{2}{3},$$

   whereas $\mathbb{P}(X>Y) = 1 - \mathbb{P}(X<Y) = 1/3$.

   g) True, since in that case $\mathbb{E}[|Y|] = \int_{\mathbb{R}} |t|\, P_Y(\mathrm{d}t) = \int_{\mathbb{R}} |t|\, P_X(\mathrm{d}t) = \mathbb{E}[|X|] < \infty$, and the same then holds without absolute values. Recall more generally that $X \sim Y \implies \mathbb{E}[g(X)] = \mathbb{E}[g(Y)]$ for g measurable $\geqslant 0$ or continuous bounded.

   h) False: take *e.g.* X, Y as in 1.b) and $Z := -X$.

2. a) False: take *e.g.* $X := N$ and $Y := N^2$, where N is a centered Gaussian r.v.

   b) False: take *e.g.* two independent Bernoulli r.v.

   c) False when $X = Y = c$ for a fixed $c \in \mathbb{R}$. (Conversely, if X, Y are independent with $\mathbb{P}(X = Y) = 1$, then $X = Y = c$ almost surely, for some $c \in \mathbb{R}$.)

   d) False: take *e.g.* X, Y i.i.d. with $\mathbb{P}(X=-1) = \mathbb{P}(X=1) = 1/2$, and $Z := XY$.

   e) True: let $F, G \colon \mathbb{R} \to [0,\infty)$ measurable. As $\tilde{F} = F \circ f$ and $\tilde{G} := G \circ g$ are also measurable nonnegative functions, we have

$$\begin{aligned}\mathbb{E}[F(f(X))\,G(g(Y))] &= \mathbb{E}[\tilde{F}(X)\,\tilde{G}(Y)] \\ &= \mathbb{E}[\tilde{F}(X)]\,\mathbb{E}[\tilde{G}(Y)] && \text{(as } X \perp\!\!\!\perp Y) \\ &= \mathbb{E}[F(f(X))]\,\mathbb{E}[G(g(Y))].\end{aligned}$$

   Since the equality holds for any pair (F,G), we conclude that $f(X) \perp\!\!\!\perp g(Y)$. ■

[1] *A fortiori*, $X \sim Y \not\Rightarrow (X,Y) \sim (Y,X)$ in general. But the statements become true if X and Y are further independent, as then $P_{(X,Y)}(\mathrm{d}u,\mathrm{d}v) = P_X(\mathrm{d}u)\,P_Y(\mathrm{d}v) = P_Y(\mathrm{d}u)\,P_X(\mathrm{d}v) = P_{(Y,X)}(\mathrm{d}u,\mathrm{d}v)$.

2.3 Computing distributions

Exercise 2.3.1. Let $p \in (0,1)$ and $X_1, X_2, \dots, Y_1, Y_2, \dots$ be i.i.d. Bernoulli(p) r.v. We define

$$N := \min\{n \in \mathbb{N} \colon X_n \neq Y_n\}$$

and set "$Z := X_N$", *i.e*

$$Z = \sum_{n=1}^{\infty} \mathbb{1}_{\{N=n\}} X_n.$$

1. Check that $N \geqslant 1$ has the geometric distribution with parameter $2p(1-p)$.

2. Show that Z has the Bernoulli($1/2$) distribution.

3. Deduce a way to simulate a fair coin toss using a potentially unfair coin.

Solution of Exercise 2.3.1.

1. Let $n \geqslant 1$. We have $N = n$ if and only if

   $$X_1 = Y_1,\ X_2 = Y_2,\ \dots,\ X_{n-1} = Y_{n-1} \ \text{ and } \ X_n \neq Y_n.$$

   Moreover, $\mathbb{P}(X_1 = Y_1) = \mathbb{P}(X_1 = Y_1 = 1) + \mathbb{P}(X_1 = Y_1 = 0) = p^2 + (1-p)^2$. Since $X_1, X_2, \dots, Y_1, Y_2, \dots$ are i.i.d. we deduce that

   $$\mathbb{P}(N = n) = q^{n-1} \cdot 2p(1-p),$$

   where $q := p^2 + (1-p)^2 = 1 - 2p(1-p)$. Thus N has the geometric distribution with parameter $2p(1-p)$.

2. It is clear by definition that $Z \in \{0,1\}$, so Z is a Bernoulli r.v. Its parameter p is

   $$\mathbb{E}[Z] = \sum_{n=1}^{\infty} \mathbb{E}\big[\mathbb{1}_{\{N=n\}} X_n\big] = \sum_{n=1}^{\infty} \mathbb{P}(N = n, X_n = 1)$$

   (the first equality resulting from monotone convergence). But $N = n$ and $X_n = 1$ if and only if $X_1 = Y_1$, $X_2 = Y_2, \dots, X_{n-1} = Y_{n-1}$ and $X_n = 1$, $Y_n = 0$. Therefore $\mathbb{P}(N = n, X_n = 1) = q^{n-1} \cdot p(1-p) = \frac{1}{2}\mathbb{P}(N = n)$, and $p = \mathbb{E}[Z] = 1/2$.

3. We toss the coin twice and repeat this step until two different faces (HT or TH) have been obtained. According to what precedes, the result of the last performed toss is a simulation of a fair coin toss. The average number of tosses (that is, the *complexity* of the simulation) is $2\mathbb{E}[N] = 1/(p(1-p))$. ∎

Exercise 2.3.2. Let X be uniformly distributed on $[-1,1]$. Find the density of $Y := X^k$ for positive integers k.

Solution of Exercise 2.3.2. The c.d.f. of X is $\mathbb{P}(X \leqslant t) = (t+1)/2$, $t \in [-1,1]$. If k is odd, then $\mathbb{P}(Y \leqslant t) = (\sqrt[k]{t} + 1)/2$, $t \in [-1,1]$, else $\mathbb{P}(Y \leqslant t) = \mathbb{P}(X \in [-\sqrt[k]{t}, \sqrt[k]{t}]) = \sqrt[k]{t}$, $t \in [0,1]$. (These distribution functions are obviously both continuous and piecewise of $\mathscr{C}^1$ class.) Hence $f_Y(y) = \mathbb{1}_{\{0<|y|<1\}} \sqrt[k]{y}^{1-k}/2k$ for k odd, and $f_Y(y) = \mathbb{1}_{\{0<y<1\}} \sqrt[k]{y}^{1-k}/k$ for k even. ∎

Exercise 2.3.3. Let X have distribution function F. What is the distribution function of $Y := |X|$? When X admits a continuous density f_X, show that Y also admits a density f_Y, and express f_Y in terms of f_X.

Solution of Exercise 2.3.3. We have $\mathbb{P}(Y \leqslant t) = F(t) - F(-t-)$, $t \geqslant 0$. If f_X exists and is continuous, then F is $\mathscr{C}^1$ with $F' = f_X$. Therefore f_Y exists, and $f_Y(y) = \mathbb{1}_{[0,\infty)}(y)(f_X(y) + f_X(-y))$. ∎

Exercise 2.3.4. Let Θ be uniformly distributed on the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

1. Find a continuous density function for $C := \tan\Theta$.
2. Find a density function for $A := (\sin\Theta)^2$ which is continuous on $(0,1)$.
3. Identify the law of $C^2 - AC^2 + A$.

Solution of Exercise 2.3.4.

1. We have $\mathbb{P}(C \leqslant t) = \mathbb{P}(\Theta \leqslant \arctan t) = (\arctan t + \pi/2)/\pi$, $t \in \mathbb{R}$, so a continuous density is $f_C(x) = 1/\big(\pi(1+x^2)\big)$ (Cauchy(0, 1) law).
2. Clearly $A \in [0,1]$ and $\mathbb{P}(A \leqslant t) = \mathbb{P}(|\Theta| \leqslant \arcsin\sqrt{t}) = (2\arcsin\sqrt{t})/\pi$, $t \in [0,1]$, thus a density $f_A(y) = \mathbb{1}_{(0,1)}(y)/\big(\pi\sqrt{y(1-y)}\big)$ continuous on $(0,1)$ (Arcsine(0, 1) law).
3. By basic trigonomometry $C^2 - AC^2 + A = 2A$, and this has density $\mathbb{1}_{(0,2)}(y) f_A(y/2)/2$. ∎

Exercise 2.3.5. Let X be Cauchy with parameters $\alpha, 1$. Let $Y := a/X$ with $a \neq 0$. Show that Y is also a Cauchy r.v. and find its parameters.

Solution of Exercise 2.3.5. Since $y \leftarrow a/x$ is a $\mathscr{C}^1$-diffeomorphism of $U := \mathbb{R} \setminus \{0\}$ with $\mathrm{Leb}(\mathbb{R} \setminus U) = 0$,

$$\mathbb{E}[f(Y)] = \int_U \frac{f(a/x)\,\mathrm{d}x}{\pi(1+(x-\alpha)^2)} = \int_U \frac{f(y)\,|a|\,\mathrm{d}y}{\pi(y^2+(a-y\alpha)^2)} = \int_{\mathbb{R}} \frac{f(y)\,\frac{|a|}{1+\alpha^2}\,\mathrm{d}y}{\pi\left[\left(y - \frac{a\alpha}{1+\alpha^2}\right)^2 + \left(\frac{|a|}{1+\alpha^2}\right)^2\right]}$$

for every f bounded. Thus Y is Cauchy with parameters $a\alpha/(1+\alpha^2)$, $|a|/(1+\alpha^2)$. ∎

Exercise 2.3.6. Let X, Y be two independent $\mathcal{N}(0,1)$ random variables. Find a density function for $Z := X^2/(X^2+Y^2)$ which is continuous on $(0,1)$.

Solution of Exercise 2.3.6. Let $g\colon \mathbb{R} \to \mathbb{R}$ be continuous and bounded. We apply two changes of variables — one using the $\mathscr{C}^1$-diffeomorphism $\varphi\colon (r,\theta) \mapsto (x,y) = (r\cos\theta, r\sin\theta)$ from $(0,\infty)\times(-\pi,\pi)$ onto

$P = \mathbb{R}^2 \setminus (-\infty, 0] \times \{0\}$, with Jacobian determinant $|J_\varphi(r,\theta)| = r$, and one using the $\mathscr{C}^1$-map $\theta \mapsto \cos^2\theta$:

$$
\begin{aligned}
\mathbb{E}[g(Z)] &= \frac{1}{2\pi} \iint_P g\left(\frac{x^2}{x^2+y^2}\right) e^{-\frac{x^2+y^2}{2}} \mathrm{d}x\mathrm{d}y && (\mathbb{P}((X,Y) \notin P) = 0) \\
&= \frac{1}{2\pi} \iint_{(0,\infty)\times(-\pi,\pi)} g\big(\cos^2\theta\big)\, r\, e^{-\frac{r^2}{2}} \mathrm{d}r\mathrm{d}\theta \\
&= \frac{1}{2\pi} \int_{(-\pi,\pi)} g\big(\cos^2\theta\big) \mathrm{d}\theta && \text{(Fubini–Lebesgue)} \\
&= \frac{2}{\pi} \int_{(0,\frac{\pi}{2})} g\big(\cos^2\theta\big) \mathrm{d}\theta && \text{(symmetry)} \\
&= \frac{1}{\pi} \int_0^1 g(z) \frac{1}{\sqrt{z(1-z)}} \mathrm{d}z.
\end{aligned}
$$

We conclude that Z has density $z \mapsto \big(\pi\sqrt{z(1-z)}\big)^{-1}$ on $(0,1)$ (Arcsine$(0,1)$ law). ■

Exercise 2.3.7. Let X be positive with a density f. Find a density for $Y := 1/(X+1)$.

Solution of Exercise 2.3.7. Since $\mathbb{P}(Y \leqslant t) = \mathbb{P}(X \geqslant -1+1/t) = \int_{-1+1/t}^{\infty} f(x)\,\mathrm{d}x = \int_0^t f(-1+1/y)/y^2\,\mathrm{d}y$ $[x \leftarrow -1+1/y]$ for $t \in (0,1)$, the map $y \mapsto \mathbb{1}_{(0,1)}(y) f(-1+1/y)/y^2$ is a density for Y. ■

Exercise 2.3.8. Let $X, X_1, X_2, \ldots$ be i.i.d. real r.v. with cumulative distribution function F and having a density function f. We set

$$N := \inf\{k \in \mathbb{N} \colon X_k > X\}.$$

1. Let $k \in \mathbb{N}$ and $t \in \mathbb{R}$. Show that

$$\mathbb{P}(N = k, X \leqslant t) = \int_{-\infty}^t F(x)^{k-1}\big(1 - F(x)\big) f(x)\,\mathrm{d}x.$$

2. Conclude that

$$\mathbb{P}(N = k) = \frac{1}{k} - \frac{1}{k+1}, \qquad k \in \mathbb{N}.$$

Solution of Exercise 2.3.8.

1. This follows easily from the fact that $X, X_1, \ldots, X_k$ are i.i.d.:

$$
\begin{aligned}
\mathbb{P}(N = k, X \leqslant t) &= \mathbb{P}(X \leqslant t, X_1 \leqslant X, \ldots, X_{k-1} \leqslant X, X_k > X) \\
&= \int \mathrm{d}F(x) \mathbb{1}_{\{x \leqslant t\}} \int\!\cdots\!\int \mathrm{d}F(x_1) \cdots \mathrm{d}F(x_k)\, \mathbb{1}_{\{x_1 \leqslant x\}} \cdots \mathbb{1}_{\{x_{k-1} \leqslant x\}} \mathbb{1}_{\{x_k > x\}} \\
&= \int_{-\infty}^t f(x)\mathrm{d}x \big(1 - F(x)\big) F(x)^{k-1}.
\end{aligned}
$$

2. Observing that $(F^k)' = kF^{k-1}f$ for all $k \in \mathbb{N}$, we deduce that

$$\mathbb{P}(N = k, X \leqslant t) = \frac{1}{k} F^k(t) - \frac{1}{k+1} F^{k+1}(t).$$

As $t \to \infty$, the left-hand-side tends to $\mathbb{P}(N = k)$ while the right-hand side tends to $\frac{1}{k} - \frac{1}{k+1}$ (because $F(t) \to 1$). Hence the result. ■

Exercise 2.3.9. Let X be a real random variable such that:

$$F_X(t) := \mathbb{P}(X \leqslant t) = \begin{cases} 0, & \text{if } t < -3, \\ 1/3, & \text{if } -3 \leqslant t < -2, \\ 7/12, & \text{if } -2 \leqslant t < 0, \\ 3/4, & \text{if } 0 \leqslant t < 4, \\ 1, & \text{if } 4 \leqslant t. \end{cases}$$

Compute $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.

Solution of Exercise 2.3.9. Clearly $X \in \{-3, -2, 0, 4\}$. Using that $\mathbb{P}(X = t) = F_X(t) - F_X(t-)$ we find $\mathbb{P}(X = -3) = 1/3$, $\mathbb{P}(X = -2) = 1/4$, $\mathbb{P}(X = 0) = 1/6$, and $\mathbb{P}(X = 4) = 1/4$. Therefore

$$\begin{aligned} \mathbb{E}[X] &= (-3) \cdot 1/3 + (-2) \cdot 1/4 + 0 \cdot 1/6 + 4 \cdot 1/4 = -1/2, \\ \mathbb{E}[X^2] &= 9 \cdot 1/3 + 4 \cdot 1/4 + 16 \cdot 1/4 = 8, \end{aligned}$$

and

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 8 - 1/4 = 31/4.$$

■

Exercise 2.3.10. Let $X, X_1, X_2, \dots$ be i.i.d. real r.v. with distribution function F and having a density function f. We set

$$N := \inf\{k \in \mathbb{N} \colon X_k > X\}.$$

1. Let $k \in \mathbb{N}$ and $t \in \mathbb{R}$. Show that

$$\mathbb{P}(N = k, X \leqslant t) = \int_{-\infty}^{t} F(x)^{k-1} (1 - F(x)) f(x) \, \mathrm{d}x.$$

2. Conclude that

$$\mathbb{P}(N = k) = \frac{1}{k} - \frac{1}{k+1}, \qquad k \in \mathbb{N}.$$

Solution of Exercise 2.3.10.

1. By definition of N and independence of $(X, X_1, \dots, X_k)$,

$$\begin{aligned} \mathbb{P}(N = k, X \leqslant t) &= \mathbb{P}(X_1, \dots, X_{k-1} \leqslant X, X_k > X, X \leqslant t) \\ &= \int_{-\infty}^{t} \mathbb{P}(X_1, \dots, X_{k-1} \leqslant x, X_k > x) f(x) \, \mathrm{d}x \\ &= \int_{-\infty}^{t} F(x)^{k-1} (1 - F(x)) f(x) \, \mathrm{d}x. \end{aligned}$$

2. Note that for each $k \in \mathbb{N}$, the function $x \mapsto F(x)^{k-1} f(x)$ is integrable on $\mathbb{R}$ with

$$\int_{-\infty}^{\infty} F(x)^{k-1} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} F(x)^{k-1} F'(x)\,\mathrm{d}x = \left[\frac{F(x)^k}{k}\right]_{x\to-\infty}^{x\to\infty} = \frac{1}{k},$$

where we used that $F' = f$ a.e. We then deduce from monotonicity and Question 1 that

$$\begin{aligned}\mathbb{P}(N=k) &= \lim_{t\to\infty}\uparrow \mathbb{P}(N=k, X \leqslant t)\\ &= \int_{-\infty}^{\infty} F(x)^{k-1} f(x)\,\mathrm{d}x - \int_{-\infty}^{\infty} F(x)^{k} f(x)\,\mathrm{d}x\\ &= \frac{1}{k} - \frac{1}{k+1}.\end{aligned}$$

■

Exercise 2.3.11. Let U, V be two independent standard uniform r.v. We set

$$X := U^2 + V^2, \quad \text{and} \quad Y := U^2/X.$$

Compute

$$\mathbb{P}(Y \leqslant t \mid X \leqslant 1) := \frac{\mathbb{P}(Y \leqslant t, X \leqslant 1)}{\mathbb{P}(X \leqslant 1)}, \qquad t \in \mathbb{R}.$$

Solution of Exercise 2.3.11. By independence, (U,V) has density function $\mathbb{1}_{(0,1)^2}$, so

$$\mathbb{P}(X \leqslant 1) = \iint_{(0,1)^2} \mathbb{1}_{\{u^2+v^2\leqslant 1\}}\,\mathrm{d}u\mathrm{d}v = \pi/4. \qquad \left(\tfrac{1}{4}\cdot\text{(area of the unit circle)}\right)$$

Similarly, for $0 \leqslant t \leqslant 1$,

$$\begin{aligned}\mathbb{P}(Y \leqslant t, X \leqslant 1) &= \iint_{(0,\infty)^2} \mathbb{1}_{\{u^2\leqslant t(u^2+v^2)\}} \mathbb{1}_{\{u^2+v^2\leqslant 1\}}\,\mathrm{d}u\mathrm{d}v\\ &= \iint_{(0,\infty)\times(0,\pi/2)} \mathbb{1}_{\{\sin^2\theta\leqslant t\}} \mathbb{1}_{\{r^2\leqslant 1\}}\, r\,\mathrm{d}r\mathrm{d}\theta && \text{(polar coordinates)}\\ &= \left(\int_0^1 r\,\mathrm{d}r\right)\left(\int_0^{\pi/2} \mathbb{1}_{\{|\sin\theta|\leqslant\sqrt{t}\}}\,\mathrm{d}\theta\right) && \text{(Fubini)}\\ &= \frac{1}{2}\arcsin\sqrt{t}.\end{aligned}$$

We conclude that $\mathbb{P}(Y \leqslant t \mid X \leqslant 1) = (2/\pi)\arcsin\sqrt{t}$ for $0 \leqslant t \leqslant 1$. (We say that *conditionally on $X \leqslant 1$, the random variable Y has the Arcsine distribution.*) ■

Exercise 2.3.12. Let X be a real r.v. in $\mathrm{L}^1(\Omega, \mathscr{A}, \mathbb{P})$.

1. Let a, b be two real numbers. Show that

$$\mathbb{E}[|X-b|] - \mathbb{E}[|X-a|] = \int_a^b \left[\mathbb{P}(X \leqslant t) - \mathbb{P}(X \geqslant t)\right]\mathrm{d}t.$$

Hint. Observe that $|b-x| - |x-a| = \int_a^b \left(\mathbb{1}_{\{x\leqslant t\}} - \mathbb{1}_{\{x\geqslant t\}}\right)\mathrm{d}t$. Use Fubini's theorem.

2. We call $m \in \mathbb{R}$ a *median* of a real r.v. Y if $\mathbb{P}(Y \leqslant m) \geqslant 1/2$ and $\mathbb{P}(Y \geqslant m) \geqslant 1/2$.

   a) Show that every real random variable admits a median. Is there uniqueness?

   b) Let m be a median of X. Deduce from Question 1 that

$$\mathbb{E}[|X-m|] = \inf_{c\in\mathbb{R}} \mathbb{E}[|X-c|].$$

   Conclude that $|\mathbb{E}[X]-m| \leqslant \sigma$ where $\sigma^2 := \mathrm{Var}(X)$.

Solution of Exercise 2.3.12.

1. We may suppose $a<b$. The observation is immediate. From the right-hand side,

$$\int_a^b \big[\mathbb{P}(X\leqslant t)-\mathbb{P}(X\geqslant t)\big]\,\mathrm{d}t = \int_a^b \Big(\int_\Omega \underbrace{\big[\mathbb{1}_{\{X(\omega)\leqslant t\}}-\mathbb{1}_{\{X(\omega)\geqslant t\}}\big]}_{=:f(\omega,t)}\mathbb{P}(\mathrm{d}\omega)\Big)\,\mathrm{d}t$$

   where $f \in \mathrm{L}^1(\Omega\times(a,b), \mathscr{A}\otimes\mathscr{B}((a,b)), \mathbb{P}\otimes\mathrm{d}t)$ — indeed,

$$\int_a^b \Big(\int_\Omega |f(\omega,t)|\,\mathbb{P}(\mathrm{d}\omega)\Big)\,\mathrm{d}t \leqslant \int_a^b \big[\mathbb{P}(X\leqslant t)+\mathbb{P}(X\geqslant t)\big]\,\mathrm{d}t \leqslant 2(b-a) < \infty.$$

   Fubini's theorem then entails

$$\begin{aligned}\int_a^b \big[\mathbb{P}(X\leqslant t)-\mathbb{P}(X\geqslant t)\big]\,\mathrm{d}t &= \int_\Omega \Big(\int_a^b f(\omega,t)\,\mathrm{d}t\Big)\mathbb{P}(\mathrm{d}\omega)\\ &= \mathbb{E}[|b-X|-|X-a|]\\ &= \mathbb{E}[|X-b|]-\mathbb{E}[|X-a|],\end{aligned}$$

   where we used the stated observation for the second equality.

2. a) Recall that $F_X(t) := \mathbb{P}(X\leqslant t) \to 0$ as $t\to-\infty$ and $F_X(t)\to 1$ as $t\to\infty$, thus $m := F_X^{-1}(1/2) := \inf\{t\in\mathbb{R} : F_X(t) > 1/2\}$ is a well-defined real number. For $m_n \downarrow m$ such that $F_X(m_n) > 1/2$ for every n, we have, by right-continuity of F_X,

$$\mathbb{P}(X\leqslant m) = F_X(m) = \lim_{n\to\infty}\downarrow F_X(m_n) \geqslant 1/2$$

   Now $F_X(m-1/n) \leqslant 1/2$ by definition of m, and by monotonicity

$$\mathbb{P}(X\geqslant m) = \lim_{n\to\infty}\downarrow \mathbb{P}(X > m-1/n) = 1 - \lim_{n\to\infty} F_X(m-1/n) \geqslant 1/2,$$

   so m is a median of X. There is no uniqueness in general. For instance if X is a Bernoulli$(1/2)$ random variable, that is $\mathbb{P}(X=0)=\mathbb{P}(X=1)=1/2$, then any real $m\in(0,1)$ is a median of X.

b) For $c > m$ we have, using the result of Question 1,

$$\begin{aligned}\mathbb{E}[|X-c|]-\mathbb{E}[|X-m|] &= \int_m^c \big[\mathbb{P}(X\leqslant t)-\mathbb{P}(X\geqslant t)\big]\,\mathrm{d}t \\ &\geqslant \int_m^c \big[\mathbb{P}(X\leqslant m)-\mathbb{P}(X> m)\big]\,\mathrm{d}t \\ &= \int_m^c \big[2\mathbb{P}(X\leqslant m)-1\big]\,\mathrm{d}t \\ &\geqslant 0.\end{aligned}$$

Likewise, for $c < m$,

$$\begin{aligned}\mathbb{E}[|X-c|]-\mathbb{E}[|X-m|] &= \int_c^m \big[\mathbb{P}(X\geqslant t)-\mathbb{P}(X\leqslant t)\big]\,\mathrm{d}t \\ &\geqslant \int_c^m \big[\mathbb{P}(X\geqslant m)-\mathbb{P}(X< m)\big]\,\mathrm{d}t \\ &= \int_c^m \big[2\mathbb{P}(X\geqslant m)-1\big]\,\mathrm{d}t \\ &\geqslant 0.\end{aligned}$$

Hence $\mathbb{E}[|X-m|]=\inf_{c\in\mathbb{R}}\mathbb{E}[|X-c|]$. In particular, for $c:=\mathbb{E}[X]$,

$$|\mathbb{E}[X]-m|\leqslant \mathbb{E}[|X-m|]\leqslant \mathbb{E}[|X-c|]\leqslant \sqrt{\mathbb{E}[|X-c|^2]}=\sigma,$$

where the last inequality stems from Cauchy–Schwarz (or Hölder). ∎

Exercise 2.3.13. For any distribution function F, we define

$$F^{-1}(u):=\inf\{t\in\mathbb{R}:\ F(t)>u\},\qquad u\in(0,1),$$

the *right-continuous inverse* of F.

1. Compute F^{-1} when F is the standard exponential distribution.
2. Show that for every $t\in\mathbb{R}$ and $u\in(0,1)$, $\;u<F(t)\implies F^{-1}(u)\leqslant t\implies u\leqslant F(t)$.
3. Let U be uniformly distributed on $(0,1)$.
   a) Show that $\lfloor \log_{1/2} U\rfloor$ has the Geometric$(1/2)$ distribution (with $\lfloor\cdot\rfloor$ = integer part).
   b) More generally, show that $F^{-1}(U)$ has law F.
4. Show that F^{-1} is non-decreasing.
5. Show that F^{-1} is right-continuous.

Consequently, the set $(0,1)\setminus\mathscr{C}(F^{-1})$ of discontinuity points of F^{-1} is at most countable.

6. Let $F, F_1, F_2, \dots$ be distribution functions such that $\forall t \in \mathscr{C}(F)$, $F_n(t) \to F(t)$. Show that $\forall u \in \mathscr{C}(F^{-1})$, $F_n^{-1}(u) \to F^{-1}(u)$.

7. Consider a convergence in distribution $X_n \Longrightarrow X$ of real r.v., and let U be a standard uniform r.v. Show that there exist Y and Y_n, $n \in \mathbb{N}$, measurable w.r.t. U such that $Y \sim X$, $Y_n \sim X_n$, and $Y_n \to Y$ a.s.

Solution of Exercise 2.3.13.

1. Here $u < F(t)$ happens for $t > 0$ such that $1 - e^{-t} > u$, *i.e*, $t \in (-\log(1-u), \infty)$. Hence $F^{-1}(u) = -\log(1-u)$ for every $u \in (0,1)$.

2. The first implication is obvious. For the second one, if $F^{-1}(u) \leqslant t$, then for every $n \in \mathbb{N}$ we can find $t_n < t + \frac{1}{n}$ such that $u < F(t_n) \leqslant F(t + \frac{1}{n})$ (F is non-decreasing). Taking $n \to \infty$ gives $u \leqslant F(t)$, by right-continuity of F.

Remark. If F is one-to-one from $F^{-1}\langle(0,1)\rangle$ into an interval, then F^{-1} coincides with the inverse function of F.

3. a) If $k \in \mathbb{N}_0$, then $\lfloor \log_{1/2} y \rfloor = k \iff k \leqslant \log_{1/2} y < k+1 \iff 2^{-k-1} < y \leqslant 2^{-k}$ for every $y \in (0,1)$, so $\mathbb{P}(\lfloor \log_{1/2} Y \rfloor = k) = 2^{-k} - 2^{-k-1} = 2^{-k-1}$, $k \in \mathbb{N}_0$, and this is indeed the Geometric$(1/2)$ distribution.

   b) By Question 2, and because U has the standard uniform distribution,

$$F(t) = \mathbb{P}\big(U < F(t)\big) \leqslant \mathbb{P}\big(F^{-1}(U) \leqslant t\big) \leqslant \mathbb{P}\big(U \leqslant F(t)\big) = F(t), \qquad t \in \mathbb{R}.$$

   As F is thus the distribution function of both X and $F^{-1}(U)$, these two r.v. are equally distributed. For instance $-\log(1-U)$ (or simply $-\log U$, since $1-U$ is also uniformly distributed on $(0,1)$) has the standard exponential distribution.

4. If $u' \geqslant u$, then $\{t \in \mathbb{R} \colon F(t) > u'\} \subseteq \{t \in \mathbb{R} \colon F(t) > u\}$, and so $F^{-1}(u') \geqslant F^{-1}(u)$.

5. Let $u_k \downarrow u$ in $(0,1)$. There exist $t_r \downarrow F^{-1}(u)$ such that $F(t_r) > u$ for all r. For r fixed, we have $F(t_r) > u_k$ for k large enough, so $F^{-1}(u_k) \leqslant t_r$, and thus $\limsup F^{-1}(u_k) \leqslant t_r$. By letting $r \to \infty$, we get $\limsup F^{-1}(u_k) \leqslant F^{-1}(u)$. According to Question 1 we also have $F^{-1}(u_k) \geqslant F^{-1}(u)$ for all k, and therefore $\liminf F^{-1}(u_k) \geqslant F^{-1}(u)$. Finally

$$\lim_{k\to\infty} F^{-1}(u_k) = F^{-1}(u)$$

   for every sequence $u_k \downarrow u$, which proves that F^{-1} is right-continuous.

6. Let $u \in \mathscr{C}(F^{-1})$. There exist $t_r \downarrow F^{-1}(u)$ in $\mathscr{C}(F)$ such that $F(t_r) > u$ for every r. For r fixed, $F_n(t_r) \to F(t_r)$ so $F_n(t_r) > u$ and then $F_n^{-1}(u) \leqslant t_r$ for n large enough, thus $\limsup F_n^{-1}(u) \leqslant t_r$. By letting $r \to \infty$, we get $\limsup F_n^{-1}(u) \leqslant F^{-1}(u)$. Now for each $u' < u$, there exist $t_r \uparrow F^{-1}(u')$ in $\mathscr{C}(F)$, so $F(t_r) \leqslant u' < u$. Hence $F_n(t_r) \leqslant u$ and then $F_n^{-1}(u) \geqslant t_r$ for n large enough, so $\liminf F_n^{-1}(u) \geqslant t_r$. By letting $r \to \infty$, we get $\liminf F_n^{-1}(u) \geqslant F^{-1}(u')$. Taking now $u' \to u$ gives $\liminf F_n^{-1}(u) \geqslant F^{-1}(u)$ because $u \in \mathscr{C}(F^{-1})$.

7. Let F and F_n, $n \in \mathbb{N}$, denote the distribution functions of X and X_n respectively. We know that $Y := F^{-1}(U) \sim X$ and $Y_n := F_n^{-1}(U) \sim X_n$. Since $X_n \Longrightarrow X$, we have $F_n(t) \to F(t)$ for all $t \in \mathscr{C}(F)$. Now $(0,1) \setminus \mathscr{C}(F^{-1})$ is at most countable so $U \in \mathscr{C}(F^{-1})$ a.s. It then follows from Question 3 that $Y_n \to Y$ almost surely.

   Remark. This is a version of Skorokhod's theorem in $\mathbb{R}$. ∎

Exercise 2.3.14. Recall that a r.v. X has a *continuous distribution* if $x \mapsto \mathbb{P}(X \leqslant x)$ is continuous.

1. Show that X has a continuous distribution if and only if $\mathbb{P}(X = x) = 0$ for all $x \in \mathbb{R}$.

2. Show that if X has a continuous distribution and Y is any random variable independent of X, then $X + Y$ has a continuous distribution.

3. Let $f : \mathbb{R} \to [0, \infty)$ measurable. We suppose that (the distribution of) X has density f, that is

$$\mathbb{P}(X \in A) = \int_A f(x)\,\mathrm{d}x$$

   for every Borel set A. Show that:

   a) f is integrable on $\mathbb{R}$.
   b) If $\mathbb{P}(X \in A) > 0$, then A has positive Lebesgue measure.
   c) X has a continuous distribution.
   d) If X has another density g, then $f = g$ almost everywhere.

Solution of Exercise 2.3.14.

1. By monotonicity of the measure:

$$\mathbb{P}(X = x) = \mathbb{P}\left(\bigcap_{n \geqslant 1} \left(x - \tfrac{1}{n}, x\right]\right) = \lim_{n \to \infty} \left(\mathbb{P}(X \leqslant x) - \mathbb{P}(X \leqslant x - \tfrac{1}{n})\right),$$

   which equals 0 for all x if and only if $x \mapsto \mathbb{P}(X \leqslant x)$ is continuous.

2. By independence,

$$\mathbb{P}(X + Y = x) = \int \underbrace{\mathbb{P}(X = x - y)}_{=0} \mathbb{P}(Y \in \mathrm{d}y) = 0$$

   for all $x \in \mathbb{R}$. We conclude with Question 1.

3. a) For $A = \mathbb{R}$ we have

$$1 = \int_{\mathbb{R}} f(x)\,\mathrm{d}x,$$

   so f is integrable on $\mathbb{R}$.

   b) If A has zero Lebesgue measure, then

$$\mathbb{P}(X \in A) = \int_A f(x)\,\mathrm{d}x = 0.$$

c) In particular $\mathbb{P}(X=x)=0$ (since $\{x\}$ has zero Lebesgue measure) for all $x\in\mathbb{R}$.

d) Let g be another density. Since $A:=\{f<g\}$ is a Borel set we have

$$\mathbb{P}(X\in A)=\int_A f(x)\,\mathrm{d}x\leqslant\int_A g(x)\,\mathrm{d}x=\mathbb{P}(X\in A),$$

so A must have zero Lebesgue measure (otherwise the inequality would be strict). Similarly, $A':=\{f>g\}$ has also zero Lebesgue measure. Hence $f=g$ a.e. ■

Exercise 2.3.15. Let L be the uniform distribution on $E:=(0,1)$, and $\mathbb{P}$ be the Arcsine distribution:

$$\mathbb{P}((0,t])=:F(t)=\frac{1}{2}+\frac{\arcsin(2t-1)}{\pi},\quad t\in E.$$

Define $X(s,t):=t\mathbb{1}_{\{s\leqslant t\}}+(1-t)\mathbb{1}_{\{s>t\}}$ for $s,t\in E$ and write $\mathbb{Q}$ for the law of X under $L\otimes\mathbb{P}$.

1. Show that for every bounded, measurable function $f\colon E\to\mathbb{R}$,

$$\int_E f(t)\,\mathbb{Q}(\mathrm{d}t)=\int_E 2tf(t)\,\mathbb{P}(\mathrm{d}t).$$

Deduce that $\mathbb{Q}$ admits w.r.t. $\mathbb{P}$ the Radon-Nikodym derivative

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}=2t,\qquad t\in E.$$

2. Conclude that X has density

$$t\longmapsto\frac{2t}{\pi\sqrt{t(1-t)}},\qquad t\in E.$$

Solution of Exercise 2.3.15.

1. Observe first that

$$\mathbb{P}((1-t,1])=1-\mathbb{P}((0,1-t])=\frac{1}{2}-\frac{\arcsin\big(2(1-t)-1\big)}{\pi}=\mathbb{P}((0,t]),\quad t\in E.$$

Since f is bounded, we can apply Fubini's theorem and get

$$\begin{aligned}\int_E f(t)\,\mathbb{Q}(\mathrm{d}t)&=\int_{E\times E} f\big(X(s,t)\big)\,L\otimes\mathbb{P}(\mathrm{d}s,\mathrm{d}t)\\&=\int_E f(t)L((0,t])\,\mathbb{P}(\mathrm{d}t)+\int_E f(1-t)L((1-t,1])\,\mathbb{P}(\mathrm{d}t)\\&=\int_E 2tf(t)\,\mathbb{P}(\mathrm{d}t)\end{aligned}$$

(the last equality following from the change of variable $t\leftarrow 1-t$ in the second integral). As this holds for every bounded measurable function f, we deduce that $\mathbb{Q}(\mathrm{d}t)=2t\mathbb{P}(\mathrm{d}t)$.

2. Since F is of class $\mathscr{C}^1$ on E, $\mathbb{P}$ has a density. Then X has density

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}t}=\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\cdot\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}t}=2t\cdot F'(t)=\frac{2t}{\pi\sqrt{t(1-t)}},\quad t\in E.$$

■

2.4 Convergence of random variables, limit theorems

Exercise 2.4.1. Let L^0 denote the space of real r.v. defined on $(\Omega, \mathbb{P})$.

1. Show that
$$d(X,Y) := \mathbb{E}[1 \wedge |X-Y|]$$
is a distance on L^0 such that
$$X_n \xrightarrow[n\to\infty]{\mathbb{P}} X \iff d(X_n, X) \xrightarrow[n\to\infty]{} 0.$$

2. Let $(X_{n,k} \colon n, k \geqslant 1)$ be elements in L^0, and $K \colon \Omega \to \mathbb{N}$ be an *independent* r.v. We suppose that for each $k \in \mathbb{N}$,
$$X_{n,k} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$
Show that
$$\sum_{k=1}^{K} X_{n,k} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Solution of Exercise 2.4.1.

1. It is clear that $d \colon E \times E \to [0, \infty)$ is symmetric and satisfies the triangle inequality. Further, if $d(X,Y) = 0$ then $1 \wedge |X-Y| = 0$ a.s., so $X = Y$ a.s. Now for $\varepsilon > 0$,
$$\mathbb{P}(|X_n - X| > \varepsilon) \leqslant \mathbb{P}(1 \wedge |X_n - X| > 1 \wedge \varepsilon) \leqslant (1 \wedge \varepsilon)^{-1} d(X_n, X)$$
(by Markov's inequality), and
$$d(X_n, X) \leqslant (1 \wedge \varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon),$$
which readily implies that $d(X_n, X) \to 0$ if and only if $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$ for all $\varepsilon > 0$. Hence the assertion.

2. Let μ denote the law of K. By independence
$$d\left(\sum_{k=1}^{K} X_{n,k}, 0\right) = \int d\left(\sum_{i=1}^{k} X_{n,i}, 0\right) \mu(\mathrm{d}k),$$
where the integrand is bounded by 1, and tends to 0 for each k (because any partial sum of $X_{n,k}$, $k \in \mathbb{N}$, converges to 0 in probability). We conclude by dominated convergence. ∎

Exercise 2.4.2. We have seen in Exercise 2.4.1 that the convergence in probability in the space $\mathrm{L}^0(\Omega, \mathscr{A}, \mathbb{P})$ of real r.v. is metrized by
$$d(X,Y) := \mathbb{E}[1 \wedge |X-Y|].$$

0. Let $(X_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathrm{L}^0(\Omega,\mathscr{A},\mathbb{P})$:

$$\forall \varepsilon>0,\ \exists k\in\mathbb{N},\ \forall m\geqslant k,\quad d(X_m,X_k)\leqslant\varepsilon.$$

   a) Construct an increasing sequence $(k_n)_{n\geqslant 0}$ of positive integers such that

$$\mathbb{P}\left(\left|X_{k_{n+1}}-X_{k_n}\right|\geqslant\frac{1}{2^n}\right)\leqslant\frac{1}{2^n}.$$

   b) Show that almost surely, there exists $N\geqslant 0$ sufficiently large such that

$$\forall n\geqslant N,\quad \left|X_{k_{n+1}}-X_{k_n}\right|\leqslant\frac{1}{2^n}.$$

   Deduce that the sequence $(X_{k_n})_{n\geqslant 0}$ converges almost surely.

1. Prove that the space $\mathrm{L}^0(\Omega,\mathscr{A},\mathbb{P})$ is complete.
2. Prove that the space $\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})$, $p\geqslant 1$, is complete.

Solution of Exercise 2.4.2.

0. a) Let $n\geqslant 0$. Taking $\varepsilon:=4^{-(n+1)}$, there exists $k'_n\in\mathbb{N}$ such that (by Markov's inequality)

$$\forall m\geqslant k'_n,\quad \mathbb{P}\left(\left|X_m-X_{k'_n}\right|\geqslant\frac{1}{2^{n+1}}\right)\leqslant 2^{n+1}\,d(X_m,X_{k'_n})\leqslant\frac{1}{2^{n+1}}.$$

   Let $k_0:=k'_0$ and, by induction, $k_n:=k'_n+k_{n-1}$ for $n\geqslant 1$. Then $(k_n)_{n\geqslant 0}$ is increasing and for all $m\geqslant k_n$ (in particular, for $m=k_{n+1}$),

$$\begin{aligned}\mathbb{P}\left(\left|X_m-X_{k_n}\right|\geqslant\frac{1}{2^n}\right)&\leqslant\mathbb{P}\left(\left\{\left|X_m-X_{k'_n}\right|\geqslant\frac{1}{2^{n+1}}\right\}\bigcup\left\{\left|X_{k_n}-X_{k'_n}\right|\geqslant\frac{1}{2^{n+1}}\right\}\right)\\&\leqslant\frac{1}{2^{n+1}}+\frac{1}{2^{n+1}}\\&=\frac{1}{2^n}.\end{aligned}$$

   b) Since the series $\sum_n 2^{-n}$ converges, it follows from the first Borel–Cantelli lemma that almost surely, there exists $N\geqslant 0$ such that

$$\forall n\geqslant N,\quad \left|X_{k_{n+1}}-X_{k_n}\right|\leqslant\frac{1}{2^n}.$$

   Thus, almost surely, the series $\sum_n(X_{k_{n+1}}-X_{k_n})$ converges absolutely. Because $\mathbb{R}$ is complete, this implies that the sequence $(X_{k_n})_{n\geqslant 0}$ converges almost surely.

1. By Question 0, if $(X_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathrm{L}^0(\Omega,\mathscr{A},\mathbb{P})$, then one can extract a subsequence $(X_{k_n})_{n\geqslant 0}$ converging almost surely, and *a fortiori* in probability. Thus $(X_n)_{n\in\mathbb{N}}$ is convergent for the metric d. Hence the completeness of $\mathrm{L}^0(\Omega,\mathscr{A},\mathbb{P})$.

2. Let $(X_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})$, $p\geqslant 1$. By Hölder's inequality
$$d(X_m,X_k)=\mathbb{E}[1\wedge|X_m-X_k|]\leqslant\|X_m-X_k\|_{\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})},\qquad m,k\in\mathbb{N},$$
which means that $(X_n)_{n\in\mathbb{N}}$ is also a Cauchy sequence in $\mathrm{L}^0(\Omega,\mathscr{A},\mathbb{P})$. By Question 0, we can extract a subsequence $(X_{k_n})_{n\geqslant 0}$ converging almost surely to some r.v. X. Now, for all $\varepsilon>0$, since
$$\|X_{k_n}-X_k\|^p_{\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})}\leqslant\varepsilon,$$
when n and k are sufficiently large, we see by Fatou's lemma that $\|X-X_k\|^p_{\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})}\leqslant\varepsilon$ for all k large enough. This means that $(X_n)_{n\in\mathbb{N}}$ converges to X in $\mathrm{L}^p(\Omega,\mathscr{A},\mathbb{P})$. ■

Exercise 2.4.3. For each $p\in(0,1)$, let $B_k^{(p)}$, $k\in\mathbb{N}$, be i.i.d. Bernoulli(p) r.v. We set
$$X^{(p)}:=\lim_{n\to\infty}X_n^{(p)},\quad\text{where}\quad X_n^{(p)}:=\sum_{k=1}^{n}B_k^{(p)}2^{-k},$$
and
$$A^{(p)}:=\left\{\sum_{k=1}^{\infty}b_k2^{-k}\;\middle|\;b_k\in\{0,1\},\ \text{and}\ \lim_{k\to\infty}\frac{b_1+\cdots+b_k}{k}=p\right\}\subset(0,1).$$

1. Show that $X^{(p)}\in A^{(p)}$ almost surely.
2. Show that for every $k,n\in\mathbb{N}_0$, $\mathbb{P}(k\leqslant 2^nX^{(p)}<k+1)\leqslant\theta^n$, with $\theta:=\max(p,1-p)$. Deduce that $X^{(p)}$ has a continuous distribution. We denote it $\mu^{(p)}$.
3. In this question we consider $p=1/2$.
   a) Let U be a standard uniform r.v. Compute the characteristic function Φ_U.
   b) Show that for every $t\in\mathbb{R}$,
   $$\Phi_{X_n}(t)=\exp\!\big(it/2-i2^{-(n+1)}t\big)\frac{\sin(t/2)}{2^n\sin\!\big(2^{-(n+1)}t\big)},$$
   and deduce that $\mu^{(1/2)}$ is the standard uniform distribution.
   Hint. Use that $(1+e^{i\theta})\sin(\theta/2)=e^{i\theta/2}\sin\theta$ to obtain a telescopic product.
4. We now consider $p\neq 1/2$.
   a) Show that $\mu^{(p)}(A^{(p)})=1$ and $\mu^{(1/2)}(A^{(p)})=0$.
   b) Deduce that $\mu^{(p)}$ has no density function.

Solution of Exercise 2.4.3.

1. This follows from the law of large numbers, $(B_k^{(p)})_{k\in\mathbb{N}}$ being i.i.d. with $\mathbb{E}[B_1^{(p)}]=p$.

2. The integer part of $2^n X^{(p)}$ is $2^n X_n^{(p)}$, whose probability mass function is clearly bounded by θ^n. Now for each $t \in \mathbb{R}$, let for every $n \in \mathbb{N}_0$, $k_n \in \mathbb{N}_0$ be the unique integer such that $k_n \leqslant 2^n t < k_n + 1$. Then
$$\mathbb{P}(X^{(p)} = t) = \lim_{n\to\infty} \mathbb{P}(k_n \leqslant 2^n X^{(p)} < k_n + 1) \leqslant \lim_{n\to\infty} \theta^n = 0$$
(because $\theta < 1$), hence $X^{(p)}$ has a continuous distribution.

3. a) For every $t \in \mathbb{R}$,
$$\Phi_U(t) \stackrel{\text{def}}{=} \mathbb{E}\left[e^{itU}\right] = \int_0^1 e^{itu}\,\mathrm{d}u = \frac{e^{it}-1}{it} = e^{it/2}\frac{\sin(t/2)}{t/2}.$$

   b) Knowing that $B_1, \dots, B_n$ are i.i.d. Bernoulli(1/2),
$$\Phi_{X_n}(t) = \prod_{k=1}^{n} \Phi_{B_k}\big(2^{-k}t\big) = \prod_{k=1}^{n} \frac{1+\exp\big(i2^{-k}t\big)}{2}.$$
   Using the indication, this equals
$$2^{-n}\prod_{k=1}^{n}\exp\Big(i2^{-(k+1)}t\Big)\frac{\sin\big(2^{-k}t\big)}{\sin\big(2^{-(k+1)}t\big)} = \exp\big(it/2 - i2^{-(n+1)}t\big)\frac{\sin(t/2)}{2^n\sin\big(2^{-(n+1)}t\big)}.$$
   Because $\sin h \sim h$ as $h \to 0$, we find
$$\Phi_{X_n}(t) \xrightarrow[n\to\infty]{} e^{it/2}\frac{\sin(t/2)}{t/2} = \Phi_U(t).$$
   As $X_n \to X$ (a.s., and *a fortiori* in distribution), X has thus the standard uniform distribution. In other words, $\mu^{(1/2)}(\mathrm{d}x) = \mathbb{1}_{[0,1]}(x)\,\mathrm{d}x$.

4. a) The sets $A^{(p)}$, $p \in (0,1)$, are disjoint. Applying Question 1, we get for $p \neq 1/2$,
$$\mu^{(p)}(A^{(p)}) = \mathbb{P}(X^{(p)} \in A^{(p)}) = 1, \text{ and } \mu^{(1/2)}(A^{(p)}) \leqslant \mathbb{P}(X^{(1/2)} \notin A^{(1/2)}) = 0.$$

   b) Suppose that $X^{(p)}$ has a density $f(x)$. Then from 3.b) and 4.a), we obtain
$$1 = \mu^{(p)}(A^{(p)}) = \int_{A^{(p)}} f(x)\,\mathrm{d}x = \int_{A^{(p)}} f(x)\,\mu^{(1/2)}(\mathrm{d}x) = 0,$$
   a contradiction. Thus $\mu^{(p)}$, $p \neq 1/2$, is a continuous law without density. ■

Exercise 2.4.4. Let μ be a probability distribution on $\mathbb{R}$ having a second moment $\sigma^2 < \infty$ such that, if X and Y are independent with law μ, then the law of $(X+Y)/\sqrt{2}$ is also μ. Show that $\mu = \mathcal{N}(0, \sigma^2)$.
Hint. Apply the central limit theorem to packs of 2^n variables.

Solution of Exercise 2.4.4. Let $X, Y, X_1, X_2, \dots$ be i.i.d. r.v. with law μ. We first show that μ has zero mean (its first moment exists because $\sigma^2 < \infty$):

$$m := \mathbb{E}[X] = \mathbb{E}\left[\frac{X+Y}{\sqrt{2}}\right] = \frac{1}{\sqrt{2}}(\mathbb{E}[X] + \mathbb{E}[Y]) = \sqrt{2}\, m, \quad \text{hence } m = 0.$$

Let $S_{0,k} := X_k$, $k \in \mathbb{N}$, and for every $k, n \in \mathbb{N}$, $S_{n,k} := (S_{n-1,2k-1} + S_{n-1,2k})/\sqrt{2}$. By an immediate induction over n, the $S_{n,k}$, $k \in \mathbb{N}$, are i.i.d. with law μ, and in particular we find that

$$S_{n,1} = \sqrt{2^n}\left(\frac{X_1 + \cdots + X_{2^n}}{2^n} - 2^n m\right) \xrightarrow[n\to\infty]{} \mathcal{N}(0, \sigma^2), \qquad \text{in distribution,}$$

using the central limit theorem. We conclude that $\mu = \mathcal{N}(0, \sigma^2)$. ■

Exercise 2.4.5. Let X_n, $n \in \mathbb{N}$, be i.i.d. standard Poisson r.v., and $S_n := X_1 + \cdots + X_n$.

Find the expression of $\mathbb{P}\left(\frac{S_n - n}{\sqrt{n}} \leqslant 0\right)$, and deduce that $\lim\limits_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$.

Solution of Exercise 2.4.5. We know from that $\mathbb{E}[X_1] = \mathrm{Var}(X_1) = 1$ (see *e.g.* Exercise 2.1.3), and it is a simple exercise to show that S_n, $n \in \mathbb{N}$, is a Poisson(n) r.v. It thus follows from the central limit theorem that

$$e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \mathbb{P}(S_n \leqslant n) = \mathbb{P}\left(\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n}} \leqslant 0\right) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathcal{N}(0,1) \leqslant 0) = \frac{1}{2}.$$

■

Exercise 2.4.6. Let $X_1, X_2, \dots$ be i.i.d. real r.v. with $\mathrm{Var}(X_1) = 1$, $\mathbb{E}[X_1] = 0$, and

$$S_n := X_1 + \cdots + X_n, \qquad n \in \mathbb{N}.$$

1. Using the central limit theorem, show that there exist $p > 0$ and $n_0 \in \mathbb{N}$ such that

$$\forall n \geqslant n_0, \quad \mathbb{P}(|S_n| \geqslant \sqrt{n}) \geqslant p.$$

2. Deduce that $\lim\limits_{n\to\infty} \mathbb{E}[|S_n|] = \infty$.

Solution of Exercise 2.4.6.

1. If N is a $\mathcal{N}(0,1)$ random variable, then $p := \mathbb{P}(N \geqslant 1) > 0$ and the central limit theorem gives

$$\mathbb{P}\left(\left|\frac{S_n}{\sqrt{n}}\right| \geqslant 1\right) \xrightarrow[n\to\infty]{} \mathbb{P}(|N| \geqslant 1) = 2p.$$

   Therefore it exists $n_0 \in \mathbb{N}$ such that $\mathbb{P}(|S_n| \geqslant \sqrt{n}) \geqslant p$ for all $n \geqslant n_0$.

2. With such $p > 0$ and $n_0 \in \mathbb{N}$, Markov's inequality entails

$$\forall n \geqslant n_0, \quad \mathbb{E}[|S_n|] \geqslant \sqrt{n}\,\mathbb{P}(|S_n| \geqslant \sqrt{n}) \geqslant p\sqrt{n},$$

   hence $\lim\limits_{n\to\infty} \mathbb{E}[|S_n|] = \infty$. ■

Exercise 2.4.7. For each $n \in \mathbb{N}$, let X_n be a $\mathcal{N}(\mu_n, \sigma_n^2)$ r.v. ($\mu_n \in \mathbb{R}$, $\sigma_n^2 > 0$). We suppose that X_n converges in distribution to some r.v. X.

1. Using characteristic functions, show that $(\sigma_n^2)_{n\in\mathbb{N}}$ converges to some $\sigma^2 \geqslant 0$.
2. Let $S(t) := \mathbb{P}(X > t)$, $t \in \mathbb{R}$.
   a) Justify that S is continuous at some $t_0 > 0$ large enough, with $S(t_0) < 1/4$.
   b) Deduce that $(\mu_n)_{n\in\mathbb{N}}$ is bounded from above (more precisely, $\limsup_{n\to\infty} \mu_n \leqslant t_0$).
   c) Deduce that $(\mu_n)_{n\in\mathbb{N}}$ is bounded.
3. Conclude that X has a normal distribution.

Solution of Exercise 2.4.7.

1. In terms of characteristic functions, the convergence in distribution $X_n \to X$ translates into

$$\forall t \in \mathbb{R}, \quad \Phi_{X_n}(t) = \exp\bigl(i\mu_n t - \sigma_n^2 t^2/2\bigr) \xrightarrow[n\to\infty]{} \Phi_X(t).$$

   Since Φ_X is continuous at 0 and $\Phi_X(0) = 1$, there is $t > 0$ small enough such that $\Phi_X(t) \neq 0$. Taking modulus above gives $\exp(-\sigma_n^2 t^2/2) \to |\Phi_X(t)| > 0$, so σ_n^2 converges to $\sigma^2 := -2\log|\Phi_X(t)|/t^2 \geqslant 0$.

2. a) Clear, since $S(t) = 1 - F_X(t) \to 0$ as $t \to \infty$, and S has at most countably many points of discontinuity.
   b) Because t_0 is a continuity point of S (*i.e*, of F_X),

$$\mathbb{P}(X_n > t_0) \xrightarrow[n\to\infty]{} S(t_0) < 1/4.$$

   If $(\mu_n)_{n\in\mathbb{N}}$ were not bounded from above, we could find n large enough such that $\mu_n > t_0$ and $\mathbb{P}(X_n > t_0) \leqslant S(t_0) + 1/4$, resulting in the contradiction

$$1/2 = \mathbb{P}(X_n > \mu_n) \leqslant \mathbb{P}(X_n > t_0) \leqslant S(t_0) + 1/4 < 1/2.$$

   Therefore $(\mu_n)_{n\in\mathbb{N}}$ must be bounded from above.
   c) Applying what precedes to the r.v. $-X_n$, $n \in \mathbb{N}$, and $-X$, we get that $(-\mu_n)_{n\in\mathbb{N}}$ is bounded from above. Hence $(\mu_n)_{n\in\mathbb{N}}$ is bounded.

3. It follows from the Bolzano–Weierstrass theorem that there exists a converging subsequence $\mu_{n_k} \to \mu \in \mathbb{R}$, as $k \to \infty$. Then

$$\forall t \in \mathbb{R}, \quad \Phi_{X_{n_k}}(t) = \exp\bigl(i\mu_{n_k} t - \sigma_{n_k}^2 t^2/2\bigr) \xrightarrow[k\to\infty]{} \exp\bigl(i\mu t - \sigma^2 t^2/2\bigr),$$

   which, because of the convergence in distribution of the subsequence $X_{n_k} \to X$, equals $\Phi_X(t)$. We conclude that X is a $\mathcal{N}(\mu, \sigma^2)$ r.v. (Consequently, $\mu_n \to \mu$.) ■

Exercise 2.4.8. We suppose that $X, X_1, X_2, \dots$ are real r.v. such that X_n converges to X in distribution.

1. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a nonnegative continuous function. Show that

$$\liminf_{n\to\infty} \mathbb{E}[f(X_n)] \geqslant \mathbb{E}[f(X)].$$

   Hint. Apply Fatou's lemma/monotone convergence theorem to some $(f_k(X))_{k\in\mathbb{N}}$.

2. Deduce that if $(\mathbb{E}[|X_n|])_{n\in\mathbb{N}}$ is bounded, then $\mathbb{E}[|X|] < \infty$.

3. Deduce that if $X_n \geqslant 0$ a.s. for every $n \in \mathbb{N}$, then $X \geqslant 0$ a.s.

Solution of Exercise 2.4.8.

1. For each $k \in \mathbb{N}$, the function $f_k\colon x \mapsto \min(f(x), k)$ is bounded and continuous. Since $f \geqslant f_k$ we then have

$$\liminf_{n\to\infty} \mathbb{E}[f(X_n)] \geqslant \lim_{n\to\infty} \mathbb{E}[f_k(X_n)] = \mathbb{E}[f_k(X)], \qquad \text{for every } k \in \mathbb{N}.$$

   But $(f_k(X))_{k\in\mathbb{N}}$ is a sequence of nonnegative r.v. converging pointwise to $f(X)$ as $k \to \infty$, so

$$\liminf_{k\to\infty} \mathbb{E}[f_k(X)] \geqslant \mathbb{E}[f(X)]$$

   (actually $\mathbb{E}[f_k(X)] \uparrow \mathbb{E}[f(X)]$, by monotone convergence). The conclusion follows.

2. Just apply Question 1 with the nonnegative continuous function $f\colon x \mapsto |x|$.

3. We apply Question 1 with the nonnegative continuous function $f\colon x \mapsto \max(-x, 0)$. We obtain $\mathbb{E}[\max(-X, 0)] = 0$, which means that $\max(-X, 0) = 0$ almost surely, or equivalently, $X \geqslant 0$ a.s. ■

Exercise 2.4.9. Let $X_1, X_2, \dots$ be i.i.d. centered, square-integrable r.v. Show that

$$\liminf_{n\to\infty} \mathbb{P}(|X_1 + \cdots + X_n| \geqslant \sqrt{n}) > 0.$$

Solution of Exercise 2.4.9. Let $S_n := X_1 + \cdots + X_n$ and $\sigma^2 := \mathrm{Var}(X_1)$. If N is a $\mathcal{N}(0, \sigma^2)$ random variable, then $p := \mathbb{P}(N \geqslant 1) > 0$ and the central limit theorem gives

$$\mathbb{P}\left(\left|\frac{S_n}{\sqrt{n}}\right| \geqslant 1\right) \xrightarrow[n\to\infty]{} \mathbb{P}(|N| \geqslant 1) = 2p.$$

Therefore it exists $n_0 \in \mathbb{N}$ such that $\mathbb{P}(|S_n| \geqslant \sqrt{n}) \geqslant p$ for all $n \geqslant n_0$. Hence the result.
Remark. In particular we deduce from Markov's inequality that

$$\liminf_{n\to\infty} \frac{\mathbb{E}[|X_1 + \cdots + X_n|]}{\sqrt{n}} > 0.$$

■

Exercise 2.4.10. Let $\lambda > 0$, and for $n > \lambda$, X_n be a random variable having the binomial distribution with parameter $(n, \lambda/n)$, that is

$$\mathbb{P}(X_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}, \qquad k = 0,1,2,\dots,n.$$

Compute $\lim_{n\to\infty} \mathbb{P}(X_n = k)$. What do you recognize?

Solution of Exercise 2.4.10. For k fixed we have, as $n \to \infty$,

$$\binom{n}{k} \sim \frac{n^k}{k!}, \quad \text{and} \quad \left(1-\frac{\lambda}{n}\right)^{n-k} = \left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{\lambda}{n}\right)^{n} \sim e^{-\lambda}.$$

Reporting these two equivalents in $\mathbb{P}(X_n = k)$ and simplifying, we get $e^{-\lambda}\frac{\lambda^k}{k!}$ at the limit. We recognize here the Poisson distribution with parameter λ (see Exercise 2.1.8). (We say that X_n converges *in distribution* to a Poisson random variable.) ■

Exercise 2.4.11. Let $f\colon [0,1] \to \mathbb{R}$ be a continuous function, $x \in [0,1]$ and $X_n := X_n(x)$ be a random variable having the binomial distribution with parameter (n, x); see Exercise 2.1.4. We define $Y_n := f(X_n/n)$, so that Y_n is a discrete random variable taking values in the set $\mathscr{Y} := \{f(k/n)\colon k = 0,1,2,\dots,n\}$.

1. Let $m \in \mathbb{N}$. Recall why there exist $C > 0$ and $\delta_m > 0$ such that

$$\forall t \in [0,1], \quad |f(t)| \leqslant C,$$

and

$$\forall (s,t) \in [0,1]^2 \text{ with } |t-s| \leqslant \delta_m, \quad |f(t)-f(s)| \leqslant \frac{1}{m}.$$

2. Check that for every $\delta > 0$,

$$\mathbb{E}[|Y_n - f(x)|] \leqslant 2C\,\mathbb{P}(|X_n - \mathbb{E}[X_n]| > n\delta) + \mathbb{E}[|f(X_n/n) - f(x)|\mathbb{1}_{\{|X_n - nx| \leqslant n\delta\}}],$$

then deduce that for every $m \in \mathbb{N}$,

$$\mathbb{E}[|Y_n - f(x)|] \leqslant 2C\frac{x(1-x)}{n\delta_m^2} + \frac{1}{m}.$$

Hint. Recall $\mathbb{E}[X_n]$, $\mathrm{Var}(X_n)$ (see Exercise 2.1.4), and apply Chebyshev's inequality.

3. We define $B_n\colon x \mapsto B_n(x) := \mathbb{E}[Y_n] = \mathbb{E}[f(X_n(x)/n)]$.

   a) Check that B_n is a polynomial function in $x \in [0,1]$.

   b) Conclude that

$$\sup_{x\in[0,1]} |B_n(x) - f(x)| \xrightarrow[n\to\infty]{} 0.$$

Conclusion. Continuous functions defined on a compact interval can be (uniformly) approximated by polynomials!

Solution of Exercise 2.4.11.

1. Being continuous on the compact set $[0,1]$, the function f is bounded — hence the existence of such a $C>0$, and (by Heine's theorem) uniformly continuous — hence the existence of such $\delta_m>0$, $m\in\mathbb{N}$.

2. We have

$$|Y_n - f(x)| = |f(X_n/n) - f(x)|\mathbb{1}_{\{|X_n-nx|>n\delta\}} + |f(X_n/n) - f(x)|\mathbb{1}_{\{|X_n-nx|\leqslant n\delta\}},$$

where for C and $\delta := \delta_m$ as in Question 1, the first term in the right-hand side is bounded by $2C\mathbb{1}_{\{|X_n-nx|>n\delta_m\}}$ and the second term is bounded by $1/m$. Now, recall the moments of the binomial distribution in Exercise 2.1.4. On the one hand, we have $\mathbb{E}[X_n]=nx$, so we deduce the first inequality by taking expectations on both sides. On the other hand, $\mathrm{Var}(X_n)=nx(1-x)$, so the second inequality then follows from Chebyshev's inequality

$$\mathbb{P}(|X_n-\mathbb{E}[X_n]|>n\delta_m)\leqslant \frac{\mathrm{Var}(X_n)}{n^2\delta_m^2} = \frac{x(1-x)}{n\delta_m^2}.$$

3. Successively,

$$\begin{aligned}
\mathbb{E}[Y_n] &\overset{\text{def}}{=} \sum_{y\in\mathscr{Y}} y\,\mathbb{P}(Y_n=y)\\
&= \sum_{y\in\mathscr{Y}} y \sum_{k=0}^{n} \mathbb{1}_{\{f(k/n)=y\}}\mathbb{P}(X_n=k)\\
&= \sum_{k=0}^{n} f(k/n)\,\mathbb{P}(X_n=k) \underbrace{\sum_{y\in\mathscr{Y}} \mathbb{1}_{\{f(k/n)=y\}}}_{=1}\\
&= \sum_{k=0}^{n} f(k/n)\binom{n}{k} x^k(1-x)^{n-k},
\end{aligned}$$

which is a polynomial function in $x\in[0,1]$.

4. Observe, by linearity of the expectation and the triangle inequality, that

$$|B_n(x)-f(x)| = |\mathbb{E}[Y_n-f(x)]| \leqslant \mathbb{E}[|Y_n-f(x)|].$$

Therefore the result of Question 2 gives, for all $n,m\in\mathbb{N}$,

$$\sup_{x\in[0,1]}|B_n(x)-f(x)|\leqslant \frac{C}{2n\delta_m^2}+\frac{1}{m},$$

because $x(1-x)\leqslant\frac{1}{4}$ for $x\in[0,1]$. We finally get

$$\limsup_{n\to\infty}\sup_{x\in[0,1]}|B_n(x)-f(x)|\leqslant\frac{1}{m}\xrightarrow[m\to\infty]{}0,$$

hence what we wanted to show.

Remark. This is a *probabilistic* proof of Weierstraß approximation theorem. ■

Exercise 2.4.12. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. r.v. with $\mathbb{E}[|X_1|] < \infty$, $\mu := \mathbb{E}[X_1]$, and

$$S_n := X_1 + \cdots + X_n, \qquad n \in \mathbb{N}.$$

1. Show that if $\mu > 0$ (resp. $\mu < 0$), then $S_n \longrightarrow \infty$ (resp. $-\infty$) almost surely.
2. We suppose here that $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$.
   a) Let $m \geqslant 2k+1$ in $\mathbb{N}$. Show that $S_{n+m} - S_n = m$ infinitely often, almost surely. Deduce that $\limsup |S_n| > k$ almost surely.
   b) Conclude that $\limsup |S_n| = \infty$ a.s.

Solution of Exercise 2.4.12.

1. This follows from the law of large numbers, since then $S_n \sim n\mu$ as $n \to \infty$, a.s.
2. a) The probability of $\Lambda_p := \{S_{pm+m} - S_{pm} = m\}$ is that of

   $$\{X_{pm+1} = 1, \ldots, X_{pm+m} = 1\},$$

   which equals the constant 2^{-m} because $(X_n)_{n\in\mathbb{N}}$ is i.i.d. with $\mathbb{P}(X_1 = 1) = 1/2$. As $(\Lambda_p)_{p\in\mathbb{N}}$ is moreover an independent sequence of events, it follows from the second part of the Borel–Cantelli lemma that Λ_p occurs infinitely often, almost surely. But when $S_{n+m} - S_n = m$, we must have either $S_{n+m} > k$ or $S_n < -k$. Hence $\limsup |S_n| > k$ almost surely.

   b) By Question 2,

   $$\mathbb{P}(\limsup |S_n| = \infty) = \mathbb{P}\left(\bigcap_{k\in\mathbb{N}} \{\limsup |S_n| > k\}\right) = 1.$$

   (In fact, we can show that $\limsup S_n = \infty$ and $\liminf S_n = -\infty$, a.s.) ■

Exercise 2.4.13. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. real r.v. with distribution function F such that $F(t)/t \longrightarrow \lambda$ as $t \to 0^+$, for some $\lambda > 0$. Let $Z_n := n \min(X_1, \ldots, X_n)$, $n \in \mathbb{N}$.

1. Check the following facts:
   a) For every $n \in \mathbb{N}$, $Z_n > 0$ almost surely.
   b) For every $t > 0$, $\mathbb{P}(Z_n > t) \longrightarrow e^{-\lambda t}$ as $n \to \infty$.
   c) For every $\varepsilon > 0$, there is $nX_n \leqslant \varepsilon$ infinitely often, almost surely.
2. Conclude that $\liminf Z_n = 0$ a.s., but that $(Z_n)_{n\in\mathbb{N}}$ does not converge a.s.

Solution of Exercise 2.4.13.

1. a) Since $F(t) \sim \lambda t$ as $t \to 0^+$, we have $F(0) = 0$ by right-continuity. Then

$$\mathbb{P}(Z_n \leqslant 0) \leqslant \mathbb{P}\left(\bigcup_{i=1}^{n} \{nX_i \leqslant 0\}\right) \leqslant nF(0) = 0,$$

because $X_1, \dots, X_n$ are all distributed according to F. Thus $Z_n > 0$ a.s.

b) Let $t > 0$. Using that the r.v. $X_1, \dots, X_n$ are mutually independent, and that $\log(1 - F(\frac{t}{n})) = -\lambda \frac{t}{n} + o(\frac{1}{n})$ as $n \to \infty$, we have

$$\mathbb{P}(Z_n > t) = \mathbb{P}\left(X_1 > \frac{t}{n}, \dots, X_n > \frac{t}{n}\right) = \left(1 - F\left(\frac{t}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{-\lambda t}.$$

(This actually shows that $(Z_n)_{n\in\mathbb{N}}$ converges *in law* toward the exponential distribution with rate λ.)

c) Let $\varepsilon > 0$. Because $\frac{1}{n} = O\big(F(\frac{\varepsilon}{n})\big)$ as $n \to \infty$, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(nX_n \leqslant \varepsilon) = \sum_{n=1}^{\infty} F\left(\frac{\varepsilon}{n}\right) = \infty,$$

where the events $\{nX_n \leqslant \varepsilon\}$, $n \in \mathbb{N}$, are independent. The result follows by the second part of the Borel–Cantelli lemma.

2. By Question 1.c),

$$1 = \mathbb{P}\left(\bigcap_{k\in\mathbb{N}} \left\{\limsup \left\{Z_n \leqslant \frac{1}{k}\right\}\right\}\right) \leqslant \mathbb{P}\left(\bigcap_{k\in\mathbb{N}} \left\{\liminf Z_n \leqslant \frac{1}{k}\right\}\right) = \mathbb{P}(\liminf Z_n \leqslant 0),$$

and by 1.a), $1 = \mathbb{P}(\bigcap_{n\in\mathbb{N}}\{Z_n \geqslant 0\}) \leqslant \mathbb{P}(\liminf Z_n \geqslant 0)$. Thus $\liminf Z_n = 0$ a.s., so if $(Z_n)_{n\in\mathbb{N}}$ were converging a.s., then the limit would be 0. This is not possible because, by Question 1.b), Z_n does not even converge to 0 in probability. ■

Exercise 2.4.14. For each $k \in \mathbb{N}$, let $(X_n^{(k)})_{n\in\mathbb{N}}$ be a sequence of real r.v. converging to 0 in probability, as $n \to \infty$. Define, for $k, n \in \mathbb{N}$,

$$Y_n^{(k)} := \sum_{i=1}^{k} X_n^{(i)},$$

and, for $\varepsilon > 0$ arbitrary,

$$f_n(k) := \mathbb{P}\left(\left|Y_n^{(k)}\right| > \varepsilon\right).$$

1. Let $k \in \mathbb{N}$. Show that $f_n(k) \longrightarrow 0$ ($Y_n^{(k)}$ converges to 0 in probability), as $n \to \infty$.
2. Let K be a $\mathbb{N}$-valued r.v. *independent* of $(X_n^{(k)})$, and $Y_n^{(K)}(\omega) := Y_n^{(K(\omega))}(\omega)$, $\omega \in \Omega$.
   a) Show that $\mathbb{P}(|Y_n^{(K)}| > \varepsilon) = \mathbb{E}[f_n(K)]$.
   b) Conclude that $Y_n^{(K)}$ converges to 0 in probability, as $n \to \infty$.

Solution of Exercise 2.4.14.

1. By the triangle inequality, if $|X_n^{(i)}| \leqslant \varepsilon/k$ for all $i = 1, \dots, k$, then $|Y_n^{(k)}| \leqslant \varepsilon$. Therefore

$$f_n(k) \leqslant \mathbb{P}\left(\bigcup_{i=1}^{k} \left\{\left|X_n^{(i)}\right| > \frac{\varepsilon}{k}\right\}\right) \leqslant \sum_{i=1}^{k} \mathbb{P}\left(\left|X_n^{(i)}\right| > \frac{\varepsilon}{k}\right),$$

where each element in the latter sum tends to 0 as $n \to \infty$, since the sequences $(X_n^{(i)})_{n\in\mathbb{N}}$, $i = 1, \dots, k$, all converge to 0 in probability.

2. a) Writing the event of interest as the disjoint union

$$\left\{\left|Y_n^{(K)}\right| > \varepsilon\right\} = \bigsqcup_{k\in\mathbb{N}} \{K = k\} \cap \left\{\left|Y_n^{(k)}\right| > \varepsilon\right\},$$

we have, using the independence of K with $(X_n^{(k)})$,

$$\mathbb{P}\left(\left|Y_n^{(K)}\right| > \varepsilon\right) = \sum_{k\in\mathbb{N}} \mathbb{P}(K = k) \underbrace{\mathbb{P}\left(\left|Y_n^{(k)}\right| > \varepsilon\right)}_{=f_n(k)} = \mathbb{E}[f_n(K)].$$

b) Let $m \in \mathbb{N}$. For every $\omega \in \Omega$, applying Question 1 with $k := K(\omega) \in \mathbb{N}$ gives $f_n(K(\omega)) \longrightarrow 0$ as $n \to \infty$. It is moreover clear that $0 \leqslant f_n \leqslant 1$ for every $n \in \mathbb{N}$. Question 2.a) and dominated convergence theorem then imply

$$\mathbb{P}\left(\left|Y_n^{(K)}\right| > \varepsilon\right) = \mathbb{E}[f_n(K)] \xrightarrow[n\to\infty]{} 0,$$

that is the convergence in probability of $Y_n^{(K)}$ toward 0, as $n \to \infty$. ∎

Exercise 2.4.15. Let X_n, $n \geqslant 1$, be centered with variance σ_n^2, such that $\sigma_n^2 \to 0$ as $n \to \infty$. Show that X_n converges to 0 in $\mathrm{L}^2(\mathbb{P})$ (and in probability).

Solution of Exercise 2.4.15. X_n is centered, so $\mathbb{E}[|X_n|^2] = \mathrm{Var}(X_n) = \sigma_n^2 \to 0$ as $n \to \infty$, *i.e*, $X_n \xrightarrow[n\to\infty]{\mathrm{L}^2(\mathbb{P})} 0$. ∎

Exercise 2.4.16. Let X_n, $n \geqslant 1$, be i.i.d. centered random variables with variance $\sigma^2 < \infty$. Show that $\frac{1}{n}\sum_{j=1}^{n} X_j$ converges to 0 in $\mathrm{L}^2(\mathbb{P})$ (and in probability).

Solution of Exercise 2.4.16. We have

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{1}{n}\sum_{j=1}^{n} X_j\right|^2\right] &= \frac{1}{n^2}\mathrm{Var}\left(\sum_{j=1}^{n} X_j\right) && (X_1, \dots, X_n \text{ centered}) \\
&= \frac{1}{n^2}\sum_{j=1}^{n} \mathrm{Var}(X_j) && (X_1, \dots, X_n \text{ independent}) \\
&= \frac{1}{n^2} \cdot n \cdot \sigma^2 && (X_1, \dots, X_n \text{ have same law}) \\
&\xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

∎

Exercise 2.4.17. Let X_j, $j \geqslant 1$, be i.i.d. with standard Laplace distribution (having common density $e^{-|x|}/2$). Show the convergence in distribution

$$\sqrt{n}\,\frac{\sum_{j=1}^n X_j}{\sum_{j=1}^n X_j^2} \xrightarrow[n\to\infty]{D} Y,$$

where Y is a $\mathscr{N}(0,1/2)$ Gaussian variable.
Hint. Use Slutsky's lemma.

Solution of Exercise 2.4.17. We have $\mathbb{E}[X_1]=0$, $\mathbb{E}[X_1^2]=2$. By the central limit theorem,

$$U_n := \frac{1}{\sqrt{n}}\sum_{j=1}^n X_j \xrightarrow[n\to\infty]{D} 2Y,$$

with Y having the $\mathscr{N}(0,1/2)$ distribution, while by the strong law of large numbers,

$$V_n := \frac{1}{n}\sum_{j=1}^n X_j^2 \xrightarrow[n\to\infty]{\text{a.s.}} 2$$

(and in probability). By Slutsky's lemma, we conclude that

$$\frac{U_n}{V_n} = \sqrt{n}\,\frac{\sum_{j=1}^n X_j}{\sum_{j=1}^n X_j^2} \xrightarrow[n\to\infty]{D} Y.$$

■

Exercise 2.4.18. Let X_j, $j \geqslant 1$, be i.i.d. with mean 1 and variance $\sigma^2 \in (0,\infty)$. Define $S_n := \sum_{j=1}^n X_j$, $n \in \mathbb{N}$. Show the convergence in distribution

$$\frac{2}{\sigma}(\sqrt{S_n}-\sqrt{n}) \xrightarrow[n\to\infty]{D} Y,$$

where Y is a $\mathscr{N}(0,1)$ Gaussian variable.

Solution of Exercise 2.4.18. We observe that

$$\frac{2}{\sigma}(\sqrt{S_n}-\sqrt{n}) = \frac{2}{1+\sqrt{\frac{S_n}{n}}}\cdot\frac{S_n-n}{\sigma\sqrt{n}}.$$

By the strong law of large numbers, the first factor of the right-hand side tends a.s. to 1, while by the central limit theorem, the second factor converges in distribution to Y, where Y has the standard $\mathscr{N}(0,1)$ distribution. We conclude by Slutsky's lemma. ■

Exercise 2.4.19. Let X_j, $j \geqslant 1$, be i.i.d. with mean 0 and variance $\sigma^2 \in (0,\infty)$. Define $S_n := \sum_{j=1}^n X_j$, $n \in \mathbb{N}$. Show that $S_n/\sigma\sqrt{n}$ does not converge in probability.

Solution of Exercise 2.4.19. Let N have the standard $\mathscr{N}(0,1)$ distribution. By the central limit theorem,

$$\mathbb{P}\big(\big|S_n/\sigma\sqrt{n}\big|>1\big) \xrightarrow[n\to\infty]{} \mathbb{P}(|N|>1).$$

Since the right-hand side is >0, the convergence in probability cannot hold. ■

Exercise 2.4.20. Let X_j, $j \geqslant 1$, be i.i.d. with mean 0 and variance $\sigma^2 < \infty$. Define $S_n := \sum_{j=1}^n X_j$, $n \in \mathbb{N}$. Show that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{|S_n|}{\sqrt{n}}\right] = \sqrt{\frac{2}{\pi}}\,\sigma.$$

Solution of Exercise 2.4.20. Let $Y_n := S_n/\sqrt{n}$. By the central limit theorem,

$$Y_n \xrightarrow[n\to\infty]{D} Y, \tag{$\star$}$$

where $Y \sim \mathcal{N}(0,\sigma^2)$. Observe that $\mathbb{E}[|Y|] = 2\mathbb{E}[Y; Y>0] = \sigma\sqrt{2/\pi}$. Thus we need to justify the convergence of first moments

$$\mathbb{E}[|Y_n|] \xrightarrow[n\to\infty]{} \mathbb{E}[|Y|]. \tag{$\star\star$}$$

This does not follow from just ($\star$), because $x \mapsto |x|$ is of course continuous, but *not* bounded. However, for every $k \in \mathbb{N}$, the map $x \mapsto |x| \wedge k$ is continuous and bounded. On the one hand, convergence ($\star$) gives

$$\mathbb{E}[|Y_n|] \geqslant \mathbb{E}[|Y_n| \wedge k] \xrightarrow[n\to\infty]{} \mathbb{E}[|Y| \wedge k],$$

so that letting now $k \to \infty$ yields $\liminf_{n\to\infty} \mathbb{E}[|Y_n|] \geqslant \mathbb{E}[|Y|]$ by monotone convergence (or Fatou's lemma). On the other hand,

$$\begin{aligned}
\mathbb{E}[|Y_n|] &\leqslant \mathbb{E}[|Y_n| \wedge k] + \mathbb{E}[|Y_n|\mathbb{1}_{\{|Y_n|>k\}}] \\
&\leqslant \mathbb{E}[|Y_n| \wedge k] + \sigma\sqrt{\mathbb{P}(|Y_n|>k)} && \text{(Cauchy–Schwarz)} \\
&\xrightarrow[n\to\infty]{} \mathbb{E}[|Y| \wedge k] + \sigma\sqrt{\mathbb{P}(|Y|>k)} && \text{(by ($\star$))} \\
&\xrightarrow[k\to\infty]{} \mathbb{E}[|Y|]. && \text{(Beppo Levi)}
\end{aligned}$$

This shows that $\limsup_{n\to\infty} \mathbb{E}[|Y_n|] \leqslant \mathbb{E}[|Y|]$, and ($\star\star$) is now proved.

Additional exercise. Use the same technique to prove that if $X_n \xrightarrow{D} X$ and $(X_n)_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^q(\mathbb{P})$ for some $q > 1$, then $\mathbb{E}[X_n^p] \to \mathbb{E}[X^p]$ for every $1 \leqslant p < q$. Find a counterexample where $X_n \xrightarrow{D} X$ but $\mathbb{E}[X_n] \not\to \mathbb{E}[X]$. ■

Exercise 2.4.21. Let $q \geqslant 1$ and $(X_n)_{n\in\mathbb{N}}$ be a sequence of real r.v. bounded in $\mathrm{L}^q(\mathbb{P})$:

$$C := \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^q] < \infty.$$

1. Suppose that X_n converges almost surely to some r.v. X as $n \to \infty$.
   a) Is X in $\mathrm{L}^q(\mathbb{P})$?
   b) Suppose $q > 1$ and $1 \leqslant p < q$. Does $\mathbb{E}[|X_n|^p]$ converge to $\mathbb{E}[|X|^p]$ as $n \to \infty$?
2. Same questions if the convergence $X_n \to X$ holds in probability.
3. Same questions if the convergence $X_n \to X$ holds in distribution.

Solution of Exercise 2.4.21.

1. a) Yes. Fatou's lemma gives $\mathbb{E}[|X|^q] \leqslant \liminf \mathbb{E}[|X_n|^q] \leqslant C < \infty$, so $X \in \mathrm{L}^q(\mathbb{P})$.

   b) Yes. There is in fact convergence in $\mathrm{L}^p(\mathbb{P})$:

$$\mathbb{E}[|X_n - X|^p] = \mathbb{E}[|X_n - X|^p \mathbb{1}_{\{|X_n - X| \leqslant 1\}}] + \mathbb{E}[|X_n - X|^p \mathbb{1}_{\{|X_n - X| > 1\}}].$$

   The first term tends to 0 by dominated convergence. For the second term, Hölder's inequality gives (for $p/q + (q-p)/q = 1$)

$$\begin{aligned}\mathbb{E}[|X_n - X|^p \mathbb{1}_{\{|X_n - X| > 1\}}] &\leqslant \mathbb{E}[|X_n - X|^q]^{p/q} \mathbb{P}(|X_n - X| > 1)^{(q-p)/q} \\ &\leqslant 2^p C^{p/q} \mathbb{P}(|X_n - X| > 1)^{(q-p)/q},\end{aligned}$$

   which also tends to 0 as $n \to \infty$.

2. Because of the convergence in probability there is a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ which converges almost surely toward X (this is a consequence of Borel–Cantelli's lemma). Applying 1.a) to this subsequence yields $X \in \mathrm{L}^q(\mathbb{P})$. Further, for $1 \leqslant p < q$, we deduce from 1.b) that for every subsequence of $(X_n)_{n\in\mathbb{N}}$, there is a subsubsequence converging to X in $\mathrm{L}^p(\mathbb{P})$. Hence $(X_n)_{n\in\mathbb{N}}$ converges to X in $\mathrm{L}^p(\mathbb{P})$.

3. a) Yes. For each $k \in \mathbb{N}$, the function $x \mapsto |x|^q \wedge k$ is bounded and continuous, so the convergence in distribution $X_n \to X$ entails

$$\mathbb{E}[|X|^q \wedge k] = \lim_{n\to\infty} \mathbb{E}[|X_n|^q \wedge k].$$

   Now, by the monotone convergence theorem,

$$\mathbb{E}[|X|^q] = \lim_{k\to\infty} \mathbb{E}[|X|^q \wedge k] = \liminf_{k\to\infty} \liminf_{n\to\infty} \mathbb{E}[|X_n|^q \wedge k] \leqslant \liminf_{n\to\infty} \mathbb{E}[|X_n|^q], \tag{2.1}$$

   where the right-hand side is bounded by $C < \infty$.

   b) Yes. As (2.1) also holds with p in place of q, it remains to show that

$$\limsup_{n\to\infty} \mathbb{E}[|X_n|^p] \leqslant \mathbb{E}[|X|^p]. \tag{2.2}$$

   But

$$\mathbb{E}[|X_n|^p] \leqslant \mathbb{E}[|X_n|^p \wedge k^p] + \mathbb{E}[|X_n|^p \mathbb{1}_{\{|X_n| > k\}}]$$

   for $k > 0$ arbitrary. As $x \mapsto |x|^p \wedge k^p$ is a continuous bounded function, the first term tends to $\mathbb{E}[|X|^p \wedge k^p]$ as $n \to \infty$ by the convergence in distribution $X_n \to X$. Using Hölder's inequality for the second term gives

$$\limsup_{n\to\infty} \mathbb{E}[|X_n|^p] \leqslant \mathbb{E}[|X|^p \wedge k^p] + C^{p/q} \limsup_{n\to\infty} \mathbb{P}(|X_n| > k)^{(q-p)/q}.$$

   Now we may choose $k > 0$ arbitrary large such that k and $-k$ are continuity points of F_X, so that $\mathbb{P}(|X_n| > k)$ converges to $\mathbb{P}(|X| > k)$ as $n \to \infty$. Then, using the monotone convergence theorem and the fact that $\mathbb{P}(|X| > k)$ tends to 0 as $k \to \infty$, we get (2.2) as desired.

Remark. Here is another solution to the exercise involving an argument of uniform integrability. The convergence of $(X_n)_{n\in\mathbb{N}}$ to X almost surely (or in probability) implies that of the sequence $(|X_n - X|^p)_{n\in\mathbb{N}}$ (towards 0), which, by Minkowski's inequality and Fatou's lemma, is bounded in $\mathrm{L}^{q/p}(\mathbb{P})$ with $q/p>1$, and thus uniformly integrable; then, $|X_n - X|^p$ converges to 0 in $\mathrm{L}^1(\mathbb{P})$ (*i.e* X_n converges to X in $\mathrm{L}^p(\mathbb{P})$), and in particular we have the convergence of $\mathbb{E}[|X_n|^p]$ towards $\mathbb{E}[|X|^p]$. When the convergence $X_n \to X$ holds only in distribution, we may apply the preceding for $Y_n \xrightarrow{\text{a.s.}} Y$ given by Skorokhod's representation theorem, with Y_n distributed as X_n, $n\in\mathbb{N}$, and Y distributed as X (defined on some other probability space): we can conclude because $\mathbb{E}[|X_n|^p]$ and $\mathbb{E}[|X|^p]$ are fully determined by the respective distributions of X_n, $n\in\mathbb{N}$, and of X. ∎

Exercise 2.4.22. Let $X, X_1, \ldots$ be random variables and $g\colon \mathbb{R}\to[0,\infty)$ measurable. We suppose that

$$X_n \xrightarrow[n\to\infty]{(d)} X, \qquad \text{and} \quad \Theta := \sup_{n\geqslant 1} \mathbb{E}[g(X_n)] < \infty.$$

Show that for every continuous function $f\colon \mathbb{R}\to\mathbb{R}$ with $f = o(g)$ at $\pm\infty$, we have

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)] \text{ in } \mathbb{R}.$$

Solution of Exercise 2.4.22. Suppose first $f\geqslant 0$. Since $f=o(g)$ at $\pm\infty$, we have, for every $\varepsilon>0$ fixed, $f(x)\leqslant \varepsilon g(x)$ whenever $|x|>K_\varepsilon$ with K_ε sufficiently large. On the one hand, for all $n\geqslant 1$,

$$\begin{aligned}\mathbb{E}[f(X_n)] &\leqslant \Theta + \mathbb{E}[f(X_n)\mathbb{1}_{\{|X_n|\leqslant K_1\}}]\\ &\leqslant \Theta + \sup_{|x|\leqslant K_1} f(x);\end{aligned}$$

hence

$$\begin{aligned}\mathbb{E}[f(X)] &\leqslant \liminf_{k\to\infty} \mathbb{E}[f(X)\wedge k] && \text{(Fatou's lemma)}\\ &= \liminf_{k\to\infty}\liminf_{n\to\infty} \mathbb{E}[f(X_n)\wedge k] && (X_n \xrightarrow[n\to\infty]{(d)} X)\\ &\leqslant \liminf_{n\to\infty} \mathbb{E}[f(X_n)]\\ &< \infty.\end{aligned}$$

On the other hand, for all $n\geqslant 1$ and $k > K_\varepsilon \vee \sup_{|x|\leqslant K_\varepsilon} f(x)$,

$$\begin{aligned}\mathbb{E}[f(X_n)] &= \mathbb{E}[f(X_n)\mathbb{1}_{\{|X_n|>k\}}] + \mathbb{E}[f(X_n)\mathbb{1}_{\{|X_n|\leqslant k\}}]\\ &\leqslant \varepsilon\Theta + \mathbb{E}[f(X_n)\wedge k];\end{aligned}$$

hence

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{E}[f(X_n)] &\leqslant \varepsilon\Theta + \mathbb{E}[f(X)\wedge k] && (X_n \xrightarrow[n\to\infty]{(d)} X)\\ &\leqslant \varepsilon\Theta + \mathbb{E}[f(X)].\end{aligned}$$

Since this is true for all $\varepsilon>0$, we conclude that

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)].$$

The result holds in all generality by writing $f = f_+ - f_-$. ∎

Exercise 2.4.23 (True or false?). Prove, or disprove (by giving a counterexample), briefly the following statements. We consider real r.v. on some general probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

1. If $|X_n - X| \to 0$ a.s., then $\mathbb{E}[|X_n - X|] \to 0$.
2. If $X_n \to X$ in probability and $(\mathbb{E}[X_n^2])_{n\in\mathbb{N}}$ is bounded, then $\mathbb{E}[|X_n - X|] \to 0$.
3. If X_n tends to 0 in probability, then so does $(X_1 + \cdots + X_n)/n$.
4. If $\mathbb{E}[|X_n - X|] \to 0$, then $|X_n - X| \to 0$ a.s.

Solution of Exercise 2.4.23.

1. False[2]: take *e.g.* $X_n := n\mathbb{1}_{\{U\leqslant 1/n\}}$, where U is uniformly distributed on $(0,1)$. Then there is $X_n \to 0$ almost surely, whereas $\mathbb{E}[X_n] = 1$ for every $n \in \mathbb{N}$.
2. True: let $C := \sup_{n\in\mathbb{N}} \mathbb{E}[X_n^2] < \infty$. Extracting a subsequence converging a.s., we have also $\mathbb{E}[X^2] \leqslant C$ by Fatou's lemma. As $1 = \mathbb{1}_{\{|X_n - X|\leqslant\varepsilon\}} + \mathbb{1}_{\{|X_n - X|>\varepsilon\}}$, the triangle inequality and Cauchy–Schwarz inequality give
$$\begin{aligned}\mathbb{E}[|X_n - X|] &\leqslant \varepsilon + \mathbb{E}[|X_n - X|^2]^{1/2}\,\mathbb{P}(|X_n - X| > \varepsilon)^{1/2} \\ &\leqslant \varepsilon + 2\sqrt{C}\,\mathbb{P}(|X_n - X| > \varepsilon)^{1/2},\end{aligned}$$
hence $\limsup \mathbb{E}[|X_n - X|] \leqslant \varepsilon$. Taking $\varepsilon > 0$ arbitrary, the conclusion follows.
3. False: take $(X_n)_{n\in\mathbb{N}}$ independent with $\mathbb{P}(X_n = n^2) = 1 - \mathbb{P}(X_n = 0) = 1/n$. Clearly, $X_n \to 0$ in probability, but $(X_1 + \cdots + X_n)/n$ does not since
$$\mathbb{P}\left(\frac{1}{n^2}\sum_{i=1}^{n^2} X_i > 1\right) \geqslant 1 - \mathbb{P}(X_i = 0,\ n < i \leqslant n^2) = 1 - \frac{1}{n} \xrightarrow[n\to\infty]{} 1.$$
4. False: take *e.g.* $(X_n)_{n\in\mathbb{N}}$ independent with $\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = 1/n$. Then $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] = 1/n \to 0$, but X_n does not converge to 0 a.s. because by the second part of the Borel–Cantelli lemma, $X_n = 1$ infinitely often, a.s. ∎

2.5 Gaussian vectors

Exercise 2.5.1. Let $\mathbf{X} := (X_1, X_2, X_3) \in \mathbb{R}^3$ be a centered random Gaussian vector such that $\mathbb{E}[X_i^2] = 1$ and $\mathbb{E}[X_i X_j] = 1/2$ for $1 \leqslant i \neq j \leqslant 3$.

1. Give the dispersion matrix and the characteristic function of $\mathbf{X}$.
2. What is the law of $X_1 - X_2 + 2X_3$?
3. Does there exist $a \in \mathbb{R}$ such that $X_1 + aX_2$ and $X_1 - X_2$ are independent?

[2] Of course the statement becomes true if there is moreover domination: $\forall n,\ |X_n| \leqslant Y \in \mathrm{L}^1(\mathbb{P})$.

4. Show that $\mathbf{X}$ admits a density and explicit it.

Solution of Exercise 2.5.1.

1. The dispersion matrix is
$$D = \begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix}.$$
The characteristic function is $\Phi(\boldsymbol{\lambda}) := \exp(-\langle \boldsymbol{\lambda}, D\boldsymbol{\lambda}\rangle/2)$, $\boldsymbol{\lambda} \in \mathbb{R}^3$, that is
$$\Phi(x,y,z) = \exp\left(-\frac{1}{2}\left(x^2+y^2+z^2+xy+xz+yz\right)\right), \qquad (x,y,z)\in\mathbb{R}^3.$$

2. As $\mathbf{X}$ is a centered Gaussian vector, the linear combination $X_1 - X_2 + 2X_3 = \langle \boldsymbol{\lambda}, \mathbf{X}\rangle$, where $\boldsymbol{\lambda} := (1,-1,2)$, is a Gaussian random variable with mean 0 and variance $\langle \boldsymbol{\lambda}, D\boldsymbol{\lambda}\rangle = 1^2 + (-1)^2 + 2^2 + 1\cdot(-1) + 1\cdot 2 + (-1)\cdot 2 = 5$.

3. Being a linear map of a Gaussian vector, $\mathbf{X}' := (X_1 + aX_2, X_1 - X_2)$ is also a Gaussian vector. Its components are independent if and only if its dispersion matrix is diagonal, that is if and only if
$$\mathbb{E}[(X_1+aX_2)(X_1-X_2)] = 1 + \frac{a}{2} - \frac{1}{2} - a = \frac{1-a}{2}$$
is zero. Hence $X_1 + aX_2$ and $X_1 - X_2$ are independent if and only if $a = 1$.

4. We have $\det D = 1/2 \neq 0$, so $\mathbf{X}$ admits a density. The inverse of D is
$$D^{-1} = \frac{1}{2}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix},$$
so a density of $\mathbf{X}$ is given by
$$(2\pi)^{-3/2}\sqrt{2}\,\exp\left(-\frac{1}{4}\left(3x^2+3y^2+3z^2-2xy-2xz-2yz\right)\right), \quad (x,y,z)\in\mathbb{R}^3.$$
■

Exercise 2.5.2. Let $a > 0$, X be a $\mathcal{N}(0,1)$ random variable, and
$$Y := \begin{cases} X, & \text{if } |X| < a, \\ -X, & \text{if } |X| \geqslant a. \end{cases}$$

1. Show that Y has the $\mathcal{N}(0,1)$ distribution.
2. Express $\mathbb{E}[XY]$ in terms of the density function $f(x) := \exp(-x^2/2)/\sqrt{2\pi}$ of X.
3. Is (X,Y) a Gaussian random vector?

Solution of Exercise 2.5.2.

1. For every $t \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}(Y \leqslant t) &= \mathbb{P}(X \leqslant t, |X| < a) + \mathbb{P}(-X \leqslant t, |X| \geqslant a)\\ &= \mathbb{P}(X \leqslant t, |X| < a) + \mathbb{P}(X \leqslant t, |X| \geqslant a)\\ &= \mathbb{P}(X \leqslant t),\end{aligned}$$

where for the second equality we used that $-X$ is distributed like X. Thus Y has the $\mathcal{N}(0,1)$ distribution.

2. Clearly,

$$\begin{aligned}\mathbb{E}[XY] &= \mathbb{E}[X^2 \mathbb{1}_{\{|X|<a\}}] - \mathbb{E}[X^2 \mathbb{1}_{\{|X|\geqslant a\}}]\\ &= 2\,\mathbb{E}[X^2 \mathbb{1}_{\{|X|<a\}}] - \mathbb{E}[X^2]\\ &= 4\int_0^a x^2 f(x)\,\mathrm{d}x - 1,\end{aligned}$$

where for the last equality we used that the function $x \mapsto x^2 f(x)$ is even.

3. We have $\mathbb{P}(X+Y=0) = \mathbb{P}(|X| \geqslant a) > 0$, so $X+Y$ cannot be a Gaussian variable. In particular, (X,Y) is not a Gaussian vector. ■

Exercise 2.5.3. Let $n \geqslant 2$ and $X_1, \dots, X_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ r.v. Prove that the empirical mean and variance

$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i \quad \text{and} \quad S_n^2 := \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$$

are independent.

Hint. Let $\mathbf{X}' := (X_1 - \bar{X}_n, \dots, X_n - \bar{X}_n)$. Check that $\mathbf{X} := (\bar{X}_n, \mathbf{X}') \in \mathbb{R}^{n+1}$ is a Gaussian vector. Express its dispersion matrix using the one of $\mathbf{X}'$ and deduce that $\bar{X}_n$ and $\mathbf{X}'$ are independent.

Solution of Exercise 2.5.3. Because the vector $(X_1, \dots, X_n)$ has independent Gaussian components, it is a Gaussian vector. By linearity, $\mathbf{X} := (\bar{X}_n, X_1 - \bar{X}_n \dots, X_n - \bar{X}_n)$ is a also a Gaussian vector. We can see that

$$\mathbb{E}[\bar{X}_n^2] - \mathbb{E}[\bar{X}_n]^2 = \mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n},$$

and for each $1 \leqslant i \leqslant n$, using that the $(\bar{X}_n, X_k)$, $1 \leqslant k \leqslant n$, are identically distributed,

$$\begin{aligned}\mathbb{E}[\bar{X}_n(X_i - \bar{X}_n)] - \mathbb{E}[\bar{X}_n]\,\mathbb{E}[X_i - \bar{X}_n] &= \mathbb{E}[\bar{X}_n X_i] - \mathbb{E}[\bar{X}_n^2] - 0\\ &= \frac{1}{n}\sum_{k=1}^n \mathbb{E}[\bar{X}_n X_k] - \mathbb{E}[\bar{X}_n^2]\\ &= \mathbb{E}[\bar{X}_n^2] - \mathbb{E}[\bar{X}_n^2]\\ &= 0.\end{aligned}$$

Thus the dispersion matrix of $\mathbf{X}$ has the form

$$\left(\begin{array}{c|ccc} \sigma^2/n & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{array}\right),$$

where we have noted $D' \in \mathbb{R}^{n\times n}$ the dispersion matrix of the (centered) Gaussian vector $\mathbf{X}' := (X_1 - \bar{X}_n, \dots, X_n - \bar{X}_n)$. We find that the characteristic function of $\mathbf{X}$ is

$$\exp\left(i\lambda_0\mu - \frac{\lambda_0^2\sigma^2}{2n}\right)\exp\left(-\frac{\langle \boldsymbol{\lambda}, D'\boldsymbol{\lambda}\rangle}{2}\right), \qquad \lambda_0 \in \mathbb{R},\ \boldsymbol{\lambda} := (\lambda_1,\dots,\lambda_n) \in \mathbb{R}^n,$$

i.e, the product of the characteristic function of $\bar{X}_n$ with that of $\mathbf{X}'$. Hence $\bar{X}_n$ and $\mathbf{X}'$ are independent, and since S_n^2 is a function of $\mathbf{X}'$, so are $\bar{X}_n$ and S_n^2. ∎

Exercise 2.5.4. Let $\mathbf{X}_1, \mathbf{X}_2, \dots$ be i.i.d. random vectors in $\mathbb{R}^2$. Apply the 2-dimensional CLT in the following cases:

1. $\mathbb{P}(\mathbf{X}_1 = (-1,-1)) = \mathbb{P}(\mathbf{X}_1 = (1,1)) = 1/2$;
2. $\mathbb{P}(\mathbf{X}_1 = (1,-1)) = \mathbb{P}(\mathbf{X}_1 = (1,1)) = \mathbb{P}(\mathbf{X}_1 = (-1,-1))/2 = 1/4$.

Solution of Exercise 2.5.4.

1. Here, $\mathbf{X}_1$ is centered with dispersion matrix

$$D = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The multidimensional central limit theorem then gives

$$\frac{\mathbf{X}_1 + \cdots + \mathbf{X}_n}{\sqrt{n}} \xRightarrow[n\to\infty]{} \mathcal{N}(0, D),$$

which is also the distribution of the degenerate Gaussian vector (N, N) where N is a standard Gaussian random variable.

Remark. A Gaussian vector $\mathbf{X}$ with *singular* dispersion matrix D has no density[3].

2. Here, $\mathbb{E}[\mathbf{X}_1] = (0, -1/2)$ and the dispersion matrix is

$$D' = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 3/4 \end{pmatrix}.$$

[3]Indeed, suppose it has a density f. As there exists a non-zero vector $\boldsymbol{\lambda}$ such that $D\boldsymbol{\lambda} = 0$ and consequently, $\mathrm{Var}(\langle \boldsymbol{\lambda}, \mathbf{X}\rangle) = \langle \boldsymbol{\lambda}, D\boldsymbol{\lambda}\rangle = 0$, the vector $\mathbf{X}$ lives almost surely in the hyperplane $\boldsymbol{\lambda}^\perp$, which yet has null Lebesgue measure (we obtain the contradiction $1 = \int P_{\mathbf{X}}(\mathrm{d}\mathbf{x}) = \int_{\boldsymbol{\lambda}^\perp} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0$).

We have $\det D' = 1/2$ so D' is invertible; its inverse is

$$D'^{-1} := \begin{pmatrix} 3/2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Therefore

$$\frac{\mathbf{X}_1 + \cdots + \mathbf{X}_n + (0, n/2)}{\sqrt{n}} \underset{n\to\infty}{\Longrightarrow} \mathcal{N}(0, D'),$$

whose a density function is

$$\frac{1}{\pi\sqrt{2}} \exp\left(-\left(\frac{3}{4}x^2 - xy + y^2\right)\right), \quad (x,y) \in \mathbb{R}^2.$$

■

Exercise 2.5.5. Let $\mathbf{X}_1, \mathbf{X}_2, \dots$ be i.i.d. vectors in $\mathbb{R}^k$ having the same distribution as $\mathbf{X} := (\xi_1, \xi_1 + \xi_2, \dots, \xi_1 + \cdots + \xi_k)$, for $\xi_1, \dots, \xi_k$ i.i.d. with $\mathbb{P}(\xi_1 = 1) = \mathbb{P}(\xi_1 = -1) = 1/2$. Show that $(\mathbf{X}_1 + \cdots + \mathbf{X}_n)/\sqrt{n}$ has a limiting distribution which one will describe in terms of a density function.

Solution of Exercise 2.5.5. The covariance matrix of the (centered) vector $\mathbf{X}$ is $D := (\min(i,j))_{1\leqslant i,j\leqslant k}$: indeed, we have $\mathrm{Var}(\xi_1 + \cdots + \xi_i) = i\,\mathrm{Var}(\xi_1) = i$ for $1 \leqslant i \leqslant k$, and

$$\mathbb{E}[(\xi_1 + \cdots + \xi_i)(\xi_1 + \cdots + \xi_i + \cdots + \xi_j)] = \mathbb{E}[(\xi_1 + \cdots + \xi_i)^2] + 0 = i = \min(i,j)$$

for $1 \leqslant i < j \leqslant k$. The multidimensional central limit theorem then yields

$$\frac{\mathbf{X}_1 + \cdots + \mathbf{X}_n}{\sqrt{n}} \underset{n\to\infty}{\Longrightarrow} \mathcal{N}(0, D).$$

Further, we can see that $D = T^\top T$, where T is the plain upper triangular matrix having all its coefficients equal to 1. Thus D is invertible ($\det D = 1$); its inverse

$$D^{-1} = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & (0) & \\ & -1 & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & (0) & & \ddots & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}$$

has associated quadratic form

$$\langle \mathbf{x}, D^{-1}\mathbf{x} \rangle = \sum_{r=1}^{k} (x_r - x_{r-1})^2, \qquad \mathbf{x} := (x_1, \dots, x_k) \in \mathbb{R}^k$$

(with the convention $x_0 := 0$). We conclude that the limiting distribution $\mathcal{N}(0, D)$ is given by the density function

$$(2\pi)^{-k/2} \exp\left(-\frac{1}{2}\sum_{r=1}^{k}(x_r - x_{r-1})^2\right).$$

Remark. This is the k-dimensional marginal distribution of the standard Brownian motion at integer times $1, 2, \dots, k$.

■

Exercise 2.5.6. Let ρ be in between -1 and 1, and μ_j, σ_j^2, $j=1,2$, be given. Construct Gaussian variables X_1, X_2 with means μ_1, μ_2, variances σ_1^2, σ_2^2, and correlation ρ.

Solution of Exercise 2.5.6. Suppose constructed Y, Y_1 two i.i.d. $\mathcal{N}(0,1)$ r.v. Then $Y_2 := \rho Y_1 + \sqrt{1-\rho^2}\, Y$ is still $\mathcal{N}(0,1)$, with $\rho_{Y_1,Y_2} = \rho$. Therefore $X_j := \mu_j + \sigma_j Y_j$, $j=1,2$, are $\mathcal{N}(\mu_j, \sigma_j^2)$ r.v. with $\rho_{X_1,X_2} = \rho$.

Remark. We have (X_1, X_2) Normal with mean (μ_1, μ_2) and covariance $\begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$. ■

Exercise 2.5.7. Let (X, Y) be bivariate normal with correlation ρ and $\sigma_X^2 = \sigma_Y^2$. Show that X and $Y - \rho X$ are independent.

Solution of Exercise 2.5.7. Since (X,Y) is Gaussian, so is $(X, Y-\rho X)$. By bilinearity and definition of ρ, we have $\mathrm{Cov}(X, Y-\rho X) = \mathrm{Cov}(X,Y) - \rho\sigma_X^2 = 0$, so X and $Y - \rho X$ are independent. ■

Exercise 2.5.8. Let $\mathbf{X} := (X_1, X_2, \dots, X_n)$ be a n-dimensional centered Gaussian vector. We suppose that there exist $k \geqslant 2$ and $0 = i_0 < \cdots < i_k = n$ such that the covariance matrix Q of $\mathbf{X}$ is a block-diagonal matrix consisting of k blocks $Q_1, \dots, Q_k$, *i.e*,

$$Q = \begin{pmatrix} Q_1 & & (0) \\ & \ddots & \\ (0) & & Q_k \end{pmatrix},$$

with respective sizes $i_1 - i_0, \dots, i_k - i_{k-1}$. Show that $\mathbf{X}_j := (X_{i_{j-1}+1}, \dots, X_{i_j})$, $1 \leqslant j \leqslant k$, are independent centered Gaussian vectors with respective covariance matrices Q_j.

Solution of Exercise 2.5.8. Write $d_j := i_j - i_{j-1}$. Let $1 \leqslant j < j' \leqslant k$ and $\boldsymbol{\lambda} \in \mathbb{R}^{d_j}, \boldsymbol{\lambda}' \in \mathbb{R}^{d_{j'}}$. We must check that

$$\mathbb{E}\left[e^{\mathrm{i}\langle \boldsymbol{\lambda}, \mathbf{X}_j\rangle} e^{\mathrm{i}\langle \boldsymbol{\lambda}', \mathbf{X}_{j'}\rangle}\right] = \exp\left(-\frac{1}{2}\langle \boldsymbol{\lambda}, Q_j \boldsymbol{\lambda}\rangle\right) \exp\left(-\frac{1}{2}\langle \boldsymbol{\lambda}', Q_{j'} \boldsymbol{\lambda}'\rangle\right).$$

We have $\langle \boldsymbol{\lambda}, \mathbf{X}_j\rangle + \langle \boldsymbol{\lambda}', \mathbf{X}_{j'}\rangle = \langle \boldsymbol{\mu}, \mathbf{X}\rangle$, where

$$\mu_r = \begin{cases} \lambda_{r-i_{j-1}}, & \text{if } i_{j-1} < r \leqslant i_j, \\ \lambda'_{r-i_{j'-1}}, & \text{if } i_{j'-1} < r \leqslant i_{j'}, \\ 0, & \text{otherwise,} \end{cases}$$

from which we easily see that $\langle \boldsymbol{\mu}, Q\boldsymbol{\mu}\rangle = \langle \boldsymbol{\lambda}, Q_j \boldsymbol{\lambda}\rangle + \langle \boldsymbol{\lambda}', Q_{j'} \boldsymbol{\lambda}'\rangle$. ■

Exercise 2.5.9. Let $\mathbf{X}$ be Gaussian $\mathcal{N}(\boldsymbol{\mu}, Q)$ in $\mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$, and $\mathbf{Y} := A\mathbf{X} + \mathbf{b}$.

1. Show that $\mathbf{Y}$ is still a Gaussian vector. Give its parameters in terms of $Q, A, \boldsymbol{\mu}, \mathbf{b}$.
2. Show that $\mathbf{Y}$ is nondegenerate if and only if $\mathbf{X}$ is nondegenerate and A is invertible.
3. We suppose $\det(Q) \neq 0$. Find A and $\mathbf{b}$ such that $\mathbf{Y}$ is standard $\mathcal{N}(\mathbf{0}, I)$.

Solution of Exercise 2.5.9.

1. Since Gaussian vectors are stable by linear maps (by definition), $\mathbf{Y}$ is still Gaussian. More precisely, its characteristic function is

$$\begin{aligned}\varphi_{\mathbf{Y}}(\mathbf{u}) &\stackrel{\text{def}}{=} \mathbb{E}\left[e^{\mathrm{i}\langle \mathbf{u}, A\mathbf{X}+\mathbf{b}\rangle}\right] \\ &= e^{\mathrm{i}\langle \mathbf{u},\mathbf{b}\rangle}\,\varphi_{\mathbf{X}}(A^{\mathsf{T}}\mathbf{u}) \\ &= \exp\left(\mathrm{i}\langle \mathbf{u}, A\boldsymbol{\mu}+\mathbf{b}\rangle - \frac{1}{2}\langle \mathbf{u}, AQA^{\mathsf{T}}\mathbf{u}\rangle\right), \qquad \mathbf{u}\in\mathbb{R}^n,\end{aligned}$$

since $\varphi_{\mathbf{X}}(\mathbf{u}) = \exp(\mathrm{i}\langle \mathbf{u},\boldsymbol{\mu}\rangle - \frac{1}{2}\langle \mathbf{u}, Q\mathbf{u}\rangle)$. This shows that $\mathbf{Y}$ is $\mathcal{N}(A\boldsymbol{\mu}+\mathbf{b}, AQA^{\mathsf{T}})$.

2. $\mathbf{Y}$ nondegenerate $\iff$ AQA^{T} invertible $\iff$ $\mathbf{X}$ nondegenerate and A invertible.

3. We have $\mathbf{Y}\sim\mathcal{N}(\mathbf{0}, I) \iff AQA^{\mathsf{T}} = I$ and $\mathbf{b} = -A\boldsymbol{\mu}$. Thus we simply need to find $A = B^{-1}$ such that Q factorizes into $Q = BB^{\mathsf{T}}$. Since Q is positive-definite, this is possible and known as a Cholesky factorization of Q. For instance, if Q has matrix $\operatorname{diag}(\lambda_1,\dots,\lambda_n)$ in some orthonormal basis, then we can take "$B = \sqrt{Q}$" as having the matrix $\operatorname{diag}(\sqrt{\lambda_1},\dots,\sqrt{\lambda_n})$ in this same basis. (Alternatively, we can take for B^{T} any orthonormal basis of $\mathbb{R}^n$ w.r.t. the inner product induced by Q.) ■

Exercise 2.5.10. Let $\mathbf{Y} := (Y_1,\dots,Y_n)$ be a nondegenerate Gaussian vector with covariance matrix Q, X be some random variable with finite variance, and $\mathbf{v} := (v_1,\dots,v_n)\in\mathbb{R}^n$. Show that

$$\operatorname{Var}\left(\sum_{i=1} v_i Y_i - X\right)$$

is minimal for $\mathbf{v} = Q^{-1}\mathbf{u}$, where $\mathbf{u} := (u_1,\dots,u_n)$ is given by $u_i = \operatorname{Cov}(Y_i, X)$, $1\leqslant i\leqslant n$.

Solution of Exercise 2.5.10. Let $A := \sum_i v_i Y_i$ with $\mathbf{v}$ as above, and $B := \sum_i w_i Y_i$ for some other $\mathbf{w}\in\mathbb{R}^n$. Then by bilinearity,

$$\begin{aligned}\operatorname{Var}(B-X) &= \operatorname{Var}(B-A+A-X) \\ &= \operatorname{Var}(B-A) + \operatorname{Var}(A-X) + 2\operatorname{Cov}(B-A, A-X),\end{aligned}$$

where, since $Qv - u = 0$,

$$\operatorname{Cov}(B-A, A-X) = \sum_{j=1}^n (w_j - v_j)\operatorname{Cov}(Y_j, A-X) = \sum_{j=1}^n (w_j-v_j)((Qv)_j - u_j) = 0.$$

Thus $\operatorname{Var}(B-X) = \operatorname{Var}(B-A) + \operatorname{Var}(A-X) \geqslant \operatorname{Var}(A-X)$: $\operatorname{Var}(A-X)$ is minimal.

Remark. This is the Pythagorean theorem applied with the inner product $\operatorname{Cov}(\cdot,\cdot)$ in the space of square-integrable r.v. (up to translation by an a.s. constant variable). ■

Exercise 2.5.11. Let (X,Y) be a nondegenerate 2-dimensional centered Gaussian vector, and

$$Z := \begin{cases} X, & \text{if } X^2+Y^2<1, \\ -X, & \text{else.}\end{cases}$$

Show that Z is Gaussian, $Z\sim X$, but that (X,Y,Z) is not a Gaussian vector.

Solution of Exercise 2.5.11. For every $f : \mathbb{R} \to \mathbb{R}$ measurable,

$$\begin{aligned}
\mathbb{E}[f(Z)] &= \mathbb{E}\big[f(X)\mathbb{1}_{\{X^2+Y^2<1\}}\big] + \mathbb{E}\big[f(-X)\mathbb{1}_{\{X^2+Y^2\geqslant 1\}}\big] \\
&= \mathbb{E}\big[f(X)\mathbb{1}_{\{X^2+Y^2<1\}}\big] + \mathbb{E}\big[f(X)\mathbb{1}_{\{(-X)^2+(-Y)^2\geqslant 1\}}\big] \\
&= \mathbb{E}[f(X)],
\end{aligned}$$

where the second equality holds because $(X,Y) \sim (-X,-Y)$. Hence Z is Gaussian with $Z \sim X$. However $Z - X$ is not Gaussian because $\mathbb{P}(Z - X = 0) = \mathbb{P}(X^2 + Y^2 < 1) > 0$ (since (X,Y) is nondegenerate). In particular (X,Y,Z) cannot be a Gaussian vector. ∎

2.6 Conditional expectations

Exercise 2.6.1. Let X, Y be two independent Poisson variables with parameters $\lambda, \mu > 0$ respectively. We set $N := X + Y$.

1. Compute $\mathbb{P}(X = k \mid N = n)$ for $k, n \in \mathbb{Z}_+$.
2. Deduce $\mathbb{E}[X \mid N = n]$ for $n \in \mathbb{Z}_+$, and then $\mathbb{E}[X \mid N]$.
3. Check that $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid N]]$.

Solution of Exercise 2.6.1.

1. We have from independence of (X,Y) that N is a Poisson$(\lambda+\mu)$ r.v., and

$$\begin{aligned}
\mathbb{P}(X = k \mid N = n) &= \frac{\mathbb{P}(X = k, Y = n-k)}{\mathbb{P}(X+Y = n)} \\
&\overset{\perp\!\!\!\perp}{=} \left(\frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{\mu^{n-k}}{(n-k)!}e^{-\mu}\right) \Big/ \left(\frac{(\lambda+\mu)^n}{n!}e^{-(\lambda+\mu)}\right) \\
&= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k}.
\end{aligned}$$

   We see that conditionally on the event $\{N = n\}$, the r.v. X is Binomial(n,p) distributed, where $p := \lambda/(\lambda+\mu)$. That is to say, conditionally on the variable N, the r.v. X has the Binomial(N,p) distribution.

2. From our knowledge of the Binomial distribution we see that $\mathbb{E}[X \mid N = n] = np$, hence $\mathbb{E}[X \mid N] = Np$.

3. On the one hand, $\mathbb{E}[X] = \lambda$ since X is a Poisson(λ) r.v. On the other hand, $\mathbb{E}[\mathbb{E}[X \mid N]] = \lambda\mathbb{E}[N]/(\lambda+\mu) = \lambda$ since N is a Poisson$(\lambda+\mu)$ r.v. ∎

Exercise 2.6.2. Let X, Y be two independent exponential r.v. with parameters $\lambda, \mu > 0$ respectively. We set $T := \min(X,Y)$.

1. What is the law of T?

2. Compute $\mathbb{E}[T \mid X]$.

   Hint. Go back to definitions. (Find $\mathbb{E}[Tf(X)]$ for $f\colon \mathbb{R} \to \mathbb{R}$ measurable bounded...)

3. Compute $\mathbb{E}[X \mid T]$.

4. Check that $\mathbb{E}[\mathbb{E}[T \mid X]] = \mathbb{E}[T]$ and $\mathbb{E}[\mathbb{E}[X \mid T]] = \mathbb{E}[X]$.

Solution of Exercise 2.6.2.

1. For every $t \geqslant 0$, $\mathbb{P}(T \leqslant t) = 1 - \mathbb{P}(X > t, Y > t) \overset{\perp\!\!\!\perp}{=} 1 - e^{-\lambda t} e^{-\mu t}$, so T has the Exponential$(\lambda + \mu)$ distribution.

2. Because Y is independent of X, the conditioning on X does not affect Y and we can thus simply integrate w.r.t. the law of Y: $\mathbb{E}[f(X,Y) \mid X] = \int f(X,y)\, P_Y(\mathrm{d}y)$. We here (re)prove this general result within the notations of the given example. As an intermediate step we have, for every $x \geqslant 0$,

$$\begin{aligned}
\mathbb{E}[Y \mathbb{1}_{\{Y \leqslant x\}}] &= \int_0^x y \mu e^{-\mu y}\, \mathrm{d}y \\
&= \left[-y e^{-\mu y}\right]_{y=0}^{y=x} + \int_0^x e^{-\mu y}\, \mathrm{d}y && \text{(i.b.p.)} \\
&= \frac{1}{\mu}\big(1 - e^{-\mu x}(1 + \mu x)\big). && (\star)
\end{aligned}$$

   Let now $f\colon \mathbb{R} \to \mathbb{R}$ be measurable and bounded. Then

$$\begin{aligned}
\mathbb{E}[Tf(X)] &= \mathbb{E}[X f(X) \mathbb{1}_{\{Y > X\}} + Y f(X) \mathbb{1}_{\{Y \leqslant X\}}] \\
&\overset{\perp\!\!\!\perp}{=} \int_0^\infty \big(x \mathbb{P}(Y > x) + \mathbb{E}[Y \mathbb{1}_{\{Y \leqslant x\}}]\big) f(x)\, P_X(\mathrm{d}x) \\
&\overset{(\star)}{=} \int_0^\infty \left(x e^{-\mu x} + \frac{1}{\mu}\big(1 - e^{-\mu x}(1 + \mu x)\big)\right) f(x)\, P_X(\mathrm{d}x) \\
&= \mathbb{E}\left[\frac{1}{\mu}\big(1 - e^{-\mu X}\big) f(X)\right],
\end{aligned}$$

   where we applied Fubini's theorem in the second equality (T is integrable!). Hence $\mathbb{E}[T \mid X] = (1 - e^{-\mu X})/\mu$ a.s.

3. As another intermediate step we have, for every $y \geqslant 0$,

$$\begin{aligned}
\mathbb{E}[X \mathbb{1}_{\{y \leqslant X\}}] &= \int_y^\infty x \lambda e^{-\lambda x}\, \mathrm{d}x \\
&= \left[-x e^{-\lambda x}\right]_{x=y}^{x \to \infty} + \int_y^\infty e^{-\lambda x}\, \mathrm{d}x && \text{(i.b.p.)} \\
&= \frac{1}{\lambda}(1 + \lambda y) e^{-\lambda y}. && (\star\star)
\end{aligned}$$

Let again $f\colon \mathbb{R} \to \mathbb{R}$ be measurable and bounded. Then, as in Question 2,

$$\begin{aligned}
\mathbb{E}[Xf(T)] &= \mathbb{E}[Xf(X)\mathbb{1}_{\{Y>X\}}] + \mathbb{E}[Xf(Y)\mathbb{1}_{\{Y\leqslant X\}}] \\
&\overset{\perp\!\!\!\perp}{=} \int_0^\infty xf(x)\mathbb{P}(Y>x)\,P_X(\mathrm{d}x) + \int_0^\infty f(y)\,\mathbb{E}[X\mathbb{1}_{\{y\leqslant X\}}]\,P_Y(\mathrm{d}y) \\
&\overset{(\star\star)}{=} \int_0^\infty \left(\frac{\lambda}{\lambda+\mu}t + \frac{\mu}{\lambda(\lambda+\mu)}(1+\lambda t)\right) f(t)\,\underbrace{(\lambda+\mu)e^{-(\lambda+\mu)t}}_{\text{density of } T \text{ that we made appear}}\mathrm{d}t \\
&= \mathbb{E}\left[\left(T + \frac{\mu}{\lambda(\lambda+\mu)}\right) f(T)\right].
\end{aligned}$$

Hence $\mathbb{E}[X \mid T] = T + \frac{\mu}{\lambda(\lambda+\mu)}$ a.s.

4. We have

$$\mathbb{E}[\mathbb{E}[T \mid X]] = \frac{1}{\mu}\left(1 - \int_0^\infty e^{-\mu x}\lambda e^{-\lambda x}\mathrm{d}x\right) = \frac{1}{\mu}\left(1 - \frac{\lambda}{\lambda+\mu}\right) = \frac{1}{\lambda+\mu} = \mathbb{E}[T],$$

and

$$\mathbb{E}[\mathbb{E}[X \mid T]] = \mathbb{E}[T] + \frac{\mu}{\lambda(\lambda+\mu)} = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda(\lambda+\mu)} = \frac{1}{\lambda} = \mathbb{E}[X].$$

■

Exercise 2.6.3. Let U, V be two independent standard uniform r.v. on $(0,1)$. Compute

$$\mathbb{E}[(U-V)^+ \mid U].$$

Solution of Exercise 2.6.3. The conditional expectation is well defined since $(U-V)^+$ is a nonnegative (or integrable) r.v. The independence between V and $\sigma(U)$ allows us to integrate with respect to the law of V:

$$\mathbb{E}[(U-V)^+ \mid U] = \int_0^1 (U-v)^+\,\mathrm{d}v = \int_0^U (U-v)\,\mathrm{d}v = \frac{1}{2}U^2.$$

(We could have proceeded as in Solution of Exercise 2.6.2.2.) ■

Exercise 2.6.4. Let $X, Y \in \mathrm{L}^1(\Omega, \mathscr{A}, \mathbb{P})$.

1. Show that if $X = Y$ a.s., then $\mathbb{E}[X \mid Y] = \mathbb{E}[Y \mid X]$ a.s.

2. Conversely, show that if $\mathbb{E}[X \mid Y] = Y$ and $\mathbb{E}[Y \mid X] = X$ a.s., then $X = Y$ a.s.

   Hint. You may only consider the case $X, Y \in \mathrm{L}^2(\mathbb{P})$ (show that $\mathbb{E}[(X-Y)^2] = 0$).

Solution of Exercise 2.6.4.

1. Note that since $X = Y$ a.s., we may identify $\mathrm{L}^1(\Omega, \sigma(X), \mathbb{P})$ and $\mathrm{L}^1(\Omega, \sigma(Y), \mathbb{P})$. Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be measurable and bounded. We have $\mathbb{E}[X\varphi(Y)] = \mathbb{E}[Y\varphi(Y)]$, so $\mathbb{E}[X \mid Y] = Y$. Likewise, $\mathbb{E}[Y \mid X] = X$. Hence $\mathbb{E}[X \mid Y] = \mathbb{E}[Y \mid X]$ a.s.

2. Suppose $X, Y \in \mathrm{L}^2(\mathbb{P})$ with $\mathbb{E}[X \mid Y] = Y$ and $\mathbb{E}[Y \mid X] = X$ a.s. Then

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}[\mathbb{E}[(X-Y)^2 \mid X]] \\
&= \mathbb{E}[X^2 + \mathbb{E}[Y^2 \mid X] - 2X\underbrace{\mathbb{E}[Y \mid X]}_{X}] \\
&= \mathbb{E}[Y^2] - \mathbb{E}[X^2] \\
&= \mathbb{E}[X^2] - \mathbb{E}[Y^2] && \text{(symmetry } X \leftrightarrow Y) \\
&= 0,
\end{aligned}$$

so $X = Y$ a.s. The general case when X, Y are only L^1 is less straightforward. First, for $t \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t\}}] &= \mathbb{E}[\mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t\}} \mid X]] \\
&= \mathbb{E}[X\mathbb{1}_{\{X \geqslant t\}} - \underbrace{\mathbb{E}[Y \mid X]}_{X}\mathbb{1}_{\{X \geqslant t\}}] \\
&= 0.
\end{aligned}$$

But $\mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t\}}] = \mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t, Y \geqslant t\}}] + \mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t > Y\}}]$, so

$$\mathbb{E}[(Y-X)\mathbb{1}_{\{X \geqslant t, Y \geqslant t\}}] = \mathbb{E}[(X-Y)\mathbb{1}_{\{X \geqslant t > Y\}}] \geqslant 0.$$

Exchanging the roles of X and Y shows that the last two expectations above are in fact 0. Since $X - Y > 0$ on $\{X \geqslant t > Y\}$, we have in particular $\mathbb{P}(X \geqslant t > Y) = 0$. Finally,

$$\begin{aligned}
\mathbb{P}(X > Y) &= \mathbb{P}\left(\bigcup_{t \in \mathbb{Q}} \{X \geqslant t > Y\}\right) \\
&\leqslant \sum_{t \in \mathbb{Q}} \mathbb{P}(X \geqslant t > Y) \\
&= 0,
\end{aligned}$$

and by symmetry, $\mathbb{P}(X < Y) = 0$. Hence $X = Y$ a.s. ■

Exercise 2.6.5. Let $\mathbf{X} := (X_1, \ldots, X_d)$ be a $\mathcal{N}(0, \Gamma)$ centered Gaussian vector in $\mathbb{R}^d$. Compute $\mathbb{E}[\langle \boldsymbol{\lambda}, \mathbf{X} \rangle \mid \langle \boldsymbol{\mu}, \mathbf{X} \rangle]$ for $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^d$ (with $\langle \cdot, \cdot \rangle$ the usual inner product in $\mathbb{R}^d$).

Solution of Exercise 2.6.5. Let $N := \langle \boldsymbol{\lambda}, \mathbf{X} \rangle$ and $N' := \langle \boldsymbol{\mu}, \mathbf{r}X \rangle$. Since N is L^2 (it is a Gaussian r.v.), the conditional expectation $Y := \mathbb{E}[N \mid N']$ coincides with the orthogonal projection of N onto $\mathrm{L}^2(\Omega, \sigma(N'), \mathbb{P})$. But since (N, N') is a Gaussian vector (because linear combinations of N and N' are also linear combinations of $X_1, \ldots, X_d$), this orthogonal projection reduces to that onto the line $\mathbb{R}N' \subset \mathrm{L}^2(\Omega, \sigma(N'), \mathbb{P})$. That is, $Y = cN'$ where $c \in \mathbb{R}$ is such that $\mathbb{E}[(N - Y)N'] = 0$, namely $c = \mathrm{Cov}(N, N')/\mathrm{Var}(N')$. In terms of the covariance matrix Γ, we conclude that

$$\mathbb{E}[\langle \boldsymbol{\lambda}, \mathbf{X} \rangle \mid \langle \boldsymbol{\mu}, \mathbf{X} \rangle] = \frac{\langle \boldsymbol{\lambda}, \boldsymbol{\mu}\Gamma \rangle}{\langle \mathbf{m}u, \boldsymbol{\mu}\Gamma \rangle} \langle \boldsymbol{\mu}, \mathbf{X} \rangle$$

(where the right-hand side is understood to be 0 when $\boldsymbol{\mu}\Gamma = 0$). ■

Exercise 2.6.6. Suppose $(B_n) \in \mathscr{A}^{\mathbb{N}}$ is a partition of Ω (that is, $\Omega = \bigcup_{n\geqslant 1} B_n$ with $B_n \neq \emptyset$ and $B_n \cap B_m = \emptyset$ for $n \neq m$), and let $\mathscr{B} := \sigma(B_n\colon\ n \geqslant 1)$. Show that for every $X \in \mathrm{L}^1(\Omega, \mathscr{A}, \mathbb{P})$,

$$\mathbb{E}[X \mid \mathscr{B}] = \sum_{n=1}^{\infty} \mathbb{E}[X \mid B_n]\, \mathbb{1}_{B_n}.$$

Solution of Exercise 2.6.6. We first recall the well-known fact that

$$\mathscr{B} = \left\{ \bigcup_{n\in S} B_n\colon\ S \subseteq \mathbb{N} \right\}$$

(indeed, the right-hand side is a σ-algebra included in $\mathscr{B}$, and contains all B_n's). It is then clear that the $\mathscr{B}$-measurable r.v. $\mathbb{E}[X \mid \mathscr{B}]$ can be written as

$$\mathbb{E}[X \mid \mathscr{B}] = \sum_{n\geqslant 1} \beta_n \mathbb{1}_{B_n},$$

with $\beta_n \in \mathbb{R}$ to be determined. Since $\mathbb{E}[\mathbb{E}[X \mid \mathscr{B}]\mathbb{1}_{B_n}] = \mathbb{E}[X\mathbb{1}_{B_n}]$ must be fulfilled for all $n \geqslant 1$, this entails $\beta_n \mathbb{P}(B_n) = \mathbb{E}[X\mathbb{1}_{B_n}]$, *i.e*, $\beta_n = \mathbb{E}[X \mid B_n]$. ∎

Exercise 2.6.7. Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space, $\mathscr{B} \subseteq \mathscr{A}$ be a sub-σ-field, and $A \in \mathscr{A}$ be an event. Show that the event $B := \{\mathbb{P}(A \mid \mathscr{B}) > 0\}$ contains a.s. A (that is, $\mathbb{P}(A \setminus B) = 0$).

Solution of Exercise 2.6.7. By definition of $B \in \mathscr{B}$,

$$0 = \mathbb{E}[\mathbb{P}(A \mid \mathscr{B})\, \mathbb{1}_{\Omega\setminus B}] = \mathbb{E}[\mathbb{P}(A \setminus B \mid \mathscr{B})] = \mathbb{P}(A \setminus B).$$ ∎

Exercise 2.6.8. Let $X \in \mathrm{L}^2(\Omega, \mathscr{A}, \mathbb{P})$ and $\mathscr{B} \subseteq \mathscr{A}$ a sub-σ-field. We define the conditional variance of X w.r.t. $\mathscr{B}$ by:

$$\mathrm{Var}(X \mid \mathscr{B}) := \mathbb{E}\big[(X - \mathbb{E}[X \mid \mathscr{B}])^2 \,\big|\, \mathscr{B}\big].$$

Prove the *law of total variance*:

$$\mathrm{Var}(X) = \mathbb{E}\big[\mathrm{Var}(X \mid \mathscr{B})\big] + \mathrm{Var}\big(\mathbb{E}[X \mid \mathscr{B}]\big).$$

Solution of Exercise 2.6.8. On the one hand, $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. On the other hand,

$$\begin{aligned}\mathbb{E}\big[\mathrm{Var}(X \mid \mathscr{B})\big] &= \mathbb{E}\big[X^2 - 2X\,\mathbb{E}[X \mid \mathscr{B}] + \mathbb{E}[X \mid \mathscr{B}]^2\big] \\ &= \mathbb{E}[X^2] - \mathbb{E}\big[X\,\mathbb{E}[X \mid \mathscr{B}]\big],\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}\big(\mathbb{E}[X \mid \mathscr{B}]\big) &= \mathbb{E}\Big[\big(\mathbb{E}[X \mid \mathscr{B}] - \mathbb{E}[X]\big)^2\Big] \\ &= \mathbb{E}\big[X\,\mathbb{E}[X \mid \mathscr{B}]\big] - \mathbb{E}[X]^2.\end{aligned}$$ ∎

Exercise 2.6.9. Let $X_1, X_2, \ldots$ be i.i.d. r.v. in $\mathrm{L}^1(\mathbb{P})$, and $S_n := X_1 + \cdots + X_n$, $n \geqslant 1$.

1. Find $\mathbb{E}[X_1 \mid X_2]$, $\mathbb{E}[S_n \mid X_1]$, and $\mathbb{E}[S_n \mid S_{n-1}]$.

2. Show that if (X, Z) and (Y, Z) have the same joint law, then for every $f : \mathbb{R} \to \mathbb{R}$ with $f(X) \in \mathrm{L}^1(\mathbb{P})$, we have $\mathbb{E}[f(X) \mid Z] = \mathbb{E}[f(Y) \mid Z]$. Deduce $\mathbb{E}[X_1 \mid S_n]$.

Solution of Exercise 2.6.9.

1. Note that $S_n \in \mathrm{L}^1(\mathbb{P})$ (because $\mathrm{L}^1(\mathbb{P})$ is a vector space). Since the X_i's are i.i.d., we have $\mathbb{E}[X_1 \mid X_2] = \mathbb{E}[X_1]$, $\mathbb{E}[S_n \mid X_1] = X_1 + (n-1)\mathbb{E}[X_1]$, and $\mathbb{E}[S_n \mid S_{n-1}] = S_{n-1} + \mathbb{E}[X_1]$.

2. Let $f : \mathbb{R} \to \mathbb{R}$ with $f(X) \in \mathrm{L}^1(\mathbb{P})$. There exists $g : \mathbb{R} \to \mathbb{R}$ measurable such that $\mathbb{E}[f(X) \mid Z] = g(Z)$. Now if (X, Z) and (Y, Z) have the same joint law, then for every $h : \mathbb{R} \to \mathbb{R}$ bounded measurable,

$$\mathbb{E}[f(Y)h(Z)] = \mathbb{E}[f(X)h(Z)] = \mathbb{E}[g(Z)h(Z)].$$

This shows that $\mathbb{E}[f(Y) \mid Z] = g(Z)$ a.s. Consequently, $(X_1, S_n), \dots, (X_n, S_n)$ having all the same (joint) distribution, we have (using linearity of the conditional expectation)

$$\mathbb{E}[X_1 \mid S_n] = \frac{1}{n}\mathbb{E}[X_1 + \cdots + X_n \mid S_n] = \frac{S_n}{n}.$$

■

Exercise 2.6.10. Let $p \in (0, 1]$, let X_n, $n \in \mathbb{N}$, be a Binomial(n, p) r.v., and, given X_n, let Y_n have a Poisson(X_n) distribution.

1. Compute the mean m_n, the variance σ_n^2, and the characteristic function Φ_n of Y_n.

2. Show that
$$\frac{Y_n - m_n}{\sigma_n} \xrightarrow[n\to\infty]{(d)} Z,$$
where $Z \sim \mathcal{N}(0, 1)$. Is there a link with the central limit theorem?

Solution of Exercise 2.6.10.

1. We have $m_n := \mathbb{E}[Y_n] = \mathbb{E}[\mathbb{E}[Y_n \mid X_n]] = \mathbb{E}[X_n] = np$ and, by Exercise 2.6.8,

$$\begin{aligned}
\sigma_n^2 &= \mathbb{E}\big[\mathrm{Var}(Y_n \mid X_n)\big] + \mathrm{Var}\big(\mathbb{E}[Y_n \mid X_n]\big) \\
&= \mathbb{E}[X_n] + \mathrm{Var}(X_n) \\
&= np + np(1-p) \\
&= np(2-p).
\end{aligned}$$

Finally,
$$\Phi_n(t) = \mathbb{E}[\mathbb{E}[e^{\mathrm{i}tY_n} \mid X_n]] = \mathbb{E}[e^{X_n(e^{\mathrm{i}t}-1)}] = \big(1 - p + p\exp(e^{\mathrm{i}t} - 1)\big)^n.$$

2. Successively,

$$
\begin{aligned}
e^{\frac{\mathrm{i}t}{\sigma_n}}-1 &= \frac{\mathrm{i}t}{\sigma_n} - \frac{t^2}{2\sigma_n^2} + o\left(\frac{1}{n}\right),\\
1-\exp\left(e^{\frac{\mathrm{i}t}{\sigma_n}}-1\right) &= -\frac{\mathrm{i}t}{\sigma_n} + \frac{t^2}{\sigma_n^2} + o\left(\frac{1}{n}\right),\\
1-p+p\exp\left(e^{\frac{\mathrm{i}t}{\sigma_n}}-1\right) &= 1 + \frac{\mathrm{i}tp}{\sigma_n} - \frac{t^2 p}{\sigma_n^2} + o\left(\frac{1}{n}\right),\\
n\log\left(1-p+p\exp\left(e^{\frac{\mathrm{i}t}{\sigma_n}}-1\right)\right) &= \frac{\mathrm{i}tnp}{\sigma_n} - \frac{t^2 \cancel{np(2-p)}}{2\cancel{\sigma_n^2}} + o(1),
\end{aligned}
$$

and thus

$$
\mathbb{E}\left[e^{\mathrm{i}t\frac{Y_n-m_n}{\sigma_n}}\right] = e^{-\frac{\mathrm{i}tnp}{\sigma_n}}\Phi_n\left(\frac{t}{\sigma_n}\right) \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}},
$$

which is the characteristic function of the standard Normal distribution. We could have concluded directly by applying the CLT (which we somehow reproved): the expression of Φ_n shows that Y_n is distributed like the sum of n i.i.d. r.v. with the same law as Y_1. ∎

Exercise 2.6.11. Let U be a uniformly distributed r.v. on $[0,1)$ and let $X_n := \lfloor nU \rfloor$ for $n \geqslant 1$. Determine the conditional law of U given X_n.

Solution of Exercise 2.6.11. For all $0 \leqslant k < n$ and $t \geqslant 0$,

$$
\mathbb{P}(U \leqslant t, X_n = k) = \begin{cases} 0, & \text{if } nt < k,\\ t - \frac{k}{n}, & \text{if } k \leqslant nt < k+1,\\ \frac{1}{n}, & \text{if } nt > k+1. \end{cases}
$$

Thus, conditionally on X_n, the r.v. U is uniformly distributed on $[\frac{X_n}{n}, \frac{X_n+1}{n})$. That is, for every bounded measurable function $f: \mathbb{R} \to \mathbb{R}$,

$$
\mathbb{E}[f(U) \mid X_n] = \int_{X_n}^{X_n+1} f\left(\frac{y}{n}\right)\mathrm{d}y.
$$

∎

Exercise 2.6.12. Let (X,Y) be a random vector in $\mathbb{R}^{n+m}$ with probability density function (p.d.f.) p.

1. Show that $Y \in \mathbb{R}^m$ admits a p.d.f. q and give its expression in terms of p.

2. For each $y \in \mathbb{R}^m$, we let $\nu(y,\cdot)$ denote the measure on $\mathbb{R}^n$ given by

$$
\nu(y,A) := \frac{1}{q(y)}\int_A p(x,y)\,\mathrm{d}x, \qquad A \in \mathscr{B}(\mathbb{R}^n)
$$

(with the convention $\nu(y,A) = 0$ if $q(y) = 0$). Prove that for every bounded measurable function $f: \mathbb{R}^{n+m} \to \mathbb{R}$,

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

Solution of Exercise 2.6.12.

1. By Fubini–Tonelli's theorem, the nonnegative function

$$y \in \mathbb{R}^m \mapsto q(y) := \int_{\mathbb{R}^n} p(x,y)\,\mathrm{d}x$$

is measurable, with integral 1, and for every $f\colon \mathbb{R}^m \to \mathbb{R}$ bounded measurable,

$$\mathbb{E}[f(Y)] = \iint_{\mathbb{R}^n\times\mathbb{R}^m} f(y)\,p(x,y)\mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}^n} f(y)\,q(y)\mathrm{d}y.$$

Thus Y admits q as probability density function.

2. We see from the definition of ν that $q(y)\,\nu(y,\mathrm{d}x) = p(x,y)\,\mathrm{d}x$. Fix $f\colon \mathbb{R}^{n+m} \to \mathbb{R}$ bounded measurable. Define

$$g(y) := \int_{\mathbb{R}^n} f(x,y)\,\nu(y,\mathrm{d}x), \qquad y \in \mathbb{R}^m.$$

Then, for every bounded measurable function $h\colon \mathbb{R}\to\mathbb{R}$,

$$\begin{aligned}
\mathbb{E}[g(Y)h(Y)] &= \int_{\mathbb{R}^m} g(y)h(y)\,q(y)\mathrm{d}y \\
&= \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)\,\nu(y,\mathrm{d}x)\right)h(y)\,q(y)\mathrm{d}y \\
&= \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)\,p(x,y)\mathrm{d}x\right)h(y)\,\mathrm{d}y \\
&= \iint_{\mathbb{R}^n\times\mathbb{R}^m} f(x,y)h(y)\,p(x,y)\mathrm{d}x\mathrm{d}y \qquad \text{(Fubini–Lebesgue)} \\
&= \mathbb{E}[f(X,Y)h(Y)],
\end{aligned}$$

Hence $\mathbb{E}[f(X,Y)\mid Y] = g(Y)$. ■

Exercise 2.6.13. Let $\{X_n\}_{n\geqslant 0} \subset \mathrm{L}^2(\mathbb{P})$ such that $S_n := X_1 + \cdots + X_n$, $n \geqslant 0$, defines a martingale. Show that $\mathbb{E}[X_iX_j] = 0$ for all $i \neq j$.

Solution of Exercise 2.6.13. Let $i<j$. Then

$$\mathbb{E}[X_iX_j] = \mathbb{E}\big[\mathbb{E}[X_iX_j \mid S_0,\ldots,S_i]\big] = \mathbb{E}\big[X_i\underbrace{\mathbb{E}[S_j - S_{j-1}\mid S_0,\ldots,S_i]}_{=0}\big] = 0.$$

■

2.7 Martingales

Exercise 2.7.1. Let $(X_n)_{n\geqslant 0}$ be a martingale and T a stopping time. Recall that $(X_{n\wedge T})_{n\geqslant 0}$ is again a martingale and that (optional stopping theorem) if $\boldsymbol{T} \in \mathrm{L}^\infty$, then

$$X_T \in \mathrm{L}^1, \quad \text{with } \mathbb{E}[X_T] = \mathbb{E}[X_0]. \tag{$\star$}$$

Show that ($\star$) also holds in the other two following cases:

1. When $T < \infty$ a.s. and $(X_n)_{n\geqslant 0}$ is dominated by some r.v. in L^1.
2. When $T \in \mathrm{L}^1$ and $(X_{n+1} - X_n)_{n\geqslant 0}$ is bounded in L^∞.

Hint. Use the dominated convergence theorem.

Solution of Exercise 2.7.1.

1. Since $n \wedge T$ always tends to T in $\mathbb{Z}_+ \cup \{\infty\}$ as $n \to \infty$ and $T < \infty$ a.s., we see that $X_{n\wedge T}$ is a.s. eventually equal to X_T (consequently $X_{n\wedge T} \to X_T$ a.s.). Further, $(X_{n\wedge T})_{n\geqslant 0}$ is a martingale, so

$$\mathbb{E}[X_{n\wedge T}] = \mathbb{E}[X_0], \tag{2.3}$$

   The hypothesis of domination says that there is $Y \in \mathrm{L}^1$ such that a.s., $|X_n| \leqslant Y$ for all n. In particular, a.s., $|X_{n\wedge T}| \leqslant Y$ for all n. The conclusion follows by applying the dominated convergence theorem.

2. Recall that $T \in \mathrm{L}^1$ implies $T < \infty$ a.s., so like in Question 1, $X_{n\wedge T} \to X_T$ as $n \to \infty$. Now, there is a constant $c > 0$ such that a.s., $|X_{n+1} - X_n| \leqslant c$ for all n, so

$$|X_{n\wedge T}| \leqslant |X_0| + \sum_{k=1}^{n\wedge T} |X_{k\wedge T} - X_{(k-1)\wedge T}| \leqslant |X_0| + cT \in \mathrm{L}^1.$$

   The dominated convergence theorem entails that $X_T \in \mathrm{L}^1$ and allows us again to take the limit in (2.3).

Remark. These results will be generalized later (stopping theorem for UI martingales). ∎

Exercise 2.7.2 (Pig). Let D_i, $i \geqslant 1$, be i.i.d. realizations of a fair 6-faced die roll. We define

$$T := \inf\{i \geqslant 1 \colon D_i = 1\},$$
$$\mathscr{F}_n := \sigma(D_1, \ldots, D_n), \qquad n \geqslant 0,$$

and

$$S_n := \sum_{i=1}^{n} D_i, \qquad n \geqslant 0.$$

1. Check that T is a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time. Compute $\mathbb{E}[T]$.
2. Show that

$$\mathbb{E}[S_n \mid T] = 4n\mathbb{1}_{\{T>n\}} + \left(\frac{7n+T}{2} - 3\right)\mathbb{1}_{\{T\leqslant n\}}.$$

3. Deduce that $\mathbb{E}[S_T] = 21$.

4. Provide an alternative solution to Question 3 using a martingale.

   Hint. Determine $m \in \mathbb{R}$ such that $(S_n - mn)_{n\geqslant 0}$ is a $(\mathscr{F}_n)_{n\geqslant 0}$-martingale.

Solution of Exercise 2.7.2.

1. As the entrance time in the Borel set $\{1\}$ of the $(\mathscr{F}_n)_{n\geqslant 0}$-adapted process $(D_n)_{n\geqslant 0}$, T is a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time. Because T has the geometric distribution on $\mathbb{N}$ with success probability $\frac{1}{6}$, we have $\mathbb{E}[T] = 6$.

2. Let $k \in \mathbb{N}$ be any possible value for T. Since $\{T = k\} = \{D_1 \neq 1, \ldots, D_{k-1} \neq 1, D_k = 1\}$, we easily observe that, conditionally on $\{T = k\}$:

   - the D_i, $i < k$, are independent uniform r.v. on $\{2, \ldots, 6\}$ (each with mean 4);
   - the D_i, $i > k$, are independent uniform r.v. on $\{1, \ldots, 6\}$ (each with mean $\frac{7}{2}$);
   - and (of course) $D_k = 1$.

   Therefore, for all $k \geqslant 1$ and $i \geqslant 0$,

$$\mathbb{E}[D_i \mid T = k] = \begin{cases} 4, & \text{if } i < k, \\ \frac{7}{2}, & \text{if } i > k, \\ 1, & \text{if } i = k. \end{cases}$$

   Then, by linearity of expectation,

$$\begin{aligned} \mathbb{E}[S_n \mid T = k] &= 4(k-1) \wedge n + \frac{7}{2}(n-k)_+ + \mathbb{1}_{\{k \leqslant n\}} \\ &= 4n\mathbb{1}_{\{k>n\}} + \left(\frac{7n+k}{2} - 3\right)\mathbb{1}_{\{k\leqslant n\}}. \end{aligned}$$

3. It follows that
$$\mathbb{E}[S_T] = \mathbb{E}[\mathbb{E}[S_T \mid T]] = \mathbb{E}[4T - 3] = 4\mathbb{E}[T] - 3 = 21.$$

4. Let $m := \mathbb{E}[D_1] = \frac{7}{2}$. Then $S_n - mn$, $n \geqslant 0$, is the partial sum of independent, integrable, and centered $\mathscr{F}_n$-measurable variables, so $(S_n - mn)_{n\geqslant 0}$ is a $(\mathscr{F}_n)$-martingale. Because T is a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time, we deduce from the stopping theorem that the process

$$M_n := S_{n\wedge T} - m(n \wedge T),\ n \geqslant 0,$$

   is also a martingale. In particular $\mathbb{E}[M_n] = \mathbb{E}[M_0] = 0$, that is $\mathbb{E}[S_{n\wedge T}] = m\mathbb{E}[n \wedge T]$. Letting now $n \to \infty$ yields (by monotone convergence) $\mathbb{E}[S_T] = m\mathbb{E}[T] = \frac{7}{2} \cdot 6 = 21$. ∎

Exercise 2.7.3 (Pokémon Go). Imagine that at each time $n = 1, 2, \ldots$, you find one of the m Pokémon™, assuming they all appear independently and uniformly at random. Let $R_0 := m$ and R_n be the number of different Pokémon you still need to capture after time n in order to complete your Pokédex.

1. Justify that conditionally on R_n, the r.v. $R_n - R_{n+1}$ is Bernoulli(R_n/m) distributed.

2. Let $h(k) := \sum_{1\leqslant i\leqslant k} 1/i$ for every $k \geqslant 0$. Deduce from Question 1 that, respectively,

$$M_n := \left(\frac{m}{m-1}\right)^n R_n, \quad \text{and} \quad L_n := \frac{n}{m} + h(R_n), \qquad n \geqslant 0,$$

define a martingale and a submartingale w.r.t. the natural filtration $(\mathscr{F}_n)_{n\geqslant 0}$.

3. Let $T := \inf\{n \geqslant 0 \colon R(n) = 0\}$ be the time you catch them all.

   a) Check that T is a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time.

   b) Show that $X_n := L_{n\wedge T}$, $n \geqslant 0$, becomes a martingale and deduce that $T \in \mathrm{L}^1$.

   c) Deduce $\mathbb{E}[T]$ (apply Exercise 2.7.1). Give an equivalent when m is large.

Solution of Exercise 2.7.3.

1. Clearly, $R_n - R_{n+1} \in \{0,1\}$ and is 1 if and only if we find a new Pokémon at time $n+1$, which conditionally on $R_n = k \in \{0,\dots,m\}$ happens with probability k/m. That is, conditionally on R_n, the r.v. $R_n - R_{n+1}$ is Bernoulli(R_n/m) distributed.

2. In particular $\mathbb{E}[R_n - R_{n+1} \mid R_n] = R_n/m$, and thus $\mathbb{E}[R_{n+1} \mid R_n] = (1-1/m)R_n$. We then readily have the martingale property for $(1-1/m)^{-n}R_n = M_n$. Next, $(L_n)_{n\geqslant 0}$ is clearly $(\mathscr{F}_n)_{n\geqslant 0}$-adapted, in L^1, and $L_{n+1} - L_n = 1/m - (R_n - R_{n+1})/R_n$. Hence

$$\mathbb{E}[L_{n+1} - L_n \mid \mathscr{F}_n] = \frac{1}{m}\mathbb{1}_{\{R_n=0\}} \geqslant 0. \tag{2.4}$$

3. a) Clear, as hitting time of the Borel set $\{0\}$ by the adapted process $(R_n)_{n\geqslant 0}$. More straightforwardly: $\{T = n\} = \{R_0 \neq 0, \dots, R_{n-1} \neq 0, R_n = 0\} \in \mathscr{F}_n$.

   Remark. Since $(R_n)_{n\geqslant 0}$ is non-increasing in $\mathbb{N}$ and the nonnegative martingale $(M_n)_{n\geqslant 0}$ must converge a.s., we already see that $R_n = 0$ eventually, so $T < \infty$ a.s.

   b) Since $\{T > n\} = \{R_n \neq 0\} \in \mathscr{F}_n$, we see by (2.4) that $(X_n)_{n\geqslant 0}$ is a martingale:

$$E[X_{n+1} - X_n \mid \mathscr{F}_n] = \mathbb{1}_{\{T>n\}}\mathbb{E}[L_{n+1} - L_n \mid \mathscr{F}_n] = 0.$$

   Then

$$h(m) = \mathbb{E}[X_0] = \mathbb{E}[X_{n\wedge T}] = \frac{1}{m}\mathbb{E}[n\wedge T] + \mathbb{E}[h(R_{n\wedge T})] \geqslant \frac{1}{m}\mathbb{E}[n\wedge T],$$

   where $n \wedge T \uparrow T$, so we deduce from Fatou's lemma that $\mathbb{E}[T] \leqslant m\,h(m) < \infty$.

   c) It is plain that $|X_n| \leqslant T/m + h(m) \in \mathrm{L}^1$ or that $(L_{n+1} - L_n)_{n\geqslant 0}$ is bounded (by 2) in L^∞. We may thus apply either of the two criteria of Exercise 2.7.1: then $X_T = L_T \in \mathrm{L}^1$, and $h(m) = \mathbb{E}[L_0] = \mathbb{E}[L_T] = \mathbb{E}[T/m]$ (note that $h(R_T) = 0$ because $R_T = 0$, a.s.). Hence $\mathbb{E}[T] = m\,h(m) \sim m \log m$ as $m \to \infty$.

Remark. This problem is known as the Coupon collector's problem. ∎

Exercise 2.7.4. Let $\theta \in \mathbb{R}$, $(X_k)_{k\in\mathbb{N}}$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ r.v., and

$$S_n := \sum_{k=1}^{n} X_k, \qquad n \geqslant 0.$$

1. Find $f : \mathbb{R} \to \mathbb{R}$ such that $M_n^{(\theta)} := \exp(\theta S_n - n f(\theta))$, $n \geqslant 0$, is a martingale.
2. Does $M_n^{(\theta)}$ converge as $n \to \infty$, almost surely? in L^1?

Solution of Exercise 2.7.4.

1. Let $\mathscr{F}_n := \sigma(X_1,\dots,X_n)$, $n \geqslant 0$. Clearly, $e^{\theta S_n}$ is $\mathscr{F}_n$-mesurable, and since by independence S_n is $\mathcal{N}(0,n)$-distributed we have $\mathbb{E}[e^{\theta S_n}] = e^{n\theta^2/2} < \infty$. Finally

$$\mathbb{E}[\exp(\theta(S_{n+1} - S_n)) \mid \mathscr{F}_n] = \mathbb{E}[\exp(\theta S_1)] = e^{\theta^2/2},$$

so that for $f(\theta) := \theta^2/2$, the process $(M_n^{(\theta)})_{n\geqslant 0}$ is a $(\mathscr{F}_n)_{n\geqslant 0}$-martingale.

2. On the one hand, the convergence theorem for nonnegative (super)martingales tells us that M_n converges almost surely as $n \to \infty$ to some r.v. $M_\infty^{(\theta)}$. On the other hand, we know from the law of large numbers that $S_n = o(n)$, so if $\theta \neq 0$ we have

$$\log M_n = n(\theta S_n/n - f(\theta)) \xrightarrow[n\to\infty]{} -\infty,$$

hence $M_\infty^{(\theta)} = 0$ a.s. Since $\mathbb{E}[M_n^{(\theta)}] = 1$ for every n, we conclude that, unless $\theta = 0$ (in which case $M^{(\theta)} \equiv 1$), $M_n^{(\theta)}$ cannot converge in L^1. ∎

Exercise 2.7.5. Let $E := \{\mathbf{A}, \mathbf{B}\}$ be a set with two elements, $m \in \mathbb{N}$, and consider an initial population $X_0 \in E^m$ of m individuals, each of which has either type $\mathbf{A}$ or type $\mathbf{B}$. Suppose that at each time $n = 1, 2, \dots$, a new population X_n is born in such a way that each individual inherits the type of one individual in the previous generation X_{n-1}, which is chosen independently and uniformly at random. Formally

$$X_n = (X_{n-1}(\sigma_{n,1}),\dots,X_{n-1}(\sigma_{n,m})) \in E^m,$$

with $(\sigma_{n,i})_{n\in\mathbb{N},1\leqslant i\leqslant m}$ an independent family of i.i.d. uniform r.v. on $\{1,\dots,m\}$.

1. What do you think will eventually happen to the population?

Let A_n, $n \geqslant 0$, denote the number of individuals in X_n which have type $\mathbf{A}$.

2. Justify that conditionally on A_n, the r.v. A_{n+1} is Binomial$(m, A_n/m)$ distributed.
3. Show that $(A_n)_{n\geqslant 0}$ is a martingale converging a.s.
4. Check that $\mathbb{E}[A_{n+1}^2 \mid A_n] = \frac{m-1}{m} A_n^2 + A_n$. Deduce that

$$M_n := (m-1)(m - A_n) + (mA_n - A_n^2)\left(\frac{m}{m-1}\right)^n, \qquad n \geqslant 0,$$

defines another martingale.

5. Prove your conjecture in Question 1.

Solution of Exercise 2.7.5.

1. We conjecture that after some time, all individuals are of the same type, *i.e*, there exists n_0 such that $X_n = X_{n_0} \in \{\mathbf{A}\}^m \cup \{\mathbf{B}\}^m$ for all $n \geqslant n_0$.

2. We see that A_{n+1} is the number of balls of type **A** we would obtain by drawing with replacement m indistinguishable balls from an urn containing A_n balls labeled by **A** and $m - A_n$ balls labeled by **B**. Thus, conditionally on A_n, the r.v. A_{n+1} is Binomial$(m, A_n/m)$ distributed.

3. In particular $\mathbb{E}[A_{n+1} \mid A_n] = m \cdot A_n/m = A_n$. Hence $(A_n)_{n\geqslant 0}$ is a nonnegative martingale. By the martingale convergence theorem, its limit A_∞ as $n \to \infty$ then exists almost surely.

4. Under $\mathbb{P}(\cdot \mid A_n)$ we have $\mathrm{Var}(A_{n+1} \mid A_n) = m \cdot (A_n/m) \cdot (1 - A_n/m)$ (variance of a Binomial$(m, A_n/m)$ r.v.), so
$$\mathbb{E}[A_{n+1}^2 \mid A_n] = \mathrm{Var}(A_{n+1} \mid A_n) + \mathbb{E}[A_{n+1} \mid A_n]^2 = \frac{m-1}{m} A_n^2 + A_n.$$
For $\mathscr{F}_n := \sigma(A_0, \dots, A_n)$, it is then plain that M_n is $\mathscr{F}_n$-measurable, integrable, and that
$$\mathbb{E}[M_{n+1} \mid \mathscr{F}_n] = (m-1)(m - A_n) + (mA_n - \tfrac{m-1}{m} A_n^2 - A_n)\left(\frac{m}{m-1}\right)^{n+1} = M_n.$$

5. Because $0 \leqslant A_n \leqslant m$, it is clear that M_n is further nonnegative. By the martingale convergence theorem, M_n then converges almost surely. In particular
$$(m-1)(mA_\infty - A_\infty^2)\left(\frac{m}{m-1}\right)^n$$
must be a.s. bounded in n, which is possible only if $mA_\infty - A_\infty^2 = 0$ a.s., *i.e*, $A_\infty \in \{0, m\}$. But A_n is integer-valued, so we can conclude that $A_n \in \{0, m\}$ when n is sufficiently large, almost surely, which proves our conjecture.

   Remark. Applying Exercise 2.7.1 for the stopping time $T := \inf\{n \geqslant 0 \colon A_n \in \{0, m\}\}$ provides the law of A_∞: we find $\mathbb{P}(A_\infty = m) = A_0/m$ and $\mathbb{P}(A_\infty = 0) = (m - A_0)/m$. ■

Exercise 2.7.6. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent r.v. We suppose that there exists a constant $C > 0$ such that the following three (deterministic) series

$$\text{(a)} \sum_{n\in\mathbb{N}} \mathbb{P}(|X_n| > C), \qquad \text{(b)} \sum_{n\in\mathbb{N}} \mathbb{E}\big[X_n \mathbb{1}_{\{|X_n|\leqslant C\}}\big], \qquad \text{(c)} \sum_{n\in\mathbb{N}} \mathrm{Var}\big(X_n \mathbb{1}_{\{|X_n|\leqslant C\}}\big),$$

all converge (in $\mathbb{R}$). Show that the series $\sum_{n\in\mathbb{N}} X_n$ converges almost surely.

Hint. Show that $M_n := \sum_{k=1}^{n} \big(X_k \mathbb{1}_{\{|X_k|\leqslant C\}} - \mathbb{E}\big[X_k \mathbb{1}_{\{|X_k|\leqslant C\}}\big]\big)$, $n \geqslant 0$, is bounded in $\mathrm{L}^2(\mathbb{P})$.

Solution of Exercise 2.7.6. The given process $(M_n)_{n\geqslant 0}$ is a martingale (defined as a sum of independent, integrable, centered r.v.). The convergence of (c) shows that this martingale is indeed bounded in $\mathrm{L}^2(\mathbb{P})$: for every $n \geqslant 0$,

$$\mathbb{E}[M_n^2] = \mathrm{Var}(M_n) \overset{\perp\!\!\!\perp}{=} \sum_{k=1}^{n} \mathrm{Var}\big(X_k \mathbb{1}_{\{|X_k|\leqslant C\}}\big) \leqslant \sum_{k=1}^{\infty} \mathrm{Var}\big(X_k \mathbb{1}_{\{|X_k|\leqslant C\}}\big) < \infty.$$

By the L^p-convergence theorem, the martingale $(M_n)_{n\geqslant 0}$ converges a.s. (and in $\mathrm{L}^2(\mathbb{P})$). Adding the convergence of (b), we deduce the a.s. convergence of the series

$$\sum_{n\in\mathbb{N}} X_n \mathbb{1}_{\{|X_n|\leqslant C\}}. \tag{$\star$}$$

Finally, by the (first) Borel–Cantelli lemma, the convergence of (a) entails that the event $\{|X_n| \leqslant C\}$ must occur for n large enough with probability 1, implying that the convergence of $\sum_{n\in\mathbb{N}} X_n$ is a.s. equivalent to that of ($\star$). The conclusion follows.

Remark. Conversely, the three series (a), (b), and (c) converge (for any $C > 0$) if the series $\Sigma_{n\in\mathbb{N}} X_n$ converges almost surely [Kolmogorov's three-series theorem]. ■

Exercise 2.7.7 (A counterexample). Let T be a r.v. in $\mathbb{N}$ and $(Y_k)_{k\in\mathbb{N}}$ be an independent family of i.i.d. r.v. with $\mathrm{Var}(Y_1) = 1$ and $\mathbb{E}[Y_1] = 0$. We set $\mathscr{F}_n := \sigma(T, Y_1, \ldots, Y_n)$ and

$$X_n := \sum_{k=1}^{n} Y_k, \qquad n \geqslant 0.$$

1. Show that $(X_n)_{n\geqslant 0}$ is a $(\mathscr{F}_n)_{n\geqslant 0}$-martingale which is not bounded in $\mathrm{L}^1(\mathbb{P})$.

2. Give an example of distribution for T such that the $(\mathscr{F}_n)_{n\geqslant 0}$-stopped martingale $(X_{n\wedge T})_{n\in\mathbb{N}}$ is still not bounded in $\mathrm{L}^1(\mathbb{P})$ (although it converges almost surely).

Solution of Exercise 2.7.7.

1. You (should) already know that $(X_n)_{n\geqslant 0}$ is a $(\mathscr{F}_n)_{n\geqslant 0}$-martingale. Besides, the unboundedness in $\mathrm{L}^1(\mathbb{P})$ is not new and follows *e.g.* from the central limit theorem (Exercise 2.4.6). More precisely, $\sqrt{n} = O(\mathbb{E}[|X_n|])$ as $n \to \infty$.

2. By independence between T and $(X_n)_{n\in\mathbb{N}}$ (and monotone convergence theorem),

$$\begin{aligned}\mathbb{E}[|X_{n\wedge T}|] &= \sum_{k=1}^{\infty} \mathbb{P}(T = k)\,\mathbb{E}[|X_{n\wedge k}|]\\ &\geqslant \sum_{k=1}^{n} \mathbb{P}(T = k)\,\mathbb{E}[|X_k|].\end{aligned}$$

Since $\sqrt{k} = O(\mathbb{E}[|X_k|])$, the latter sum diverges if $k^{-3/2} = O(\mathbb{P}(T = k))$, $k \to \infty$. Hence, if we take for instance the probability distribution

$$\mathbb{P}(T = k) = \frac{1}{\sqrt{k}} - \frac{1}{\sqrt{k+1}}, \qquad k \in \mathbb{N},$$

then $(X_{n\wedge T})_{n\in\mathbb{N}}$ converges to X_T a.s., while it is not bounded in $\mathrm{L}^1(\mathbb{P})$. ■

Exercise 2.7.8. Let S, T be two $(\mathscr{F}_n)_{n\geqslant 0}$-stopping times and $X \in \mathrm{L}^1(\mathbb{P})$. Show that

$$\mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_S] \mid \mathscr{F}_T] = \mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_T] \mid \mathscr{F}_S] = \mathbb{E}[X \mid \mathscr{F}_{S\wedge T}].$$

Hint. Apply the stopping theorem... a few times.

Solution of Exercise 2.7.8. On the one hand, we observe that $X_n := \mathbb{E}[X \mid \mathscr{F}_n]$, $n \geqslant 0$, defines a uniformly integrable $(\mathscr{F}_n)_{n\geqslant 0}$-martingale. On the other hand, applying the stopping theorem for the stopping time $n \wedge T$, $n \in \mathbb{N}$, entails that $X_{n\wedge T} = \mathbb{E}[X \mid \mathscr{F}_{n\wedge T}]$, so the $(\mathscr{F}_n)_{n\geqslant 0}$-martingale $(X_{n\wedge T})_{n\geqslant 0}$ is also uniformly integrable. Applying again the stopping theorem, but to this stopped martingale and for the stopping time S, yields

$$X_{S\wedge T} = \mathbb{E}[X_T \mid \mathscr{F}_S].$$

Now, two other applications of the stopping theorem show that $X_T = \mathbb{E}[X \mid \mathscr{F}_T]$ and $X_{S\wedge T} = \mathbb{E}[X \mid \mathscr{F}_{S\wedge T}]$. The conclusion follows (exchanging the roles of S and T). ∎

Exercise 2.7.9 (Other counterexamples). Let $f \colon \mathbb{N} \to \mathbb{R}$ measurable and T be a $\mathbb{N}$-valued r.v. such that $f(T) \in \mathrm{L}^1(\mathbb{P})$. For every $n \geqslant 0$, define $\mathscr{F}_n := \sigma(\{T = k\}, k \leqslant n)$ and

$$X_n := \mathbb{1}_{\{T\leqslant n\}} f(T) + \mathbb{1}_{\{T>n\}} r(n), \quad \text{where} \;\; r(n) := \frac{\mathbb{E}[\mathbb{1}_{\{T>n\}} f(T)]}{\mathbb{P}(T > n)}.$$

1. Check that T is a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time and that $(X_n)_{n\geqslant 0}$ is a uniformly integrable $(\mathscr{F}_n)_{n\geqslant 0}$-martingale.

2. In this question we suppose that $f(k) = 2^k k^{-2}$ and that $\mathbb{P}(T = k) = 2^{-k}$, $k \in \mathbb{N}$.

   a) Show that $X_{T-1} \notin \mathrm{L}^1(\mathbb{P})$. *What is wrong regarding the stopping theorem?*

   b) Deduce that $(X_n)_{n\geqslant 0}$ is not dominated in $\mathrm{L}^1(\mathbb{P})$.

3. In this question we suppose that $f(k) = \log k$, $k \in \mathbb{N}$, and that, as $k \to \infty$,

$$\mathbb{P}(T = k) = \frac{1}{k^2(\log k)^2} + O\left(\frac{1}{k^2(\log k)^3}\right).$$

   a) Check that $T \in \mathrm{L}^1(\mathbb{P})$, while $T \notin \mathrm{L}^2(\mathbb{P})$.

   b) Show that $(X_{n+1} - X_n)_{n\geqslant 0}$ is bounded in $\mathrm{L}^\infty(\mathbb{P})$.

   Hint. $\displaystyle\sum_{k>n} \frac{1}{k^2(\log k)^p} = \frac{1}{n(\log n)^p} + O\left(\frac{1}{n(\log n)^{p+1}}\right)$ as $n \to \infty$, for $p \in \{1, 2\}$.

   c) Show that $\displaystyle\sum_{k=1}^{T} X_k \notin \mathrm{L}^1(\mathbb{P})$.

Solution of Exercise 2.7.9.

1. Clearly T is a stopping time. We further observe that $X_n = \mathbb{E}[f(T) \mid \mathscr{F}_n]$, $n \geqslant 0$. By the equivalence theorem, $(X_n)_{n\geqslant 0}$ is a uniformly integrable $(\mathscr{F}_n)_{n\geqslant 0}$-martingale.

2. Note that in this question, indeed, $f(T) \in \mathrm{L}^1(\mathbb{P})$: $\mathbb{E}[2^T T^{-2}] = \sum_{k\geqslant 1} k^{-2} < \infty$.

a) We have $X_{T-1} = r(T-1)$, where

$$r(n) = 2^n \sum_{k>n} \frac{1}{k^2} \sim \frac{2^n}{n}, \qquad \text{as } n \to \infty.$$

Therefore $X_{T-1} \notin \mathrm{L}^1(\mathbb{P})$:

$$\mathbb{E}[X_{T-1}] = \mathbb{E}[r(T-1)] = \frac{1}{2} \sum_{k=0}^{\infty} 2^{-k} r(k) = \infty.$$

There is of course no contradiction with the stopping theorem: we cannot apply it because $T-1$ is *not* a $(\mathscr{F}_n)_{n\geqslant 0}$-stopping time!

b) Obviously $\sup_{n\geqslant 0} X_n \geqslant X_{T-1}$, so $\mathbb{E}[\sup_{n\geqslant 0} X_n] = \infty$ by 2.a). This shows that $(X_n)_{n\geqslant 0}$ is not dominated in $\mathrm{L}^1(\mathbb{P})$.

3. a) By Bertrand's test, $(n\mathbb{P}(T = n))_{n\in\mathbb{N}}$ is summable whereas $(n^2\mathbb{P}(T = n))_{n\in\mathbb{N}}$ is not. Thus $T \in \mathrm{L}^1(\mathbb{P})$ (in particular $f(T) = \log T \in \mathrm{L}^1(\mathbb{P})$), but $T \notin \mathrm{L}^2(\mathbb{P})$.

b) By triangle inequality,

$$|X_{n+1} - X_n| \leqslant \big|\log(n+1) - r(n)\big| + \big|r(n+1) - r(n)\big|.$$

But as $n \to \infty$ (using the indications),

$$\mathbb{P}(T > n) = \sum_{k>n} \left(\frac{1}{k^2(\log k)^2} + O\left(\frac{1}{k^2(\log k)^3}\right)\right) = \frac{1}{n(\log n)^2} + O\left(\frac{1}{n(\log n)^3}\right),$$

and

$$\mathbb{E}[\mathbb{1}_{\{T>n\}} f(T)] = \sum_{k>n} \left(\frac{1}{k^2 \log k} + O\left(\frac{1}{k^2(\log k)^2}\right)\right) = \frac{1}{n\log n} + O\left(\frac{1}{n(\log n)^2}\right),$$

so $r(n) = \log n + O(1)$. It follows that $(X_{n+1} - X_n)_{n\geqslant 0}$ is bounded in $\mathrm{L}^\infty(\mathbb{P})$.

c) We deduce from the monotone convergence theorem, our computations in 3.b) and Bertrand's test that

$$\mathbb{E}\left[\sum_{k=1}^{T} X_k\right] = \sum_{k=1}^{\infty} \mathbb{E}\big[X_k \mathbb{1}_{\{T\geqslant k\}}\big] \geqslant \sum_{k=1}^{\infty} r(k)\mathbb{P}(T > k) = \infty.$$

∎

Exercise 2.7.10. On the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbb{P})$, let $(X_n)_{n\geqslant 0}$ be a martingale and T be a stopping time. We suppose that

$$\mathbb{P}(T < \infty) = 1, \quad \mathbb{E}[|X_T|] < \infty, \quad \text{and} \quad \lim_{n\to\infty} \mathbb{E}[|X_n| \mathbb{1}_{\{T>n\}}] = 0.$$

Show that $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

Solution of Exercise 2.7.10. By the stopping theorem $(X_{n\wedge T})_{n\geqslant 0}$ is also a martingale, so

$$\mathbb{E}[X_0] = \mathbb{E}[X_{n\wedge T}] = \mathbb{E}[X_T \mathbb{1}_{\{T\leqslant n\}}] + \mathbb{E}[X_n \mathbb{1}_{\{T>n\}}].$$

On the one hand, the last term tends to 0 as $n \to \infty$ by the third assumption. On the other hand, we observe that $X_T \mathbb{1}_{\{T\leqslant n\}}$ converges a.s. to X_T (because $T < \infty$ a.s.), and is also dominated by $X_T \in \mathrm{L}^1(\Omega, \mathscr{F}, \mathbb{P})$. Therefore $\mathbb{E}[X_T \mathbb{1}_{\{T\leqslant n\}}] \to \mathbb{E}[X_T]$ by dominated convergence, and we thus conclude that $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

∎

Exercise 2.7.11. On $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbb{P})$, let $(X_n)_{n\geqslant 0}$ be an adapted, integrable process,

$$A_n := \sum_{k=1}^{n} \mathbb{E}[X_k - X_{k-1} \mid \mathscr{F}_{k-1}], \qquad n \geqslant 0,$$

and

$$M_n := X_n - A_n, \qquad n \geqslant 0.$$

1. Show that $(A_n)_{n\geqslant 0}$ is a predictable process.

2. Show that $(M_n)_{n\geqslant 0}$ is a martingale.

3. Suppose we are given a predictable process $(A'_n)_{n\geqslant 0}$ with $A'_0 = 0$ and a martingale $(M'_n)_{n\geqslant 0}$ such that $X_n = M'_n + A'_n$, $n \geqslant 0$. Show that $A'_n = A_n$ and $M'_n = M_n$ a.s. for all $n \geqslant 0$.

4. We suppose in this question that $(X_n)_{n\geqslant 0}$ is a nonnegative submartingale.

   a) Show that $A_n \leqslant A_{n+1}$ a.s. for all $n \geqslant 0$. We write $A_\infty := \lim_{n\to\infty} A_n \in [0, \infty]$.
   b) Show that if $\mathbb{E}[A_\infty] < \infty$, then $(X_n)_{n\geqslant 0}$ converges a.s.
   c) For $a > 0$, let $T_a := \inf\{n \geqslant 0\colon\ A_{n+1} > a\}$.

      i- Check that T_a is a stopping time, and that $\mathbb{E}[X_{n\wedge T_a}] \leqslant a + \mathbb{E}[X_0]$.
      ii- Deduce that $(X_n)_{n\geqslant 0}$ converges a.s. on the event $\{T_a = \infty\}$.
      iii- Conclude that $(X_n)_{n\geqslant 0}$ converges a.s. on the event $\{A_\infty < \infty\}$.

   d) We suppose that the increments of $(X_n)_{n\geqslant 0}$ are dominated in $\mathrm{L}^1(\Omega, \mathscr{F}, \mathbb{P})$:

$$\mathbb{E}[S] < \infty, \quad \text{where } S := \sup_{n\geqslant 1} |X_n - X_{n-1}|. \tag{$\star$}$$

   Show that $\limsup_{n\to\infty} X_n = \infty$ a.s. on the event $\{A_\infty = \infty\}$.

5. We suppose in this question that $(X_n)_{n\geqslant 0}$ is a martingale satisfying to $(\star)$. Show that a.s. as $n \to \infty$, X_n either converges or oscillates, that is

$$\lim_{n\to\infty} X_n \text{ exists in } \mathbb{R} \quad \text{or} \quad \left(\liminf_{n\to\infty} X_n = -\infty \text{ and } \limsup_{n\to\infty} X_n = \infty\right).$$

6. Prove the *conditional Borel–Cantelli lemma*: if $E_n \in \mathscr{F}_n$, $n \in \mathbb{N}$, then (up to a $\mathbb{P}$-null set)

$$\{\limsup E_n\} = \left\{\sum_{n=1}^{\infty} \mathbb{P}(E_n \mid \mathscr{F}_{n-1}) = \infty\right\}.$$

Solution of Exercise 2.7.11.

1. Clearly, $A_0 = 0$ and A_n is $\mathscr{F}_{n-1}$-measurable for every $n \geqslant 1$.

2. Since X_n and A_n are $\mathscr{F}_n$-integrable r.v., so is M_n, and plainly $\mathbb{E}[M_{n+1} - M_n \mid \mathscr{F}_n] = 0$.

3. We see that $M_n - M'_n = A'_n - A_n$, $n \geqslant 0$, is a predictable martingale:

$$M_n - M'_n = \mathbb{E}[M_n - M'_n \mid \mathscr{F}_{n-1}] = M_{n-1} - M'_{n-1}, \qquad n \geqslant 1.$$

Because $M_0 - M'_0 = A'_0 - A_0 = 0$, it follows that $M'_n = M_n$ and $A'_n = A_n$ for all $n \geqslant 0$.

4. a) By the submartingale property $\mathbb{E}[X_n - X_{n-1} \mid \mathscr{F}_{n-1}] \geqslant 0$, so $A_{n-1} \leqslant A_n$, $n \geqslant 1$.

   b) By monotone convergence, $\mathbb{E}[A_\infty] < \infty \iff \sup_{n\geqslant 0} \mathbb{E}[X_n] < \infty$. In this case the submartingale $(X_n)_{n\geqslant 0}$ is bounded in L^1, and therefore converging a.s.

   c) i- As entrance time in the Borel set (a,∞) of the $(\mathscr{F}_n)_{n\geqslant 0}$-adapted process $(A_{n+1})_{n\geqslant 0}$, the r.v. T_a is a stopping time. By the stopping theorem,

$$\mathbb{E}[X_{n\wedge T_a}] = \mathbb{E}\Big[\underbrace{A_{n\wedge T_a}}_{\leqslant a \text{ (by def. of } T_a)}\Big] + \mathbb{E}[M_{n\wedge T_a}] \leqslant a + \mathbb{E}[M_{0\wedge T_a}] = a + \mathbb{E}[X_0],$$

   ii- Because the submartingale $(X_{n\wedge T_a})_{n\geqslant 0}$ is bounded in L^1, it a.s. converges (in $\mathbb{R}$). It remains to note that $X_n = X_{n\wedge T_a}$ on the event $\{T_a = \infty\}$.

   iii- Plainly $\{A_\infty \leqslant a\} = \{T_a = \infty\}$ (monotonicity), so $\{A_\infty < \infty\} = \bigcup_{a\in\mathbb{N}}\{T_a = \infty\}$. By Question 4.c)ii- and σ-additivity, $(X_n)_{n\geqslant 0}$ a.s. converges on $\{A_\infty < \infty\}$.

   d) Introduce the stopping time $R_a := \inf\{n \geqslant 0 \colon X_n > a\}$. Then $(M_{n\wedge R_a})_{n\geqslant 0}$ is a martingale. We have (setting $X_{-1} := X_0$)

$$\mathbb{E}[X_{n\wedge R_a}] = \mathbb{E}[X_{n\wedge R_a - 1}] + \mathbb{E}[X_{n\wedge R_a} - X_{n\wedge R_a - 1}] \leqslant a + \mathbb{E}[X_0] + \mathbb{E}[S] < \infty,$$

   so

$$\mathbb{E}[A_{n\wedge R_a}] = \mathbb{E}[X_{n\wedge R_a}] - \mathbb{E}[M_{n\wedge R_a}] \leqslant a + \mathbb{E}[S] < \infty.$$

   Hence $\mathbb{E}[A_{\infty\wedge R_a}] < \infty$ by monotone convergence. This implies in particular that $\mathbb{P}(A_\infty = \infty, R_a = \infty) = 0$ for all $a \in \mathbb{N}$, that is $R_a < \infty$ a.s. on the event $\{A_\infty = \infty\}$. By σ-additivity, $\sup_{n\geqslant 0} X_n = \infty$ a.s. on the event $\{A_\infty = \infty\}$. Because $X_n < \infty$ a.s. for all n, this is equivalent to $\limsup_{n\to\infty} X_n = \infty$ a.s.

5. We know that X_n^+, $n \geqslant 0$, and X_n^-, $n \geqslant 0$, are nonnegative submartingales. Further, $\mathbb{E}[X_k^+ - X_{k-1}^+ \mid \mathscr{F}_{k-1}] = \mathbb{E}[X_k^- - X_{k-1}^- \mid \mathscr{F}_{k-1}]$ (since $X_n = X_n^+ - X_n^-$, $n \geqslant 0$, is a martingale). By Question 4.c)iii-, a.s. on $\{A_\infty < \infty\}$, the sequences $(X_n^+)_{n\geqslant 0}$ and $(X_n^-)_{n\geqslant 0}$ converge, *i.e*, $\lim_{n\to\infty} X_n$ exists in $\mathbb{R}$. By Question 4.d), a.s. on $\{A_\infty = \infty\}$, we have $\limsup_{n\to\infty} X_n^\pm = \infty$, *i.e*, $\liminf_{n\to\infty} X_n = -\infty$ and $\limsup_{n\to\infty} X_n = \infty$.

6. Clearly, $X_n := \sum_{k=1}^n I_{E_k} - \sum_{k=1}^n \mathbb{P}(E_k \mid \mathscr{F}_{k-1})$, $n \geqslant 0$, is a martingale whose increments are dominated in L^1 (we have $\mathbb{E}[S] \leqslant 2$ in ($\star$)). It thus follows from Question 5 that with probability 1, either X_n oscillates (which means that both nonnegative sums in the definition of X_n diverge to ∞) or X_n converges (which means that both sums converge). Hence, with probability 1, both sums in the definition of X_n a.s. have the same nature. In other words, for $\mathbb{P}$-a.e. $\omega \in \Omega$, $\sum_n \mathbb{P}(E_n \mid \mathscr{F}_{n-1})(\omega) = \infty$ if and only if $\omega \in E_n$ for infinitely many n (*i.e*, $\omega \in \{\limsup E_n\}$). *(This extends the Borel–Cantelli lemmas!)* ∎

Exercise 2.7.12. Let $U_1, U_2, \dots$ be i.i.d. Uniform$(0,1)$ r.v. Let X_0 be any r.v. on $(0,1)$ independent of $(U_i)_{i\geqslant 1}$, and define by induction

$$X_n := tX_{n-1} + (1-t)\mathbb{1}_{\{U_n \leqslant X_{n-1}\}}, \qquad n \geqslant 1,$$

where $t \in (0,1)$ is fixed.

1. Show that $(X_n)_{n\geqslant 0}$ is a martingale converging a.s. and in L^p for every $p \geqslant 1$.
2. Determine the law of $X_\infty := \lim_{n\to\infty} X_n$.
   Hint. Compute $\mathbb{E}[(X_{n+1} - X_n)^2]$.

Solution of Exercise 2.7.12.

1. Clearly, $(X_n)_{n\geqslant 0}$ is adapted to the filtration $\mathscr{F}_n := \sigma(X_0, U_1, \dots, U_n)$, $n \geqslant 0$. It is further plain by induction on n that $X_n \in (0,1)$. Finally, for every $n \geqslant 1$,

$$\mathbb{E}[X_n - X_{n-1} \mid \mathscr{F}_{n-1}] = (1-t) \cdot \Big[\underbrace{\mathbb{P}(U_n \leqslant X_{n-1} \mid \mathscr{F}_{n-1})}_{=X_{n-1}} - X_{n-1}\Big] = 0.$$

   Thus $(X_n)_{n\geqslant 1}$ is a bounded martingale. In particular it converges a.s. and in L^p, $p \geqslant 1$.

2. We have (by independence of U_{n+1} and X_n)

$$\begin{aligned}\mathbb{E}[(X_{n+1} - X_n)^2] &= (1-t)^2\,\mathbb{E}\Big[\big(\mathbb{1}_{\{U_{n+1}\leqslant X_n\}} - X_n\big)^2\Big] \\ &= (1-t)^2\,\mathbb{E}\big[X_n(1-X_n)^2 + (1-X_n)X_n^2\big] \\ &= (1-t)^2\,\mathbb{E}[X_n(1-X_n)].\end{aligned}$$

   By passing to the limit in L^2, we obtain $\mathbb{E}[X_\infty(1-X_\infty)] = 0$, and since $X_\infty \in [0,1]$ a.s., it follows that $X_\infty \in [0,1]$ a.s. Besides, $\mathbb{E}[X_\infty] = \lim_{n\to\infty}\mathbb{E}[X_n] = \mathbb{E}[X_0]$. We conclude that X_∞ is a Bernoulli$(\mathbb{E}[X_0])$ random variable. ■

Exercise 2.7.13. Let X_n, $n \geqslant 1$, be independent nonnegative r.v. with mean 1, and

$$M_n := \prod_{i=1}^{n} X_i, \qquad n \geqslant 0.$$

1. Show that $M_\infty := \lim_{n\to\infty} M_n$ exists almost surely, and $\mathbb{E}[M_\infty] \leqslant 1$.
2. Let $a_n := \mathbb{E}[\sqrt{X_n}]$, $n \geqslant 1$. Prove that the following conditions are equivalent:
   (a) $\mathbb{E}[M_\infty] = 1$;
   (b) $M_n \to M_\infty$ in $\mathrm{L}^1(\mathbb{P})$;
   (c) $(M_n)_{n\geqslant 0}$ is uniformly integrable;
   (d) $\prod_{k\geqslant 1} a_k > 0$;

(e) $\sum_{k\geqslant 1}(1-a_k)<\infty$.

3. Show that if one of the above condition is not satisfied, then $M_\infty = 0$ a.s.

4. Express M_∞ in the particular case where the X_n, $n\in\mathbb{N}$, are i.i.d.

Solution of Exercise 2.7.13.

1. For all $n\geqslant 0$, $\mathbb{E}[|M_n|] \overset{\perp\!\!\perp}{=} \prod_{k=1}^n \mathbb{E}[X_k] = 1 < \infty$ and $\mathbb{E}[M_{n+1}\mid\mathscr{F}_n] = M_n\mathbb{E}[X_{n+1}] = M_n$, so $(M_n)_{n\geqslant 0}$ is a nonnegative, therefore a.s. converging martingale. By Fatou's lemma its a.s. limit M_∞ satisfies to $\mathbb{E}[M_\infty] = \mathbb{E}[\liminf M_n] \leqslant \liminf \mathbb{E}[M_n] = 1$.

2. Let us introduce the auxiliary process
$$N_n = \prod_{k=1}^n \frac{\sqrt{X_k}}{a_k}, \quad n\geqslant 0.$$
Then $(N_n)_{n\geqslant 0}$ is a product of 1-mean independent r.v., so it is a martingale. Being nonnegative, it converges a.s. towards an a.s. finite r.v. N_∞, which fulfills
$$N_\infty = \frac{\sqrt{M_\infty}}{\prod_{k\geqslant 1} a_k}.$$
Let us show the equivalence of all stated conditions.
(b) $\Longrightarrow$ (a) is obvious by passing to the limit in $\mathrm{L}^1(\mathbb{P})$. Further, as $M_n\to M_\infty$ in $\mathbb{P}$, we have (b) $\Longleftrightarrow$ (c) by the L^1-convergence theorem. The equivalence (d) $\Longleftrightarrow$ (e) is quite elementary : we have $0<a_k=\mathbb{E}[\sqrt{X_k}]\leqslant\sqrt{\mathbb{E}[X_k]}=1$ by Cauchy–Schwarz' inequality, and
$$1-\sum_{k=1}^n(1-a_k)\leqslant\prod_{k=1}^n a_k\leqslant\exp\left(-\sum_{k=1}^n(1-a_k)\right)$$
(write
$$\prod_{k=1}^n(1-(1-a_k)) = 1-\sum_{k=1}^n(1-a_k)+\sum_{m=2}^n(-1)^m\sum_{1\leqslant k_1<\cdots<k_m\leqslant n}(1-a_{k_1})\cdots(1-a_{k_m})$$
to obtain the first inequality; the second one follows from the concavity of the logarithm). Now if $\mathbb{E}[M_\infty]=1$, then there exists a $\mathbb{P}$-nonnegligible set on which $M_\infty>0$ and $N_\infty<\infty$, so $\prod_{k\geqslant 1}a_k>0$. Thus (a) $\Longrightarrow$ (d). Conversely, if $\prod_{k\geqslant 1}a_k>0$, then
$$\mathbb{E}[N_n^2] = \frac{1}{\prod_{k=1}^n a_k} \underset{n\to+\infty}{\leqslant} \frac{1}{\prod_{k\geqslant 1}a_k}<\infty,$$
so the martingale $(N_n)_{n\geqslant 0}$, bounded in $\mathrm{L}^2(\mathbb{P})$, converges a.s. and in L^2 toward N_∞. By passing to the limit (in L^2),
$$\frac{1}{\left(\prod_{k\geqslant 1}a_k\right)^2} = \frac{\mathbb{E}[M_\infty]}{\left(\prod_{k\geqslant 1}a_k\right)^2},$$
i.e, $\mathbb{E}[M_\infty]=1$. Thus (d) $\Longrightarrow$ (a). Finally, (a) $\Longrightarrow$ (b) is a consequence of Riesz–Scheffé's lemma (Exercise 1.1.24): because M_∞ and M_n, $n\geqslant 0$, belong to $\mathrm{L}^1(\mathbb{P})$ and $\mathbb{E}[M_\infty]=1=\lim_{n\to+\infty}\mathbb{E}[M_n]$, we have $M_n\to M_\infty$ in $\mathrm{L}^1(\mathbb{P})$.

3. Suppose one of the above conditions is not fulfilled. Then $\prod_{k\geqslant 1} a_k = 0$ and, a.s.,
$$M_\infty = N_\infty^2 \prod_{k\geqslant 1} a_k = 0.$$

4. If the $X_n, n \geqslant 1$, are i.i.d., then $a_k = a_1$ for all $k \geqslant 1$, and (e) is not fulfilled if and only if $a_1 < 1$; in this case $M_\infty = 0$ a.s. by Question 3. The case $a_1 = 1$ implies $X_1 = 1$ a.s. (because there is equality in Cauchy–Schwarz' inequality), and therefore $M_\infty = 1$ a.s.

   Remark. We can derive this last result more directly using the law of large numbers. Suppose first $\mathbb{P}(X_1 = 0) = 0$. By Jensen's inequality, $\ell := \mathbb{E}[\log(X_1)] \leqslant \log(\mathbb{E}[X_1]) = 0$, with equality if and only if $X_1 = 1$ a.s. (log is a *strictly* concave function). But the law of large numbers entails
$$\frac{\log(M_n)}{n} \xrightarrow[n\to\infty]{} \ell, \quad \text{a.s.}$$
   Thus $M_\infty = 0$ a.s. when $\mathbb{P}(X_1 = 1) < 1$ (and so $\ell < 0$). If $\mathbb{P}(X_1 = 0) > 0$, then we also have $M_\infty = 0$ a.s. since $\mathbb{P}(\bigcup_{n\geqslant 0}\{M_n = 0\}) = \lim_{n\to\infty}\uparrow \mathbb{P}(M_n = 0) = \lim_{n\to+\infty} 1 - \mathbb{P}(X_1 > 0)^n = 1$. ∎

Exercise 2.7.14 (Counterexamples).

1. Let U be a Uniform$(0,1)$ r.v. and $X_n := n\mathbb{1}_{\{nU<1\}}$, $n \geqslant 1$. Show that $(X_n)_{n\geqslant 1}$ is bounded in $\mathrm{L}^1(\mathbb{P})$ but not uniformly integrable.

2. Show that the two following families are uniformly integrable but not dominated in $\mathrm{L}^1(\mathbb{P})$ (that is, $\mathbb{E}[\sup_{X\in\mathscr{X}} |X|] = \infty$):

   a) $\mathscr{X} := \{X_{n,k}\}_{\substack{n\geqslant 0\\ 0\leqslant k<2^{2n}}}$ with $X_{n,k} := 2^n \mathbb{1}_{\{k\leqslant 2^{2n}U<k+1\}}$ and $U \sim$ Uniform$(0,1)$;

   b) $\mathscr{X} := \{X_n\}_{n\geqslant 1}$ with $X_n := nA_nB_n$, A_n, B_n, $n \geqslant 1$, Bernoulli$(\frac{1}{n})$ r.v., all independent.
   Hint. Use Borel–Cantelli lemmas to prove that $X_n \to 0$ a.s., and that $\mathbb{E}[X_n \mid \mathscr{F}] \to 0$ in $\mathrm{L}^1(\mathbb{P})$ but not a.s., where $\mathscr{F} := \sigma(A_n : n \geqslant 1)$.

3. Let X_n, $n \geqslant 1$, be independent r.v. with $\mathbb{P}(X_n = 1 - n^2) = 1 - \mathbb{P}(X_n = 1) = n^{-2}$. Show that $S_n := X_1 + \cdots + X_n$, $n \geqslant 0$, defines a martingale converging a.s. to $+\infty$. Is this in contradiction with Exercise 2.7.11.5?

Solution of Exercise 2.7.14.

1. We have $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] = n\mathbb{P}(nU < 1) = 1$, so $(X_n)_{n\geqslant 1}$ is bounded in $\mathrm{L}^1(\mathbb{P})$. But it is not uniformly integrable because $\mathbb{E}[|X_n|\mathbb{1}_{\{|X_n|>a\}}] = n\mathbb{P}(nU < 1) = 1$ for all $n \geqslant a > 0$.

2. a) If $a > 0$, then $\mathbb{E}[|X_{n,k}|\mathbb{1}_{\{|X_{n,k}|\geqslant 2^a\}}] = 2^{-n}\mathbb{1}_{\{n\geqslant a\}} \leqslant 2^{-a}$ for all $n \geqslant 0$ and $0 \leqslant k < 2^{2n}$. Therefore
$$\limsup_{a\to\infty} \sup_{\substack{n\geqslant 0\\ 0\leqslant k<2^{2n}}} \mathbb{E}[|X_{n,k}|\mathbb{1}_{\{|X_{n,k}|\geqslant a\}}] = 0,$$
   whereas
$$\sup_{\substack{n\geqslant 0\\ 0\leqslant k<2^{2n}}} |X_{n,k}| \geqslant \sup_{n\geqslant 0} X_{n,\lfloor 2^{2n}U\rfloor} = \sup_{n\geqslant 0} 2^n = \infty.$$

b) On the one hand, $X_n \in \{0, n\}$ with $\mathbb{P}(X_n = n) = \mathbb{P}(A_n = B_n = 1) = n^{-2}$, so by the first Borel–Cantelli lemma, $X_n = 0$ eventually, a.s. On the other hand, by the second Borel–Cantelli lemma, the independent variables $\mathbb{E}[X_n \mid \mathscr{F}] = A_n$, $n \geqslant 1$, equal 1 infinitely often, a.s., so they converge in $\mathrm{L}^1(\mathbb{P})$, but not a.s. Therefore $(X_n)_{n\geqslant 1}$ cannot be dominated in $\mathrm{L}^1(\mathbb{P})$ because the conditional dominated convergence theorem would entail that $\mathbb{E}[X_n \mid \mathscr{F}]$ tends to 0 a.s. It is however uniformly integrable because it converges a.s. and in $\mathrm{L}^1(\mathbb{P})$.

3. The X_n are independent, integrable r.v. with $\mathbb{E}[X_n] = n^{-2}(1-n^2)+1-n^{-2} = 0$, so $(S_n)_{n\geqslant 0}$ is a martingale. Now the first Borel–Cantelli lemma entails that a.s., $X_n = 1$ for all but finitely many n, hence $S_n \to +\infty$ a.s. This is not in contradiction with Exercise 2.7.11.5 because $\mathbb{E}[\sup_{n\geqslant 0}|X_n|] = \infty$ (eventhough $\sup_{n\geqslant 0}|X_n| < \infty$ a.s.). Indeed, the probabilities

$$\mathbb{P}\Big(\sup_{n\geqslant 0}|X_n| \geqslant k\Big) = 1 - \mathbb{P}\big(\forall i \geqslant \sqrt{k+1},\, X_i = 1\big) = 1 - \prod_{i\geqslant\sqrt{k+1}}\Big(1-\frac{1}{i^2}\Big) = \frac{1}{\lfloor\sqrt{k+1}\rfloor},$$

for $k \geqslant 2$, are not summable. ■

Exercise 2.7.15. Let $\mathscr{S} := \bigcup_{n\geqslant 1}\mathscr{S}_n$, where $\mathscr{S}_n := \{\pi\colon \mathbb{N}\to\mathbb{N}$ bijective with $\pi(k) = k$ for all $k > n\}$. Suppose $X := (X_n)_{n\geqslant 1}$ is a stochastic process such that for every $\pi \in \mathscr{S}$, $X^\pi := (X_{\pi(n)})_{n\geqslant 1}$ has the same law as X. Define the *exchangeable σ-algebra* $\mathscr{E} := \bigcap_{n\geqslant 1}\mathscr{E}_n$, where

$$\mathscr{E}_n := \Big\{\{X \in A\}\colon A \subseteq \mathbb{R}^{\mathbb{N}} \text{ measurable s.t. } \{X\in A\} = \{X^\pi \in A\} \text{ for all } \pi \in \mathscr{S}_n\Big\}.$$

1. Show that for every $f\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ bounded measurable,

$$\mathbb{E}[f(X) \mid \mathscr{E}_n] = \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} f(X^\pi), \quad n \geqslant 1,$$

and that as $n\to\infty$, this sequence converges to $\mathbb{E}[f(X)\mid\mathscr{E}]$ a.s. and in $\mathrm{L}^1(\mathbb{P})$.

2. Let the *tail σ-algebra* $\mathscr{T} := \bigcap_{n\geqslant 1}\sigma(X_k\colon k\geqslant n)$. Show that $\mathscr{T} \subseteq \mathscr{E}$ and that for all $f\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ bounded measurable, $\mathbb{E}[f(X)\mid\mathscr{E}] = \mathbb{E}[f(X)\mid\mathscr{T}]$.

3. Show that if $A \in \mathscr{E}$, then there is $B \in \mathscr{T}$ such that $A = B$ up to a $\mathbb{P}$-null set.

   Hint. Show that $\mathbb{P}(A\mid\mathscr{T}) = \mathbb{1}_A$.

4. Suppose $X \in \{0,1\}^{\mathbb{N}}$. Compute $\mathbb{P}(X_1 = x_1, \dots, X_k = x_k \mid \mathscr{E}_n)$ for all $n, k \geqslant 1$, $x \in \{0,1\}^k$. Deduce that given $P := \mathbb{P}(X_1 = 1\mid\mathscr{E})$, the X_n, $n\geqslant 1$, are i.i.d. Bernoulli(P) r.v.

Solution of Exercise 2.7.15.

1. For every $n \geqslant 1$, the map

$$f_{(n)}\colon x \in \mathbb{R}^{\mathbb{N}} \mapsto \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} f(x^\pi)$$

is measurable, n-symmetric ($f_{(n)}(x^\pi) = f_{(n)}(x)$ for all $\pi \in \mathscr{S}_n$), and bounded (by $\|f\|_\infty$). It follows that $f_{(n)}(X)$ is a $\mathscr{E}_n$-measurable, integrable r.v. Moreover, for every $h\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ n-symmetric and bounded, we have

$$\begin{aligned}
\mathbb{E}[f_{(n)}(X)h(X)] &= \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} \mathbb{E}[f(X^\pi)h(X)] \\
&= \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} \mathbb{E}[f(X^\pi)h(X^\pi)] && (h\ n\text{-symmetric}) \\
&= \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} \mathbb{E}[f(X)h(X)] && (X^\pi \overset{(d)}{=} X) \\
&= \mathbb{E}[f(X)h(X)].
\end{aligned}$$

Hence $\mathbb{E}[f(X) \mid \mathscr{E}_n] = f_{(n)}(X)$. Finally, by the tower property for $\mathscr{E}_{n+1} \subseteq \mathscr{E}_n$, we have $\mathbb{E}[f_{(n)}(X) \mid \mathscr{E}_{n+1}] = f_{(n+1)}(X)$, which means that $f_{(n)}(X) = \mathbb{E}[f(X) \mid \mathscr{E}_n]$, $n \geqslant 1$, is a backwards martingale; it converges a.s. and in $\mathrm{L}^1(\mathbb{P})$ towards $\mathbb{E}[f(X) \mid \mathscr{E}]$.

2. Let $B \in \mathscr{T}$. Then $B \in \mathscr{E}_n$, $n \geqslant 1$, since $B \in \sigma(X_k\colon k \geqslant n+1)$. Thus $B \in \cap_{n\geqslant 1}\mathscr{E}_n = \mathscr{E}$. We must show that $\mathbb{E}[f(X) \mid \mathscr{E}]$ is $\mathscr{T}$-measurable for $f\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ bounded measurable. Let us first make the extra assumption that f is continuous. Since f is then the limit of $x \mapsto f(x_1,\ldots,x_k,0,\ldots)$ as $k\to\infty$, we may also assume by the conditional dominated convergence theorem that f is a function of the first $k \geqslant 1$ coordinates only. In this case, for all $n \geqslant k+p$, the number of permutations $\pi \in \mathscr{S}_n$ with $\{\pi(1),\ldots,\pi(k)\} \cap \{1,\ldots,p\} = \emptyset$ is $(n-k)!(n-p)!/(n-k-p)! \sim n!$, which implies that the a.s. limit $\mathbb{E}[f(X) \mid \mathscr{E}]$ of $f_{(n)}(X)$ is a $\sigma(X_i\colon i > p)$-measurable r.v. Since this is true for all p, we conclude that this limit is in fact $\mathscr{T}$-measurable.

   Consider now the linear space $\mathscr{H} := \{f\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ measurable ; $\mathbb{E}[f(X) \mid \mathscr{E}]$ is $\mathscr{T}$-measurable$\}$. By the conditional monotone convergence theorem, this space is stable by non-increasing limits of nonnegative functions, and further contains the indicator of any open set $O \subseteq \mathbb{R}^{\mathbb{N}}$ (precisely because $\mathbb{1}_O$ is the non-decreasing limit as $k\to\infty$ of the continuous functions $f_k\colon x \in \mathbb{R}^{\mathbb{N}} \mapsto 1 \wedge \big(k\cdot d(x,O^c)\big)$, which by the previous point belong to $\mathscr{H}$ — here, d is a distance in $\mathbb{R}^{\mathbb{N}}$ which metrizes the product topology). It follows from the monotone class theorem that $\mathscr{H}$ contains all bounded measurable functions $f\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$.

3. Let $A \in \mathscr{E}$. Since $A \in \sigma(X)$, there is $E \subseteq \mathbb{R}^{\mathbb{N}}$ measurable with $A = \{X \in E\}$. Then $f := \mathbb{1}_E$ defines a bounded measurable function $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, so

$$\mathbb{1}_A = f(X) = \mathbb{E}[f(X) \mid \mathscr{E}] = \mathbb{E}[f(X) \mid \mathscr{T}] = \mathbb{P}(A \mid \mathscr{T}),$$

   by Question 1. Hence, for $B := \{\mathbb{P}(A \mid T) = 1\} \in \mathscr{T}$, we have $A = B$ up to a $\mathbb{P}$-null set.

4. Applying Question 1 with $f := \mathbb{1}_{\{x_1\}\times\cdots\times\{x_k\}}$ yields

$$\mathbb{P}(X_1 = x_1,\ldots,X_k = x_k \mid \mathscr{E}_n) = \frac{1}{n!}\sum_{\pi\in\mathscr{S}_n} \mathbb{1}_{\{X_{\pi(1)}=x_1,\ldots,X_{\pi(k)}=x_k\}}.$$

Let $y_k := \sum_{i=1}^k x_i$ and $Y_n := \sum_{i=1}^n X_i$. To construct a permutation $\pi \in \mathscr{S}_n$ contributing to the sum in the right-hand side, we must map injectively y_k of its first k values to indexes $i \in \{1,\dots,n\}$ among the Y_n ones with $X_i = 1$ and the $k - y_k$ other to indexes $i \in \{1,\dots,n\}$ among the $n - Y_n$ ones with $X_i = 0$, so

$$\begin{aligned}\mathbb{P}(X_1 = x_1,\dots,X_k = x_k \mid \mathscr{E}_n) &= y_k!\binom{Y_n}{y_k}\cdot(k-y_k)!\binom{n-Y_n}{k-y_k}\cdot\frac{(n-k)!}{n!}\\ &\underset{n\to\infty}{\sim}(Y_n)^{y_k}\cdot(n-Y_n)^{k-y_k}\cdot n^{-k}.\end{aligned}$$

Hence, for $P := \mathbb{P}(X_1 = 1 \mid \mathscr{E}) = \lim_{n\to\infty}\frac{Y_n}{n}$ a.s.,

$$\mathbb{P}(X_1 = x_1,\dots,X_k = x_k \mid \mathscr{E}) = P^{y_k}(1-P)^{k-y_k}, \qquad k \geqslant 1,$$

which means that given P, the X_n, $n \geqslant 1$, are i.i.d. Bernoulli(P) r.v. ∎

Exercise 2.7.16 (0-1 laws). Let $(\Omega,\mathscr{F},(\mathscr{F}_n),\mathbb{P})$ be a filtred probability space and $\mathscr{F}_\infty := \bigvee_{n\geqslant 0}\mathscr{F}_n$.

1. a) Show that for every $X \in \mathrm{L}^1(\Omega,\mathscr{F},\mathbb{P})$,

$$\mathbb{E}[X \mid \mathscr{F}_n] \xrightarrow[n\to\infty]{} \mathbb{E}[X \mid \mathscr{F}_\infty], \quad \text{a.s. and in } \mathrm{L}^1.$$

   b) Deduce *Lévy's 0-1 law*: for every $A \in \mathscr{F}_\infty$,

$$\mathbb{P}(A \mid \mathscr{F}_n) \xrightarrow[n\to\infty]{} \mathbb{1}_A, \quad \text{a.s.}$$

2. Let $(X_n)_{n\geqslant 1}$ be a sequence of i.i.d. real r.v.

   a) Show *Kolmogorov's 0-1 law*: the tail σ-algebra $\mathscr{T} := \bigcap_{n\geqslant 1}\sigma(X_k\colon k \geqslant n)$ is $\mathbb{P}$-trivial:

$$\forall A \in \mathscr{T}, \quad \mathbb{P}(A) \in \{0,1\}.$$

   Hint. Use Lévy's 0-1 law.

   b) Use Kolmogorov's 0-1 law and Exercise 2.7.15.3 to reprove *Hewitt–Savage's 0-1 law*: the exchangeable σ-algebra $\mathscr{E} := \sigma(f(X)\colon f \in \mathbf{S})$, where $\mathbf{S} := \{f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R} \text{ symmetric}\}$, is $\mathbb{P}$-trivial.

Solution of Exercise 2.7.16.

1. a) The process $X_n := \mathbb{E}[X \mid \mathscr{F}_n]$, $n \geqslant 0$, defines a closed martingale, so X_n converges a.s. and in $\mathrm{L}^1(\mathbb{P})$ towards some r.v. X_∞. It remains to show that $X_\infty = \mathbb{E}[X \mid \mathscr{F}_\infty]$. The dominated convergence theorem shows that $\{A \in \mathscr{F} : \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X_\infty\mathbb{1}_A]\}$ is a monotone class containing the π-system $\bigcup_{k\geqslant 0}\mathscr{F}_k$, and thus also $\mathscr{F}_\infty$ by the monotone class theorem. Because X_∞ is $\mathscr{F}_\infty$-measurable, we conclude that $X_\infty = \mathbb{E}[X \mid \mathscr{F}_\infty]$.

   b) In particular $\mathbb{1}_A \in \mathrm{L}^1(\Omega,\mathscr{F},\mathbb{P})$ for every $A \in \mathscr{F}_\infty$, so by Question 1.a),

$$\mathbb{P}(A \mid \mathscr{F}_n) = \mathbb{E}[\mathbb{1}_A \mid \mathscr{F}_n] \xrightarrow[n\to\infty]{} \mathbb{E}[\mathbb{1}_A \mid \mathscr{F}_\infty] = \mathbb{1}_A, \quad \text{a.s.}$$

2. a) Let $\mathscr{F}_n := \sigma(X_1, \dots, X_n)$, $n \geqslant 1$. Then for every $A \in \mathscr{T}$, we have $A \in \mathscr{F}_\infty$ and also $A \perp\!\!\!\perp \mathscr{F}_n$ because $A \in \sigma(X_k : k \geqslant n+1)$. By Lévy's 0-1 law (Question 1.b)),

$$\mathbb{P}(A) = \mathbb{P}(A \mid \mathscr{F}_n) \xrightarrow[n\to\infty]{} \mathbb{1}_A, \qquad \text{a.s.,}$$

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

b) Note first that, in the notation of Exercise 2.7.15, the distribution of X^π for any $\pi \in \mathscr{S}$ is equal to that of X because the r.v. X_n, $n \geqslant 1$, are i.i.d. Next, Question 3 there also shows that for every $A \in \mathscr{E}$, there exists $B \in \mathscr{T}$ such that $A = B$ up to a $\mathbb{P}$-null set, and thus $\mathbb{P}(A) = \mathbb{P}(B) \in \{0,1\}$ using Kolmogorov's 0-1 law. Hence $\mathscr{E}$ is $\mathbb{P}$-trivial. ■

2.8 Markov chains

Exercise 2.8.1. Let $p \in (0,1)$, $X_1, X_2, \dots$ i.i.d. Bernoulli(p) r.v., and $S_n := X_1 + \cdots + X_n$. Justify whether each of the following processes is a Markov chain or not; if it is, give the corresponding state space E and the transition matrix Q.

1. X_n, $n \geqslant 0$;
2. S_n, $n \geqslant 0$;
3. $T_n := S_1 + \cdots + S_n$, $n \geqslant 0$;
4. $\mathbf{V}_n := (S_n, T_n)$, $n \geqslant 0$.

Solution of Exercise 2.8.1.

1. It is a Markov chain because by independence, the law of X_{n+1} given $(X_0, \dots, X_n)$ is Bernoulli(p). We have $E = \{0,1\}$ and $Q(x,y) = (1-p)\mathbb{1}_{\{y=0\}} + p\mathbb{1}_{\{y=1\}}$.
2. It is a Markov chain. The law of S_{n+1} given $S_0, \dots, S_n$ is $(Q(S_n, y))_{y\in E}$ with $E = \mathbb{Z}_+$ and $Q(x,y) = p\mathbb{1}_{\{y=x+1\}} + (1-p)\mathbb{1}_{\{x=y\}}$, $x, y \in E$.
3. It is not a Markov chain because

$$\mathbb{P}(T_4 = 4 \mid T_3 = 3, T_2 = 2, T_1 = 1) = \frac{\mathbb{P}(X_1 = 1, X_2 = X_3 = X_4 = 0)}{\mathbb{P}(X_1 = 1, X_2 = X_3 = 0)} = 1 - p,$$

whereas

$$\mathbb{P}(T_4 = 4 \mid T_3 = 3, T_2 = 1, T_1 = 0) = \frac{\mathbb{P}(X_1 = 0, X_2 = X_3 = 1, X_4 = -1)}{\mathbb{P}(X_1 = 0, X_2 = X_3 = 1)} = 0.$$

4. It is a Markov chain. The law of $\mathbf{V}_{n+1}$ given $\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_n = (S_n, T_n)$ is the law of $(S_n + X_{n+1}, T_n + S_n + X_{n+1})$, *i.e*, $(Q(\mathbf{V}_n, \mathbf{v}))_{\mathbf{v}\in E}$ with $E = \mathbb{Z}_+^2$ and

$$Q(\mathbf{x}, \mathbf{y}) = p\mathbb{1}_{\{s'=s+1, t'=s+t+1\}} + (1-p)\mathbb{1}_{\{s'=s, t'=s+t\}}, \quad \mathbf{x} := (s,t), \mathbf{y} := (s', t') \in E.$$ ■

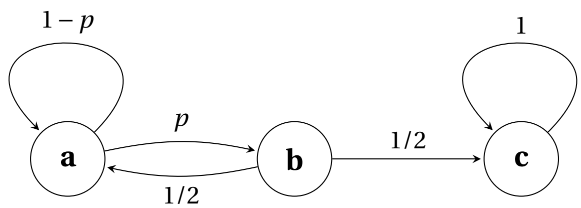
Exercise 2.8.2. Let $p \in (0,1)$ and $(X_n)_{n\geqslant 0}$ be a Markov chain on $E := \{\mathbf{a},\mathbf{b},\mathbf{c}\}$ with transition matrix

$$Q := \begin{pmatrix} 1-p & p & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

1. Draw its transition graph.
2. Compute the probability $\mathbb{P}(X_n = \mathbf{b} \mid X_0 = \mathbf{a})$, $n \in \mathbb{N}$. Find its limit as $n \to \infty$.

Solution of Exercise 2.8.2.

1.



2. We have $\mathbb{P}(X_n = \mathbf{b} \mid X_0 = \mathbf{a}) = Q_n(\mathbf{a},\mathbf{b})$, that is the element at position $(1,2)$ (first row, second column) in Q^n, which we can compute by reducing Q. But we may also observe that $Q_n = Q \cdot Q_{n-1}$ for every $n \geqslant 1$ (that is just the Markov property at time 1). In particular, if we let $a_n := \mathbb{P}(X_n = \mathbf{b} \mid X_0 = \mathbf{a})$ and $b_n := \mathbb{P}(X_n = \mathbf{b} \mid X_0 = \mathbf{b})$, then

$$\begin{cases} a_n = (1-p)\, a_{n-1} + p\, b_{n-1}, \\ b_n = \dfrac{1}{2}\, a_{n-1} \end{cases}$$

(this can also be derived by reasoning on the transition graph). Thus $(a_n)_{n\geqslant 0}$ fulfills the linear, homogeneous, second order, recurrence system

$$\begin{cases} a_n - (1-p)\, a_{n-1} - \dfrac{p}{2}\, a_{n-2} = 0, \quad n \geqslant 2, \\ a_0 = 0,\ a_1 = p. \end{cases}$$

This is easily solved to

$$a_n = p\,\frac{\left(1-p+\sqrt{1+p^2}\right)^n - \left(1-p-\sqrt{1+p^2}\right)^n}{2^n\sqrt{1+p^2}}.$$

We find that $a_n \to 0$, which is not surprising by looking at the transition graph: the chain will eventually reach the *absorbing state* $\mathbf{c}$ and stay there forever. ■

Exercise 2.8.3. Let $f\colon E \to F$ be a function between countable sets, and let $(X_n)_{n\geqslant 0}$ be a Markov chain on E with transition matrix P.

1. Find a simple counterexample showing that $Y_n := f(X_n)$, $n \geqslant 0$, is not necessarily a Markov chain on F.

2. We suppose that whenever $f(x) = f(y)$, then $P(x, A) = P(y, A)$ for every $A \subseteq E$. Show that $(Y_n)_{n\geqslant 0}$ is a Markov chain; express its transition matrix using P and f.

Solution of Exercise 2.8.3.

1. Consider for instance the symmetric random walk: $X_n := Y_1 + \cdots + Y_n$, $n \geqslant 0$, where $Y_1, Y_2, \ldots$ are i.i.d. r.v. with $\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_1 = -1) = 1/2$. This is a Markov chain on $E = \mathbb{Z}$ with transition matrix $P(x, y) = (\mathbb{1}_{\{y=x+1\}} + \mathbb{1}_{\{y=x-1\}})/2$. Then $(X_n^+)_{n\geqslant 0}$ is not a Markov chain on $F = \mathbb{Z}_+$ because
$$\mathbb{P}(X_3^+ = 0 \mid X_2^+ = 0, X_1^+ = 0) = \mathbb{P}\big((Y_2, Y_3) \neq (1, 1)\big) = 3/4,$$
whereas
$$\mathbb{P}(X_3^+ = 0 \mid X_2^+ = 0, X_1^+ = 1) = \mathbb{P}(Y_3 = -1) = 1/2.$$

2. Define $[y] := \{x \in E\colon f(x) = y\}$, $y \in F$, and, thanks to the assumption on P, $P(y, \cdot) := P(x, \cdot)$ for any $x \in [y]$ (with the convention $P(y, \cdot) \equiv 0$ when $[y] = \varnothing$). Let $y_0, \ldots, y_n, y_{n+1} \in F$, and note $\mathscr{P} := \mathbb{P}(\cdot \mid Y_n = y_n, \ldots, Y_0 = y_0)$. Then
$$\begin{aligned}\mathscr{P}(Y_{n+1} = y_{n+1}) &= \sum_{\substack{x_n \in [y_n],\\ x_{n+1} \in [y_{n+1}]}} \mathscr{P}(X_{n+1} = x_{n+1} \mid X_n = x_n) \mathscr{P}(X_n = x_n)\\ &= P(y_n, [y_{n+1}]) \sum_{x_n \in [y_n]} \mathscr{P}(X_n = x_n)\\ &= P(y_n, [y_{n+1}]),\end{aligned}$$
so $(Y_n)_{n\geqslant 0}$ is a Markov chain on F with transition matrix $Q(y, y') := P(y, [y'])$.

Remark. In particular the assumption holds if f is injective; then $(Y_n)_{n\geqslant 0}$ is a Markov chain on $f(E)$ with transition matrix $Q(y, y') := P(f^{-1}(y), f^{-1}(y'))$. ∎

Exercise 2.8.4. Let $(U_n)_{n\geqslant 1}$ be i.i.d. uniform r.v. on $(0, 1)$ and X_0 an independent r.v. on E.

1. Let $f\colon E \times (0, 1) \to E$ and define $X_{n+1} := f(X_n, U_{n+1})$, $n \geqslant 0$. Show that $(X_n)_{n\geqslant 0}$ is a Markov chain on E. Express its transition matrix in terms of f and U_1.

2. Conversely, let P be a given transition matrix. Find a function $f\colon E \times (0, 1) \to E$ such that the Markov chain $(X_n)_{n\geqslant 0}$ above has transition matrix P.

Solution of Exercise 2.8.4.

1. The law of X_{n+1} given $(X_0, \ldots, X_n)$ is the law of $f(X_n, U_{n+1})$, so it is a Markov chain on E with transition matrix $P(x, y) := \mathbb{P}(f(x, U_1) = y)$.

2. We fix $\{y_1, y_2, \ldots\}$ an enumeration of E and define, for every $(x, u) \in E \times (0, 1)$, $f(x, u) := y_k$, where $k \in \mathbb{N}$ is the unique integer such that
$$\sum_{i=1}^{k-1} P(x, y_i) \leqslant u < \sum_{i=1}^{k} P(x, y_i).$$
(Such an integer always exists because P is a transition matrix.) This indeed defines $f\colon E \times (0, 1) \to E$ so that $\mathbb{P}(f(x, U_1) = y) = P(x, y)$ for every $x, y \in E$. ∎

Exercise 2.8.5. Let E be a *finite* set of cardinal $k \geqslant 2$, and P be a transition matrix on E such that $\alpha := \inf\{P(x,y) \colon x, y \in E\} > 0$ (note then that $0 < \alpha \leqslant 1/2$).

1. We fix $y \in E$ and set $p_n(x) := P_n(x,y)$, $x \in E$.

   a) Show that for every $n \geqslant 0$,

   $$\begin{cases} \inf p_{n+k} \geqslant \alpha \sup p_n + (1-\alpha) \inf p_n, \\ \sup p_{n+k} \leqslant \alpha \inf p_n + (1-\alpha) \sup p_n. \end{cases}$$

   Hint. Use that $\sum_{x \in X} P_k(\cdot, x) + \sum_{x \in E \setminus X} P_k(\cdot, x) = 1$ for $X := \{x \in E \colon p_n(x) = \sup p_n\}$.

   b) Deduce that $d_n := \sup p_n - \inf p_n$ converges to 0 as $n \to \infty$.

2. Conclude that there exists a probability distribution $(p(y))_{y \in E}$ on E such that

$$\forall x \in E, \quad p(y) = \lim_{n \to \infty} P_n(x,y).$$

Solution of Exercise 2.8.5.

1. a) For $\inf P = \alpha$,

   $$\forall (x,y) \in E^2, \quad P_2(x,y) = \sum_{z \in E} P(x,z)P(z,y) \geqslant \alpha \sum_{z \in E} P(x,z) = \alpha$$

   (P is a transition matrix), we see that $\inf P_2 \geqslant \alpha$ and by immediate induction, $\inf P_n \geqslant \alpha$ for all $n \in \mathbb{N}$. Also, note that $X := \{x \in E \colon p_n(x) = \sup p_n\} \neq \emptyset$ because E is finite. Then, for every $x' \in E$,

   $$\begin{aligned} p_{n+k}(x') &= \sum_{x \in E} p_n(x) P_k(x', x) \\ &= \sup p_n \sum_{x \in X} P_k(x', x) + \sum_{x \in E \setminus X} p_n(x) P_k(x', x) \\ &\geqslant \sup p_n \sum_{x \in X} P_k(x', x) + \inf p_n \sum_{x \in E \setminus X} P_k(x', x) \\ &= (\sup p_n - \inf p_n) \sum_{x \in X} P_k(x', x) + \inf p_n \\ &\geqslant \alpha (\sup p_n - \inf p_n) + \inf p_n, \end{aligned}$$

   which yields the first of the two desired inequalities. We proceed similarly for the second (exchanging sup and inf).

   b) Substracting both inequalities in 1.a), we find $0 \leqslant d_{n+k} \leqslant (1-2\alpha) d_n$, $n \geqslant 0$. Thus, since $1 - 2\alpha \geqslant 0$,

   $$0 \leqslant \limsup_{n \to \infty} d_n \leqslant (1-2\alpha) \limsup_{n \to \infty} d_n,$$

   and using that $1 - 2\alpha < 1$, we deduce that d_n converges to 0 as $n \to \infty$.

2. Note that for every $x' \in E$,

$$p_{n+1}(x') = \sum_{x \in E} P(x', x)\, p_n(x) \leqslant \sup p_n \underbrace{\sum_{x \in E} P(x', x)}_{=1} = \sup p_n,$$

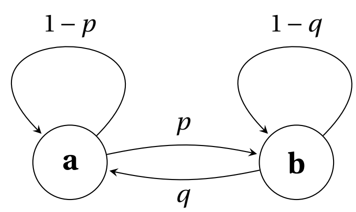
   so the nonnegative sequence $(\sup p_n)_{n \geqslant 0}$ is non-increasing. It has therefore a limit $p(y) \in [0,1]$, which by 1.b) is also the limit of $(\inf p_n)_{n \geqslant 0}$. Since $\inf p_n \leqslant p_n \leqslant \sup p_n$, we have thus proved that for every $x, y \in E$, $p_n(x) = P_n(x,y)$ converges to $p(y)$ as $n \to \infty$. That $(p(y))_{y \in E}$ is indeed a probability distribution follows by taking the limit in the finite sum

$$\sum_{y \in E} P_n(x,y) = 1,$$

   for some $x \in E$.

Conclusion. Any Markov chain $(X_n)_{n \geqslant 0}$ on a finite state space and whose transition matrix has non-zero coefficients admits a limiting distribution which does not depend on the law of the initial state X_0. ∎

Exercise 2.8.6. Let $p, q \in [0,1]$ and $(X_n)_{n \geqslant 0}$ be a Markov chain on $E := \{\mathbf{a}, \mathbf{b}\}$ with graph



.

1. For which values of p, q is $(X_n)_{n \geqslant 0}$ irreducible? Give the state classification.

2. Give the transition matrix of $(X_n)_{n \geqslant 0}$ and find the invariant probability measures.

3. Determine explicitly the law of X_n under $\mathbb{P}_{\mathbf{a}}$, for all $n \geqslant 0$.

4. Does $(X_n)_{n \geqslant 0}$ converge in law?

Solution of Exercise 2.8.6.

1. Clearly, the chain is irreducible if and only if $p, q > 0$. State $\mathbf{a}$ (resp. $\mathbf{b}$) is transient if and only if $q = 0$ and $p > 0$ (resp. $p = 0$ and $q > 0$).

2. The transition matrix is

$$Q := \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}.$$

   A distribution μ on E is invariant (for Q) if and only if $\mu Q = \mu$, that is

$$\begin{cases} \mu(\mathbf{a}) + \mu(\mathbf{b}) = 1, \\ p \cdot \mu(\mathbf{a}) - q \cdot \mu(\mathbf{b}) = 0. \end{cases}$$

   When $p + q > 0$, this determines μ uniquely: $\mu(\mathbf{a}) = q/(p+q)$ and $\mu(\mathbf{b}) = p/(p+q)$.

3. Let $p_n := \mathbb{P}_{\mathbf{a}}(X_n = \mathbf{b}) = Q_n(\mathbf{a},\mathbf{b})$. Then $p_0 = 0$ and, since $1 - p_n = \mathbb{P}_{\mathbf{a}}(X_n = \mathbf{a})$,
$$\begin{aligned} p_{n+1} = (Q_n \cdot Q)(\mathbf{a},\mathbf{b}) &= Q_n(\mathbf{a},\mathbf{a})\,Q(\mathbf{a},\mathbf{b}) + Q_n(\mathbf{r}a,\mathbf{b})\,Q(\mathbf{b},\mathbf{b}) \\ &= (1-p_n)p + p_n(1-q) \\ &= (1-p-q)p_n + p. \end{aligned}$$
Hence $p_n = p(1-(1-p-q)^n)/(p+q)$, $n \geqslant 0$ if $p+q>0$, and $p_n \equiv 0$ if $p+q=0$.

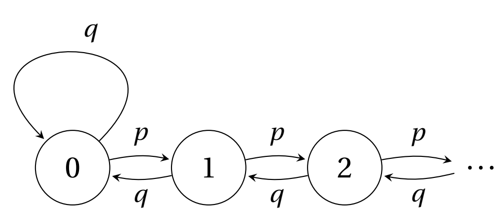
4. If $p+q=0$, then $X_n = X_0$ for all $n \geqslant 0$. If $0 < p+q < 2$, then $(X_n)_{n\geqslant 0}$ converges in distribution to the unique invariant law (Question 2). If $p = q = 1$, then $X_{2n} = X_0 \neq X_1 = X_{2n+1}$: unless $X_1 \sim X_0$ (*i.e* X_0 has already the uniform distribution on E), there cannot be convergence ($(X_n)_{n\geqslant 0}$ is *periodic*)! ■

Exercise 2.8.7. Let $p \in [0,1]$, $q := 1-p$, and $(Y_k)_{k\geqslant 1}$ be a sequence of i.i.d. r.v. with $\mathbb{P}(Y_1 = 1) = p$, $\mathbb{P}(Y_1 = -1) = q$. Define $(X_n)_{n\geqslant 0}$ by $X_0 \in \mathbb{Z}_+$ and
$$X_{n+1} := (X_n + Y_{n+1})^+, \qquad n \geqslant 0,$$
where $x^+ := \max(x,0)$.

1. Prove that $(X_n)_{n\geqslant 0}$ is a Markov chain. Give its state space and transition graph.
2. Is $(X_n)_{n\geqslant 0}$ irreducible? Give the state classification. (Discuss according to p.)
   Hint. Compare $(X_n)_{n\geqslant 0}$ to the random walk $\widetilde{X}_n := Y_1 + \cdots + Y_n$, $n \geqslant 0$, on $\mathbb{Z}$.
3. Determine all invariant measures of $(X_n)_{n\geqslant 0}$. Is there some invariant law?

Solution of Exercise 2.8.7.

1. Since X_{n+1} is given as a (measurable) function of X_n and Y_{n+1}, where $(Y_n)_{n\geqslant 0}$ is i.i.d., the process $(X_n)_{n\geqslant 0}$ is a Markov chain. It has values in $E := \mathbb{Z}_+$ and transition matrix $Q(x,y) := \mathbb{P}((x+Y_1)^+ = y) = p\mathbb{1}_{\{y=x+1\}} + q\mathbb{1}_{\{y=(x-1)^+\}}$, $x, y \in E$.



2. Clearly, $(X_n)_{n\geqslant 0}$ is irreducible (that is, the above graph is strongly connected) if and only if $p \notin \{0,1\}$. If $p=0$, then $X_n = (X_0 - n)^+$ so the state 0 is recurrent while all other states are transient. If $p=1$, then $X_n = X_0 + n$ so all states are transient. For $0<p<1$, that is when the chain is irreducible, all states are of the same type as state 0 and we may suppose $X_0 = 0$. Then $X_n = \widetilde{X}_n - \inf_{i\leqslant n} \widetilde{X}_i$, where by the law of large numbers, $\lim_{n\to\infty} \widetilde{X}_n/n = \mathbb{E}[Y_1] = p-q$, almost surely. Thus, if $q<p$, then $X_n \to \infty$ a.s. and 0 is transient, and if $p<q$, then $X_n = 0$ infinitely often a.s. and 0 is recurrent. In the last case $p=q=1/2$, we can also conclude that 0 is recurrent: we know that the symmetric random walk $(\widetilde{X}_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}$ fulfills $\liminf_{n\to\infty} \widetilde{X}_n = -\infty$ a.s. (combine, for instance, Exercise 2.4.12 and Exercise 2.7.11.5), thus $X_n = 0$ infinitely often.

3. Let μ be a nontrivial measure on E with $\mu(x) < \infty$ for all $x \in E$. It is invariant for Q (*i.e*, $\mu Q = \mu$) if and only if
$$\begin{cases} q \cdot \mu(1) = p \cdot \mu(0), \\ q \cdot (\mu(k+1) - \mu(k)) = p \cdot (\mu(k) - \mu(k-1)), \quad k \geqslant 1. \end{cases}$$
This implies $p < 1$ (otherwise $\mu \equiv 0$), and in this case $(X_n)_{n\geqslant 0}$ admits invariant measures which are all of the form
$$\mu(k) = \mu(0) \left(1 - \sum_{i=0}^{k-1} \left(1 - \frac{p}{q}\right)\left(\frac{p}{q}\right)^i\right), \qquad k \geqslant 0.$$
Further, this can be normalized to a probability measure if and only if $0 < p < q$, *i.e*, $0 < p < 1/2$, and the (unique) invariant law is then the geometric distribution with parameter $1 - p/q$:
$$\mu(k) = \left(1 - \frac{p}{q}\right)\left(\frac{p}{q}\right)^k, \qquad k \geqslant 0.$$
Remark. In particular the chain has no invariant law when $1/2 \leqslant p < 1$, although it is irreducible. The chain is *null recurrent* when $p = 1/2$. ■

Exercise 2.8.8. Let $(X_n)_{n\geqslant 0}$ be a Markov chain on a finite or countable state space E, and μ be a probability distribution on E.

1. Show that if X_n converges in law to μ as $n \to \infty$, then μ is an invariant measure.

2. Show that if μ is an invariant measure, then $\mu(x) = 0$ for all transient state $x \in E$.

Solution of Exercise 2.8.8. We let Q denote the transition matrix of $(X_n)_{n\geqslant 0}$.

1. Let μ_n, $n \geqslant 0$, be the law of X_n, so that $\mu_n(x) \to \mu(x)$ as $n \to \infty$, for all $x \in E$. For every $n \geqslant 0$ we have $\mu_{n+1} = \mu_n Q$, that is
$$\mu_{n+1}(x) = \sum_{y \in E} \mu_n(y)\, Q(y,x), \qquad x \in E. \tag{$\star$}$$
By Fatou's lemma $\mu \geqslant \mu Q$, but we can also apply it in "$1 - \mu_{n+1} = \mu_n(1-Q)$" to obtain the converse inequality $\mu \leqslant \mu Q$. Hence μ is an invariant measure.

   Remark. We could also apply Portmanteau's theorem since ($\star$) is nothing else than $\mathbb{E}[Q(X_n, x)]$, where $Q(\cdot, x)$ is a continuous bounded function on E.

2. Suppose that μ is an invariant probability measure. Then $\mu = \mu Q_n$ for every $n \geqslant 0$, so
$$\mu(x) = \sum_{y \in E} \mu(y)\, Q_n(y,x), \qquad x \in E.$$
If x is transient, then we know that $Q_n(y,x) \to 0$ as $n \to \infty$ for all $y \in E$, and the dominated convergence theorem (or Fatou's lemma!) therefore yields $\mu(x) = 0$. ■

Exercise 2.8.9. Suppose that we shuffle a traditional deck of 52 cards in the following way: at each time $n \in \mathbb{N}$, we choose two cards uniformly at random and exchange them.

1. Model this process by a Markov chain. (Give its state space and transition matrix.)

2. Show that this chain is irreducible and find its unique invariant distribution.

Solution of Exercise 2.8.9.

1. We can see this process as a Markov chain $(X_n)_{n\geqslant 0}$ on the state space $E := \mathfrak{S}_{52}$ of permutations of $\{1,\dots,52\}$. The transition matrix is $Q(\sigma,\sigma') := \mathbb{1}_{\{\sigma\sim\sigma'\}}/\binom{52}{2}$, where $\sigma \sim \sigma'$ means that $\sigma'\sigma^{-1}$ is one of the $\binom{52}{2}$ transpositions in E.

2. Since the set of transpositions span the symmetric group, the chain is irreducible (eventually the chain will move from any deck configuration to any other with positive probability). It has therefore at most one invariant law. But we notice that Q is a symmetric matrix, so nonzero constant measures are trivially reversible and in particular invariant. We conclude that the uniform distribution on E is the unique invariant law.

Remark. The chain is irreducible, positive recurrent, but 2-periodic: it always moves from an even permutation to an odd one and *vice versa*, *i.e*, $\operatorname{sign}(X_n) = (-1)^n \operatorname{sign}(X_0)$, where $\operatorname{sign}\colon \mathfrak{S}_{52} \to \{-1,1\}$ is the signature morphism. However, if we allow at each step that the two uniformly chosen cards can be equal, the chain becomes *aperiodic* (since then $Q(\sigma,\sigma) > 0$, for all $\sigma \in E$), and we then have the convergence of X_n toward the uniform distribution as $n \to \infty$ (whatever the law of X_0 be). ■

Exercise 2.8.10. Let $(X_n)_{n\geqslant 0}$ be a Markov chain on a finite or countable state space E. Recall that $H_x := \inf\{n \geqslant 1\colon X_n = x\}$, $x \in E$, and, when x is recurrent, that

$$\nu_x(y) := \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}_{\{X_n=y\}}\right], \qquad y \in E$$

(the mean number of visits of y before returning to x), defines an invariant measure.

1. We suppose in this question that $(X_n)_{n\geqslant 0}$ is the symmetric random walk on $E = \mathbb{Z}$. Show that $\nu_0 \equiv 1$ (the mean number of visits of $y \in \mathbb{Z}$ before returning to 0 is 1).

2. We suppose in this question that $(X_n)_{n\geqslant 0}$ is irreducible and positive recurrent. Show that for every $x, y \in E$,
$$\nu_x(y) = \frac{\mathbb{E}_x[H_x]}{\mathbb{E}_y[H_y]}.$$

Solution of Exercise 2.8.10.

1. Recall that the symmetric random walk on $\mathbb{Z}$ is a recurrent Markov chain with transition matrix $Q(x,y) := (\mathbb{1}_{\{y=x-1\}} + \mathbb{1}_{\{y=x+1\}})/2$. Thus, ν_0 is an invariant measure ($\nu Q = \nu$) and therefore fulfills
$$\nu_0(k) = \frac{1}{2}\big(\nu_0(k-1) + \nu_0(k+1)\big), \qquad k \in \mathbb{Z},$$

that is to say,

$$\nu_0(k+1) - \nu_0(k) = \nu_0(k) - \nu_0(k-1), \qquad k \in \mathbb{Z},$$

with $\nu_0(0) = 1$ (by definition of H_0). We get that $\nu_0(k) = 1 + kC$, $k \in \mathbb{Z}$, for some constant C. Since ν_0 is a positive measure, we must have $C = 0$. Hence $\nu_0 \equiv 1$.

2. Since $(X_n)_{n\geqslant 0}$ is irreducible and positive recurrent, it admits a unique invariant probability measure given by

$$\mu(y) = \frac{1}{\mathbb{E}_y[H_y]}, \qquad y \in E,$$

and any other invariant measure must be proportional to μ. Consequently, for every $x \in E$, there exists $C > 0$ such that $\nu_x = C\mu$, and thus, for every $y \in E$,

$$\frac{\mathbb{E}_x[H_x]}{\mathbb{E}_y[H_y]} = \frac{\mu(y)}{\mu(x)} = \frac{\nu_x(y)}{\nu_x(x)} = \nu_x(y),$$

as stated. ∎

Exercise 2.8.11. Let $(X_n)_{n\geqslant 0}$ be a Markov chain on a finite or countable state space E. Recall that $H_x := \inf\{n \geqslant 1 \colon X_n = x\}$, $x \in E$, and, when x is recurrent, that

$$\nu_x(y) := \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}_{\{X_n=y\}}\right], \qquad y \in E$$

(the mean number of visits of y before returning to x), defines an invariant measure.

1. We suppose in this question that $(X_n)_{n\geqslant 0}$ is the symmetric random walk on $E = \mathbb{Z}$. Show that $\nu_0 \equiv 1$ (the mean number of visits of $y \in \mathbb{Z}$ before returning to 0 is 1).

2. We suppose in this question that $(X_n)_{n\geqslant 0}$ is irreducible and positive recurrent. Show that for every $x, y \in E$,

$$\nu_x(y) = \frac{\mathbb{E}_x[H_x]}{\mathbb{E}_y[H_y]}.$$

Solution of Exercise 2.8.11.

1. The symmetric random walk on $\mathbb{Z}$, with matrix $Q(x,y) := (\mathbb{1}_{\{y=x-1\}} + \mathbb{1}_{\{y=x+1\}})/2$, $x, y \in \mathbb{Z}$, is (as we know) recurrent, so ν_0 is invariant ($\nu_0 = \nu_0 Q$) and thus fulfills

$$\nu_0(k) = (\nu_0(k-1) + \nu_0(k+1))/2, \qquad k \in \mathbb{Z},$$

that is,

$$\nu_0(k+1) - \nu_0(k) = \nu_0(k) - \nu_0(k-1), \qquad k \in \mathbb{Z},$$

with $\nu_0(0) = 1$ (by definition of H_0). We get that $\nu_0(k) = 1 + Ck$, $k \in \mathbb{Z}$, for some constant C. Since ν_0 is a positive measure, we must have $C = 0$. Hence $\nu_0 \equiv 1$.

2. Being positive recurrent, $(X_n)_{n\geqslant 0}$ admits $\mu(y) := 1/\mathbb{E}_y[H_y]$, $y \in E$, as unique stationary law, and any other invariant measure differs by a multiplicative constant. Hence

$$\frac{\mathbb{E}_x[H_x]}{\mathbb{E}_y[H_y]} = \frac{\mu(y)}{\mu(x)} = \frac{\nu_x(y)}{\nu_x(x)} = \nu_x(y), \qquad x, y \in E.$$

■

Exercise 2.8.12. Let $(X_n)_{n\geqslant 0}$ be a Markov chain on $E := \mathbb{Z}$ with transition matrix

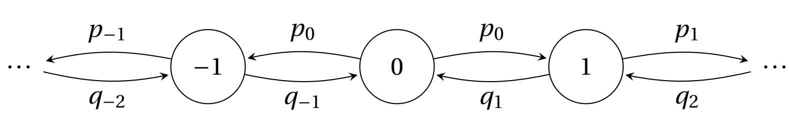
$$Q(i,j) := \begin{cases} p_i, & \text{if } j^+ = i^+ + 1 \text{ or } j^- = i^- + 1, \\ q_i, & \text{if } j^+ = i^+ - 1 \text{ or } j^- = i^- - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $p_i \in (0,1)$, $q_i := 1 - p_i$, for every $i \in E$.

1. Check that $(X_n)_{n\geqslant 0}$ is irreducible. (Sketch the transition graph.)
2. We suppose that
$$\limsup_{|k|\to\infty} p_k < \frac{1}{2}.$$
Show that $(X_n)_{n\geqslant 0}$ is (positive) recurrent.
Hint. Apply Foster–Lyapunov's criterion.

Solution of Exercise 2.8.12.

1. It is clear that $Q_{|j-i|}(i,j) > 0$ for all $i, j \in E$:



2. Let $f(i) := |i|$, $i \in E$. Note that, for $i \neq 0$,
$$\sum_{j\in E} Q(i,j)f(j) = Q(i,i-1)|i-1| + Q(i,i+1)|i+1| = f(i) + p_i - q_i. \tag{$\star$}$$
By assumption, there exists $k > 0$ such that, for every $i \notin F := \{i \in E : |i| \leqslant k\}$ (which is a finite subset of E), we have $p_i \leqslant 1/2$, so ($\star$) shows that $Qf(i) \leqslant f(i)$, *i.e*, $f : E \to \mathbb{R}_+$ is superharmonic on $E \setminus F$. Since further $\{x \in E : f(x) < M\}$ is finite for every $M > 0$, we conclude by Foster's criterion that $(X_n)_{n\geqslant 0}$ is recurrent.
Remark. Even better, the assumption says that $\sup_{|i|>k} p_i < 1/2$ for k large enough, so we have $\varepsilon > 0$, $f : E \to \mathbb{R}_+$, and $F \subset E$ finite (as before) such that, thanks to ($\star$),
$$\text{(i)}\ \ \forall i \in E \setminus F,\ Qf(i) \leqslant f(i) - \varepsilon; \qquad \text{(ii)}\ \ \forall i \in F,\ Qf(i) < \infty.$$

By another criterion of Foster, the irreducible chain $(X_n)_{n\geqslant 0}$ is *positive* recurrent. Indeed, for $H_F := \inf\{n \geqslant 1 : X_n \in F\}$, $Y_n := f(X_n)\mathbb{1}_{\{H_F>n\}}$, $n \geqslant 0$, and $i \notin F$, we have

$$\mathbb{E}_i[Y_{n+1} \mid X_0,\dots,X_n] \leqslant Qf(X_n)\mathbb{1}_{\{H_F>n\}} \leqslant Y_n - \varepsilon\mathbb{1}_{\{H_F>n\}}$$

(using $\{H_F > n+1\} \subseteq \{H_F > n\}$ for the first inequality, and (i) for the second). Since $\mathbb{E}_i[Y_n] \geqslant 0$, it follows by taking expectations above that $\mathbb{P}_i(H_F > n)$, $n \geqslant 0$, is summable; more precisely $\mathbb{E}_i[H_F] \leqslant \mathbb{E}_i[Y_0]/\varepsilon = f(i)/\varepsilon$. Now for $i \in F$, the Markov property at first step yields

$$\mathbb{E}_i[H_F] = 1 + \mathbb{E}_i\big[(H_F \circ \theta_1)\mathbb{1}_{\{X_1\notin F\}}\big] \leqslant 1 + Qf(i)/\varepsilon,$$

which is bounded, by (ii) and finiteness of F. The positive recurrence follows. ■

Exercise 2.8.13. Let $(X_n)_{n\geqslant 0}$ be an irreducible Markov chain on E. We suppose that there exist a finite subset $F \subseteq E$ and a function $f : E \to \mathbb{R}$ such that

$$\text{(i)}\ \ \forall x \in F,\ f(x) > 0; \quad \text{(ii)}\ \ \inf_{x\in E} f(x) = 0; \quad \text{(iii)}\ \ \forall x \in E\setminus F,\ \mathbb{E}_x[f(X_1)] \leqslant f(x).$$

Show that $(X_n)_{n\geqslant 0}$ is transient.

Hint. Introduce the hitting time $T_F := \inf\{n \geqslant 0 : X_n \in F\}$...

Solution of Exercise 2.8.13. By (ii), f is nonnegative. By (i) and finiteness of F, we have $\alpha := \inf_{x\in F} f(x) > 0$. Fix $x \in E$ arbitrary. By (iii), we also know that the process $Y_n := f(X_{n\wedge T_F})$, $n \geqslant 0$, is a nonnegative $\mathbb{P}_x$-supermartingale (constant if $x \in F$); in particular,

$$\forall n \geqslant 0, \quad \mathbb{E}_x[Y_n] \leqslant \mathbb{E}_x[Y_0] = f(x).$$

Now, suppose that $(X_n)_{n\geqslant 0}$ is recurrent. In this case $\mathbb{P}_x(T_F < \infty) = 1$, $X_{T_F} \in F$, and thus

$$\lim_{n\to\infty} Y_n = f(X_{T_F}) \geqslant \alpha.$$

Fatou's lemma then yields $f(x) \geqslant \alpha$, for any $x \in E$. Since this contradicts (ii), we conclude that $(X_n)_{n\geqslant 0}$ must be transient. ■

Exercise 2.8.14. Let Q be a *symmetric*, irreducible, aperiodic transition matrix on E, and μ be a probability measure on E such that $\mu(x) > 0$ for all $x \in E$. We set

$$P(x,y) := Q(x,y)\min\left(1, \frac{\mu(y)}{\mu(x)}\right), \qquad \text{for } x \neq y \in E.$$

1. Check that P extends to a transition matrix which is also irreducible and aperiodic.

We consider a Markov chain $(X_n)_{n\geqslant 0}$ on E with transition matrix P.

2. Show that μ is an invariant measure for P, and that $(X_n)_{n\geqslant 0}$ is positive recurrent.
3. Let U_n, $n \in \mathbb{N}$, be a r.v. independent of $(X_k)_{k\geqslant 0}$, with $\mathbb{P}(U_n = i) = 1/n$, $0 \leqslant i < n$. Show that for every $x \in E$,
$$\sum_{x\in E}\big|\mathbb{P}(X_{U_n} = x) - \mu(x)\big| \xrightarrow[n\to\infty]{} 0.$$

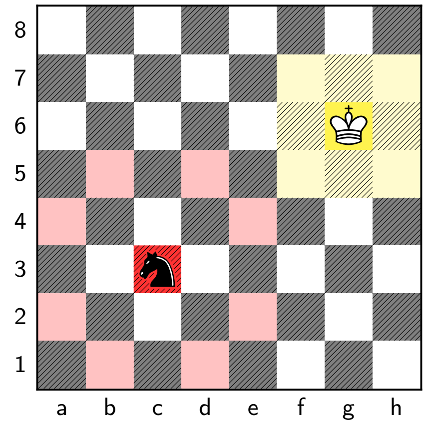
Solution of Exercise 2.8.14.

1. Because Q is a transition matrix we have $\sum_{y\neq x} P(x,y) \leqslant 1$ for all $x \in E$, so P extends to a transition matrix by setting $P(x,x)$ appropriately. Further, since $\mu(x) > 0$ for every $x \in E$, we can see that
$$\{n \geqslant 0\colon Q_n(x,y) > 0\} \subseteq \{n \geqslant 0\colon P_n(x,y) > 0\}. \qquad x,y \in E.$$
Since Q is irreducible and aperiodic, so is P.

2. We check that μ is reversible for P: for $x \neq y \in E$, using that Q is symmetric,
$$\mu(x)\,P(x,y) = Q(x,y)\,\min\big(\mu(x),\mu(y)\big) = \mu(y)\,P(y,x)$$
It follows that μ is an invariant measure for P. Since P is irreducible, μ is the unique stationary distribution and $(X_n)_{n\geqslant 0}$ is positive recurrent.

3. On the one hand, since $(X_n)_{n\geqslant 0}$ is an irreducible, recurrent and aperiodic Markov chain with invariant probability measure μ, we have the convergence to equilibrium:
$$a_n := \sum_{x\in E} \big|\mathbb{P}(X_n = x) - \mu(x)\big| \xrightarrow[n\to\infty]{} 0.$$
On the other hand, by independence between U_n and $(X_k)_{k\geqslant 0}$,
$$b_n(x) := \mathbb{P}(X_{U_n} = x) = \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{P}(X_k = x), \qquad n \in \mathbb{N}.$$
Therefore
$$\begin{aligned}\sum_{x\in E} \big|\mathbb{P}(X_{U_n} = x) - \mu(x)\big| &\leqslant \frac{1}{n}\sum_{x\in E}\sum_{k=0}^{n-1} \big|\mathbb{P}(X_k = x) - \mu(x)\big| && (\Delta\text{-inequality})\\ &= \frac{1}{n}\sum_{k=0}^{n-1} a_k && (\text{Fubini})\\ &\xrightarrow[n\to\infty]{} 0. && (\text{Cesàro})\end{aligned}$$
Remark. Without assumption of aperiodicity, we still have $b_n(x) \xrightarrow[n\to\infty]{} \mu(x)$ for every $x \in E$: indeed, by positive recurrence,
$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{\{X_k = x\}} \xrightarrow[n\to\infty]{} \mu(x), \qquad \text{a.s.},$$
and the dominated convergence theorem allows us to take the expectations. ∎

Exercise 2.8.15. We consider the simple random walks of the knight and the king on a classical chessboard, $E := \{\mathsf{a},\dots,\mathsf{h}\} \times \{1,\dots,8\}$. Authorized moves are recalled below.



1. Starting in a8, what is the expected time for the king to return to a8? In the meantime, how many visits in the four squares {d4, e4, e5, d5} will he have performed?

   Hint. Use Exercise 2.8.11.2.

2. At which frequency does the knight visit square g6, as time tends to infinity?

Solution of Exercise 2.8.15. It is plain that both random walks are irreducible. Since the state space is finite, they are positive recurrent. Further, the (unique) invariant distribution of a simple random walk on a finite graph is reversible; in our setting, it is given by

$$\frac{1}{\mathbb{E}_x[H_x]} =: \mu(x) = \frac{\operatorname{card} A_x}{\sum\limits_{y\in E} \operatorname{card} A_y}, \qquad x \in E,$$

where the card A_y, giving the number of squares accessible in a single move from $y \in E$ by the king (respectively, the knight), are recorded below.

8	3	5	5	5	5	5	5	3
7	5	8	8	8	8	8	8	5
6	5	8	8	8	8	8	8	5
5	5	8	8	8	8	8	8	5
4	5	8	8	8	8	8	8	5
3	5	8	8	8	8	8	8	5
2	5	8	8	8	8	8	8	5
1	3	5	5	5	5	5	5	3
	a	b	c	d	e	f	g	h

(a) Accessible squares for the king.

8	2	3	4	4	4	4	3	2
7	3	4	6	6	6	6	4	3
6	4	6	8	8	8	8	6	4
5	4	6	8	8	8	8	6	4
4	4	6	8	8	8	8	6	4
3	4	6	8	8	8	8	6	4
2	3	4	6	6	6	6	4	3
1	2	3	4	4	4	4	3	2
	a	b	c	d	e	f	g	h

(b) Accessible squares for the knight.

1. For the king, using the table (a) we find $\mu(\mathsf{a8}) = 3/420 = 1/140$. The king thus returns to its starting point in $\mathbb{E}_{\mathsf{a8}}[H_{\mathsf{a8}}] = 1/\mu(\mathsf{a8}) = 140$ time units on average. Using the notations of Exercise 2.8.11, the average time spent on the four central squares between two visits of a8 is

$$\nu_{\mathsf{a8}}(\{\mathsf{d4}, \mathsf{e4}, \mathsf{e5}, \mathsf{d5}\}) = \frac{1}{\mu(\mathsf{a8})}\big(\mu(\mathsf{d4}) + \mu(\mathsf{e4}) + \mu(\mathsf{e5}) + \mu(\mathsf{d5})\big) = 140\left(\frac{2}{105} \times 4\right) = \frac{32}{3}.$$

2. For the knight, using the table (b) we find $\mu(\texttt{g6}) = 6/336 = 1/56$. By the ergodic theorem, this is also the a.s. asymptotic frequency of visits to g6. ■

Exercise 2.8.16. Let $(X_n)_{n\geqslant 0}$ be a Markov chain on a *finite* state space E, with transition matrix Q. We call a state $x \in E$ *absorbing*, and we write $x \in A$, if $Q(x,x) = 1$. We suppose $r := \sharp A \geqslant 1$ and A *accessible*: $\forall x \in E$, $\exists n \in \mathbb{N}$, $Q^n(x,A) > 0$.

1. Is $(X_n)_{n\geqslant 0}$ irreducible?
2. Let I_r denote the $r \times r$ identity matrix. Check that we may write Q in the form
$$Q := \left(\begin{array}{c|c} P & T \\ \hline 0 & I_r \end{array}\right).$$
3. Let $H_A := \inf\{n > 0 \colon X_n \in A\}$.
   a) Show that for all $i, j \notin A$, $P^n(i,j) \leqslant \mathbb{P}_i(H_A > n)$.
   b) Show that there exists $M \geqslant 1$ such that
   $$p := \sup_{i\notin A} \mathbb{P}_i(H_A > M) < 1.$$
   Hint. You can take $M := \sup_{i\notin A} m_i$, where $m_i := \inf\{n > 0 \colon \mathbb{P}_i(X_n \in A) > 0\}$.
   c) Deduce that $\mathbb{P}_i(H_A = \infty) = 0$ and $P^n(i,j) \to 0$ for all $i, j \notin A$.
   Hint. Check that $\sup_{i\notin A} \mathbb{P}_i(H_A \geqslant Mn) \leqslant p^n$ (use the Markov property).
4. Let $s := \sharp E - r$. Show that $I_s - P$ is invertible and that, for $F := (I_s - P)^{-1}$,
$$\lim_{n\to\infty} Q^n = \left(\begin{array}{c|c} 0 & FT \\ \hline 0 & I_r \end{array}\right).$$
   Hint. Prove that 1 is not an eigenvalue of P.
5. a) Check that for all $i, j \notin A$,
   $$F(i,j) = \mathbb{E}_i\left[\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = j\}}\right].$$
   b) Show that $\sum_{j\notin A} F(i,j) = \mathbb{E}_i[H_A]$ for all $i \notin A$.
   c) Show that $FT(i,j) = \mathbb{P}_i(X_{H_A} = j)$ for all $i \notin A$ and $j \in A$.

Solution of Exercise 2.8.16.

1. No (unless $r = \sharp E = 1$): from $i \in A$, no other state $j \neq i$ is accessible.
2. We order the absorbing states, so that Q is triangular by blocks with the bottom-right block $I_r = Q_{|A\times A}$ corresponding to the transition probabilities of the absorbing states.

3. a) Since $(X_n)_{n\geqslant 0}$ cannot escape A and $j \notin A$, $\{X_n = j\} \subseteq \{X_n \notin A\} = \{H_A > n\}$, so $P^n(i,j) \leqslant \mathbb{P}_i(X_n = j) \leqslant \mathbb{P}_i(H_A > n)$.

   b) With M as given, the chain can reach A from any state $i \notin A$ in at most M steps. Indeed:

$$\mathbb{P}_i(H_A \leqslant M) \geqslant \mathbb{P}_i(H_A \leqslant m_i) \geqslant \mathbb{P}_i(X_{m_i} \in A) > 0.$$

   Because $E \setminus A$ is finite, it follows that $p < 1$.

   c) Let $n \geqslant 2$ and $i \notin A$. By the Markov property at time M,

$$\begin{aligned}\mathbb{P}_i(H_A > Mn) &= \mathbb{P}_i(X_{Mn} \notin A)\\ &= \mathbb{E}_i\big[\mathbb{P}_i(X_M \notin A, X_{M+(n-1)M} \notin A \mid \mathscr{F}_M)\big]\\ &= \mathbb{E}_i\big[\mathbb{1}_{X_M \notin A}\mathbb{P}_{X_M}(X_{(n-1)M} \notin A)\big]\\ &\leqslant p\,\mathbb{P}_i(H_A > M(n-1)).\end{aligned}$$

   Hence $\mathbb{P}_i(H_A > Mn) \leqslant p^n$ for all $i \notin A$. Letting $n \to \infty$, we get $\mathbb{P}_i(H_A = \infty) = 0$ and also, by Question 1, $P^n(i,j) \to 0$ for all $j \notin A$.

4. Let $x \in \mathbb{R}^s$ with $Px = x$. Then $x = P^n x \to 0$ as $n \to \infty$, by Question 3.c). Hence 1 is not an eigenvalue of P, *i.e*, $I_s - P$ is invertible. The stated limit for Q^n then follows from the facts that $F := (I_s - P)^{-1} = \sum_{n=0}^{\infty} P^n$ and (by immediate induction on n)

$$Q^n = \left(\begin{array}{c|c} P^n & (I_s + P + \cdots + P^{n-1})T \\ \hline 0 & I_r \end{array}\right).$$

5. a) For all $i, j \notin A$, we have by Fubini–Tonelli's theorem,

$$F(i,j) = \sum_{n=0}^{\infty} P^n(i,j) = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n = j) = \mathbb{E}_i\Bigg[\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = j\}}\Bigg].$$

   b) For all $i \notin A$,

$$\sum_{j\notin A} F(i,j) = \sum_{n=0}^{\infty}\sum_{j\notin A} \mathbb{P}_i(X_n = j) = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n \notin A) = \sum_{n=0}^{\infty} \mathbb{P}_i(H_A > n) = \mathbb{E}_i[H_A].$$

   c) Let $i \in A$ and $j \notin A$. We see from the simple Markov property at time $n \geqslant 0$ that

$$\mathbb{P}_i(H_A = n+1, X_{n+1} = j) = \mathbb{P}_i(X_n \notin A, X_{n+1} = j) = \sum_{k\notin A} \mathbb{P}_i(X_n = k)\mathbb{P}_k(X_1 = j).$$

   Therefore $FT(i,j)$ equals

$$\begin{aligned}\sum_{k\notin A} F(i,k)\,T(k,j) &= \sum_{n=0}^{\infty}\sum_{k\notin A} \mathbb{P}_i(X_n = k)\mathbb{P}_k(X_1 = j)\\ &= \sum_{n=0}^{\infty} \mathbb{P}_i(H_A = n+1, X_{n+1} = j)\\ &= \mathbb{P}_i(X_{H_A} = j).\end{aligned}$$

■

Exercise 2.8.17. Consider the Markov chain on $E := \{1,2,3,4,5,6\}$ with transition matrix:

$$Q := \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 & 0 & 0 \\ 0.3 & 0 & 0.4 & 0.2 & 0.1 & 0 \\ 0 & 0 & 0 & 0.3 & 0.7 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0.8 & 0 & 0.2 \end{pmatrix}.$$

1. Draw the transition graph.

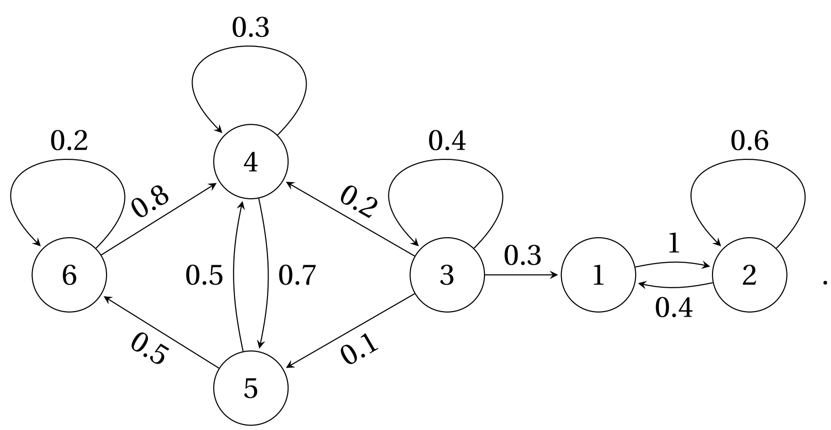
2. Give the recurrence/transience classes.

3. Compute $\mathbb{P}_3(X_n \in \{4,5,6\} \text{ eventually})$.

   Hint. Use the Markov property.

Solution of Exercise 2.8.17.

1. The transition graph is



2. We easily see that the recurrence classes are $\{1,2\}$ and $\{4,5,6\}$, and that 3 is transient.

3. We have $\mathbb{P}_3(X_n \in \{4,5,6\} \text{ eventually}) = 1 - \mathbb{P}_3(X_n = 1 \text{ eventually})$ and, by the Markov property at time 1,

$$\begin{aligned} p := \mathbb{P}_3(X_n = 1 \text{ eventually}) &= \mathbb{P}_3(X_1 = 1) + \mathbb{P}_3(X_1 = 3, X_n = 1 \text{ eventually}) \\ &= 0.3 + \mathbb{E}_3\big[\mathbb{1}_{\{X_1=3\}}\,\mathbb{P}_{X_1}(X_n = 1 \text{ eventually})\big] \\ &= 0.3 + 0.4\,p, \end{aligned}$$

which solves to $p = \frac{0.3}{1-0.4} = 0.5$. Hence $\mathbb{P}_3(X_n \in \{4,5,6\} \text{ eventually}) = 1 - p = 0.5$. ∎

Exercise 2.8.18. Let $(Y_n)_{n\geqslant 0}$ be the symmetric random walk on $\mathbb{Z}$, that is $Y_n = Y_0 + \sum_{i=1}^n \xi_i$, $n \geqslant 0$, with ξ, $i \geqslant 1$, i.i.d. uniform ± 1 r.v. independent of $Y_0 \in \mathrm{L}^1$. Let $H_{-1} := \inf\{n > 0 \colon Y_n = -1\}$.

1. Let $k \in \mathbb{Z}_+$.

   a) What is $\mathbb{P}_0(H_{-1} = 2k)$?

   b) Compute $\mathbb{P}_0(Y_{2k+1} = -1)$.
   Hint. Under $\mathbb{P}_0$, $\{Y_{2k+1} = -1\}$ means that exactly k of the $\xi_1, \dots, \xi_{2k+1}$ equal $+1$...

   c) Let $(x_i) \in \{\pm 1\}^{2k+1}$ with $x_1 + \cdots + x_{2k+1} = -1$. Check that there is one and only one $1 \leqslant r \leqslant 2k+1$ such that, if we set $\widetilde{x} := (x_{r+1}, \dots, x_{2k+1}, x_1, \dots, x_r)$, then

$$\forall\, j \leqslant 2k, \quad \sum_{i=1}^{j} \widetilde{x}_i \geqslant 0.$$

   Suggestion. Do a drawing.

   d) Deduce that $\mathbb{P}_0(H_{-1} = 2k+1) = \frac{1}{2k+1} \mathbb{P}_0(Y_{2k+1} = -1)$.

2. Give an equivalent of $\mathbb{P}_0(H_{-1} = 2k+1)$ as $k \to \infty$.
   Hint. Use Stirling's formula.

3. Conclude that $\mathbb{E}_0[H_{-1}] = \infty$.

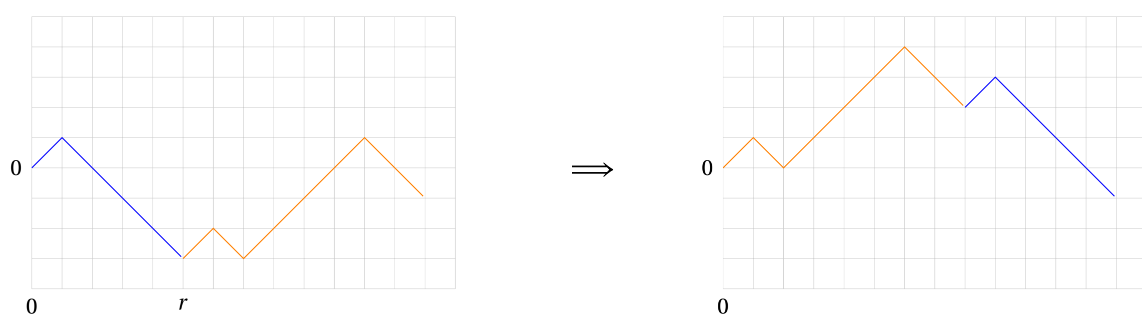
Solution of Exercise 2.8.18.

1. a) From 0, the chain can only reach even states at even times, so $\mathbb{P}_0(H_{-1} = 2k) = 0$.

   b) Under $\mathbb{P}_0$, $\{Y_{2k+1} = -1\}$ means that among $\xi_1, \dots, \xi_{2k+1}$, exactly k variables equal $+1$, while the $k+1$ other equal -1. It follows that $\mathbb{P}_0(Y_{2k+1} = -1)$ is the probability of k successes of a Binomial$(2k+1, \frac{1}{2})$ r.v., so

$$\mathbb{P}_0(Y_{2k+1} = -1) = 2^{-(2k+1)} \binom{2k+1}{k}.$$

   c) Let $r \in \{1, \dots, 2k+1\}$ be the first index realizing the minimum in the sequence of partial sums of $x_1 + \cdots + x_{2k+1}$:



   It is clear from the figure that the $2k$ first partial sums of $(x_{r+1}, \dots, x_{2k+1}, x_1, \dots, x_r)$ are all nonnegative. Consider now any other shift $(x_{j+1}, \dots, x_{2k+1}, x_1, \dots, x_j)$ of x. If $j < r$, then $x_{j+1} + \cdots + x_r < 0$ by definition of r as being the *first* index realizing the minimum. If $j > r$, then $x_{j+1} + \cdots + x_{2k+1} + x_1 \cdots + x_r = -1 - \sum_{i=r+1}^{j} \xi_i < 0$ because $\sum_{i=1}^{j} \xi_i \geqslant \sum_{i=1}^{r} \xi_i$. In any case, at least one partial sum is negative, which proves the uniqueness of r as above.

d) Let $x \in \{\pm 1\}^{2k+1}$ and $\widetilde{x}$ as above. The family $(\xi_i)_{i \geqslant 1}$ is i.i.d., so

$$\mathbb{P}_0\big((\xi_1,\dots,\xi_{2k+1}) = x\big) = \mathbb{P}_0\big((\xi_1,\dots,\xi_{2k+1}) = \widetilde{x}\big).$$

Summing over all possible x, we obtain $\mathbb{P}_0(Y_{2k+1} = -1)$ for the left-hand side. For the right-hand side, we note that in the union $\bigcup_x \{(\xi_1,\dots,\xi_{2k+1}) = \widetilde{x}\} = \{H_{-1} = 2k+1\}$, each event is repeated exactly $(2k+1)$ times, so we get $(2k+1)\,\mathbb{P}_0(H_{-1} = 2k+1)$. Hence the result.

2. By Stirling's formula,

$$\mathbb{P}_0(H_{-1} = 2k+1) = \frac{2^{-(2k+1)}}{2k+1}\binom{2k+1}{k} = \frac{2^{-(2k+1)}}{k+1} \cdot \frac{(2k)!}{k!^2} \sim \frac{2^{-(\cancel{2k}+1)}}{k} \frac{\big(\frac{\cancel{2k}}{\cancel{e}}\big)^{2k} \sqrt{\cancel{4}\pi k}}{\big(\frac{\cancel{k}}{\cancel{e}}\big)^{2k} \cancel{2}\pi k},$$

that is

$$\mathbb{P}_0(H_{-1} = 2k+1) \sim \frac{1}{2k\sqrt{\pi k}}, \qquad k \to \infty.$$

3. We deduce by comparison with a Riemann sum that

$$\sum_{k=0}^{n} (2k+1)\,\mathbb{P}_0(H_{-1} = 2k+1) \sim \frac{2\sqrt{n}}{\sqrt{\pi}}, \qquad n \to \infty,$$

so $\mathbb{E}_0[H_{-1}] = \infty$. (Actually, the argument even shows that $\mathbb{E}_0[\sqrt{H_{-1}}] = \infty$.) ∎

Exercise 2.8.19. Using Exercise 2.7.11.5, give a simple proof that every irreducible, centered, finite-range random walk on $\mathbb{Z}$ is recurrent.

Solution of Exercise 2.8.19. Such a random walk can be written $X_n = X_0 + \sum_{i=1}^n \xi_i$, $n \geqslant 0$, where the variables ξ_i, $i \geqslant 1$, are i.i.d., independent of X_0, with $\mathbb{P}(|\xi_1| > 2K) = 0$ for some $K > 0$ (finite range) and $\mathbb{E}[\xi_1] = 0$ (centered). The process $(X_n)_{n\geqslant 0}$ is clearly a Markov chain with transition matrix $Q(x,y) := \mathbb{P}(\xi_1 = y - x)$, $x, y \in \mathbb{Z}$, but it is also a martingale whose increments are independent and dominated in L^1, so we may apply Exercise 2.7.11.5 together with Kolmogorov's 0-1 law. Then, either X_n converges a.s. to some r.v. $X_\infty \in \mathbb{Z}$ or oscillates a.s. In the first case, the law of X_∞ is an invariant probability measure[4] for the irreducible Markov chain $(X_n)_{n\geqslant 0}$, which is then recurrent. In the second case, the set $\{-K, -K+1, \dots, K\}$ is visited infinitely often a.s., which also entails by irreducibility that $(X_n)_{n\geqslant 0}$ is recurrent. ∎

[4]Indeed, passing the equality $\mathbb{P}(X_{n+1} = x) = \mathbb{E}[\mathbb{P}(X_{n+1} = x \mid X_n)] = \mathbb{E}[Q(X_n, x)]$ to the limit as $n \to \infty$ yields (by dominated convergence) $\mu(x) := \mathbb{P}(X_\infty = x) = \mathbb{E}[Q(X_\infty, x)]$ for all $x \in \mathbb{Z}$, that is $\mu = \mu Q$.

CHAPTER 3

COMBINATORICS OF INTEGER PARTITIONS

The following exercises are due to Jehanne Dousse.

3.1 Generating functionology

Exercise 3.1.1. List all partitions of 6.

Solution of Exercise 3.1.1. There are 11 partitions of 6:

$$
\begin{aligned}
6 &= 6 & &= 3+1+1+1\\
&= 5+1 & &= 2+2+2\\
&= 4+2 & &= 2+2+1+1\\
&= 4+1+1 & &= 2+1+1+1+1\\
&= 3+3 & &= 1+1+1+1+1+1.\\
&= 3+2+1
\end{aligned}
$$

■

Exercise 3.1.2.

1. List all partitions of 6 into even parts, and those in which each part occurs an even number of times. What do you notice?
2. Explain why, for n odd,
$$p(n \mid \text{even parts}) = p(n \mid \text{each part occurs an even number of times}) = 0.$$
3. Show that for all $n \in \mathbb{N}$,
$$p(n \mid \text{even parts}) = p(n \mid \text{each part occurs an even number of times}).$$

Solution of Exercise 3.1.2.

1. We notice that there are 3 partitions of 6 into even parts (highlighted in green above), and also 3 partitions of 6 in which each part occurs an even number of times (highlighted in red).

2. A partition with only even parts must sum to an even number. Likewise, a partition in which each part occurs an even number of times must sum to an even number.

3. Clearly, the map $(2k_1 + \cdots + 2k_r) \mapsto (k_1 + k_1) + \cdots + (k_r + k_r)$ forms a bijection between the set of partitions of n into even parts and the set of partitions of n in which each part occurs an even number of times.

   Remark. The generating function for partitions into even parts is

$$\sum_{n\geqslant 0} p(n \mid \text{even parts})\, q^n = \prod_{k \text{ even}} (1 + q^k + q^{2k} + \cdots) = \prod_{k \text{ even}} \frac{1}{1-q^k},$$

   while the generating function for partitions in which each part occurs an even number of times is

$$\begin{aligned}\sum_{n\geqslant 0} p(n \mid \text{each part occurs an even number of times})\, q^n &= \prod_{k=1}^{\infty} (1 + q^{2k} + q^{4k} + \cdots) \\ &= \prod_{k=1}^{\infty} \frac{1}{1-q^{2k}}.\end{aligned}$$

   (They are the same!) ■

Exercise 3.1.3.

1. What is the generating function for partitions into distinct parts equal to 2, 5 or 7?

2. What is the generating function for partitions into parts equal to 2, 5 or 7, such that each part occurs at most d times ($d \in \mathbb{N}$)?

3. What is the generating function for partitions into parts equal to 2, 5 or 7?

Solution of Exercise 3.1.3.

1. The generating function for the partitions of n into distinct parts equal to 2, 5 or 7 is

$$\begin{aligned}\sum_{n\geqslant 0} Q(n \mid \text{parts} = 2,5,7)\, q^n &= \sum_{i,j,k\in\{0,1\}} q^{2i+5j+7k} \\ &= (1+q^2)\cdot(1+q^5)\cdot(1+q^7) \\ &= 1 + q^2 + q^5 + 2q^7 + q^9 + q^{12} + q^{14}.\end{aligned}$$

2. The generating function for the partitions of n into distinct parts equal to 2, 5, or 7 and where each part occurs at most $d \in \mathbb{N}$ times is

$$\begin{aligned}
&\sum_{n\geqslant 0} p(n \mid \text{parts} = 2,5,7;\ \text{each part occurs} \leqslant d \text{ times})\, q^n \\
&\quad = \sum_{i,j,k\in\{0,\dots,d\}} q^{2i+5j+7k} \\
&\quad = (1+q^2+q^4+\cdots+q^{2d})\cdot(1+q^5+q^{10}+\cdots+q^{5d})\cdot(1+q^7+q^{14}+\cdots+q^{7d}) \\
&\quad = \frac{1-q^{2(d+1)}}{1-q^2}\cdot\frac{1-q^{5(d+1)}}{1-q^5}\cdot\frac{1-q^{7(d+1)}}{1-q^7}.
\end{aligned}$$

3. The generating function for the partitions of n into parts equal to 2, 5, or 7 is

$$\begin{aligned}
\sum_{n\geqslant 0} p(n \mid \text{parts} = 2,5,7)\, q^n &= \sum_{i,j,k\geqslant 0} q^{2i+5j+7k} \\
&= \left(\sum_{i\geqslant 0} q^{2i}\right)\cdot\left(\sum_{j\geqslant 0} q^{5j}\right)\cdot\left(\sum_{k\geqslant 0} q^{7k}\right) \\
&= \frac{1}{1-q^2}\cdot\frac{1}{1-q^5}\cdot\frac{1}{1-q^7}.
\end{aligned}$$

■

Exercise 3.1.4. What generating function would you compute and what coefficient would you extract if you wanted to know the number of ways of changing a 100 CHF bill into coins of 1, 2 and 5 CHF and bills of 10 and 20 CHF?

Solution of Exercise 3.1.4. We want the coefficient in q^{100} of the generating function for the partitions of n into parts equal to $1,2,5,10,20$, that is

$$[q^{100}]\,\frac{1}{1-q}\cdot\frac{1}{1-q^2}\cdot\frac{1}{1-q^5}\cdot\frac{1}{1-q^{10}}\cdot\frac{1}{1-q^{20}} \qquad (=4111).$$

■

Exercise 3.1.5. What is the generating function for partitions into parts $\leqslant 2k$ ($k \in \mathbb{N}$) where odd parts cannot repeat?

Solution of Exercise 3.1.5.

$$\begin{aligned}
\sum_{n=0}^{\infty} p(n \mid \text{parts} \leqslant 2k;\ \text{odd parts cannot repeat})\, q^n &= \frac{(1+q)(1+q^3)\cdots(1+q^{2k-1})}{(1-q^2)(1-q^4)\cdots(1-q^{2k})} \\
&= \frac{(-q;q^2)_k}{(q^2;q^2)_k}.
\end{aligned}$$

■

Exercise 3.1.6. Give the generating function for

$$\left(\frac{n^2+4n+5}{n!}\right)_{n\geqslant 0}.$$

Solution of Exercise 3.1.6.

$$\begin{aligned}\sum_{n\geqslant 0}\frac{n^2+4n+5}{n!}q^n &= \sum_{n\geqslant 0}\frac{n(n-1)+5n+5}{n!}q^n\\ &= q^2\sum_{n\geqslant 2}\frac{q^{n-2}}{(n-2)!}+5q\sum_{n\geqslant 1}\frac{q^{n-1}}{(n-1)!}+5\sum_{n\geqslant 0}\frac{q^n}{n!}\\ &= (q^2+5q+5)\,e^q.\end{aligned}$$

■

Exercise 3.1.7.

1. Show that if f is the generating function for $(a_n)_{n\geqslant 0}$, then $\frac{f}{1-X}$ is the generating function for $(\sum_{j=0}^n a_j)_{n\geqslant 0}$.
2. Give the generating function for
$$\left(\sum_{j=0}^n j\right)_{n\geqslant 0}.$$
3. Show that if f is the generating function for $(a_n)_{n\geqslant 0}$, then f^k is the generating function for
$$\left(\sum_{n_1+\cdots+n_k=n} a_{n_1}\cdots a_{n_k}\right)_{n\geqslant 0}.$$
4. Recover the classical formula
$$\sum_{j=0}^n j = \frac{n(n+1)}{2}.$$

Solution of Exercise 3.1.7.

1. Two methods. By applying the product formula:
$$\frac{f(X)}{1-X} = \left(\sum_{n=0}^\infty a_n X^n\right)\left(\sum_{n=0}^\infty X^n\right) = \sum_{n=0}^\infty\left(\sum_{k=0}^n a_k\right)X^n,$$
or by inverting sums:
$$\sum_{n=0}^\infty\left(\sum_{k=0}^n a_k\right)X^n = \sum_{0\leqslant k\leqslant n} a_k X^n = \sum_{k=0}^\infty a_k\left(\sum_{n=k}^\infty X^n\right) = \sum_{k=0}^\infty a_k\frac{X^k}{1-X} = \frac{f(X)}{1-X}.$$
2. We know that $(1-X)^{-1}$ is the generating function for the sequence $n\mapsto 1$. Applying twice Question 1, we deduce that $(1-X)^{-2}$ is the generating function for
$$n\mapsto\sum_{j=0}^n 1 = n+1,$$

and then that $(1-X)^{-3}$ is the generating function for

$$n \longmapsto \sum_{j=0}^{n}(j+1) = (n+1) + \sum_{j=0}^{n} j.$$

By difference, $(1-X)^{-3} - (1-X)^{-2} = X(1-X)^{-3}$ is the generating function for

$$\left(\sum_{j=0}^{n} j\right)_{n\geqslant 0}.$$

Remark. We can conclude with just one application of Question 1 to

$$\frac{X}{(1-X)^2} = X \cdot \left(\frac{1}{1-X}\right)' = \sum_{n=0}^{\infty} nX^n.$$

3. Either by induction on k, or directly:

$$\begin{aligned}\sum_{n=0}^{\infty}\left(\sum_{n_1+\cdots+n_k=n} a_{n_1}\cdots a_{n_k}\right)X^n &= \sum_{n=0}^{\infty}\sum_{n_1+\cdots+n_k=n}(a_{n_1}X^{n_1})\cdots(a_{n_k}X^{n_k})\\ &= \sum_{n_1,\ldots,n_k\geqslant 0}(a_{n_1}X^{n_1})\cdots(a_{n_k}X^{n_k})\\ &= \left(\sum_{n=0}^{\infty} a_n X^n\right)^k\\ &= f(X)^k.\end{aligned}$$

4. On the one hand,

$$\frac{X}{(1-X)^3} = \sum_{n=0}^{\infty}\left(\sum_{j=0}^{n} j\right)X^n.$$

On the other hand, using Question 3,

$$\frac{X}{(1-X)^3} = \sum_{n=0}^{\infty}\underbrace{\left(\sum_{n_1+n_2+n_3=n} 1\right)}_{=\binom{n+2}{2}} X^{n+1} = \sum_{n=0}^{\infty}\frac{n(n+1)}{2}X^n. \qquad (\star)$$

We conclude by identifying the terms in X^n.

Remark. We can also obtain ($\star$) by derivation:

$$\frac{X}{(1-X)^3} = X\cdot\frac{1}{2}\left(\frac{1}{1-X}\right)'' = \sum_{n=0}^{\infty}\frac{n(n+1)}{2}X^n. \qquad \blacksquare$$

Exercise 3.1.8. Prove that for all $n \geqslant 0$,

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Hint. Compute a well-chosen generating function.

Solution of Exercise 3.1.8. Fix $n \in \mathbb{N}$. Since $a_k := \binom{n}{k}$ equals $\binom{n}{n-k}$ for $0 \leqslant k \leqslant n$, we have

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \sum_{k=0}^{n} a_k\, a_{n-k}.$$

By the product formula, this is the term in X^n of the square of the series

$$\sum_{k=0}^{\infty} a_k X^k = \sum_{k=0}^{n} \binom{n}{k} X^k = (1+X)^n$$

(using that $a_k = 0$ for $k > n$, and the Binomial theorem). Thus

$$\sum_{k=0}^{n} \binom{n}{k}^2 = [X^n]\,(1+X)^{2n} = [X^n] \sum_{k=0}^{2n} \binom{2n}{k} X^k = \binom{2n}{n}$$

(by another use of the Binomial theorem). ∎

Exercise 3.1.9. Let $(a_n)_{n\geqslant 0}$ be a sequence defined by $a_0 = 0$ and for all $n \geqslant 1$,

$$a_n = 2a_{n-1} + 1. \tag{$\star$}$$

What is the generating function for $(a_n)_{n\geqslant 0}$?

Solution of Exercise 3.1.9. The recurrence relation ($\star$) gives

$$\sum_{n\geqslant 1} a_n X^n = 2X \sum_{n\geqslant 1} a_{n-1} X^{n-1} + \sum_{n\geqslant 1} X^n,$$

that is (since $a_0 = 0$)

$$A(X) = 2XA(X) + \frac{X}{1-X}.$$

Solving this equation in $A(X)$ yields

$$A(X) = \frac{X}{(1-2X)(1-X)} = \frac{1}{1-2X} - \frac{1}{1-X}.$$

Remark. Hence $a_n = [X^n]\,A(X) = 2^n - 1$. (This is obvious because ($\star$) says that $(a_n + 1) = 2(a_{n-1} + 1)$ for all $n \geqslant 1$.) ∎

Exercise 3.1.10. Let $(b_n)_{n\geqslant 0}$ be a sequence defined by $b_0 = 1$ and for all $n \geqslant 1$,

$$b_n = 2b_{n-1} + n - 1. \tag{$\star$}$$

1. What is the generating function for $(b_n)_{n\geqslant 0}$?
2. Give an explicit formula for b_n.

Solution of Exercise 3.1.10.

1. The recurrence relation ($\star$) gives

$$\sum_{n\geqslant 1} b_n X^n = 2X \sum_{n\geqslant 1} b_{n-1} X^{n-1} + \sum_{n\geqslant 1} (n-1) X^n,$$

that is (since $b_0 = 1$)

$$B(X) - 1 = 2XB(X) + \frac{X}{(1-X)^2}.$$

Solving this equation in $B(X)$ yields

$$B(X) = \frac{1}{1-2X} + \frac{X^2}{(1-X)^2(1-2X)} = \frac{2}{1-2X} - \frac{1}{(1-X)^2}.$$

2. Hence $b_n = [X^n]\, B(X) = 2^{n+1} - n - 1$. ■

Exercise 3.1.11. We saw that if f is the generating function for $(a_n)_{n\geqslant 0}$, then $f/(1-X)$ is the generating function for $(\sum_{j=0}^{n} a_j)_{n\geqslant 0}$.
Use this to prove that the Fibonacci numbers f_n satisfy, for all $n \geqslant 0$,

$$f_0 + f_1 + \cdots + f_n = f_{n+2} - 1. \tag{$\star$}$$

Solution of Exercise 3.1.11. Recall the generating function F for the Fibonacci numbers:

$$F(X) = \frac{X}{1-X-X^2}.$$

On the one hand, the generating function for $(f_0 + \cdots + f_n)_{n\geqslant 0}$ is

$$\frac{F(X)}{1-X} = \frac{X}{(1-X)(1-X-X^2)}.$$

On the other hand, the generating function for $(f_{n+2} - 1)_{n\geqslant 0}$ is

$$\sum_{n\geqslant 0} (f_{n+1} + f_n - 1)\, X^n = \left(\frac{1}{X} + 1\right) F(X) - \frac{1}{1-X} = \frac{X}{(1-X)(1-X-X^2)}.$$

Since both generating functions are equal, this proves ($\star$). ■

Exercise 3.1.12. Let $c_{n,k}$ denote the number of compositions of n into k (nonzero) parts.

1. What is the (univariate) generating function for $(c_{n,k})_{n\geqslant 0}$?
2. Give an exact formula for $c_{n,k}$. You may use the formula

$$\sum_{n\geqslant 0}\binom{n}{k}X^n = \frac{X^k}{(1-X)^{k+1}}.$$

   Can you give a combinatorial interpretation?

3. What is the (bivariate) generating function for $(c_{n,k})_{k\geqslant 0}$?
4. Deduce an exact formula for c_n, the number of compositions of n. Can you give a combinatorial interpretation?

Solution of Exercise 3.1.12.

1. Fix k. A composition of n into k parts is described by an *ordered* k-tuple of *positive* (nonzero) integers $(n_1,\dots,n_k)\in\mathbb{N}^k$ such that $n_1+\cdots+n_k=n$. Thus, the combinatorial class $\mathscr{C}_k$ of compositions of n into k parts corresponds to the k^{th} power of the combinatorial class $\mathscr{N}:=\mathbb{N}$ of positive integers (with size function $|n|:=n$ for every $n\in\mathscr{N}$). The generating function for $\mathscr{N}$ is

$$N(X) := \sum_{n\in\mathscr{N}} X^{|n|} = \sum_{n=1}^{\infty} X^n = \frac{X}{1-X}.$$

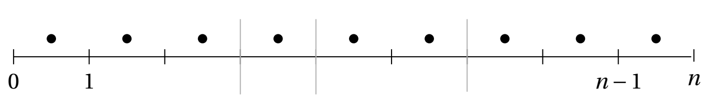
   Therefore, by the product principle, the generating function for $\mathscr{C}_k$ is

$$C_k(X) = N(X)^k = \frac{X^k}{(1-X)^k}.$$

2. Then

$$\sum_{n\geqslant 0} c_{n,k}\,X^n =: C_k(X) = X\cdot\frac{X^{k-1}}{(1-X)^k} = \sum_{n\geqslant 0}\binom{n}{k-1}X^{n+1} = \sum_{n\geqslant 0}\binom{n-1}{k}X^n.$$

   Identifying the terms in X^n yields $c_{n,k}=\binom{n-1}{k-1}$. One can interpret this combinatorially as the number of ways to divide a sequence of n "balls" into k (non-empty) parts by placing $k-1$ "separators" among $n-1$ possible emplacements, e.g.



   corresponds to the composition $(3,1,2,3)$ of $n=9$ into $k=4$ parts, placing the $k-1=3$ separators (in grey) at positions $\{3,4,6\}$ among $\{1,\dots,n-1\}$.

3. The bivariate generating function for $(c_{n,k})_{n,k\geqslant 0}$ is

$$C(X,Y) := \sum_{n,k\geqslant 0} c_{n,k}\,X^nY^k = \sum_{k\geqslant 0} C_k(X)\,Y^k = \sum_{k\geqslant 0}\left(\frac{XY}{1-X}\right)^k = \frac{1-X}{1-X-XY}.$$

4. In particular
$$C(X) := \sum_{n\geqslant 0} c_n X^n = C(X,1) = \frac{1-X}{1-2X}.$$
Thus $c_n = [X^n]\, C(X) = 2^n - 2^{n-1} = 2^{n-1}$ $(n \geqslant 1)$. Combinatorially, we divide a sequence of n "balls" by fixing a subset of $\{1,\ldots,n-1\}$ corresponding to the emplacements of the "separators".
■

Exercise 3.1.13.

1. Show that
$$\sum_{n,k\geqslant 0} p(n \mid k \text{ parts, parts} \equiv j \bmod m)\, q^n z^k = \frac{1}{(zq^j;q^m)_\infty},$$
and
$$\sum_{n,k\geqslant 0} Q(n \mid k \text{ parts, parts} \equiv j \bmod m)\, q^n z^k = (-zq^j;q^m)_\infty.$$

2. For n,k,m nonnegative integers, let $a(n,k,m)$ denote the number of partions of n into k distinct parts congruent to 2 mod 3 and m parts congruent to 1 mod 6, such that 2 is not a part. What is the (triviariate) generating function for $a_{n,k,m}$?

Solution of Exercise 3.1.13.

1. Parts congruent to j mod m are of the form $j+rm$, $r \geqslant 0$. Thus
$$\begin{aligned}\sum_{n,k\geqslant 0} p(n \mid k \text{ parts; parts} \equiv j \bmod m)\, q^n z^k &= \prod_{r\geqslant 0}(1+zq^{j+rm}+z^2q^{2(j+rm)}+\cdots)\\ &= \prod_{r\geqslant 0}\sum_{k\geqslant 0}(zq^{j+rm})^k\\ &= \prod_{r\geqslant 0}\frac{1}{1-zq^jq^{rm}}\\ &= \frac{1}{(zq^j;q^m)_\infty},\end{aligned}$$
and
$$\begin{aligned}\sum_{n,k\geqslant 0} Q(n \mid k \text{ parts; parts} \equiv j \bmod m)\, q^n z^k &= \prod_{r\geqslant 0}(1+zq^{j+rm})\\ &= (-zq^j;q^m)_\infty.\end{aligned}$$

2. Parts congruent to 2 mod 3, but not equal to 2, are of the form $5+3k$, $k \geqslant 0$, and parts congruent to 1 mod 6 are of the form $1+6k$, $k \geqslant 0$. Thus
$$\begin{aligned}\sum_{n,k,m\geqslant 0} a(n,k,m)\, q^n z^k t^m &= \prod_{k\geqslant 0}\frac{1+zq^{5+3k}}{1-tq^{1+6k}}\\ &= \frac{(-zq^5;q^3)_\infty}{(tq;q^6)_\infty}.\end{aligned}$$
■

Exercise 3.1.14. Using generating functions, show that the number of partitions of n into parts congruent to $\pm 1 \bmod 6$ equals the number of partitions of n into distinct parts congruent to $\pm 1 \bmod 3$.

Solution of Exercise 3.1.14. On the one hand,

$$\sum_{n\geqslant 0} p(n \mid \text{parts} \equiv \pm 1 \bmod 6)\, q^n = \prod_{k\equiv \pm 1 \bmod 6} \frac{1}{1-q^k}.$$

On the other hand,

$$\sum_{n\geqslant 0} Q(n \mid \text{parts} \equiv \pm 1 \bmod 3)\, q^n = \prod_{k\equiv \pm 1 \bmod 3} (1+q^k).$$

But both products are equal:

$$\begin{aligned}
\prod_{k\equiv \pm 1 \bmod 3} (1+q^k) &= \prod_{k\equiv \pm 1 \bmod 3} (1+q^k)\cdot \frac{1-q^k}{1-q^k} \\
&= \prod_{k\equiv \pm 1 \bmod 3} \frac{1-q^{2k}}{1-q^k} \\
&= \prod_{\substack{k\equiv \pm 1 \bmod 3 \\ \text{and } k \text{ odd}}} \frac{1}{1-q^k} \\
&= \prod_{k\equiv \pm 1 \bmod 6} \frac{1}{1-q^k}.
\end{aligned}$$

■

Exercise 3.1.15 (a bit challenging). Prove that the number of partitions of n such that each part appears 2, 3 or 5 times equals the number of partitions of n into parts congruent to ± 2, ± 3, or 6 mod 12.

Solution of Exercise 3.1.15. On the one hand (using that $1+x^2+x^3+x^5 = (1+x^2)(1+x^3)$),

$$\begin{aligned}
\sum_{n\geqslant 0} p(n \mid \text{each part appears } 2,3, \text{or } 5 \text{ times})\, q^n &= \prod_{k=1}^{\infty} (1+q^{2k}+q^{3k}+q^{5k}) \\
&= \prod_{k=1}^{\infty} (1+q^{2k})(1+q^{3k}).
\end{aligned}$$

On the other hand,

$$\sum_{n\geqslant 0} p(n \mid \text{parts} \equiv \pm 2, \pm 3, \text{or } 6 \bmod 12)\, q^n = \prod_{k\equiv \pm 2,\pm 3, 6 \bmod 12} \frac{1}{1-q^k}.$$

Again, both products are equal:

$$\begin{aligned}
\prod_{k=1}^{\infty}(1+q^{2k})(1+q^{3k}) &= \prod_{k=1}^{\infty}(1+q^{2k})(1+q^{3k})\cdot\frac{(1-q^{2k})(1-q^{3k})}{(1-q^{2k})(1-q^{3k})} \\
&= \prod_{k=1}^{\infty}\frac{(1-q^{4k})(1-q^{6k})}{(1-q^{2k})(1-q^{3k})} \\
&= \prod_{\substack{k\equiv 2 \bmod 4\\ \text{or}\\ k\equiv 3 \bmod 6}} \frac{1}{1-q^k} \\
&= \prod_{k\equiv\pm 2,\pm 3,6 \bmod 12} \frac{1}{1-q^k}.
\end{aligned}$$

■

Exercise 3.1.16. Show that for all $n,k\geqslant 1$,

$$p(n\mid k\text{ parts}) = p(n-1\mid k-1\text{ parts}) + p(n-k\mid k\text{ parts}).$$

Solution of Exercise 3.1.16. Recall the generating function for partitions (*which you should know by heart*):

$$\sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^n z^k = \prod_{i=1}^{\infty}\frac{1}{1-zq^i} =: \frac{1}{(zq;q)_\infty}. \qquad (\heartsuit)$$

We first observe that

$$\begin{aligned}
\frac{1}{(zq;q)_\infty} &= \frac{zq+(1-zq)}{(zq;q)_\infty} \\
&= \frac{zq}{(zq;q)_\infty} + \frac{1}{(zq^2;q)_\infty}.
\end{aligned}$$

Then, using ($\heartsuit$) (substituting zq to z for the second term)

$$\begin{aligned}
\sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^n z^k &= zq\sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^n z^k + \sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, (zq)^n z^k \\
&= \sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^{n+1} z^{k+1} + \sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^{n+k} z^k.
\end{aligned}$$

Performing the changes of variables $\{n\leftarrow n+1, k\leftarrow k+1\}$ in the first sum and $n\leftarrow n+k$ in the second sum finally leads to

$$\sum_{n,k\geqslant 0} p(n\mid k\text{ parts})\, q^n z^k = \sum_{n,k\geqslant 0} p(n-1\mid k-1\text{ parts})\, q^n z^k + \sum_{n,k\geqslant 0} p(n-k\mid k\text{ parts})\, q^n z^k.$$

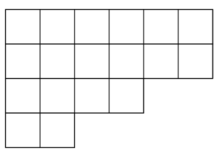
We conclude by identifying the coefficients in $q^n z^k$. ■

3.2 Ferrers diagrams and q-series identities

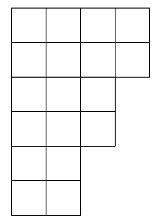
Exercise 3.2.1. Find the conjugates of the following partitions:

- $6+6+4+2$,
- $3+3+2+1$,
- $6+1$.

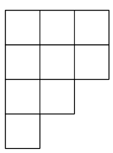
Solution of Exercise 3.2.1. Given a partition λ, its conjugate λ' is the partition whose Ferrers diagram is obtained from that of λ by exchanging rows and columns. Thus:



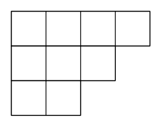
Ferrers diagram of $\lambda_1 = 6+6+4+2$.



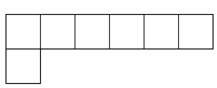
Ferrers diagram of $\lambda_1' = 4+4+3+3+2+2$.



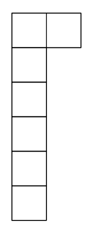
Ferrers diagram of $\lambda_2 = 3+3+2+1$.



Ferrers diagram of $\lambda_2' = 4+3+2$.



Ferrers diagram of $\lambda_3 = 6+1$.



Ferrers diagram of $\lambda_3' = 2+1+1+1+1+1$. ∎

Exercise 3.2.2. Use conjugation to show that for all n,

$$p(n \mid \text{distinct parts}) = p(n \mid \text{parts of every size from 1 to the largest part}).$$

For example, for $n=5$, the partitions of the first type are 5, $4+1$ and $3+2$, and the partitions of the second type are $1+1+1+1+1$, $2+1+1+1$ and $2+2+1$.

Solution of Exercise 3.2.2. Fix $n \geqslant 0$. By considering the Ferrers diagrams, we see that the map $\lambda \mapsto \lambda'$ induces a one-to-one correspondance between the set $\mathcal{Q}_n$ of partitions of n into distinct parts and the set $\mathscr{S}_n$ of partitions of n with parts of every size from 1 to the largest part. Indeed (denoting by $F(\lambda)$ the Ferrers diagram of a partition λ),

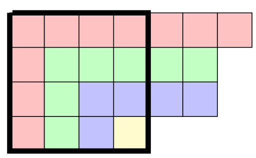
$$\begin{aligned}\lambda \in \mathscr{Q}_n &\iff \text{no two rows of } F(\lambda) \text{ are equal}\\ &\iff \text{no two columns of } F(\lambda') \text{ are equal}\\ &\iff \lambda' \text{ contains the parts 1, 2, and so on, up to the largest part}\\ &\iff \lambda' \in \mathscr{S}_n,\end{aligned}$$

for every partition λ of n. In particular $|\mathscr{Q}_n| = |\mathscr{S}_n|$, *i.e*,

$$p(n \mid \text{distinct parts}) = p(n \mid \text{parts of every size from 1 to the largest part}). \qquad \blacksquare$$

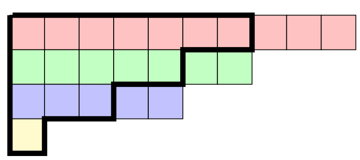
Exercise 3.2.3. Show that for all n, the number of partitions of n which have nothing under the Durfee square equals the number of partitions of n such that consecutive parts differ by at least 2.

Solution of Exercise 3.2.3. We have seen a simple transformation h on the set of partitions which bijectively maps self-conjugate partitions onto partitions into distinct odd parts (given λ, take the successive left-top hooks of its Ferrers diagram as the parts of $h(\lambda)$). The map h also induces a bijection between partitions of n which have nothing under the Durfee square and partitions of n such that consecutive parts differ by at least 2. For instance,



Ferrers diagram of $\lambda = 7+6+6+4$

(note that the number of hooks equals the width of the Durfee square), is mapped to

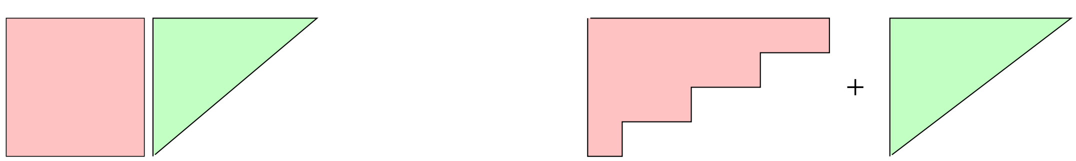


Ferrers diagram of $h(\lambda) = 10+7+5+1$.

Hence

$$p(n \mid \text{nothing under the Durfee square}) = p(n \mid \text{consecutive parts differ by at least 2}).$$

Remark. Another way to view this bijection is to transform the $n \times n$ Durfee square into a 2-staircase of height n (indeed, $n^2 = 1+3+\cdots+(2n-1)$):

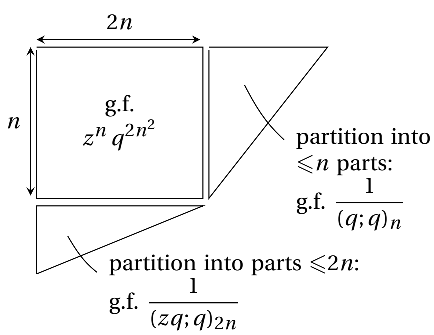


Ferrers diagram of a partition with nothing under the Durfee square. $\longleftrightarrow$ Ferrers diagram of a partition where consecutive parts differ by at least 2. $\blacksquare$

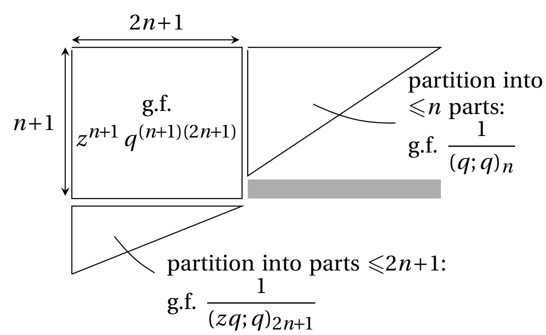
Exercise 3.2.4. Using Ferrers diagrams, show that

$$\frac{1}{(zq;q)_\infty} = \sum_{n\geqslant 0}\left(\frac{z^n\, q^{2n^2}}{(q;q)_n\,(zq;q)_{2n}} + \frac{z^{n+1}\, q^{(n+1)(2n+1)}}{(q;q)_n\,(zq;q)_{2n+1}}\right).$$

Solution of Exercise 3.2.4. For every partition λ, let n be the largest nonnegative integer such that its Ferrers diagram $F(\lambda)$ contains the upper-left rectangle with size $n\times 2n$ (so no upper-left rectangle with size $(n+1)\times 2(n+1)$). This allows us to partition the set of partitions w.r.t. this rectangle-parameter n. (The case $n=0$ includes the empty partition and the partition of 1.) Depending on whether it contains the rectangle with size $(n+1)\times(2n+1)$ or not, the Ferrers diagram of a partition with rectangle-parameter n looks like



or



(no box allowed in the grey area, as this would otherwise contradict the definition of n). By the union and product principles, we conclude that the generating function $(zq;q)_\infty^{-1}$ of $(p(n \mid k \text{ parts}))_{n,k\geqslant 0}$ fulfills the stated identity. ∎

Exercise 3.2.5. Show that for every $n\in\mathbb{N}$,

$$\sum_{k\in\mathbb{Z}} (-1)^k\, Q\!\left(n-\frac{k(3k+1)}{2}\right) = \begin{cases}(-1)^j, & \text{if } n = j(3j+1),\ j\in\mathbb{Z},\\ 0, & \text{otherwise.}\end{cases}$$

Hint. Use the pentagonal numbers theorem.

Solution of Exercise 3.2.5. Let us compute the generating function for the sequence

$$a_n := \sum_{k\in\mathbb{Z}} (-1)^k\, Q\!\left(n-\frac{k(3k+1)}{2}\right), \qquad n\geqslant 0.$$

Then (with the change of variable $n \leftarrow n + \frac{k(3k+1)}{2}$),

$$\begin{aligned}
\sum_{n\geqslant 0} a_n\, q^n &= \sum_{k\in\mathbb{Z}} \sum_{n\geqslant 0} (-1)^k\, Q(n)\, q^{n+\frac{k(3k+1)}{2}} \\
&= \left(\sum_{k\in\mathbb{Z}} (-1)^k\, q^{\frac{k(3k+1)}{2}}\right)\cdot\left(\sum_{n\geqslant 0} Q(n)\, q^n\right) \\
&= (q;q)_\infty \cdot (-q;q)_\infty \\
&= (q^2;q^2)_\infty \\
&= \sum_{j\in\mathbb{Z}} (-1)^j \left(q^2\right)^{\frac{j(3j+1)}{2}} \\
&= \sum_{j\in\mathbb{Z}} (-1)^j\, q^{j(3j+1)},
\end{aligned}$$

where we used Euler's pentagonal numbers theorem twice (in the 3rd and 5th equalities). We conclude by identifying the coefficient in q^n. ■

Exercise 3.2.6. Show that

$$\sum_{n\geqslant 0} q^{n(n+1)/2} = \frac{(q^2;q^2)_\infty}{(q;q^2)_\infty}.$$

Hint. Use Jacobi's triple product identity.

Solution of Exercise 3.2.6. The map $f\colon n \in \mathbb{Z} \mapsto n(n+1)/2$ is two-to-one (we have $f(-n-1) = f(n)$, $n \geqslant 0$), so

$$\begin{aligned}
\sum_{n\geqslant 0} q^{n(n+1)/2} &= \frac{1}{2} \sum_{n\in\mathbb{Z}} q^{n(n+1)/2} \\
&= \frac{1}{2}(-1;q)_\infty \cdot (-q;q)_\infty \cdot (q;q)_\infty \\
&= (-q;q)_\infty \cdot (q^2;q^2)_\infty \\
&= \frac{(q^2;q^2)_\infty}{(q;q^2)_\infty},
\end{aligned}$$

where we applied Jacobi's triple product identity (with $z = 1$) in the 2nd equality and Euler's "odd parts/distinct parts" theorem in the 4th equality. ■

Exercise 3.2.7. We will now use Euler's Pentagonal numbers theorem to find the recurrence relation for $p(n)$ that was mentioned in class.

1. Show that

$$\left(\sum_{k\in\mathbb{Z}} (-1)^k\, q^{k(3k+1)/2}\right)\cdot\left(\sum_{n\geqslant 0} p(n)\, q^n\right) = 1.$$

2. Deduce that for every $n \in \mathbb{N}$,

$$p(n) = \sum_{k\geqslant 1} (-1)^{k-1}\, p\left(n - \frac{k(3k\pm 1)}{2}\right).$$

3. Write $p(10)$ as a sum of smaller values of $p(n)$.

(This method, discovered by Leonhard Euler in the 18$^{\text{th}}$ century, is still the fastest way to compute $p(n)$ and is used in computing softwares such as Maple, Mathematica, etc.)

Solution of Exercise 3.2.7.

1. By Euler's pentagonal numbers theorem,

$$\left(\sum_{k\in\mathbb{Z}} (-1)^k\, q^{k(3k+1)/2}\right)\cdot\left(\sum_{n\geqslant 0} p(n)\, q^n\right) = (q;q)_\infty \cdot \frac{1}{(q;q)_\infty} = 1.$$

2. Taking the coefficient in q^n on both sides yields

$$\sum_{\substack{k\in\mathbb{Z},\, m\geqslant 0 \text{ such that} \\ \frac{k(3k+1)}{2}+m=n}} (-1)^k\, p(m) \quad = \quad 0,$$

that is (isolating the term for $k=0$, *i.e*, $m=n$)

$$p(n) = \sum_{k\geqslant 1} (-1)^{k-1}\, p\left(n - \frac{k(3k\pm 1)}{2}\right).$$

3. For $n=10$, the formula gives $p(10) = p(9)+p(8)-p(5)-p(3)$ $(=42)$. ∎

Exercise 3.2.8. Prove the second q-analogue of Pascal's triangle.

Solution of Exercise 3.2.8. Recall that $\begin{bmatrix} n+m \\ n \end{bmatrix}_q$ is the generating function for the partitions whose Ferrers diagrams fit into a rectangle of size $n\times m$ (partitions into at most n parts, which are all at most m). They consist of two types: those who actually fit in the smaller rectangle $n\times(m-1)$ (partitions into at most n parts, all at most $m-1$), which are generated by $\begin{bmatrix} n+m-1 \\ n \end{bmatrix}_q$, and those who do not (partitions into at most n parts, all at most m, with at least one part m), generated by $q^m\cdot\begin{bmatrix} n+m-1 \\ n-1 \end{bmatrix}_q$ (decomposition as the first row, which has size m, and the remaining which is a partition fitting in a rectangle with size $(n-1)\times m$). Thus

$$\begin{bmatrix} n+m \\ n \end{bmatrix}_q = \begin{bmatrix} n+m-1 \\ n \end{bmatrix}_q + q^m \begin{bmatrix} n+m-1 \\ n-1 \end{bmatrix}_q,$$

which is the second q-analogue of Pascal's triangle.

Remark 1. The first q-analogue of Pascal's triangle is

$$\begin{bmatrix} n+m \\ n \end{bmatrix}_q = \begin{bmatrix} n+m-1 \\ n-1 \end{bmatrix}_q + q^n \begin{bmatrix} n+m-1 \\ n \end{bmatrix}_q.$$

If we exchange the roles of m and n in this formula and use that $\begin{bmatrix} a+b \\ a \end{bmatrix}_q = \begin{bmatrix} a+b \\ b \end{bmatrix}_q$, then we recover the second q-analogue.

Remark 2. We could also use the exact formula for the q-binomial coefficients: setting $(a)_n := (a;q)_n$ for any a, we have

$$\begin{aligned} \begin{bmatrix} n+m \\ n \end{bmatrix}_q - \begin{bmatrix} n+m-1 \\ n \end{bmatrix}_q &= \frac{(q)_{n+m}}{(q)_n\,(q)_m} - \frac{(q)_{n+m-1}}{(q)_n\,(q)_{m-1}} \\ &= \frac{(q)_{n+m-1}}{(q)_n\,(q)_m} \cdot \underbrace{\big((1-q^{n+m})-(1-q^m)\big)}_{=q^m(1-q^n)} \\ &= q^m \begin{bmatrix} n+m-1 \\ n-1 \end{bmatrix}_q . \end{aligned}$$

■

Exercise 3.2.9. Give an analytic proof of the q-binomial series

$$\frac{1}{(zq;q)_n} = \sum_{m\geqslant 0} z^m q^m \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q .$$

Solution of Exercise 3.2.9. Call $f_n(q;z)$ the right-hand side, and proceed by induction over n. Clearly $f_1(q;z) = (1-zq)^{-1} = (zq;q)_1^{-1}$ (geometric series), so the identity is true for $n=1$. Suppose it is true for some $n \geqslant 1$. Then, using the first q-analogue of Pascal's triangle

$$\begin{bmatrix} n+m \\ m \end{bmatrix}_q = \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q + q^n \begin{bmatrix} n+m-1 \\ m-1 \end{bmatrix}_q ,$$

we obtain

$$\begin{aligned} f_{n+1}(q;z) &= \sum_{m\geqslant 0} z^m q^m \left(\begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q + q^n \begin{bmatrix} n+m-1 \\ m-1 \end{bmatrix}_q \right) \\ &= f_n(q;z) + \sum_{m\geqslant 0} z^m q^{m+n} \begin{bmatrix} n+m-1 \\ m-1 \end{bmatrix}_q . \end{aligned}$$

Thus, by the induction hypothesis (and observing that the term in $m=0$ of the last sum above is zero),

$$\begin{aligned} f_{n+1}(q;z) &= \frac{1}{(zq;q)_n} + zq^{n+1} \sum_{m\geqslant 0} z^m q^m \begin{bmatrix} n+m \\ m \end{bmatrix}_q \\ &= \frac{1}{(zq;q)_n} + zq^{n+1} f_{n+1}(q;z), \end{aligned}$$

which solves to

$$f_{n+1}(q;z) = \frac{1}{1-zq^{n+1}} \cdot \frac{1}{(zq;q)_n} = \frac{1}{(zq;q)_{n+1}},$$

as expected. ■

Exercise 3.2.10. Show that for all integers $m, n \geqslant 0$,

$$\sum_{j=0}^{n} q^j \begin{bmatrix} m+j \\ m \end{bmatrix}_q = \begin{bmatrix} n+m+1 \\ m+1 \end{bmatrix}_q .$$

Solution of Exercise 3.2.10. We use the second q-analogue of Pascal's triangle:

$$q^j \begin{bmatrix} m+j \\ m \end{bmatrix}_q = \begin{bmatrix} m+j+1 \\ m+1 \end{bmatrix}_q - \begin{bmatrix} m+j \\ m+1 \end{bmatrix}_q .$$

Then (telescopic summation)

$$\begin{aligned} \sum_{j=0}^{n} q^j \begin{bmatrix} m+j \\ m \end{bmatrix}_q &= \sum_{j=0}^{n} \left(\begin{bmatrix} m+j+1 \\ m+1 \end{bmatrix}_q - \begin{bmatrix} m+j \\ m+1 \end{bmatrix}_q \right) \\ &= \begin{bmatrix} m+n+1 \\ m+1 \end{bmatrix}_q - \underbrace{\begin{bmatrix} m \\ m+1 \end{bmatrix}_q}_{=0} \\ &= \begin{bmatrix} m+n+1 \\ m+1 \end{bmatrix}_q . \end{aligned}$$

■

Exercise 3.2.11. Let

$$H_n(t) := \sum_{j=0}^{n} \begin{bmatrix} n \\ j \end{bmatrix}_q t^j .$$

Prove that

$$\sum_{n\geqslant 0} \frac{H_n(t)\, x^n}{(q;q)_n} = \frac{1}{(x;q)_\infty (xt;q)_\infty} .$$

Solution of Exercise 3.2.11. Using the exact formula

$$\begin{bmatrix} n \\ j \end{bmatrix}_q = \frac{(q;q)_n}{(q;q)_j (q;q)_{n-j}} ,$$

the left-hand side is

$$\begin{aligned} \sum_{n\geqslant 0} \left(\sum_{j=0}^{n} \frac{(xt)^j}{(q;q)_j} \cdot \frac{x^{n-j}}{(q;q)_{n-j}} \right) &= \left(\sum_{n\geqslant 0} \frac{(xt)^n}{(q;q)_n} \right) \cdot \left(\sum_{n\geqslant 0} \frac{x^n}{(q;q)_n} \right) \\ &= F(q;xt) \cdot F(q;x), \end{aligned}$$

where

$$F(q;y) = \sum_{n\geqslant 0} \frac{y^n}{(q;q)_n} .$$

Since $(q;q)_n^{-1}$ is the generating function for partitions into at most n parts, we see that F is the generating function for partitions of n into exactly n parts, where zero parts are allowed:

$$\begin{aligned}\sum_{n\geqslant 0}\frac{y^n}{(q;q)_n} &= \sum_{n\geqslant 0}\sum_{k=0}^{n} p(n\mid k\text{ parts})\, q^n\, y^{k+\underbrace{n-k}_{\text{“zero parts”}}}\\ &= \sum_{n\geqslant 0}\sum_{k=0}^{n} p(n\mid n\text{ parts}\geqslant 0;\, k\text{ parts}>0)\, q^n y^n\\ &= \sum_{n\geqslant 0} p(n\mid n\text{ parts}\geqslant 0)\, q^n y^n\\ &= \prod_{k=0}^{\infty}\frac{1}{1-yq^k}\\ &= \frac{1}{(y;q)_\infty}\end{aligned}$$

(for the second equality, observe that, by adding/removing zero parts, there is an easy bijection between the partitions of n into k (positive) parts and the partitions of n into n nonnegative parts of which k are positive). We conclude that

$$\sum_{n\geqslant 0}\frac{H_n(t)\, x^n}{(q;q)_n} = \frac{1}{(x;q)_\infty\,(xt;q)_\infty}.$$

■

Exercise 3.2.12. Show that, if n is odd,

$$\sum_{j=0}^{n}(-1)^j\begin{bmatrix} n\\ j\end{bmatrix}_q = 0.$$

What happens if n is even?

Solution of Exercise 3.2.12. The left-hand side is $H_n(-1)$, where, by Exercise 3.2.11,

$$\frac{H_n(-1)}{(q;q)_n} = [x^n]\,\frac{1}{(x;q)_\infty\,(-x;q)_\infty} = [x^n]\,\frac{1}{(x^2;q^2)_\infty}.$$

Since the right-hand side is an even function in x, this proves that $H_n(-1)=0$ if n is odd. If however n is even, $n=2p$, then the right-hand side is the generating function (in the variable q^2) for partitions into exactly p parts, where parts 0 are allowed. By discarding the 0 parts, this is also the generating function (in the variable q^2) for partitions into at most p (positive) parts, *i.e* $(q^2;q^2)_p^{-1}$. Then

$$H_n(-1) = \frac{(q;q)_{2p}}{(q^2;q^2)_p} = (q;q^2)_p = (1-q)(1-q^3)\cdots(1-q^{n-1}).$$

Remark. The change of variable $j\leftarrow n-j$ also shows that

$$H_n(-1) = \sum_{j=0}^{n}(-1)^{n-j}\begin{bmatrix} n\\ j\end{bmatrix}_q = (-1)^n\sum_{j=0}^{n}(-1)^j\begin{bmatrix} n\\ j\end{bmatrix}_q = (-1)^n\, H_n(-1).$$

So indeed, if n is odd, then $H_n(-1)=0$. ■

Exercise 3.2.13. Let n tend to infinity in the q-binomial series

$$\frac{1}{(zq;q)_n} = \sum_{m\geqslant 0} z^m q^m \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q .$$

1. What do we obtain?

2. Give a combinatorial interpretation of the obtained formula.

Solution of Exercise 3.2.13.

1. As $n \to \infty$, the left-hand side tends to $(zq;q)_\infty^{-1}$ (at least when $|q| < 1$ and $|zq| < 1$), while the right-hand side tends to
$$\sum_{m\geqslant 0} \frac{z^m q^m}{(q;q)_m}.$$
Indeed, for $|q| < 1$, we have seen that
$$\lim_{n\to\infty} \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q = \frac{1}{(q;q)_m}.$$
Moreover (using the inequalities $1-|q| \leqslant |1-q| \leqslant 1+|q|$)
$$\forall n \in \mathbb{N}, \qquad \left| z^m q^m \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q \right| \leqslant \left| zq \cdot \frac{1+|q|}{1-|q|} \right|^m =: \eta^m,$$
so that, for $|q|$ and $|z|$ small enough, we have $\eta < 1$ and thus
$$\lim_{N\to\infty} \limsup_{n\to\infty} \sum_{m\geqslant N} z^m q^m \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q = 0.$$
We conclude that
$$\frac{1}{(zq;q)_\infty} = \lim_{n\to\infty} \frac{1}{(zq;q)_n} = \lim_{n\to\infty} \sum_{m\geqslant 0} z^m q^m \begin{bmatrix} n+m-1 \\ m \end{bmatrix}_q = \sum_{m\geqslant 0} \frac{z^m q^m}{(q;q)_m}.$$

2. We can interpret this formula combinatorially as partitioning the set of integer partitions w.r.t. their number m of parts: their Ferrers diagram are exactly determined by one first column of height m, generated by $z^m q^m$, and the remaining which corresponds to a partition into at most m parts, generated by $(q;q)_m^{-1}$. Thus:
$$\frac{1}{(zq;q)_\infty} = \sum_{m\geqslant 0} \frac{z^m q^m}{(q;q)_m}.$$
■

Exercise 3.2.14. Show that

$$\sum_{n\geqslant 0} (-1)^n (2n+1)\, q^{\frac{n(n+1)}{2}} = (q;q)_\infty^3.$$

Solution of Exercise 3.2.14. Recall Jacobi's triple product identity: for $|q| < |z| < |q|^{-1}$,

$$\sum_{n\in\mathbb{Z}} z^n q^{\frac{n(n+1)}{2}} = (q;q)_\infty (-zq;q)_\infty (-z^{-1};q)_\infty.$$

Fix $|q| < 1$. In the annulus $\mathscr{A}_q := \{z \in \mathbb{C} : |q| < |z| < |q|^{-1}\}$, the left-hand side is a converging Laurent series, while the infinite product of the right-hand side converges locally uniformly, so we may differentiate both sides with respect to z. Hence

$$\sum_{n\in\mathbb{Z}} n\, z^{n-1} q^{\frac{n(n+1)}{2}} = (q;q)_\infty \frac{\mathrm{d}}{\mathrm{d}z}\Big[(-zq;q)_\infty (-z^{-1};q)_\infty\Big].$$

We naturally aim at evaluating this at $z = -1 \in \mathscr{A}_q$. On the one hand, the left-hand side at $z = -1$ is

$$\sum_{n\in\mathbb{Z}} n(-1)^{n-1} q^{\frac{n(n+1)}{2}} = -\sum_{n\geqslant 0} (-1)^n (2n+1)\, q^{\frac{n(n+1)}{2}}$$

(using the change of variable $n \leftarrow -n-1$ in the sum over $n \leqslant -1$). On the other hand, writing $(-z^{-1};q)_\infty = (-z^{-1}q;q)_\infty (1+z^{-1})$ and observing that $1+z^{-1} = 0$ at $z=-1$, the right-hand side at $z=-1$ reduces to

$$(q;q)_\infty^3 \frac{\mathrm{d}}{\mathrm{d}z}\left(1+\frac{1}{z}\right)\Big|_{z=-1} = -(q;q)_\infty^3.$$

We thus conclude that

$$\sum_{n\geqslant 0} (-1)^n (2n+1)\, q^{\frac{n(n+1)}{2}} = (q;q)_\infty^3.$$

■

3.3 Congruence identities

Exercise 3.3.1. Prove the second Ramanujan congruence: for every $n \geqslant 0$,

$$p(7n+5) \equiv 0 \mod 7.$$

Solution of Exercise 3.3.1. Recalling that, modulo 7, $(a+b)^7 \equiv a^7 + b^7$ for all integers a, b, we have

$$\sum_{n\geqslant 0} p(n)\, q^n = \frac{1}{(q;q)_\infty} = \frac{(q;q)_\infty^6}{(q;q)_\infty^7} \equiv \frac{(q;q)_\infty^6}{(q^7;q^7)_\infty},$$

where by Exercise 3.2.14

$$(q;q)_\infty^6 = \left((q;q)_\infty^3\right)^2 = \left(\sum_{n\geqslant 0} (-1)^n (2n+1)\, q^{\frac{n(n+1)}{2}}\right)^2.$$

Therefore, modulo 7,

$$\sum_{n\geqslant 0} p(7n+5)\, q^{7n+5} \equiv \sum_{\substack{i,j,k\geqslant 0 \text{ s.t.} \\ \binom{i+1}{2}+\binom{j+1}{2}+\cancel{7k}\equiv 5}} (-1)^{i+j}(2i+1)(2j+1)\, p(7k)\, q^{\binom{i+1}{2}+\binom{j+1}{2}+7k}.$$

$i \backslash j$	0	1	2	3	4	5	6
0	0	1	3	6	3	1	0
1		2	4	0	4	2	1
2			6	2	6	4	3
3				5	2	0	6
4					6	4	3
5						2	1
6							0

— Table of $\binom{i+1}{2}+\binom{j+1}{2}$ mod 7. —

The following table shows that, for $i, j \geqslant 0$ to fulfill $\binom{i+1}{2}+\binom{j+1}{2} \equiv 5 \bmod 7$, both i and j must be congruent to 3 mod 7:
But if $i, j \equiv 3 \bmod 7$, then $2i+1, 2j+1 \equiv 0 \bmod 7$, so the right-hand side above is 0. We conclude that for every $n \geqslant 0$,

$$p(7n+5) \equiv 0 \mod 7.$$

■

Exercise 3.3.2. For $k \geqslant 2$, the number of partitions of n into parts not divisible by k equals the number of partitions of n where each part occurs at most $k-1$ times. Prove this:

1. analytically,
2. bijectively.

Solution of Exercise 3.3.2.

1. The generating function for partitions of n into parts not divisible by k is, setting $R_k := \{n \in \mathbb{N} \colon n \text{ not divisible by } k\}$,

$$\prod_{n \in R_k} \frac{1}{1-q^n} = \prod_{n=1}^{\infty} \frac{1-q^{nk}}{1-q^n},$$

and the one for partitions of n where each part occurs at most $k-1$ times is

$$\prod_{n=1}^{\infty} (1+q^n+\cdots+q^{(k-1)n}) = \prod_{n=1}^{\infty} \frac{1-(q^n)^k}{1-q^n}.$$

They are indeed the same.

2. Given a partition where each part occurs at most $k-1$ times, divide any part m divisible by k into k parts each equal to m/k, and repeat until there is no more part multiple of k. Conversely, given a partition into parts not divisible by k, merge any sequence of k same parts equal to m into one single part equal to km, and repeat unti there is no more part occurring at least k times. For example, for $k=3$, the partition $18+6+4+4$ becomes

$$(18, 6, 4^2) \to (6^4, 4^2) \to (6^3, 4^2, 2^3) \to (6^2, 4^2, 2^6) \to (6, 4, 2^9) \to (4^2, 2^{12}),$$

and conversely

$$(4^2,2^{12}) \to (6,4^2,2^9) \to (6^2,4^2,2^6) \to (6^3,4^2,2^3) \to (18,4^2,2^3) \to (18,6,4^2).$$

■

Exercise 3.3.3 (Lemma for Schur's theorem). Let $\pi_m(n)$ count the number of partitions $\lambda := \lambda_1 + \cdots + \lambda_s$ of n such that

$$\lambda_1 \leqslant m \quad \text{and, for all } 1 \leqslant i < s, \quad \lambda_i - \lambda_{i+1} \geqslant \begin{cases} 4, & \text{if } \lambda_i \text{ divisible by 3,} \\ 3, & \text{otherwise.} \end{cases} \tag{C_m}$$

Then we have the relations

(i) $\pi_{3m+1}(n) = \pi_{3m}(n) + \pi_{3m-2}(n-3m-1)$,

(ii) $\pi_{3m+2}(n) = \pi_{3m+1}(n) + \pi_{3m-1}(n-3m-2)$,

(iii) $\pi_{3m+3}(n) = \pi_{3m+2}(n) + \pi_{3m-1}(n-3m-3)$.

Prove (ii) and (iii).

Solution of Exercise 3.3.3.

Let λ be a partition of n satisfying (C_{3m+2}). If $\lambda_1 = 3m+2$, then $\lambda_2 \leqslant 3m-1$, and thus $\lambda = (3m+2)+\widetilde{\lambda}$ where $\widetilde{\lambda}$ is a partition of $n-3m-2$ satisfying (C_{3m-1}). If otherwise $\lambda_1 \leqslant 3m+1$, then λ is in fact a partition of n satisfying (C_{3m+1}). Since both cases are exhaustive and disjoint, we deduce that

$$\pi_{3m+2}(n) = \pi_{3m-1}(n-3m-2) + \pi_{3m+1}(n),$$

which is (ii).

Let now λ be a partition of n satisfying (C_{3m+3}). If $\lambda_1 = 3m+3$, then $\lambda_2 \leqslant 3m-1$, and thus $\lambda = (3m+3)+\widetilde{\lambda}$ where $\widetilde{\lambda}$ is a partition of $n-3m-3$ satisfying (C_{3m+1}). If otherwise $\lambda_1 \leqslant 3m+2$, then λ is in fact a partition of n satisfying (C_{3m+2}). Again, both cases are exhaustive and disjoint, so

$$\pi_{3m+3}(n) = \pi_{3m-1}(n-3m-3) + \pi_{3m+2}(n),$$

which is (iii). ■

Exercise 3.3.4. Let $(a_n)_{n\geqslant 0}$ be a sequence such that $\lim\limits_{n\to\infty} a_n$ exists. Prove Abel's lemma:

$$\lim_{x\to 1^-} (1-x) \sum_{n\geqslant 0} a_n x^n = \lim_{n\to\infty} a_n.$$

Solution of Exercise 3.3.4. We may suppose without loss of generality that $a_n \geqslant 0$ (the real/complex case will follow by considering positive/real and negative/imaginary parts). On the one hand,

$$\forall N \in \mathbb{N}, \qquad (1-x) \sum_{k<N} a_k x^k \leqslant N(1-x) \sup_{k<N} a_k \xrightarrow[x\to 1^-]{} 0.$$

On the other hand,

$$\forall N \in \mathbb{N}, \qquad (1-x)\sum_{k\geqslant N} a_k x^k \leqslant (1-x)\left(\sup_{k\geqslant N} a_k\right)\sum_{k\geqslant N} x^k = x^N \sup_{k\geqslant N} a_k,$$

and also

$$(1-x)\sum_{k\geqslant N} a_k x^k \geqslant x^N \inf_{k\geqslant N} a_k.$$

Putting the two pieces together, we obtain the chain of inequalities

$$\begin{array}{c}\limsup\limits_{x\to 1^-}(1-x)\sum\limits_{n\geqslant 0} a_n x^n \leqslant \limsup\limits_{n\to\infty} a_n \\ \rotatebox{90}{$\leqslant$} \\ \liminf\limits_{x\to 1^-}(1-x)\sum\limits_{n\geqslant 0} a_n x^n \geqslant \liminf\limits_{n\to\infty} a_n.\end{array}$$

Now, if $\ell := \lim_{n\to\infty} a_n$ exists, then all members of the above display are equal to ℓ. This proves Abel's lemma. ∎

Exercise 3.3.5 (Reverse bijection for Schur's theorem).

1. Show that the transformation from P_1 to P_4 in Schur's theorem is equivalent to the following process: As long as there exists some number that is not at least 3 greater than the number below, subtract 3 from this number, add 3 to the number below, and exchange these two numbers. Example:

$$P_1 = \begin{array}{l}11\\18\\5\\3\end{array} \to \begin{array}{l}21\\8\\6\\2\end{array} \to \begin{array}{l}21\\9\\5\\2\end{array} = P_1'.$$

2. Show that the following process is the reverse bijection of the above: Start by splitting parts of P_4 that are multiple of 3 into pairs of parts differing by 1 or 2. Example:

$$P_4 = \begin{array}{l}21\\9\\5\\2\end{array} \to \begin{array}{l}11+10\\5+4\\5\\2\end{array} = P_4'.$$

We obtain a partition P_4' with no multiples of 3. Now as long as the smallest part of some pair is less than 3 greater than the part below, subtract 3 from the largest part of the pair, add 3 to the part below, and switch their positions. This process ends with a partition into parts that are not multiples of 3, where parts differing by at most two are paired up, starting from the smallest part. Example:

$$P_4' = \begin{array}{l}11+10\\5+4\\5\\2\end{array} \to \begin{array}{l}11+10\\8\\4+2\\2\end{array} \to \begin{array}{l}11\\10+8\\5\\2+1\end{array} = \begin{array}{l}11\\18\\5\\3\end{array} = P_1''.$$

Solution of Exercise 3.3.5.

1. Let $P_1' := [p_1; \cdots; p_m]$ be the result of the stated process (applied to P_1). Writing $\sigma(k)$, $1 \leqslant k \leqslant m$, for the position in P_1' of the part corresponding to the k^{th} row r_k of P_1, the table P_1' is thus given by
$$p_{\sigma(k)} = r_k + 3(k - \sigma(k)) = r_k - 3(m-k) + 3(m - \sigma(k)), \qquad 1 \leqslant k \leqslant m,$$
where $p_1 \geqslant p_2 \geqslant \cdots \geqslant p_m$, and consecutive rows differ by at least 3. In particular $\sigma(k)$ is also the position of the row corresponding to r_k after subtracting the 3-staircase $[3(m-1); \cdots; 3; 0]$ to P_1 and rearranging the table in descending order. We thus see that P_1' consists in subtracting the 3-staircase to P_1, reordering, and adding back the 3-staircase, *i.e*, $P_1' = P_4$ as described in Schur's theorem.

2. Let P_1'' be the result of the second process (applied to P_1'). The transformations occurring in the process $P_1' \mapsto P_1''$ are of the form $[a+b; c] \to [c+3; b+(a-3)]$, where $a+b \equiv 0 \bmod 3$ (with $a-b \in \{1,2\}$), $c \equiv \pm 1 \bmod 3$, and $b - c \leqslant 2$. Since $x := c+3 \equiv c \equiv \pm 1 \bmod 3$ and $y := b + (a-3) \equiv a+b \equiv 0 \bmod 3$, with $(c+3) - (b+a-3) = 6 + c - (a+b) \leqslant 2$ (because $a \geqslant b+1 \geqslant 2$ and $b - c \leqslant 2$), the pair $[x; y]$ is reverted back to $[y+3; x-3] = [a+b; c]$ by the first process.

   Conversely, recall that P_1 originates from a partition into *distinct* parts which are $\pm 1 \bmod 3$, by merging pairs of parts differing by at most 2 (necessarily into a multiple of 3), *starting from the lowest part*. Thus, if $[x; y]$ are two consecutive rows in P_1 differing by at most 2, then $x \equiv \pm 1 \bmod 3$ and $y \equiv 0 \bmod 3$. Since y results from the merging of two parts $a, b \leqslant x$ with $a, -b \equiv \pm 1 \bmod 3$ and $a - b \in \{1,2\}$, the pair $[x; y]$ is transformed into $[(b+3)+a; x-3]$ where $(b+3) - a \in \{1,2\}$ and $a - (x-3) \leqslant 2$ (since $x \geqslant a$ and x, a are distinct). These are the conditions for the second process to revert this pair back to $[x-3+3; (b+3)+a-3] = [x; y]$.

   Remark. We have thus a bijection $P \mapsto P_4$ from the set $\mathscr{C}$ of partitions into distinct parts congruent to $\pm 1 \bmod 3$ and the set $\mathscr{D}$ of partitions where parts differ by at least 3 and no consecutive multiples of 3 appear. If this bijection preserves the size, it does however not strictly preserve the number of parts. It will do if parts divisible by 3 are counted *twice*: indeed,
$$\#\text{parts}(P) = \#\text{parts}(P_4) + \#\text{parts-divisible-by-3}(P_4).$$
   (An analytic argument for this refinement is given in Exercise 3.3.6.) ∎

Exercise 3.3.6 (Refinement of Schur's theorem, Gleissberg). The goal of this exercise is to prove the following refinement of Schur's theorem due to Gleissberg. Let $C(m,n)$ denote the number of partitions if n into m distinct parts congruent to 1 or 2 mod 3. Let $D(m,n)$ denote the number of partitions of n into m parts (*counting parts divisible by* 3 *twice*), where parts differ by at least 3 and no two consecutive multiples of 3 appear. Then for all $m, n \geqslant 0$, $C(m,n) = D(m,n)$.

1. Let $\pi_\ell(m,n)$ denote the number of partitions counted by $D(m,n)$ such that the largest part does not exceed ℓ. Prove that for all ℓ, m, n positive integers,
$$\pi_{3\ell+1}(m,n) = \pi_{3\ell}(m,n) + \pi_{3\ell-2}(m-1, n-3\ell-1),$$
$$\pi_{3\ell+2}(m,n) = \pi_{3\ell+1}(m,n) + \pi_{3\ell-1}(m-1, n-3\ell-2),$$
$$\pi_{3\ell+3}(m,n) = \pi_{3\ell+2}(m,n) + \pi_{3\ell-1}(m-2, n-3\ell-3).$$

2. Define, for $|q|<1$, $|t|<1$,
$$a_\ell(t,q) := \sum_{m,n\geqslant 0} \pi_\ell(m,n)\, t^m q^n.$$
What is $\lim_{\ell\to\infty} a_\ell(t,q)$?

3. Prove that
$$a_{3\ell-1}(tq^3,q) = (1+tq^{3\ell+1}+tq^{3\ell+2})\, a_{3\ell-4}(tq^3,q) + t^2q^{3\ell+3}(1-q^{3\ell-3})\, a_{3\ell-7}(tq^3,q).$$

4. Show that
$$a_{3\ell+3}(t,q) = (1+tq^{3\ell+1}+tq^{3\ell+2})\, a_{3\ell}(t,q) + t^2q^{3\ell+3}(1-q^{3\ell-3})\, a_{3\ell-3}(t,q).$$

5. What are the initial values $a_{-1}(tq^3,q)$, $a_2(tq^3,q)$, $a_3(t,q)$, $a_6(t,q)$? Verify that
$$a_3(t,q) = (1+tq)(1+tq^2)\, a_{-1}(tq^3,q),$$
and
$$a_6(t,q) = (1+tq)(1+tq^2)\, a_2(tq^3,q).$$

6. Deduce that for all $\ell\geqslant 0$,
$$a_{3\ell+3}(t,q) = (1+tq)(1+tq^2)\, a_{3\ell-1}(tq^3,q).$$

7. Conclude by finding $\lim_{\ell\to\infty} a_\ell(t,q)$.

Solution of Exercise 3.3.6 (after G. E. Andrews, *On a theorem of Schur and Gleissberg,* Arch. Math. (Basel) **22** (1971), 165–167; MR: 0286767*).*

1. Let $\lambda := \lambda_1+\cdots+\lambda_s$ be a partition of n with m counted parts and $\lambda_1\leqslant 3\ell+1$. If $\lambda_1 = 3\ell+1$, then $\lambda = (3\ell+1)+\widetilde{\lambda}$ where $\widetilde{\lambda} := \lambda_2+\cdots+\lambda_s$ is a partition of $n-3\ell-1$ with $m-1$ counted parts and $\widetilde{\lambda}_1 = \lambda_2 \leqslant 3\ell-2$. If otherwise $\lambda_1\leqslant 3\ell$, then λ is in fact a partition of n with m counted parts whose largest part does not exceed 3ℓ. As both cases are exhaustive and disjoint, we deduce the first identity
$$\pi_{3\ell+1}(m,n) = \pi_{3\ell}(m,n) + \pi_{3\ell-2}(m-1,n-3\ell-1). \tag{3.1}$$
The two other identities
$$\pi_{3\ell+2}(m,n) = \pi_{3\ell+1}(m,n) + \pi_{3\ell-1}(m-1,n-3\ell-2), \tag{3.2}$$
$$\pi_{3\ell+3}(m,n) = \pi_{3\ell+2}(m,n) + \pi_{3\ell-1}(m-2,n-3\ell-3), \tag{3.3}$$
are obtained similarly. (Beware of the 2 in (3.3), as multiples of 3 are counted *twice*.)

2. By definition, $D(m,n) = \lim\uparrow_{\ell\to\infty} \pi_\ell(m,n)$. Since for every $t, q \geqslant 0$ and $K \in \mathbb{N}$,
$$\pi(t,q) := \sum_{m,n\geqslant 0} D(m,n)\, t^m q^n \geqslant \sum_{m,n\geqslant 0} \pi_\ell(m,n)\, t^m q^n \geqslant \sum_{K\geqslant m,n\geqslant 0} \pi_\ell(m,n)\, t^m q^n,$$
we obtain
$$\pi(t,q) \geqslant \limsup_{\ell\to\infty} a_\ell(t,q) \geqslant \liminf_{\ell\to\infty} a_\ell(t,q) \geqslant \sum_{K\geqslant m,n\geqslant 0} D(m,n)\, t^m q^m.$$
As this is true for all $K \in \mathbb{N}$, it easily follows that $\lim_{\ell\to\infty} a_\ell(t,q) = \pi(t,q)$.

3. Multiplying (3.1), (3.2), (3.3) by $t^m q^n$ and summing over all $m, n \geqslant 0$, we obtain
$$a_{3\ell+1}(t,q) = a_{3\ell}(t,q) + tq^{3\ell+1}\, a_{3\ell-2}(t,q), \tag{a}$$
$$a_{3\ell+2}(t,q) = a_{3\ell+1}(t,q) + tq^{3\ell+2}\, a_{3\ell-1}(t,q), \tag{b}$$
$$a_{3\ell+3}(t,q) = a_{3\ell+2}(t,q) + t^2 q^{3\ell+3}\, a_{3\ell-1}(t,q). \tag{c}$$
Rewrite (b) and (c) as
$$a_{3\ell+1}(t,q) = a_{3\ell+2}(t,q) - tq^{3\ell+2}\, a_{3\ell-1}(t,q), \tag{b_1'}$$
$$a_{3\ell-2}(t,q) = a_{3\ell-1}(t,q) - tq^{3\ell-1}\, a_{3\ell-4}(t,q), \tag{b_2'}$$
$$a_{3\ell}(t,q) = a_{3\ell-1}(t,q) + t^2 q^{3\ell}\, a_{3\ell-4}(t,q). \tag{c'}$$
and substitute (b_1'), (b_2'), (c') into (a):
$$\begin{aligned} &a_{3\ell+2}(t,q) - tq^{3\ell+2}\, a_{3\ell-1}(t,q) \\ &\quad = \Big(a_{3\ell-1}(t,q) + t^2 q^{3\ell}\, a_{3\ell-4}(t,q)\Big) + tq^{3\ell+1}\Big(a_{3\ell-1}(t,q) - tq^{3\ell-1}\, a_{3\ell-4}(t,q)\Big).\end{aligned}$$
That is
$$a_{3\ell+2}(t,q) = (1 + tq^{3\ell+1} + tq^{3\ell+2})\, a_{3\ell-1}(t,q) + t^2 q^{3\ell}(1-q^{3\ell})\, a_{3\ell-4}(t,q). \tag{d}$$
Substituting $\ell - 1$ to ℓ yields
$$a_{3\ell-1}(t,q) = (1 + tq^{3\ell-2} + tq^{3\ell-1})\, a_{3\ell-4}(t,q) + t^2 q^{3\ell-3}(1-q^{3\ell-3})\, a_{3\ell-7}(t,q), \tag{e}$$
which is the desired identity if we replace t by tq^3:
$$a_{3\ell-1}(tq^3,q) = (1 + tq^{3\ell+1} + tq^{3\ell+2})\, a_{3\ell-4}(tq^3,q) + t^2 q^{3\ell+3}(1-q^{3\ell-3})\, a_{3\ell-7}(tq^3,q).$$

4. We perform (d) + $t^2 q^{3\ell+3} \times$ (e):
$$\begin{aligned} &\Big[a_{3\ell+2}(t,q) + t^2 q^{3\ell+3}\, a_{3\ell-1}(t,q)\Big] \\ &\quad = (1 + tq^{3\ell+1} + tq^{3\ell+2})\, a_{3\ell-1}(t,q) + t^2 q^{3\ell}(1-q^{3\ell})\, a_{3\ell-4}(t,q) \\ &\qquad + t^2 q^{3\ell+3}\Big((1 + tq^{3\ell-2} + tq^{3\ell-1})\, a_{3\ell-4}(t,q)\, t^2 q^{3\ell-3}(1-q^{3\ell-3})\, a_{3\ell-7}(t,q)\Big) \\ &\quad = (1 + tq^{3\ell+1} + tq^{3\ell+2})\Big[a_{3\ell-1}(t,q) + t^2 q^{3\ell}\, a_{3\ell-4}(t,q)\Big] \\ &\qquad + t^2 q^{3\ell+3}(1-q^{3\ell-3})\Big[a_{3\ell-4}(t,q) + t^2 q^{3\ell-3}\, a_{3\ell-7}(t,q)\Big].\end{aligned}$$

Recalling (c'), the three terms in brackets are $a_{3\ell+3}(t,q)$, $a_{3\ell}(t,q)$, and $a_{3\ell-3}(t,q)$ respectively. Thus

$$a_{3\ell+3}(t,q) = (1+tq^{3\ell+1}+tq^{3\ell+2})\,a_{3\ell}(t,q) + t^2q^{3\ell+3}(1-q^{3\ell-3})\,a_{3\ell-3}(t,q).$$

5. Let us list the partitions whose largest part is at most ℓ, where consecutive parts differ by at least 3 and no two consecutive multiples of 3 appear:

   (i) for $\ell=-1$: $\varnothing$; hence $a_{-1}(t,q)=1$;

   (ii) for $\ell=2$: $\varnothing$, 1, 2; hence $a_2(t,q)=1+tq+tq^2$;

   (iii) for $\ell=3$: $\varnothing$, 1, 2, 3; hence $a_3(t,q)=1+tq+tq^2+t^2q^3=(1+tq)(1+tq^2)$;

   (iv) for $\ell=6$: $\varnothing$, 1, 2, 3, 4, 5, 6, 4+1, 5+2, 5+1, 6+2, 6+1; hence $a_6(t,q)=1+tq+tq^2+t^2q^3+tq^4+tq^5+t^2q^6+t^2q^5+t^2q^7+t^2q^6+t^{2+1}q^8+t^{2+1}q^7$
   $= (1+tq)(1+tq^2)(1+tq^4+tq^5) = (1+tq)(1+tq^2)\,a_2(tq^3,q)$.

6. We proceed by (double) induction on ℓ. The cases $\ell=0$ and $\ell=1$ have been checked in Question 5. Let $\ell\geqslant 2$ and assume that the identities

$$a_{3\ell}(t,q) = (1+tq)(1+tq^2)\,a_{3\ell-4}(tq^3,q),$$
$$a_{3\ell-3}(t,q) = (1+tq)(1+tq^2)\,a_{3\ell-7}(tq^3,q),$$

   hold true. Then, using the recurrence relation of Question 4,

$$\begin{aligned} a_{3\ell+3}(t,q) &= (1+tq^{3\ell+1}+tq^{3\ell+2})\,a_{3\ell}(t,q)+t^2q^{3\ell+3}(1-q^{3\ell-3})\,a_{3\ell-3}(t,q)\\ &= (1+tq)(1+tq^2)\Big[(1+tq^{3\ell+1}+tq^{3\ell+2})\,a_{3\ell-4}(tq^3,q)\\ &\qquad\qquad + t^2q^{3\ell+3}(1-q^{3\ell-3})\,a_{3\ell-7}(tq^3,q)\Big].\end{aligned}$$

   Hence, by Question 3,

$$a_{3\ell+3}(t,q) = (1+tq)(1+tq^2)\,a_{3\ell-1}(tq^3,q). \tag{f}$$

7. Letting $\ell\to\infty$ in (f) we find, using Question 2,

$$\pi(t,q) = (1+tq)(1+tq^2)\,\pi(tq^3,q),$$

   for $|q|<1$, $|t|<1$. By immediate induction,

$$\begin{aligned}\pi(t,q) &= \pi(tq^{3r},q)\prod_{k=0}^{r-1}(1+tq^{3k+1})(1+tq^{3k+2})\\ &= \pi(tq^{3r},q)\,(-tq;q^3)_r\,(-tq^2;q^3)_r,\end{aligned}$$

   for every $r\in\mathbb{N}$. Letting $r\to\infty$ we end up with (since $\pi(0,q)\equiv 1$)

$$\pi(t,q) = (-tq;q^3)_\infty\,(-tq^2;q^3)_\infty.$$

   That is

$$\sum_{m,n\geqslant 0} D(m,n)\,t^m q^n = \sum_{m,n\geqslant 0} C(m,n)\,t^m q^n. \qquad \blacksquare$$

Exercise 3.3.7. Let $M(k,r,n)$ denote the number of partitions of n with crank congruent to k modulo r. Show that for all $n \geqslant 0$,

$$M(0,7,7n+5) = \cdots = M(6,7,7n+5).$$

Solution of Exercise 3.3.7. Write $\zeta := e^{\mathrm{i}2\pi/7}$ and $(a)_\infty := (a;q)_\infty$.

Step 1. We show that for every $n \geqslant 0$,

$$a_n := [q^{7n+5}]\,\frac{(q)_\infty}{(\zeta q)_\infty\,(q/\zeta)_\infty} = 0.$$

Indeed, using that $\prod_{|k|\leqslant 3}(1-\zeta^k q) = 1-q^7$, we have

$$\begin{aligned}\frac{(q)_\infty}{(\zeta q)_\infty\,(q/\zeta)_\infty} &= \frac{(q)_\infty}{\prod_{|k|\leqslant 3}(\zeta^k q)_\infty}\cdot(\zeta^{-3}q)_\infty\,(\zeta^{-2}q)_\infty\,(q)_\infty(\zeta^2 q)_\infty\,(\zeta^3 q)_\infty\\ &= \frac{1}{(q^7;q^7)_\infty}\cdot(q)_\infty\,(\zeta^2 q)_\infty\,(\zeta^{-2}q)_\infty\cdot(q)_\infty\,(\zeta^3 q)_\infty\,(\zeta^{-3}q)_\infty,\end{aligned}$$

where by Jacobi's triple product identity,

$$(q)_\infty\,(\zeta^k q)_\infty\,(\zeta^{-k}q)_\infty = \sum_{m\geqslant 0}(-1)^m\zeta^{km}q^{\frac{m(m+1)}{2}}\cdot\frac{1-\zeta^{k(2m+1)}}{1-\zeta^k}, \tag{e_k}$$

for $k=2,3$. Hence a_n arises in the product $(e_2)\cdot(e_3)$ from terms indexed by m,m' such that $\frac{m(m+1)}{2}+\frac{m'(m'+1)}{2} \equiv 5 \bmod 7$, that is $m \equiv m' \equiv 3$ (as seen in a previous exercise). But then $2m+1 \equiv 0 \bmod 7$, so $1-\zeta^{k(2m+1)} = 0$ and these terms are 0.

Step 2. Recalling the bivariate generating function for the crank, we have

$$\begin{aligned}\frac{(q)_\infty}{(\zeta q)_\infty\,(q/\zeta)_\infty} - 1 + \left(1-\zeta-\zeta^2\right)q &= \sum_{\substack{n\geqslant 2\\ m\in\mathbb{Z}}} M(m,n)\,\zeta^m q^n\\ &= \sum_{\substack{n\geqslant 2\\ m\in\mathbb{Z}}}\sum_{k=0}^{6} M(7m+k,n)\,\zeta^{\cancel{7m}+k}q^n\\ &= \sum_{n\geqslant 2}\left(\sum_{k=0}^{6} M(k,7,n)\,\zeta^k\right)q^n.\end{aligned}$$

By the first step, the coefficient in q^{7n+5}, $n \geqslant 0$, of the left-hand side is zero. We deduce that ζ is a root of the polynomial

$$f_n(X) := \sum_{k=0}^{6} M(k,7,7n+5)\,X^k \in \mathbb{Z}[X].$$

Since $\deg f_n \leqslant 6$ and $\mu(X) := 1+X+X^2+\cdots+X^6$ is the minimal polynomial of ζ, there exists a constant $C_n \in \mathbb{Z}$ such that $f_n(X) = C_n\,\mu(X)$, which proves that

$$M(0,7,7n+5) = \cdots = M(6,7,7n+5)\ (= C_n).$$

■

INDEX