One Hundred Exercises in Advanced Probability and Applications

Olivier Lévêque, IC–LTHI, EPFL

May 30, 2020

Contents

1 σ-fields and random variables

Exercise 1.1. Let $\Omega = \{1, \dots, 6\}$ et $\mathcal{A} = \{\{1, 2, 3\}, \{1, 3, 5\}\}$.

a) Describe $\mathcal{F} = \sigma(\mathcal{A})$, the σ-field generated by $\mathcal{A}$.

Hint: For a finite set Ω, the number of elements of a σ-field on Ω is always a power of 2.

b) Give the list of non-empty elements G of $\mathcal{F}$ such that

$$\text{if } F \in \mathcal{F} \text{ and } F \subset G, \text{ then } F = \emptyset \text{ or } G.$$

These elements are called the *atoms* of the σ-field $\mathcal{F}$ (cf. course). Equivalently, an event $G \in \mathcal{F}$ is *not* an atom if there exists $F \in \mathcal{F}$ such that $F \neq \emptyset$, $F \subset G$ and $F \neq G$.

The atoms of a $\mathcal{F}$ form a *partition* of the set Ω and they also generate the σ-field $\mathcal{F}$ in this case. (note also that if m is the number of atoms of $\mathcal{F}$, then the number of elements of $\mathcal{F}$ equals 2^m)

c) Let $X_1(\omega) = 1_{\{1,2,3\}}(\omega)$, $X_2 = 1_{\{1,3,5\}}(\omega)$ and $Y(\omega) = X_1(\omega) + X_2(\omega)$. Does it hold that $\sigma(Y) = \sigma(X_1, X_2)$?

Exercise 1.2. Let $\Omega = \{1, \dots, n\}$ and $\mathcal{A} = \{A_1, \dots, A_m\}$ be a collection of subsets of Ω with $m = O(\log(n))$. Design an algorithm that outputs the list of atoms of the σ-field $\sigma(\mathcal{A})$. What is the worst-case time-complexity of your algorithm?

Exercise 1.3. Let now $\Omega = [0, 1]$ and $\mathcal{F} = \mathcal{B}([0, 1])$ be the Borel σ-field on $[0, 1]$.

a) What are the atoms of $\mathcal{F}$?

b) Is it true in this case that the σ-field $\mathcal{F}$ is generated by its atoms?

c) Describe the σ-field $\sigma(\{x\},\ x \in [0, 1])$.

Exercise 1.4. Let $\Omega = \{(i, j) \,:\, i, j \in \{1, \dots, 6\}\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and define the random variables $X_1(i, j) = i$ and $X_2(i, j) = j$.

a) What are $\sigma(X_1)$, $\sigma(X_2)$?

b) Is $X_1 + X_2$ measurable with respect to one of these two σ-fields?

Exercise 1.5. Let $\mathcal{F}$ be a σ-field on a set Ω and X_1, X_2 be two $\mathcal{F}$-measurable random variables taking a finite number of values in $\mathbb{R}$. Let also $Y = X_1 + X_2$. From the course, we know that it always holds that $\sigma(Y) \subset \sigma(X_1, X_2)$, i.e., that X_1, X_2 carry together at least as much information as Y, but that the reciprocal statement is not necessarily true.

a) Provide a non-trivial example of random variables X_1, X_2 such that $\sigma(Y) = \sigma(X_1, X_2)$.

b) Provide a non-trivial example of random variables X_1, X_2 such that $\sigma(Y) \neq \sigma(X_1, X_2)$.

c) Assume that there exists $\omega_1 \neq \omega_2$ and $a \neq b$ such that $X_1(\omega_1) = X_2(\omega_2) = a$ and $X_1(\omega_2) = X_2(\omega_1) = b$. Is it possible in this case that $\sigma(Y) = \sigma(X_1, X_2)$?

Exercise 1.6. Let $\Omega =]-1, 1]$ and $(X_i, i = 1, \dots, 4)$ be a family of random variables on Ω defined as

$$X_i(\omega) = \begin{cases} 1 & \text{if } \frac{i-1}{4} < \omega \leq \frac{i}{4}, \\ (-1)^i & \text{if } -\frac{i}{4} < \omega \leq -\frac{i-1}{4}, \\ 0 & \text{otherwise.} \end{cases}$$

Describe the σ-field $\mathcal{F} = \sigma(X_i, i = 1, \dots, 4)$ using its atoms.

Exercise 1.7. Using a traditional balance, three people (say A, B and C) try to measure the weight of an object, which we assume not to exceed 100g.

For this measure, A has weights of 20g and 50g; B has weights of 20g only and C has weights of 10g only. On the other hand, the number of weights available for each of them is unlimited.

Determine the amount of information that each person has on the weight of the object, and order theses informations. In particular, determine who is able to decide whether the weight of the object is between 40g and 50g or not.

Remark: One assumes that when measuring, all weights are on the same side of the balance, with the object on the other side.

2 Probability measures and distributions

Exercise 2.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Using only the axioms given in the definition of a probability measure, show the following properties:

a) If $A, B \in \mathcal{F}$ and $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$ and $\mathbb{P}(B \backslash A) = \mathbb{P}(B) - \mathbb{P}(A)$. Also, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

b) If $A, B \in \mathcal{F}$, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

c) If $(A_n,\, n \geq 1)$ is a sequence of events in $\mathcal{F}$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

d) If $(A_n,\, n \geq 1)$ is a sequence of events in $\mathcal{F}$ such that $A_n \subset A_{n+1}$, $\forall n \geq 1$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$.

e) If $(A_n,\, n \geq 1)$ is a sequence of events in $\mathcal{F}$ such that $A_n \supset A_{n+1}$, $\forall n \geq 1$, then $\mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$.

Exercise 2.2. a) Which of the following are cdfs?

1. $F_1(t) = \exp(-e^{-t})$, $t \in \mathbb{R}$
2. $F_2(t) = \frac{1}{1-e^{-t}}$, $t \in \mathbb{R}$
3. $F_3(t) = 1 - \exp(-1/|t|)$, $t \in \mathbb{R}$
4. $F_4(t) = 1 - \exp(-e^{t})$, $t \in \mathbb{R}$

b) Let now F be a generic cdf. Which of the following functions are guaranteed to be also cdfs?

5. $F_5(t) = F(t^2)$, $t \in \mathbb{R}$
6. $F_6(t) = F(t)^2$, $t \in \mathbb{R}$
7. $F_7(t) = F(1 - \exp(-t))$, $t \in \mathbb{R}$
8. $F_8(t) = \begin{cases} 1 - \exp(-F(t)/(1 - F(t))) & \text{if } F(t) < 1 \\ 1 & \text{if } F(t) = 1 \end{cases} \quad t \in \mathbb{R}$

Exercise 2.3. Let $\lambda > 0$ and $X \sim \mathcal{E}(\lambda)$, i.e., X is a random variable with the exponential distribution, whose cdf is given by

$$F_X(t) = \mathbb{P}(\{X \le t\}) = \begin{cases} 1 - \exp(-\lambda t) & \text{if } t \ge 0 \\ 0 & \text{if } t < 0 \end{cases}$$

a) Compute the cdf of $Y = X^a$, where $a \in \mathbb{R}$.

b) Deduce the expression for the pdf of Y, when it exists.

c) In particular, study the cases $a = 2$ and $a = -1$, and draw the corresponding pdfs for a fixed value of λ.

d) Still for a fixed value of λ, which of the following two expressions is the largest when $t > 0$ gets large?

$$\mathbb{P}(\{X^2 \ge t\}) \quad \text{or} \quad \mathbb{P}(\{X^{-1} \ge t\}) \quad ?$$

Exercise 2.4. Let $X_1, \ldots, X_n$ be i.i.d. $\sim \mathcal{E}(1)$ random variables (i.e., they are independent and identically distributed, all with the exponential distribution of parameter $\lambda = 1$).

a) Compute the cdf of $Y_n = \min(X_1, \ldots, X_n)$.

b) How do $\mathbb{P}(\{Y_n \le t\})$ and $\mathbb{P}(\{X_1 \le t\})$ compare when n is large and t is such that $t \ll \frac{1}{n}$?

c) Compute the cdf of $Z_n = \max(X_1, \ldots, X_n)$.

d) How do $\mathbb{P}(\{Z_n \ge t\})$ and $\mathbb{P}(\{X_1 \ge t\})$ compare when n is large and t is such that $t \mathcal{G} \log(n)$?

Exercise 2.5. Let $n \ge 1$ and $x_1, \ldots, x_n$ be arbitrary real numbers (not necessarily ordered). Their *empirical cdf* is the function $F_n : \mathbb{R} \to \mathbb{R}$ defined as

$$F_n(t) = \frac{1}{n} \sharp\{1 \le j \le n \,:\, x_j \le t\} \quad t \in \mathbb{R}$$

where $\sharp A$ denotes the cardinality (= number of elements) of the set A.

a) Show that irrespective of the values of $x_1, \ldots, x_n$, the function F_n is a cdf. Does this cdf admit a pdf? What are the values possibly taken by F_n?

Assume now that $x_1, \ldots, x_n$ are i.i.d. samples from a distribution with cdf F.

Note: The function F_n becomes in this case a *random* cdf.

b) If $t \in \mathbb{R}$ is such that $F(t) = 0$, is it possible that: b1) $F_n(t) > 0$? b2) $F_n(s) = 0$ for some $s \ge t$?

c) If $t \in \mathbb{R}$ is such that $F(t) = 1$, is it possible that: c1) $F_n(t) < 1$? c2) $F_n(s) = 1$ for some $s \le t$?

Note: We will see later in the course that in this case, F_n approaches F as $n \to \infty$.

Let now $a, b > 0$ be two fixed real numbers and let us make two hypotheses (with an equal prior):

H_0: F is the cdf of a random variable X uniformly distributed in the interval $[0, a]$.

H_1: F is the cdf of a random variable Y uniformly distributed in the interval $[0, b]$.

d) Let $0 < t < \min(a, b)$. Given that $F_n(t) = 1$, what is the probability that Hypothesis H_0 holds? Does this probability depend on the value of t?

Hint. Use Bayes' rule.

Exercise 2.6. Let $X_1, \ldots, X_n$ be n independent and identically distributed (i.i.d.) random variables with common cdf F_X.

a) Let $Y = \max(X_1, \ldots, X_n)$. Express the cdf F_Y of Y in terms of F_X.

b) Let $Z = \min(X_1, \ldots, X_n)$. Express the cdf F_Z of Z in terms of F_X.

c) Application: Compute F_Y and F_Z when $X_1, \ldots, X_n$ are i.i.d.$\sim \mathcal{U}([0,1])$ random variables.
Compute also the corresponding pdfs p_Y and p_Z.

d) Compute finally the cdf and pdf of $1 - Y$ in this last example. What do you observe?

Exercise 2.7. Let F be a generic cdf. Which of the following are guaranteed to be also cdfs?

a) $F_a(t) = F(t)^{17}$, $t \in \mathbb{R}$.

b) $F_b(t) = F(t^{17})$, $t \in \mathbb{R}$.

c) $F_c(t) = F(\exp(t))$, $t \in \mathbb{R}$.

d) $F_d(t) = 1 - \exp\left(-\frac{F(t)}{1-F(t)}\right)$, $t \in \mathbb{R}$.

Exercise 2.8. Let $\Omega = [0,1]^2$, $\mathcal{F} = \mathcal{B}([0,1]^2)$ and $\mathbb{P}$ be the probability measure defined on $(\Omega, \mathcal{F})$ defined as

$$\mathbb{P}(\,]a,b[\times]c,d[\,) = (b-a)\cdot(d-c), \quad \text{for } 0 \le a < b \le 1 \text{ and } 0 \le c < d \le 1$$

which can be extended uniquely to all Borel sets in $\mathcal{B}([0,1]^2)$, according to Caratheodory's extension theorem.

Let us now consider the following two random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$:

$$X(\omega_1, \omega_2) = \omega_1 \quad \text{and} \quad Y(\omega_1, \omega_2) = \frac{\omega_1 + \omega_2}{2}$$

a) Compute the cdf F_X of X, as well as the cdf F_Y of Y.

b) Draw at random n independent copies $X_1, \ldots, X_n$ of X, as well as n independent copies $Y_1, \ldots, Y_n$ of Y. Draw then graphically the two (random) functions:

$$F_X^{(n)}(t) = \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_j \le t\}} \quad \text{and} \quad F_Y^{(n)}(t) = \frac{1}{n}\sum_{j=1}^{n} 1_{\{Y_j \le t\}}$$

What do you observe as n gets large?

Exercise 2.9. Let X be a random variable with cdf F_X. Express the cdf of the following random variables in terms of F_X, and deduce their pdf when X is a continuous random variable.

a) $Y_1 = aX$, for some $a \in \mathbb{R}\backslash\{0\}$.

b) $Y_2 = X + c$, for some $c \in \mathbb{R}$.

c) $Y_3 = X^2$.

d) $Y_4 = e^X$.

Exercise 2.10. Let F be a generic cdf. Which of the following are guaranteed to be also cdfs?

a) $F_1(t) = F(at)$, $t \in \mathbb{R}$, for some $a > 0$

b) $F_2(t) = F(bt)$, $t \in \mathbb{R}$, for some $b < 0$

c) $F_3(t) = F(t+c)$, $t \in \mathbb{R}$, for some $c \in \mathbb{R}$

d) $F_4(t) = F(t^2)$, $t \in \mathbb{R}$

e) $F_5(t) = F(t^3)$, $t \in \mathbb{R}$

f) $F_6(t) = F(t)^2$, $t \in \mathbb{R}$

g) $F_7(t) = F(t)^3$, $t \in \mathbb{R}$

h) $F_8(t) = \frac{1}{2}(1 + \tanh(t))$, $t \in \mathbb{R}$

i) $F_9(t) = \lim_{a \to +\infty} \frac{1}{2}(1 + \tanh(at))$, $t \in \mathbb{R}$

j) $F_{10}(t) = \lim_{a \to 0^+} \frac{1}{2}(1 + \tanh(at))$, $t \in \mathbb{R}$

Exercise 2.11. Let X be a random variable whose cdf F is the devil's staircase.

a) Write a code that allows you to sample (approximately) from F.

b) Consider n i.i.d. samples $X_1, \dots, X_n$ distributed according to F and draw on the same graph F and

$$F_n(t) = \frac{1}{n} \sharp\{1 \le j \le n \,:\, X_j \le t\}, \quad t \in \mathbb{R}$$

F_n is called the *empirical cdf* of the n samples: please note that it is itself a random function!

c) Consider now n i.i.d. samples $Y_1, \dots, Y_n$ distributed according the uniform distribution on $[0, 1]$ and with corresponding cdf G. Draw again on the same graph $G(t) = t$ and

$$G_n(t) = \frac{1}{n} \sharp\{1 \le j \le n \,:\, Y_j \le t\}, \quad t \in \mathbb{R}$$

d) Imagine a simple statistical test based on the empirical cdf of a set of n i.i.d. samples that allows to tell whether these samples are distributed according to F or G (the null hypothesis being that they are distributed according to G).

e) Bayesian analysis:
Say there is a probability $0 < \alpha < 1$ (=prior) that all your samples are distributed according to F and a probability $1 - \alpha$ that they are distributed according to G. As mentioned above, G is the null hypothesis, which we will assume to be also more likely, so that α is assumed to be small. Compute the probabilities of your test leading to 1) a false positive, 2) a false negative, i.e.,

1) $\mathbb{P}$(the samples are distributed according to $G \,|\,$ the test is positive)

2) $\mathbb{P}$(the samples are distributed according to $F \,|\,$ the test is negative)

3 Independence and convolution

Exercise 3.1. Let $n \ge 1$, $\Omega = \{1, 2, \dots, n\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and $\mathbb{P}$ be the probability measure on $(\Omega, \mathcal{F})$ defined by $\mathbb{P}(\{\omega\}) = \frac{1}{n}$ on the singletons and extended by additivity to all subsets of Ω.

a) Consider first $n = 4$. Find three subsets $A_1, A_2, A_3 \subset \Omega$ such that

$$\mathbb{P}(A_j \cap A_k) = \mathbb{P}(A_j) \cdot \mathbb{P}(A_k) \quad \forall j \neq k \quad \text{but} \quad \mathbb{P}(A_1 \cap A_2 \cap A_3) \neq \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3)$$

b) Consider now $n = 6$. Find three subsets $A_1, A_2, A_3 \subset \Omega$ such that

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) \quad \text{but} \quad \exists j \neq k \text{ such that } \mathbb{P}(A_j \cap A_k) \neq \mathbb{P}(A_j) \cdot \mathbb{P}(A_k)$$

c) Consider finally a generic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and three events $A_1, A_2, A_3 \in \mathcal{F}$ such that

$$\mathbb{P}(A_j \cap A_k) = \mathbb{P}(A_j) \cdot \mathbb{P}(A_k) \quad \forall j \neq k \quad \text{and} \quad \mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3)$$

Show that A_1, A_2, A_3 are independent according to the definition given in the course.

Exercise 3.2. Let X, Y be two discrete random variables, each with values in $\{0, 1\}$.

a) Show that $X \perp\!\!\!\perp Y$ if $\mathbb{P}(\{Y = 1\}|\{X = 0\}) = \mathbb{P}(\{Y = 1\}|\{X = 1\})$.

Let moreover $Z = X \oplus Y = \begin{cases} 1, & \text{if } X = 1, Y = 0 \text{ or } X = 0, Y = 1, \\ 0, & \text{otherwise.} \end{cases}$

b) Show that $X \perp\!\!\!\perp Z$ if $\mathbb{P}(\{Y = 1\}|\{X = 0\}) = \mathbb{P}(\{Y = 0\}|\{X = 1\})$.

c) Which assumption guarantees that both $X \perp\!\!\!\perp Y$ and $X \perp\!\!\!\perp Z$?

d) Assume that none of the 3 random variables X, Y, Z is constant (i.e., takes a single value with probability 1). Can it be that the collection of the three random variables (X, Y, Z) is independent? Justify your answer.

Exercise 3.3. Let X_1, X_2 be two independent and identically distributed (i.i.d.) $\mathcal{N}(0, 1)$ random variables. Compute the pdf of $X_1 + X_2$ (using convolution).

Exercise 3.4. Let X_1, X_2 be two i.i.d. random variables such that $\mathbb{P}(\{X_i = +1\}) = \mathbb{P}(\{X_i = -1\}) = 1/2$ for $i = 1, 2$. Let also $Y = X_1 + X_2$ and $Z = X_1 - X_2$.

a) Are Y and Z independent?

b) Same question with X_1, X_2 i.i.d.$\sim \mathcal{N}(0, 1)$ random variables (use here the change of variable formula in order to compute the joint distribution of Y and Z).

Exercise 3.5. Let $\Omega = \mathbb{R}^2$ and $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$. Let also $X_1(\omega) = \omega_1$ and $X_2(\omega) = \omega_2$ for $\omega = (\omega_1, \omega_2) \in \Omega$ and let finally μ be a probability distribution on $\mathbb{R}$. We consider below two different probability measures defined on $(\Omega, \mathcal{F})$, defined on the "rectangles" $B_1 \times B_2$ (Caratheodory's extension theorem then guarantees that these probability measures can be extended uniquely to $\mathcal{B}(\mathbb{R}^2)$).

a) $\mathbb{P}^{(1)}(B_1 \times B_2) = \mu(B_1) \cdot \mu(B_2)$

b) $\mathbb{P}^{(2)}(B_1 \times B_2) = \mu(B_1 \cap B_2)$

In each case, describe what is the relation between the random variables X_1 and X_2.

Exercise 3.6. Let $\lambda, \mu > 0$ and let X, Y be two independent random variables such that $X \sim \mathcal{E}(\lambda)$, $Y \sim \mathcal{E}(\mu)$.

a) Compute the distribution of $Z = \frac{X}{X+Y}$ (determining first the range of possible values of Z).

b) What do you obtain in the particular case $\lambda = \mu$? Does the result depend on the value of λ?

c) *Application:* You just missed the bus and are now waiting at the bus stop. But the schedule on this line is strange: the waiting time between any two buses is an exponential random variable of parameter $\lambda = 1\,[\text{min}^{-1}]$ (so the average waiting time between any two buses is $\frac{1}{\lambda} = 1\,\text{min}$) and these waiting times are independent of each other. Someone from the bus company comes and tells you that he knows for a fact that the second bus will arrive at this stop 3 minutes from now. What information do you have on the arrival time of the first bus?

4 Expectation

Exercise 4.1. Let $\lambda > 0$ and $X \sim \mathcal{E}(\lambda)$, and let us define, as in Exercise 3 of Homework 2, $Y = X^a$, where $a \in \mathbb{R}$.

a) For what values of $a \in \mathbb{R}$ does it hold that $\mathbb{E}(Y) < +\infty$?

b) For what values of $a \in \mathbb{R}$ does it hold that $\mathbb{E}(Y^2) < +\infty$?

c) For what values of $a \in \mathbb{R}$ is $\text{Var}(Y)$:

c1) well-defined and finite? c2) well-defined but infinite? c3) ill-defined?

d) Compute $\mathbb{E}(Y)$ and $\text{Var}(Y)$ for the values of $a \in \mathbb{Z}$ such that these quantities are well-defined.

Hint: Use integration by parts, recursively.

Exercise 4.2. Let $\mu \in \mathbb{R}$ and $\lambda > 0$ be fixed numbers, and let X be a Cauchy random variable with pdf

$$p_X(x) = \frac{C}{\lambda^2 + (x-\mu)^2}, \quad x \in \mathbb{R}$$

a) Compute the constant C and draw the pdf p_X for some values of μ and λ.

Hint: A primitive of $\frac{1}{1+x^2}$ is $\arctan(x)$.

b) Are $\mathbb{E}(X)$ and $\mathrm{Var}(X)$ well-defined?

c) What is your interpretation of the parameters μ and λ ?

d) Assume now $\mu = 0$ and define $Y = \frac{1}{X}$. Compute the pdf of Y (neglecting the "problem" that the function $1/x$ exploses in $x = 0$; this is actually not a problem, as $\mathbb{P}(\{X = 0\}) = 0$).

Exercise 4.3. For a generic non-negative random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, it holds that $\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(\{X \geq t\})\, dt$. Indeed, for every $\omega \in \Omega$, we have

$$X(\omega) = \int_0^{X(\omega)} 1\, dt = \int_0^{+\infty} 1_{\{t \leq X(\omega)\}}\, dt$$

so taking expectation on both sides (and swapping expectation and integral, which is allowed here, as everything is positive), we obtain the above formula.

a) Use this formula to compute $\mathbb{E}(X)$ for $X \sim \mathcal{E}(\lambda)$ and check that the result is in accordance with Exercise 1 above.

b) Particularize the above formula for $\mathbb{E}(X)$ to the case where X takes values in $\mathbb{N}$ only.

c) Use this new formula to compute $\mathbb{E}(X)$ for $X \sim \mathrm{Bern}(p)$ and $X \sim \mathrm{Geom}(p)$ for some $0 < p < 1$. Check in both cases that the result is in accordance with the classical computation of $\mathbb{E}(X)$.

Reminders: $X \sim \mathrm{Bern}(p)$ means $\mathbb{P}(\{X = 1\}) = p = 1 - \mathbb{P}(\{X = 0\})$.
$X \sim \mathrm{Geom}(p)$ means $\mathbb{P}(\{X = k\}) = (1-p)\, p^k$ for $k \in \mathbb{N}$.

Exercise 4.4. Check that the distributions below are well defined distributions and compute, when they exist, the mean and the variance of these distributions.

A) Discrete distributions:

a) Bernoulli $\mathcal{B}(p)$, $p \in [0,1]$: $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$.

b) binomial $\mathcal{B}i(n,p)$, $n \geq 1$, $p \in [0,1]$: $\mathbb{P}(X = k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k\, (1-p)^{n-k}, \quad 0 \leq k \leq n.$

c) Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$: $\mathbb{P}(X = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad k \geq 0.$

B) Continuous distributions:

d) uniform $\mathcal{U}([a,b])$, $a < b$: $p_X(x) = \frac{1}{b-a}\, 1_{[a,b]}(x), \quad x \in \mathbb{R}.$

e) Gaussian $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$: $p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$

f) Cauchy $\mathcal{C}(\lambda)$, $\lambda > 0$: $p_X(x) = \frac{1}{\pi}\, \frac{\lambda}{\lambda^2 + x^2}, \quad x \in \mathbb{R}.$

g) exponential $\mathcal{E}(\lambda)$, $\lambda > 0$: $p_X(x) = \lambda\, e^{-\lambda x}, \quad x \in \mathbb{R}_+.$

h) Gamma $\Gamma(t, \lambda)$, $t, \lambda > 0$: $p_X(x) = \frac{(\lambda x)^{t-1}\, \lambda\, e^{-\lambda x}}{\Gamma(t)}, \quad x \in \mathbb{R}_+, \quad$ where $\Gamma(t) := \int_0^\infty dx\; x^{t-1}\, e^{-x}.$

Exercise 4.5. a) Let X be a Poisson random variable with parameter $\lambda > 0$, i.e., $\mathbb{P}(\{X = k\}) = \frac{\lambda^k}{k!}\, e^{-\lambda}$, $k \geq 0$. Compute successively:

i) $\mathbb{E}(X)$ ii) $\mathbb{E}(X(X-1))$ iii) $\mathrm{Var}(X)$

b) Let X be a centered Gaussian random variable of variance σ^2. Compute successively:

i) $\mathbb{E}(X^4)$ ii) $\mathbb{E}(\exp(X))$ iii) $\mathbb{E}(\exp(-X^2))$

Exercise 4.6. Let $\mu \in \mathbb{R}$, $\sigma > 0$ and X be a continuous random variable whose pdf is given by

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

a) Draw M i.i.d. samples $X_1, \ldots, X_M$ according to p_X, compute numerically

$$\mu_M = \frac{1}{M} \sum_{j=1}^{M} X_j \quad \text{and} \quad \sigma_M = \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} (X_j - \mu_M)^2}$$

and plot these two quantities as a function of M on a graph. What do you observe as M increases?

b) Plot the empirical cdf of $X_1, \ldots, X_M$ for three different values of M (e.g., M =10, 100 and 1'000). What do you observe as M increases?

c) Consider now $M \times K$ i.i.d. samples $\{X_{m,k},\ 1 \leq m \leq M,\ 1 \leq k \leq K\}$, with a fixed K =1'000, as well as the empirical means

$$E_{M,k} = \frac{1}{M}\,(X_{1,k} + \ldots + X_{M,k}), \quad 1 \leq k \leq K$$

Plot the empirical cdf of $E_{M,1}, \ldots, E_{M,K}$ for the same three different values of M as above. Again, what do you observe as M increases?

Let again $\mu \in \mathbb{R}$, $\sigma > 0$ and Y be a random variable whose pdf is given by

$$p_Y(y) = \frac{1}{\pi}\, \frac{\sigma}{\sigma^2 + (y-\mu)^2}, \quad y \in \mathbb{R}$$

d-e-f) Same questions as in a-b-c) with the X samples replaced by the Y samples. (*NB:* In order to sample from Y, you should first compute its cdf).

Exercise 4.7. Let $\lambda > 0$ and X be an $\mathcal{E}(\lambda)$ random variable, whose cdf is given by

$$F_X(t) = \begin{cases} 1 - \exp(-\lambda t) & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases}$$

Let also $Y = e^X$. We would like first to estimate numerically the expectation and the variance of Y for various values of $\lambda > 0$. Here is how to proceed:

a) Assuming that drawing a uniform random variable U in the interval $[0, 1]$ is given, propose a method to draw $X \sim \mathcal{E}(\lambda)$ (and subsequently $Y = e^X$).

b) Draw n independent copies $Y_1, \ldots Y_n$ of Y and compute their empirical average and standard deviation

$$\hat{\mu}_n = \frac{1}{n} \sum_{j=1}^{n} Y_j \quad \text{and} \quad \hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{j=1}^{n} (Y_j - \hat{\mu}_n)^2$$

for various values of n (say, n =10'000, 100'000 and 1'000'000, for example). Repeat the experiment for multiple values of $\lambda > 0$. What do you observe?

c) Choose now a fixed large value of n and represent both $\hat{\mu}_n$ and $\hat{\sigma}_n$ as a function of $\lambda > 0$. Again, what do you observe?

d) Compute finally the theoretical distribution of Y, as well as its expectation and variance. Relate this to the numerical results you have obtained above.

5 Characteristic function

Exercise 5.1. The aim of the present exercise is to answer the following question:

$$\text{Is it true that if} \quad \phi_{X+Y}(t) = \phi_X(t)\,\phi_Y(t) \quad \forall t \in \mathbb{R}, \quad \text{then} \quad X \perp\!\!\!\perp Y \quad ? \tag{1}$$

Let X, Y be two independent random variables with the same characteristic function $\phi(t) = \exp(-|t|)$, $t \in \mathbb{R}$.

a) Compute the characteristic function of $X + Y$.

b) Compute the characteristic function of $2X$.

c) Conclude about question (2).

Exercise 5.2. a) Let X be a Poisson random variable with parameter $\lambda > 0$. Compute its characteristic function ϕ_X.

b) Show that for a discrete random variable X with values in $\mathbb{Z}$, the following inversion formula holds:

$$\mathbb{P}(\{X = k\}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk}\,\phi_X(t)\,dt, \quad \forall k \in \mathbb{Z}$$

c) Use the above formula to deduce the distribution of the random variable X with values in $\mathbb{Z}$ whose characteristic function is given by

$$\phi_X(t) = \cos(t), \quad t \in \mathbb{R}$$

d) Without solving part c), how could you be sure that ϕ_X is indeed a characteristic function?

Exercise 5.3. Let X be a discrete random variable with values in a countable set C. Show that for all $x \in C$,

$$p_x = \mathbb{P}(\{X = x\}) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx}\,\phi_X(t)\,dt.$$

Exercise 5.4. Let $\lambda > 0$ and X be a random variable whose characteristic function ϕ_X is given by

$$\phi_X(t) = \exp(-\lambda|t|), \quad t \in \mathbb{R}$$

a) What can you deduce on the distribution of X from each of the following facts?

i) ϕ_X is not differentiable in $t = 0$.

ii) $\int_{\mathbb{R}} |\phi_X(t)|\,dt < +\infty$.

b) Use the inversion formula seen in class to compute the distribution of X.

c) Let $Y = \frac{1}{X}$. Using the change of variable formula (not worrying about the fact that X might take the value 0, as this is a negligible event), compute the distribution of Y.

d) Let now $X_1, \ldots, X_n$ be n independent copies of the random variable X. What are the distributions of

$$Z_n = \frac{X_1 + \ldots + X_n}{n} \quad \text{and} \quad W_n = \frac{n}{\frac{1}{X_1} + \ldots + \frac{1}{X_n}} \quad ?$$

e) What oddities do you observe in the results of part d)? (there are at least two)

Exercise 5.5. Compute the characteristic function of the following random variables:

a) $X \sim \text{Bi}(n,p)$, with $\mathbb{P}(\{X=k\}) = \binom{n}{k} p^k (1-p)^{n-k}$, $k \in \{0,\dots,n\}$.

b) $X \sim \text{Geom}(p)$, with $\mathbb{P}(\{X=k\}) = (1-p)^{k-1}p$, $k \ge 1$.

c) X has the exponential distribution on $\mathbb{R}$ with pdf $p_X(x) = \frac{\lambda}{2}\exp(-\lambda|x|)$, $x \in \mathbb{R}$.

d*) X has the Cauchy distribution on $\mathbb{R}$ with pdf $p_X(x) = \frac{\lambda}{\pi(\lambda^2+x^2)}$, $x \in \mathbb{R}$.

6 Random vectors and Gaussian random vectors

Exercise 6.1. a) Let X be a Gaussian random vector of dimension d with expectation μ and covariance matrix A. Let B be another $d \times d$ matrix. What is the distribution of $Y = BX$?

b) Deduce from a) that X and VX have the same distribution when V is an orthogonal matrix and X is a centered Gaussian random vector whose covariance matrix is a multiple of the identity matrix (i.e., $A = \lambda I$ for some $\lambda \in \mathbb{R}$).

Exercise 6.2. Let X_1, X_2 be two i.i.d. $\sim \mathcal{N}(0,1)$ random variables. Consider the pair of random variables (R,Θ) defined by the change of variable:

$$X_1 = R\cos(\Theta) \quad \text{and} \quad X_2 = R\sin(\Theta)$$

a) Compute the joint pdf of R and Θ using the change of variable formula:

$$p_{R,\Theta}(r,\theta) = p_{X_1,X_2}(r\cos(\theta), r\sin(\theta)) \cdot |J(r,\theta)|$$

where $J(r,\theta)$ is the Jacobian of the transformation. What do you observe?

b) Compute the marginals $p_R(r)$ and $p_\Theta(\theta)$.

c) Compute the joint distribution of $(\frac{X_1}{R}, \frac{X_2}{R}) = (\cos(\Theta), \sin(\Theta))$.

d) Compute the pdf of R^2 and that of $\frac{X_1}{R} = \cos(\Theta)$.

Exercise 6.3. a) Let X_1, X_2 be two independent Gaussian random variables such that $\text{Var}(X_1) = \text{Var}(X_2)$. Show, using characteristic functions or a result from the course, that $X_1 + X_2$ and $X_1 - X_2$ are also independent Gaussian random variables.

b) Let X_1, X_2 be two independent square-integrable random variables such that $X_1 + X_2, X_1 - X_2$ are also independent random variables. Show that X_1, X_2 are jointly Gaussian random variables such that $\text{Var}(X_1) = \text{Var}(X_2)$.

Note. Part b), also known as Darmois-Skitovic's theorem, is considerably more challenging than part a)! Here are the steps to follow in order to prove the result (but please skip the first two).

Step 1.* (needs the dominated convergence theorem, which outside the scope of this course)
If X is a square-integrable random variable, then ϕ_X is twice continuously differentiable.

Step 2.* (quite technical) Under the assumptions made, ϕ_{X_1} and ϕ_{X_2} have no zeros (so $\log\phi_{X_1}$ and $\log\phi_{X_2}$ are also twice continuously differentiable, according to the previous step).

Step 3. Let $f_1 = \log\phi_{X_1}$ and $f_2 = \log\phi_{X_2}$. Show that there exist functions g_1, g_2 satisfying

$$f_1(t_1+t_2) + f_2(t_1-t_2) = g_1(t_1) + g_2(t_2) \quad \forall t_1, t_2 \in \mathbb{R}$$

Step 4. If f_1, f_2 are twice continuously differentiable and there exist functions g_1, g_2 satisfying

$$f_1(t_1+t_2)+f_2(t_1-t_2)=g_1(t_1)+g_2(t_2) \quad \forall t_1,t_2\in\mathbb{R}$$

then f_1, f_2 are polynomials of degree less than or equal to 2. *Hint:* differentiate!

Step 5. If X is square-integrable and $\log\phi_X$ is a polynomial of degree less than or equal to 2, then X is a Gaussian random variable.

Hint. If X is square-integrable, then you can take for granted that $\phi_X(0)=1$, $\phi'_X(0)=i\mathbb{E}(X)$ and $\phi''_X(0)=-\mathbb{E}(X^2)$.

Step 6. From the course, deduce that X_1, X_2 are jointly Gaussian and that $\mathrm{Var}(X_1)=\mathrm{Var}(X_2)$.

7 Inequalities

Exercise 7.1. a) Let X be a square-integrable random variable such that $\mathbb{E}(X)=0$ and $\mathrm{Var}(X)=\sigma^2$. Show that

$$\mathbb{P}(\{X\geq t\})\leq\frac{\sigma^2}{\sigma^2+t^2} \quad \text{for } t>0$$

Hint: You may try various versions of Chebyshev's inequality here, but not all of them work. A possibility is to use the function $\psi(x)=(x+b)^2$, where b is a free parameter to optimize (but watch out that only some values of $b\in\mathbb{R}$ lead to a function ψ that satisfies the required hypotheses).

b) Deduce from a) that for any square-integrable random variable X with expectation μ and variance σ^2, the following inequality holds:

$$\mathbb{P}(\{X\geq\mu+\sigma\})\leq\frac{1}{2}$$

c) *Numerical application:* Check the inequality in b) for $X\sim\mathrm{Bern}(\frac{1}{2})$.

d) (Paley-Zygmund's inequality) Let X be a square-integrable random variable such that $\mathbb{E}(X)>0$. Show that

$$\mathbb{P}(\{X>t\})\geq\frac{(\mathbb{E}(X)-t)^2}{\mathbb{E}(X^2)} \quad \forall 0\leq t\leq\mathbb{E}(X)$$

Hint: Use first Cauchy-Schwarz' inequality with the random variables X and $Y=1_{\{X>t\}}$.

e) Deduce from d) that for any square-integrable random variable X with expectation $\mu>0$ and variance σ^2 satisfying $0\leq\sigma\leq\mu$, the following inequality holds:

$$\mathbb{P}(\{X>\mu-\sigma\})\geq\frac{\sigma^2}{\sigma^2+\mu^2}$$

f) *Numerical application:* Check the inequality in e) for $X\sim\mathrm{Bern}(\frac{1}{2})$.

Exercise 7.2. (Kingman's bound)
Let D_1, X be independent and square-integrable random variables such that $\mathbb{E}(X)=\mu\in\mathbb{R}$, $\mathrm{Var}(X)=\sigma^2>0$ and let us define

$$D_2=(D_1+X)^+=\max(D_1+X,0)$$

a) Show that if $\mathbb{E}(X)=\mu<0$ *and* D_1 and D_2 are identically distributed, then

$$d=\mathbb{E}(D_1)=\mathbb{E}(D_2)\leq\frac{\sigma^2}{2|\mu|}$$

Hint: Define $Y=(D_1+X)^-=\max(-(D_1+X),0)$ and use the fact that $D_2\cdot Y=0$.

b) Show that if on the contrary $\mathbb{E}(X) = \mu > 0$ and $\mathbb{E}(D_1) \geq 0$, then $\mathbb{E}(D_2) \geq \mathbb{E}(D_1) + \mu$.

Application: D_1, D_2 can be interpreted as the queuing delays for two consecutive customers in a queue and X as the difference between the service time of the first customer and the inter-arrival time between the two customers. If $\mathbb{E}(X) < 0$ (i.e., if the service time is on average smaller than the inter-arrival time between the two customers), then the queue is stable and we expect in this case that in the long run, each customer experiences a delay with the same distribution. The above result provides then an upper bound on the average delay of each customer. If on the contrary $\mathbb{E}(X) > 0$, then the expected queuing delay will increase steadily as more customers arrive.

Exercise 7.3. a) Let X be a random variable such that $X(\omega) \in [0,1]$ for all $\omega \in \Omega$. Show that

$$\text{if} \quad \mathbb{E}(X) \geq \frac{1}{2} \quad \text{then} \quad \mathbb{P}(\{X \leq \frac{1}{4}\}) \leq \frac{2}{3}$$

b) Find a random variable X with values in $[0,1]$ satisfying both $\mathbb{E}(X) = \frac{1}{2}$ and $\mathbb{P}\left(\left\{X \leq \frac{1}{4}\right\}\right) = \frac{2}{3}$.

c) Let now X be a square-integrable and non-negative random variable. Show that

$$1. \quad \mathbb{P}(\{X = 0\}) \leq \frac{\mathrm{Var}(X)}{\mathbb{E}(X)^2}$$

$$2. \quad \mathbb{P}(\{X = 0\}) \leq \frac{\mathrm{Var}(X)}{\mathbb{E}(X^2)}$$

$$3. \quad \mathbb{P}(\{X > 0\}) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}$$

Hint for 2 and 3: Use Cauchy-Schwarz' inequality with the random variables X and $Y = 1_{\{X>0\}}$.

d) Verify the claims of part c) for $X \sim \mathrm{Poisson}(\lambda)$, where $\lambda > 0$.

Exercise 7.4. Let X be a centered random variable with variance σ^2. Using Chebyshev's inequality, show that:

a) $\mathbb{P}(\{|X| \geq a\}) \leq \dfrac{\sigma^2}{a^2}$ and $\mathbb{P}(\{|X| \geq a\}) \leq \dfrac{2\sigma^2}{a^2 + \sigma^2}$.

b) $\mathbb{P}(\{X \geq a\}) \leq \dfrac{\sigma^2}{a^2 + \sigma^2}$ (use $\psi(x) = (x+b)^2$ with $b \geq 0$, then minimize over b).

Note that in general, there is no guarantee that $\mathbb{P}(\{X \geq a\}) = \frac{1}{2}\,\mathbb{P}(\{|X| \geq a\})$, so that the inequality in b) is not a simple consequence of the second one in a).

8 Convergence in probability, almost sure convergence and the laws of large numbers

Exercise 8.1. Let $(X_n,\, n \geq 1)$ be independent random variables such that $X_n \sim \mathrm{Bern}(1 - \frac{1}{(n+1)^\alpha})$ for $n \geq 1$, where $\alpha > 0$ is some constant. Let also $Y_n = \prod_{j=1}^n X_j$ for $n \geq 1$.

a) What minimal condition on the parameter $\alpha > 0$ ensures that $Y_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$?

Hint: Use the approximation $1 - x \simeq \exp(-x)$ for x small.

b) Under the same condition as that found in a), does it also hold that $Y_n \underset{n\to\infty}{\overset{L^2}{\to}} 0$?

c) Under the same condition as that found in a), does it also hold that $Y_n \underset{n\to\infty}{\to} 0$ almost surely ?

Hint: If $Y_n = 0$, what can you deduce on Y_m for $m \geq n$?

Exercise 8.2. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $\mathbb{P}(\{X_n = 2\}) = \mathbb{P}(\{X_n = 0\}) = \frac{1}{2}$ for every $n \geq 1$. Let also $(Y_n,\, n \geq 1)$ be the sequence of random variables defined as

$$Y_n = \sum_{j=1}^{n} \frac{X_j}{3^j} \qquad n \geq 1$$

a) Run a numerical simulation illustrating the fact that there exists a limiting random variable Y such that $Y_n \underset{n\to\infty}{\to} Y$ almost surely.

b) Prove theoretically that such a random variable Y exists.

c) Run a numerical simulation illustrating the fact that $\mathbb{E}((Y_n - Y)^2) \underset{n\to\infty}{\to} 0$.

Hint. The aim here is not to compute the expectation, but to estimate it via multiple runs.

d) Prove theoretically that $Y_n \underset{n\to\infty}{\overset{L^2}{\to}} Y$.

e) Run a numerical simulation allowing to draw the empirical distribution of Y (using either a histogram or its empirical cdf). Can you guess what this distribution is?

Exercise 8.3. a) Show that if $(A_n,\, n \geq 1)$ are *independent* events in $\mathcal{F}$ and $\sum_{n\geq 1} \mathbb{P}(A_n) = \infty$, then

$$\mathbb{P}\Big(\bigcup_{n\geq 1} A_n\Big) = 1$$

Hints: - Start by observing that the statement is equivalent to $\mathbb{P}\left(\bigcap_{n\geq 1} A_n^c\right) = 0$.
- Use the inequality $1 - x \leq e^{-x}$, valid for all $x \in \mathbb{R}$.

b) From the same set of assumptions, reach the following stronger conclusion with a little extra effort:

$$\mathbb{P}(\{\omega \in \Omega : \omega \in A_n \text{ infinitely often}\}) = \mathbb{P}\Big(\bigcap_{N\geq 1} \bigcup_{n\geq N} A_n\Big) = 1$$

which is actually the statement of the *second Borel-Cantelli lemma.*

c) *Application:* Let $(X_n,\, n \geq 1)$ be a sequence of *independent* random variables such that for some $\varepsilon > 0$, $\sum_{n\geq 1} \mathbb{P}(\{|X_n| > \varepsilon\}) = +\infty$. What can you conclude on the almost sure convergence of the sequence X_n towards the limiting value 0?

Exercise 8.4. a) Let $(X_n,\, n \geq 1)$ be a sequence of independent random variables such that $\mathbb{P}(\{X_n = n\}) = p_n = 1 - \mathbb{P}(\{X_n = 0\})$ for $n \geq 1$.

What minimal condition on the sequence $(p_n, n \geq 1)$ ensures that a1) $X_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$? a2) $X_n \underset{n\to\infty}{\overset{L^2}{\to}} 0$? a3) $X_n \underset{n\to\infty}{\to} 0$ almost surely?

b) Let $(Y_n,\, n \geq 1)$ be a sequence of independent random variables such that $Y_n \sim \mathcal{U}([0, u_n])$ for $n \geq 1$.

What minimal condition on the sequence $(u_n,\, n \geq 1)$ ensures that b1) $Y_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$? b2) $Y_n \underset{n\to\infty}{\overset{L^2}{\to}} 0$? b3) $Y_n \underset{n\to\infty}{\to} 0$ almost surely?

c) Let $(Z_n,\, n \geq 1)$ be a sequence of independent random variables such that $Z_n \sim \text{Cauchy}(\lambda_n)$ for $n \geq 1$.

What minimal condition on the sequence $(\lambda_n,\, n \geq 1)$ ensures that c1) $Z_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$? c2) $Z_n \underset{n\to\infty}{\overset{L^2}{\to}} 0$? c3) $Z_n \underset{n\to\infty}{\to} 0$ almost surely?

Hint: Use the first and second Borel-Cantelli lemmas to answer questions about almost sure convergence.

Exercise 8.5. (extended law of large numbers)
Let $(\mu_n,\, n \geq 1)$ be a sequence of real numbers such that

$$\lim_{n\to\infty} \frac{\mu_1 + \ldots + \mu_n}{n} = \mu \in \mathbb{R}$$

Let $(X_n,\, n \geq 1)$ be a sequence of square-integrable random variables such that

$$\mathbb{E}(X_n) = \mu_n, \quad \forall n \geq 1 \quad \text{and} \quad \mathrm{Cov}(X_n, X_m) \leq C_1\, \exp(-C_2\, |m-n|) \quad \forall m, n \geq 1$$

for some contants $C_1, C_2 > 0$ (the random variables X_n are said to be *weakly* correlated). Let finally $S_n = X_1 + \ldots + X_n$.

a) Show that

$$\frac{S_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} \mu$$

b) Is it also true that

$$\frac{S_n}{n} \underset{n\to\infty}{\to} \mu \quad \text{almost surely?}$$

In order to check this, you need to go through the proof of the strong law of large numbers made in class. Does that proof need the fact that the random variables X_n are independent?

c) Application: Let $(Z_n,\, n \geq 1)$ be a sequence of i.i.d. $\sim \mathcal{N}(0,1)$ random variables, $x, a \in \mathbb{R}$ and $(X_n,\, n \geq 1)$ be the sequence of random variables defined recursively as

$$X_1 = x, \quad X_{n+1} = aX_n + Z_{n+1} \quad n \geq 1$$

For what values of $x, a \in \mathbb{R}$ does the sequence $(X_n,\, n \geq 1)$ satisfy the assumptions made in a)? Compute μ in this case.

Exercise 8.6. (another extension of the weak law of large numbers)
Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. square-integrable random variables such that $\mathbb{E}(X_1) = \mu \in \mathbb{R}$ and $\mathrm{Var}(X_1) = \sigma^2 > 0$.

Let $(T_n,\, n \geq 1)$ be another sequence of random variables, independent of the sequence $(X_n,\, n \geq 1)$, with all T_n taking values in the set of natural numbers $\mathbb{N}^* = \{1, 2, 3, \ldots\}$. Define

$$p_k^{(n)} = \mathbb{P}(\{T_n = k\}) \quad \text{for } n, k \geq 1 \qquad \left(\text{so } \sum_{k\geq 1} p_k^{(n)} = 1 \quad \forall n \geq 1\right)$$

a) Find a sufficient condition on the numbers $p_k^{(n)}$ guaranteeing that

$$\frac{X_1 + \ldots + X_{T_n}}{T_n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} \mu \tag{2}$$

Hint: You should use the law of total probability here: if A is an event and the events $(B_k,\, k \geq 1)$ form a partition of Ω, then:

$$\mathbb{P}(A) = \sum_{k\geq 1} \mathbb{P}(A \,|\, B_k)\, \mathbb{P}(B_k)$$

b) Apply the above criterion to the following case: each T_n is the sum of two independent geometric random variables $G_{n1} + G_{n2}$, where both G_n are distributed as

$$\mathbb{P}(\{G_n = k\}) = q_n^{k-1}\,(1 - q_n) \quad k \geq 1$$

where $0 < q_n < 1$.

b1) Compute first the distribution of T_n, as well as $\mathbb{E}(T_n)$, for each $n \geq 1$.

b2) What condition on the sequence $(q_n,\, n \geq 1)$ ensures that conclusion (2) holds?

Hint: Solving question b1) above may help you guessing what the answer to b2) should be.

Exercise 8.7. (strong law of large numbers in a "simple" setup)
Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such that $\mathbb{E}(X_1) = 0$, $\mathbb{E}(X_1^2) = 1$ and $\mathbb{E}(X_1^4) = C < \infty$. Let also $S_n = X_1 + \ldots + X_n$. Without relying on the proof of the strong law made in class, show that

$$\frac{S_n}{n} \underset{n\to\infty}{\to} 0 \quad \text{almost surely.}$$

Hint: As often, Chebyshev's inequality and the (first) Borel-Cantelli lemma will be useful here.

Exercise 8.8. Let $(X_n,\, n \geq 1)$ be a sequence of independent random variables such that $X_1 = 0$ and

$$\mathbb{P}(\{X_n = +n\}) = \mathbb{P}(\{X_n = -n\}) = \frac{1}{2n \log n} \quad \mathbb{P}(\{X_n = 0\}) = 1 - \frac{1}{n \log n} \quad n \geq 2$$

Let also $S_n = X_1 + \ldots + X_n$. Show that

$$\frac{S_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0 \quad \text{but} \quad \frac{S_n}{n} \underset{n\to\infty}{\not\to} 0 \quad \text{a.s.}$$

NB: For the proof of the second statement (which is more challenging than the first one), you will need the second Borel-Cantelli lemma.

Exercise 8.9. a) Let $Y \sim \mathcal{U}([0,1])$ and $X_n = \sqrt{n}\, 1_{\{Y \leq 1/n\}}$ for $n \geq 1$.

Does the sequence of random variables $(X_n,\, n \geq 1)$ converge in L^2? in probability? almost surely?

b) Let Y_n be i.i.d. $\sim \mathcal{U}([0,1])$ random variables and $X_n = \sqrt{n}\, 1_{\{Y_n \leq 1/n\}}$ for $n \geq 1$.

Does the sequence of random variables $(X_n,\, n \geq 1)$ converge in L^2? in probability? almost surely?

Exercise 8.10. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. non-negative random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $\mathbb{E}(|\log(X_1)|) < +\infty$. Let also $(Y_n,\, n \geq 1)$ be the sequence defined as

$$Y_n = \left(\prod_{j=1}^{n} X_j \right)^{1/n}, \quad n \geq 1$$

a) Show that there exists a constant $\mu > 0$ such that $Y_n \underset{n\to\infty}{\to} \mu$ almost surely.

b) Compute the value of μ in the case where $X_n = \exp(N_n)$ and $(N_n,\, n \geq 1)$ are i.i.d. $\sim \mathcal{N}(0,1)$ random variables.

c) In this case, look for the tightest possible upper bound on $\mathbb{P}(\{Y_n > t\})$ for $n \geq 1$ fixed and $t > \mu$.

Hint. You have two options here. One is to use Chebyshev's inequality with the function $\psi(x) = x^p$ and $p > 0$ (and then optimize over p) in order to upperbound

$$\mathbb{P}(\{Y_n > t\}) = \mathbb{P}\left(\left\{ \prod_{j=1}^{n} X_j > t^n \right\} \right)$$

for $t > \mu$. The other option is left to your imagination...

Exercise 8.11. a) Let $X_1, \dots, X_n$ be i.i.d. random variables with common cdf F. Express the cdf of $S_n = \max(X_1, \dots, X_n)$ in terms of F.

b) Let X be a discrete random variable with values in $\mathbb{N} = \{0, 1, 2, \dots\}$. Show that

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

Let now $X_1, \dots, X_n$ be i.i.d. random variables such that

$$\mathbb{P}(\{X_1 = k\}) = \frac{1}{2^{k+1}}, \quad k \geq 0.$$

and let $S_n = \max(X_1, \dots, X_n)$.

c) Show that if $k \geq c \log_2 n$ with $c > 1$, then

$$\mathbb{P}(\{S_n \geq k\}) \underset{n\to\infty}{\to} 0.$$

Hint: $\lim_{n\to\infty}(1 + f(n))^n = \begin{cases} 1 & \text{if } f(n) = o(1/n) \\ e^x & \text{if } f(n) = x/n \end{cases}$

d) Deduce from there that there exist positive constants $C_1 < C_2$ such that

$$C_1 \log_2 n \leq \mathbb{E}(S_n) \leq C_2 \log_2 n.$$

e) Show that

$$\frac{S_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0.$$

f) Show that

$$\frac{S_n}{n} \underset{n\to\infty}{\to} 0 \quad \text{almost surely.}$$

Hint: Chebyshev's inequality and the Borel-Cantelli lemma will be useful here.

Exercise 8.12. Let $(X_n, n \geq 0)$ be a sequence of random variables and X be another random variable, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

a) Show that if $X_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} X$, then there exists a subsequence $(X_{n_k}, k \geq 1)$ such that $X_{n_k} \underset{k\to\infty}{\to} X$ almost surely.

Hint: The Borel-Cantelli lemma may again be useful here...

Let us now define for two random variables X and Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$:

$$d(X, Y) = \mathbb{E}\left(\frac{|X - Y|}{1 + |X - Y|}\right).$$

This is known as the *Ky-Fan metric*: it is a metric for convergence in probability (notice that the above expectation is always finite, whatever X and Y).

b) Show that the triangle inequality is satisfied, namely that

$$d(X, Z) \leq d(X, Y) + d(Y, Z), \quad \text{for any triple of random variables } X, Y, Z.$$

c) Show that $X_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} X$ if and only if $d(X_n, X) \underset{n\to\infty}{\to} 0$.

9 Convergence in distribution and the central limit theorem

Exercise 9.1. (the birthday problem)
Let $(X_n, n \geq 1)$ be a sequence of i.i.d. random variables, each uniform on $\{1, \ldots, N\}$. Let also

$$T_N = \min\{n \geq 1 \,:\, X_n = X_m \text{ for some } m < n\}$$

(notice that whatever happens, $T_N \in \{2, \ldots, N+1\}$). Show that

$$\mathbb{P}\left(\left\{\frac{T_N}{\sqrt{N}} \leq t\right\}\right) \underset{N\to\infty}{\to} 1 - e^{-t^2/2}, \quad \forall t \geq 0$$

Remarks:
- Approximations are allowed here!

- Please observe that the limit distribution is *not* the Gaussian distribution!

Numerical application: Use this to obtain a rough estimate of $\mathbb{P}(\{T_{365} \leq 22\})$ and $\mathbb{P}(\{T_{365} \leq 50\})$ (i.e., what is the probability that among 22 / 50 people, at least two share the same birthday?)

Exercise 9.2. Someone proposes you to play the following game: start with an initial amount of $S_0 > 0$ francs, of your choice. Then toss a coin: if it falls on heads, you win $S_0/2$ francs; while if it falls on tails, you lose $S_0/2$ francs. Call S_1 your amount after this first coin toss. Then the game goes on, so that your amount after coin toss number $n \geq 1$ is given by

$$S_n = \begin{cases} S_{n-1} + \frac{S_{n-1}}{2} & \text{if coin number } n \text{ falls on heads} \\ \\ S_{n-1} - \frac{S_{n-1}}{2} & \text{if coin number } n \text{ falls on tails} \end{cases}$$

We assume moreover that the coin tosses are independent and fair, i.e., with probability 1/2 to fall on each side. Nevertheless, you should *not* agree to play such a game: explain why!

Hints:

First, to ease the notation, define $X_n = +1$ if coin n falls on heads and $X_n = -1$ if coin n falls on tails. That way, the above recursive relation may be rewritten as $S_n = S_{n-1}\,(1 + \frac{X_n}{2})$ for $n \geq 1$.

a) Compute recursively $\mathbb{E}(S_n)$; if it were only for expectation, you could still consider playing such a game, but…

b) Define now $Y_n = \log(S_n/S_0)$, and use the central limit theorem to approximate $\mathbb{P}(\{Y_n > t\})$ for a fixed value of $t \in \mathbb{R}$ and a relatively large value of n. Argue from there why it is definitely not a good idea to play such a game! (computing for example an approximate value of $\mathbb{P}(\{S_{100} > S_0/10\})$)

Exercise 9.3. Let $\lambda > 0$ be fixed. For a given $n \geq \lceil \lambda \rceil$, let $X_1^{(n)}, \ldots, X_n^{(n)}$ be i.i.d. Bernoulli(λ/n) random variables and let $S_n = X_1^{(n)} + \ldots + X_n^{(n)}$.

a) Compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$ for a fixed value of $n \geq \lceil \lambda \rceil$.

b) Deduce the value of $\mu = \lim_{n\to\infty} \mathbb{E}(S_n)$ and $\sigma^2 = \lim_{n\to\infty} \mathrm{Var}(S_n)$.

c) Compute the limiting distribution of S_n (as $n \to \infty$).

Hint: Use characteristic functions. You might also have a look at tables of characteristic functions of some well known distributions in order to solve this exercise.

For a given $n \geq 1$, let now $(Y_m^{(n)},\, m \geq 1)$ be a sequence of i.i.d. Bernoulli$(1/n)$ random variables and let

$$T_n = Y_1^{(n)} + \ldots + Y_{\lceil \lambda n \rceil}^{(n)}$$

where $\lambda > 0$ is the same as above.

d) Compute the limiting distribution of T_n (as $n \to \infty$).

e) Is it possible to talk about convergence in probability or almost sure convergence of any of the two sequences S_n or T_n? Justify your answer!

Exercise 9.4. (application of Lindeberg's principle to non i.i.d. random variables)
Let $(\sigma_n,\, n \geq 0)$ be a sequence of (strictly) positive numbers and $(X_n,\, n \geq 1)$ be a sequence of independent random variables such that $\mathbb{E}(X_n) = 0$, $\mathrm{Var}(X_n) = \sigma_n^2$ and $\mathbb{E}(|X_n|^3) \leq K\, \sigma_n^3$ for every $n \geq 1$ (note that the constant K is uniform over all values of n).

For $n \geq 1$, define also $V_n = \mathrm{Var}(X_1 + \ldots + X_n) = \sigma_1^2 + \ldots + \sigma_n^2$.

Using Lemma 9.12 in the lecture notes (equivalently, Lemma 2 in the video lecture 9.2b), find a sufficient condition on the sequence $(\sigma_n,\, n \geq 1)$ guaranteeing that

$$\frac{1}{\sqrt{V_n}}\,(X_1 + \ldots + X_n) \underset{n\to\infty}{\overset{d}{\to}} Z \sim \mathcal{N}(0,1)$$

Note: From the course, you already know a sufficient condition: $\sigma_n = 1$ for all $n \geq 1$, but this is too strong! The aim here is to find a sufficient condition which is most general possible.

b) Which of the following sequences $(\sigma_n,\, n \geq 1)$ satisfy the condition you have found in a)?

b1) $\sigma_n = n$ b2) $\sigma_n = \frac{1}{n}$ b3) $\sigma_n = 2^n$

Hint: The following might be useful: $$\sum_{j=1}^n j^\alpha = \begin{cases} \Theta(n^{\alpha+1}) & \text{if } \alpha > -1 \\ \Theta(\log(n)) & \text{if } \alpha = -1 \\ \Theta(1) & \text{if } \alpha < -1 \end{cases}$$

Exercise 9.5. Let $(X_n,\, n \geq 1)$ be a sequence of independent random variables such that

$$\mathbb{P}\left(\{X_n = +1/\sqrt{n}\}\right) = \mathbb{P}\left(\{X_n = -1/\sqrt{n}\}\right) = \frac{1}{2}$$

and let, for $n \geq 1$,

$$Y_n = X_1 + \ldots + X_n \quad \text{and} \quad Z_n = X_{n+1} + \ldots + X_{2n}$$

a) Run multiple times the process Y and draw a histogram of Y_n for $n = 100$, $n = 1'000$ and $n = 10'000$, respectively. Draw also the graphs of the empirical mean and standard deviation of Y_n as a function of n. Do you observe that the histogram of Y_n converges as n grows large (i.e., that Y_n converges in distribution)?

b) Same questions for the process Z.

c) In the case(s) where you observed convergence in distribution, prove that the sequence of random variables indeed converges to a limit, using characteristic functions. What is the limiting distribution?

Hint: You may use approximations here, as well as the following:

$$\sum_{j=n_1+1}^{n_2} j^\alpha \simeq \int_{n_1}^{n_2} dx\, x^\alpha$$

as n_2 gets large (and n_1 is either fixed or getting large also).

Exercise 9.6. (why it is not a good idea to play at roulette too many times)
On a classical roulette game with 38 numbers (including the 0 and the 00), a player bets uniquely on red, 361 times in a row. At each turn, he bets exactly one franc (he therefore wins one franc if red comes out and loses one franc if this is not the case). Assuming that the roulette wheel is balanced and that the turns are independent from each other, give a rough estimate of:

a) the average player's fortune at the end of the 361 games;

b) the probability that he has actually won some money.

NB: Remember that the numbers 0 and 00 are neither red nor black on a classical roulette.

Exercise 9.7. Let $\lambda > 0$ and $(X_n, n \geq 1)$ be a sequence of i.i.d. random variables with common characteristic function ϕ_{X_1} given by

$$\phi_{X_1}(t) = \exp(-\lambda\,|t|) \quad t \in \mathbb{R}$$

a) Compute the distribution of X_1 using the inversion formula. Does X_1 admit a pdf?

b) Compute $\mathbb{P}(\{|X_1| \leq \lambda\})$.

Let now $S_n = X_1 + \ldots + X_n$.

c) Compute the characteristic function of S_n/n.

d) To what random variable Z does S_n/n converge in distribution as $n \to \infty$?

e*) Does S_n/n converge also in probability to Z?

Exercise 9.8. Let us consider the following experiments.

A) Two dice are rolled independently; the result X is the sum of the two.

B) Two *different* numbers are chosen uniformly in $\{1, \ldots, 6\}$; the result Y is the sum of the two.

C) Same as B), but one number is chosen between 1 and 3, while the other is chosen between 4 and 6; the result Z is the sum of the two.

D) Same as B), but one number is chosen to be odd, while the other is chosen to be even; the result W is the sum of the two.

Assume now that each of these experiments is run n times independently, and let S_n/n denote the average of these n runs.

a) For each experiment, compute $\mathbb{E}(S_n/n)$. Does this value depend on the experiment?

b) For which experiment is S_n/n the closest to its actual expectation (on average)?

c) For this experiment in particular and $n = 1'000$, estimate approximately the value $t > 0$ such that

$$\mathbb{P}(|S_n/n - \mathbb{E}(S_n/n)| \leq t) = 95\%$$

10 Moments and Carleman's theorem

Exercise 10.1. Let X be a bounded random variable.

a) Show that $\mathbb{E}(|X|^k) < +\infty$ for every $k \geq 1$.

b) Show that if $\ell \geq k \geq 1$, then $\mathbb{E}(|X|^\ell)^{1/\ell} \geq \mathbb{E}(|X|^k)^{1/k}$.

c) Show that if $\ell, k \geq 1$, then $\mathbb{E}(|X|^{k+\ell}) \leq \sqrt{\mathbb{E}(X^{2k})\,\mathbb{E}(X^{2\ell})}$.

d) Show that if $X(\omega) \in [0,1]$ for every $\omega \in \Omega$ and $\ell \geq k \geq 1$, then $\mathbb{E}(X^\ell) \leq \mathbb{E}(X^k) \leq 1$.

e) Given the relations in b), c) and d), for which of the following sequences of non-negative numbers $(m_k, k \geq 1)$ does there possibly exist a random variable X taking values in the interval $[0,1]$ such that $m_k = \mathbb{E}(X^k)$ for every $k \geq 1$?

1. $(m_k = \frac{1}{k+1}, k \geq 0)$ 2. $(m_k = 1 - \frac{1}{k+1}, k \geq 0)$ 3. $(m_k = \frac{1}{2^k}, k \geq 0)$ 4. $(m_k = \frac{1}{k^k}, k \geq 0)$

In the cases for which the answer is affirmative, can you guess what the corresponding distribution of X is?

Exercise 10.2. a) Let $X \sim \mathcal{N}(0,1)$. Compute all the moments of the random variable $Y = \exp(X)$. Do these satisfy Carleman's condition?

b) Let W be the discrete random variable such that

$$\mathbb{P}(\{W = j\}) = C \exp(-j^2/2) \quad j \in \mathbb{Z}$$

where $C = 1/\sum_{j\in\mathbb{Z}} \exp(-j^2/2)$. Compute all the moments of the random variable $Z = \exp(W)$. Do these satisfy Carleman's condition?

Exercise 10.3. a) Let $X \sim \mathcal{N}(0,\sigma^2)$ with $\sigma > 0$ and $f \in C^1(\mathbb{R})$ be such that $\exists C > 0$ and $q > 0$ such that $|f(x)|, |f'(x)| \leq C\,(1+x^2)^q$ for all $x \in \mathbb{R}$. Show that

$$\mathbb{E}(X f(X)) = \sigma^2\, \mathbb{E}(f'(X))$$

b) Use a) to deduce the value of $\mathbb{E}(X^{2k})$ for $k \geq 1$. Do these moments satisfy Carleman's condition?

c) Let $Y \sim \mathcal{P}(\lambda) > 0$ with $\lambda > 0$ and $g : \mathbb{N} \to \mathbb{R}$ be such that $\exists C > 0$ and $q > 0$ such that $|g(k)| \leq C\,(1+k^2)^q$ for all $k \in \mathbb{N}$. Show that

$$\mathbb{E}(Y g(Y)) = \lambda\, \mathbb{E}(g(Y+1))$$

d) Use c) to deduce the value of $\mathbb{E}(Y(Y-1)(Y-2)\cdots(Y-p+1))$ for $p \geq 1$.

Exercise 10.4. (proof of the central limit theorem using moments)
Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. square-integrable random variables such that

$$\mathbb{E}(X_1) = 0 \quad \text{and} \quad \mathbb{E}(X_1^2) = 1$$

Let $S_n = \sum_{j=1}^n X_j$. The central limit theorem asserts that $\frac{S_n}{\sqrt{n}} \xrightarrow{d} Z$ as $n \to \infty$, where $Z \sim \mathcal{N}(0,1)$.

Prove this theorem *using moments*, under the following stronger assumptions:

a) X_1 is a bounded random variable (so all moments exist).

b) $\mathbb{E}(X_1^{2k+1}) = 0$, for all $k \geq 0$.

Hint: For the even moments, use the multinomial expansion

$$(x_1 + \ldots + x_n)^{2k} = \sum_{\substack{j_1,\ldots j_n \geq 0 \\ j_1+\ldots+j_n = 2k}} \binom{2k}{j_1, \ldots, j_n} x_1^{j_1} \cdots x_n^{j_n}$$

and divide the sum into two parts as follows: $\sum_{\substack{j_1,\ldots j_n \geq 0 \\ j_1+\ldots+j_n=2k}} = \sum_{\substack{j_1,\ldots j_n \in\{0,2\} \\ j_1+\ldots+j_n=2k}} + \sum_{\substack{\exists 1\leq i\leq n\,:\, j_i \notin\{0,2\} \\ j_1+\ldots+j_n=2k}}$.

11 Concentration inequalities

Exercise 11.1. Regarding the "balls into bins" problem, let us define $S_j^{(m)}$ to be the number of balls having landed in bin $j \in \{1, \ldots, n\}$ after m throws (with all balls landing independently and uniformly in one of the n bins).

a) Using Chebyshev's inequality with $\psi(x) = \exp(sx)$ and optimizing over $s > 0$, find an upper bound on

$$\mathbb{P}(\{S_1^{(m)} \geq k\}) \quad \text{for } n \text{ large, } m = \lambda n \text{ and } k > \lambda$$

with $\lambda > 0$ a fixed parameter.

Hint: At some point, you may use the inequality $1 + x \leq e^x$, valid for $x \in \mathbb{R}$, in order to get a nice expression for the upper bound.

b) Let now $S^{(m)} = \max\{S_1^{(m)}, \ldots, S_n^{(m)}\}$ (considering still $m = \lambda n$). Show that

$$\mathbb{P}(\{S^{(m)} > \log(n)\}) = O\left(\frac{1}{n^2}\right)$$

Exercise 11.2. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. $\mathcal{E}(\lambda)$ random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., X_1 admits the following pdf:

$$p_{X_1}(x) = \begin{cases} \lambda \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Let also $S_n = X_1 + \ldots + X_n$. Using the large deviations principle, find a tight upper bound on

$$\mathbb{P}\left(\{S_n \geq nt\}\right) \quad \text{for } t > \mathbb{E}(X_1) = \frac{1}{\lambda}$$

as well as a tight upper bound on

$$\mathbb{P}\left(\{S_n \leq nt\}\right) \quad \text{for } t < \mathbb{E}(X_1) = \frac{1}{\lambda}$$

Watch out that there is a slight asymmetry between the two problems!

Exercise 11.3. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such that

$$\mathbb{P}(\{X_1 = +1\}) = \mathbb{P}(\{X_1 = -1\}) = \frac{1}{2}$$

Let also $S_n = X_1 + \ldots + X_n$ for $n \geq 1$. For a fixed value of n, draw on the same graph the following functions:

$$\begin{aligned} f(t) &= -\frac{1}{n} \log \mathbb{P}(\{S_n > nt\}) \\ g(t) &= \Lambda^*(t) = \max_{s \in \mathbb{R}} (st - \Lambda(s)) \quad \text{where } \Lambda(s) = \log \mathbb{E}(e^{sX_1}) \\ h(t) &= t^2/2 \end{aligned}$$

NB: On these plots, $t \in [0, +1]$.

In order to draw the function $f(t)$, you should use Monte-Carlo simulation, that is, draw i.i.d. samples $X_1^{(m)}, \ldots, X_n^{(m)}$ for $m = 1, \ldots, M$ (with M reasonably large) and approximate $f(t)$ as

$$f(t) \simeq -\frac{1}{n} \log\left(\frac{1}{M}\, \sharp\{1 \leq m \leq M \,:\, S_n^{(m)} > nt\}\right)$$

As you will see, considering even moderate values of n requires considering quite large values of M.

Exercise 11.4. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let also $S_n = X_1 + \ldots + X_n$. Find the *exact value* of

$$\mathbb{P}\left(\{S_n \geq nt\}\right), \quad \text{for } t > 0.$$

Exercise 11.5. For this exercise, you will need a generalization of the Cauchy-Schwartz inequality: Hölder's inequality (written here in a slightly unsual form to help you with the exercise). This inequality says that if X, Y are two integrable random variables, then for every $\alpha \in [0,1]$,

$$\mathbb{E}\left(|X|^\alpha\, |Y|^{1-\alpha}\right) \leq \mathbb{E}(|X|)^\alpha\, \mathbb{E}(|Y|)^{1-\alpha}.$$

Preliminary: show that for $\alpha = 1/2$, this is nothing but Cauchy-Schwarz' inequality.

Let now X be a random variable such that $\mathbb{E}(\exp(sX)) < \infty$ for every $s \in \mathbb{R}$.

a) Show that the function $\Lambda(s) = \log(\mathbb{E}(\exp(sX))$ is convex.

b) Show that the function $\Lambda^*(t) = \sup_{s\in\mathbb{R}}(st - \Lambda(s))$ is also convex.

12 Conditional expectation

Exercise 12.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, X be an square-integrable random variable defined on this space and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Relying only on the definition of conditional expectation, show the following properties:

a) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$.

b) If X is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ a.s.

c) If X is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$ a.s.

d) If Y is $\mathcal{G}$-measurable and bounded, then $\mathbb{E}(XY|\mathcal{G}) = \mathbb{E}(X|\mathcal{G})\, Y$ a.s.

e) If $\mathcal{H}$ is a sub-σ-field of $\mathcal{G}$, then $\mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H})$ a.s.

Hint for parts b) to e): According to the course definition, in order to check that some candidate random variable Z is the conditional expectation of X given $\mathcal{G}$, you should check the following two conditions:

(i) $Z \in G$, i.e., Z is $\mathcal{G}$-measurable and square-integrable;

(ii) Z satisfies $\mathbb{E}((Z - X)\, U) = 0$ for every $U \in G$.

Exercise 12.2. Let X be a random variable such that $\mathbb{P}(\{X = +1\}) = \mathbb{P}(\{X = -1\}) = \frac{1}{2}$ and $Z \sim \mathcal{N}(0,1)$ be independent of X. Let also $a > 0$ and $Y = aX + Z$. We propose below four possible estimators of the variable X given the noisy observation Y:

$$\widehat{X}_1 = \frac{Y}{a} \qquad \widehat{X}_2 = \frac{aY}{a^2+1} \qquad \widehat{X}_3 = \operatorname{sign}(aY) \qquad \widehat{X}_4 = \tanh(aY)$$

a) Which estimator among these four minimizes the mean square error (MSE) $\mathbb{E}((\widehat{X} - X)^2)$?

In order to answer the question, draw on the same graph the four curves representing the MSE as a function of $a > 0$. For this, you may use either the exact mathematical expression of the MSE or the one obtained via Monte-Carlo simulations.

b) Provide a theoretical justification for your conclusion.

c) For which of the four estimators above does it hold that $\mathbb{E}((\widehat{X} - X)^2) = \mathbb{E}(X^2) - \mathbb{E}(\widehat{X}^2)$?

Exercise 12.3. Let X, Y be two discrete random variables (with values in a countable set D). Let us moreover assume that X is integrable.

a) Show that the random variable $\psi(Y)$, where ψ is defined as

$$\psi(y) = \sum_{x \in D} x\, \mathbb{P}(\{X = x\}|\{Y = y\})$$

matches the theoretical definition of conditional expectation $\mathbb{E}(X|Y)$.

b) *Application:* One rolls two independent and balanced dice (say Y and Z), each with four faces. What is the conditional expectation of the maximum of the two, given the value of one of them?

Exercise 12.4. (particular case of a proposition seen in the course)
Let X, Y be two independent discrete random variables and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be a Borel-measurable function such that $\mathbb{E}(|\varphi(X,Y)|) < \infty$.

a) Show that

$$\mathbb{E}(\varphi(X,Y)|Y) = \psi(Y), \quad \text{where} \quad \psi(y) = \mathbb{E}(\varphi(X,y)).$$

b) Reconsider the application of the previous exercise with this formula.

Exercise 12.5. (Borel's paradox)
Let Z be a two-dimensional random variable uniformly distributed on the unit disc $B(0,1)$ in $\mathbb{R}^2$. Z has two possible representations:

(i) $Z = (X,Y)$, where $X \in [-1,1]$ and $Y \in [-1,1]$ are the horizontal and vertical coordinates of Z respectively, with joint pdf

$$f_{X,Y}(x,y) = \frac{1}{\pi}\, 1_{x^2+y^2 \leq 1}.$$

(ii) $Z = (R,\Theta)$, where $R \in [0,1]$ is the radius of Z and $\Theta \in]-\pi,\pi]$ is its angle with respect to the horizontal axis. Their joint pdf is given by

$$f_{R,\Theta}(r,\theta) = \frac{1}{\pi}\, r\, 1_{0 \leq r \leq 1}\, 1_{-\pi < \theta \leq \pi},$$

where the factor r comes from the Jacobian of the change of coordinates.

a) For $t \in [0,1]$, compute $\lim_{\varepsilon \to 0} \mathbb{P}(\{0 < X \leq t\}|\{X \geq 0, -\varepsilon < Y < \varepsilon\})$.

b) For $t \in [0,1]$, compute $\lim_{\varepsilon \to 0} \mathbb{P}(\{0 < R \leq t\}|\{-\varepsilon < \Theta < \varepsilon\})$.

c) What is the paradox here? Can you resolve it?

13 Martingales and stopping times

Exercise 13.1. Let $(M_n,\, n \in \mathbb{N})$ be a submartingale with respect to a filtration $(\mathcal{F}_n,\, n \in \mathbb{N})$ and $\varphi : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable and convex function such that $\mathbb{E}(\varphi(M_n)^2) < +\infty$, $\forall n \in \mathbb{N}$.

a) What additional property of φ ensures that the process $(\varphi(M_n),\, n \in \mathbb{N})$ is also a submartingale?

b) In particular, which of the following two processes is ensured to be a submartingale: $(M_n^2,\, n \in \mathbb{N})$ and/or $(\exp(M_n),\, n \in \mathbb{N})$?

Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such $\mathbb{P}(\{X_1 = +1\}) = \mathbb{P}(\{X_1 = -1\}) = \frac{1}{2}$; let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$ for $n \geq 1$; finally, let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ for $n \geq 1$.

c) For which value of $c > 0$ is the process $(S_n^2 - cn,\, n \in \mathbb{N})$ is a square-integrable martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

d) For which value of $c > 0$ is the process $\left(\frac{\exp(S_n)}{c^n},\, n \in \mathbb{N}\right)$ a square-integrable martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

Assume now that $\mathbb{P}(\{X_1 = +1\}) = p = 1 - \mathbb{P}(\{X_1 = -1\})$ for some $0 < p < 1$ with $p \neq \frac{1}{2}$.

e) Does there exist a number $c > 0$ such that the process $(S_n^2 - cn,\, n \in \mathbb{N})$ is a square-integrable martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? If yes, compute the value of c; otherwise, justify why it is not the case.

f) Does there exist a number $c > 0$ such that the process $\left(\frac{\exp(S_n)}{c^n},\, n \in \mathbb{N}\right)$ is a square-integrable martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? If yes, compute the value of c; otherwise, justify why it is not the case.

Exercise 13.2. Let $0 < p < 1$ and $x > 0$ be fixed real numbers and $(X_n,\, n \in \mathbb{N})$ be the process defined recursively as

$$X_0 = x, \quad X_{n+1} = \begin{cases} X_n^2 + 1 & \text{with probability } p \\ X_n/2 & \text{with probability } 1-p \end{cases} \quad \text{for } n \in \mathbb{N}$$

a) What *minimal* condition on $0 < p < 1$ guarantees that the process X is a submartingale (with respect to its natural filtration)? Justify your answer.

Hint: The inequality $a^2 + b^2 \geq 2ab$ may be useful here.

b) For the values of p respecting the condition found in part a), derive a lower bound on $\mathbb{E}(X_n)$.

Hint: Proceed recursively.

c) Does there exist a value of $0 < p < 1$ such that the process X is a martingale? a supermartingale? Again, justify your answer.

Exercise 13.3. Part I. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. centered and bounded random variables; let $(\mathcal{F}_n,\, n \geq 1)$ be the filtration defined as $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$. Among the following processes $(Y_n,\, n \geq 1)$, which are martingales with respect to $(\mathcal{F}_n,\, n \geq 1)$? (just a short justification suffices here)

a) $Y_n = X_n$, $n \geq 1$.

b) $Y_1 = X_1$, $Y_{n+1} = a\, Y_n + X_{n+1}$, $n \geq 1$ ($a > 0$ fixed).

c) $Y_1 = X_1$, $Y_{n+1} = X_n + X_{n+1}$, $n \geq 1$.

d) $Y_n = \max(X_1, \ldots, X_n)$, $n \geq 1$.

e) $Y_1 = X_1$, $Y_n = \sum_{i=1}^n (X_1 + \ldots + X_{i-1})\, X_i$, $n \geq 1$.

Part II. Let now $(S_n,\, n \in \mathbb{N})$ be the symmetric random walk and $(\mathcal{F}_n,\, n \in \mathbb{N})$ be its natural filtration. Among the following random times, which are stopping times with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? which are bounded? (no justification required here)

a) $T = \sup\{n \geq 0 \,:\, S_n \geq a\}$ ($a > 0$ is fixed)

b) $T = \inf\{n \geq 1 \,:\, S_n = \max_{0 \leq k \leq n} S_k\}$

c) $T = \inf\{n \geq 0 \,:\, S_n = \max_{0 \leq m \leq N} S_m\}$ ($N \geq 1$ is fixed)

d) $T = \inf\{n \geq 0 \,:\, S_n \geq a \text{ or } n \geq N\}$ ($a > 0$ and $N \geq 1$ are fixed)

e) $T = \inf\{n \geq 0 \,:\, |S_n| \geq a\}$ ($a > 0$ is fixed)

Exercise 13.4. a) Let $(M_n,\, n \in \mathbb{N})$ be an *non-decreasing* martingale, that is, $M_{n+1} \geq M_n$ a.s. for all $n \in \mathbb{N}$. Show that $M_n = M_0$ a.s., for all $n \in \mathbb{N}$.

b) Let $(M_n,\, n \in \mathbb{N})$ be a square-integrable martingale such that $(M_n^2,\, n \in \mathbb{N})$ is also a martingale. Show that $M_n = M_0$ a.s., for all $n \in \mathbb{N}$.

Exercise 13.5. Let $(X_n,\, n \geq 1)$ be a family of independent square-integrable random variables such that $\mathbb{E}(X_n) = 0$ for all $n \geq 1$. Let $M_0 = 0$, $M_n = X_1 + \ldots + X_n$, $n \geq 1$.

The process $(M_n,\, n \in \mathbb{N})$ is a martingale, but it is also a process with independent increments. Show that $(M_n^2 - \mathbb{E}(M_n^2),\, n \in \mathbb{N})$ is also a martingale (hence the process A in the Doob decomposition of the submartingale $(M_n^2,\, n \in \mathbb{N})$ is a deterministic process in this case).

Exercise 13.6. ("The" martingale)
A player bets on a sequence of i.i.d. (and balanced) coin tosses: at each turn, the player wins twice his bet if the coin falls on "heads" or loses his bet if the coin falls on "tails".

Assume now that the player adopts the following strategy: he starts by betting 1 franc. If he wins his bet (that is, if the outcome is "heads"), he quits the game and does not bet anymore. If he loses (that is, if the outcome is "tails"), he plays again and doubles his bet for the next turn. He then goes on with the same strategy for the rest of the game.

We assume here that the player can borrow any money he wants in order to bet. Of course, we also assume that he has no information on the outcome of the next coin toss while betting on it.

a) Is the process of gains of the player a martingale (by convention, we set the gain of the player at time zero to be equal to zero)?

b) What is the gain of the player at the first time "heads" comes out?

c) Isn't there a contradiction between a) and b)?

Exercise 13.7. Let $(S_n,\, n \in \mathbb{N})$ be the simple symmetric random walk on $\mathbb{Z}$ and $(\mathcal{F}_n,\, n \in \mathbb{N})$ be its natural filtration.

a) Is the process $(S_n^4,\, n \in \mathbb{N})$ a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? Justify your answer.

b) Is the process $(S_n^4 - n,\, n \in \mathbb{N})$ a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? Justify your answer.

Hint: Recall that $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$.

c) Show that $\mathbb{E}(S_{n+1}^4) = \mathbb{E}(S_n^4) + 6n + 1$ and deduce the value of $\mathbb{E}(S_n^4)$ by induction on n.

Hint: Recall that $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$.

d) Compute $\displaystyle\lim_{n\to\infty} \frac{\mathbb{E}(S_n^4)}{n^2}$. Can you make a parallel with something you already know?

Exercise 13.8. (if one cannot win on a game, then it is a martingale)
Let $(\mathcal{F}_n,\, n \in \mathbb{N})$ be a filtration and $(M_n,\, n \in \mathbb{N})$ be a process adapted to $(\mathcal{F}_n,\, n \in \mathbb{N})$ such that $\mathbb{E}(|M_n|) < \infty$, for all $n \in \mathbb{N}$.

Show that if for any predictable process $(H_n,\, n \in \mathbb{N})$ such that H_n is a bounded random variable $\forall n \in \mathbb{N}$, we have

$$\mathbb{E}((H \cdot M)_N) = 0, \quad \forall N \in \mathbb{N},$$

then $(M_n,\, n \in \mathbb{N})$ is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

Exercise 13.9. Let $(S_n,\, n \in \mathbb{N})$ be the simple symmetric random walk, $(\mathcal{F}_n,\, n \in \mathbb{N})$ be its natural filtration and

$$T = \inf\{n \geq 1 \,:\, |S_n| \geq a\},$$

where $a \geq 1$ is an integer number.

a) Show that T is a stopping time with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

Let now $(M_n,\, n \in \mathbb{N})$ be defined as $M_n = S_n^2 - n$, for all $n \in \mathbb{N}$.

b) Show that the process $(M_n,\, n \in \mathbb{N})$ is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

c) Apply the optional stopping theorem to compute $\mathbb{E}(T)$.

Remark: Even though T is an unbounded stopping time, the optional stopping theorem applies here. Notice that the theorem would *not* apply if one would consider the following stopping time:

$$T' = \inf\{n \geq 1 \,:\, S_n \geq a\}.$$

14 Martingale convergence theorems

Exercise 14.1. Part I. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such that $\mathbb{P}(\{X_n = +1\}) = p$ and $\mathbb{P}(\{X_n = -1\}) = 1 - p$ for some fixed $0 < p < 1/2$.

Let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$, $n \geq 1$. Let also $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$.

Let now $(Y_n,\, n \in \mathbb{N})$ be the process defined as $Y_n = \lambda^{S_n}$ for some $\lambda > 0$ and $n \in \mathbb{N}$.

a) *Using Jensen's inequality only*, for what values of λ can you conclude that the process Y is a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

b) Identify now the values of $\lambda > 0$ for which it holds that the process $(Y_n = \lambda^{S_n},\, n \in \mathbb{N})$ is a martingale / submartingale / supermartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

c) Compute $\mathbb{E}(|Y_n|)$ and $\mathbb{E}(Y_n^2)$ for every $n \in \mathbb{N}$ (and every $\lambda > 0$).

d) For what values of $\lambda > 0$ does it hold that $\sup_{n\in\mathbb{N}} \mathbb{E}(|Y_n|) < +\infty$? $\quad \sup_{n\in\mathbb{N}} \mathbb{E}(Y_n^2) < +\infty$?

e) Run the process Y numerically. For what values of $\lambda > 0$ do you observe that there exists a random variable Y_∞ such that $Y_n \underset{n\to\infty}{\to} Y_\infty$ a.s.? Justify why this is the case and compute the random variable Y_∞ when it exists (this computation might depend on λ, of course).

f) For what values of $\lambda > 0$ does it hold that $Y_n \underset{n\to\infty}{\overset{L^1}{\to}} Y_\infty$?

g) Finally, for what values of $\lambda > 0$ does it hold that $\mathbb{E}(Y_\infty|\mathcal{F}_n) = Y_n$, $\forall n \in \mathbb{N}$?

Part II. Consider now the (interesting) value λ for which the process Y is a martingale.
(Spoiler: there is a unique such value of λ, and it is greater than 1.)

Let $a \geq 1$ be an integer and consider the stopping time $T_a = \inf\{n \in \mathbb{N} : Y_n \geq \lambda^a \quad \text{or} \quad Y_n \leq \lambda^{-a}\}$.

a) Estimate numerically $\mathbb{P}(\{Y_{T_a} = \lambda^a\})$ for some values of a. Explain your method.

b) Is it true that $\mathbb{E}(Y_{T_a}) = \mathbb{E}(Y_0)$? Justify your answer.

c) If possible, use the previous statement to compute $P = \mathbb{P}(\{Y_{T_a} = \lambda^a\})$ theoretically. How fast does this probability decay with a?

Consider finally the other stopping time $T'_a = \inf\{n \in \mathbb{N} : Y_n \geq \lambda^a\}$.

d) Estimate numerically $\mathbb{P}(\{Y_{T'_a} = \lambda^a\})$ for some values of a. Explain your method.

e) Is it true that $\mathbb{E}(Y_{T'_a}) = \mathbb{E}(Y_0)$? Justify your answer.

f) If possible, use the above statement to compute $P' = \mathbb{P}(\{Y_{T'_a} = \lambda^a\})$ theoretically. Is this probability P' greater or smaller than P?

Exercise 14.2. Let $0 < p < 1$ and $M = (M_n,\, n \in \mathbb{N})$ be the process defined recursively as

$$M_0 = x \in]0,1[, \quad M_{n+1} = \begin{cases} p\, M_n, & \text{with probability } 1 - M_n \\ (1-p) + pM_n, & \text{with probability } M_n \end{cases}$$

and $(\mathcal{F}_n,\, n \in \mathbb{N})$ be the filtration defined as $\mathcal{F}_n = \sigma(M_0, \ldots, M_n)$, $n \in \mathbb{N}$.

a) For what value(s) of $0 < p < 1$ is the process M is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? Justify your answer.

b) In the case(s) M is a martingale, compute $\mathbb{E}(M_{n+1}\,(1 - M_{n+1})\,|\,\mathcal{F}_n)$ for $n \in \mathbb{N}$.

c) Deduce the value of $\mathbb{E}(M_n\,(1 - M_n))$ for $n \in \mathbb{N}$.

d) Does there exist a random variable M_∞ such that

(i) $M_n \underset{n\to\infty}{\to} M_\infty$ a.s. ? (ii) $M_n \underset{n\to\infty}{\overset{L^2}{\to}} M_\infty$? (iii) $\mathbb{E}(M_\infty|\mathcal{F}_n) = M_n$, $\forall n \in \mathbb{N}$?

e) What can you say more about M_∞? (No formal justification required here; an intuitive argument will do.)

Exercise 14.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Let $U \sim \mathcal{U}([-1,+1])$ be a random variable independent of $\mathcal{G}$ and M be a positive, integrable and $\mathcal{G}$-measurable random variable.

a) Compute the function $\psi : \mathbb{R}_+ \to \mathbb{R}$ satisfying

$$\psi(M) = \mathbb{E}(|M+U| \mid \mathcal{G})$$

Let now $(U_n,\, n \geq 1)$ be a sequence of i.i.d.$\sim \mathcal{U}([-1,+1])$ random variables, all defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(U_1, \ldots, U_n)$, $n \geq 1$. Let finally $(M_n,\, n \geq 1)$ be the process defined recursively as

$$M_0 = 0, \quad M_{n+1} = |M_n + U_{n+1}|, \quad n \in \mathbb{N}$$

b) Show that the process $(M_n,\, n \in \mathbb{N})$ is a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

c) Is it true that the process $(M_n^2,\, n \in \mathbb{N})$ is also a submartingale with respect to $(\mathcal{F}_n,,\, n \in \mathbb{N})$? Justify your answer.

d) Determine the value of $c > 0$ such that the process $(N_n = M_n^2 - cn, n \in \mathbb{N})$ is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

e) Does there exist a random variable M_∞ such that $M_n \underset{n\to\infty}{\to} M_\infty$ almost surely? (Again, no formal justification required here; an intuitive argument will do.)

Exercise 14.4. Let $(a_n,\, n \geq 1)$ be a decreasing sequence of positive numbers and $(X_n,\, n \in \mathbb{N})$ be a sequence of independent random variables such that

$$\mathbb{P}(\{X_n = +a_n\}) = \mathbb{P}(\{X_n = -a_n\}) = \frac{1}{2} \quad \forall n \geq 1$$

Let also $(M_n,\, n \in \mathbb{N})$ be the process defined as $M_0 = 0$ and $M_n = \sum_{j=1}^n X_j$ for $n \geq 1$.

a) Find a tight upper bound on

$$\mathbb{P}(\{|M_n| \geq nt\})$$

Hint: For this, you may use the following (which is a small adaptation of what we have seen in the course): if X is a random variable taking values $+a$ and $-a$ with probability $1/2$, then

$$\mathbb{E}(\exp(sX)) = \frac{e^{sa} + e^{-sa}}{2} = \cosh(sa) \leq \exp(s^2a^2/2) \quad \forall s \in \mathbb{R}$$

b) Application: assume now that $a_n = \frac{1}{n}$ for $n \geq 1$. Find the least value of $\alpha > 0$ for which the upper bound found in a) allows to conclude that

$$\frac{M_n}{n^\alpha} \underset{n\to\infty}{\to} 0 \quad \text{almost surely}$$

c) Back to the general case: Considering the process M as a martingale (and letting $(\mathcal{F}_n,\, n \in \mathbb{N})$ be its natural filtration), what condition on the decreasing sequence $(a_n,\, n \geq 1)$ ensures that there exists a random variable M_∞ such that

$$M_n \underset{n\to\infty}{\to} M_\infty \quad \text{almost surely} \quad \text{and} \quad \mathbb{E}(M_\infty \,|\, \mathcal{F}_n) = M_n \quad \forall n \in \mathbb{N}?$$

d) Is there a simple condition on the decreasing sequence $(a_n,\, n \geq 1)$ ensuring only that

$$M_n \underset{n\to\infty}{\to} M_\infty \quad \text{almost surely?}$$

[Please pay attention: this question is a (big) trap!]

e) Which of the above two conditions is satisfied by the sequence $a_n = 1/n$?

Exercise 14.5. Let $(S_n,\, n \in \mathbb{N})$ be the simple symmetric random walk and $(\mathcal{F}_n,\, n \in \mathbb{N})$ be its natural filtration. Let also $\alpha > 0$ and $(M_n,\, n \in \mathbb{N})$ be the process defined as

$$M_n = \exp(S_n - \alpha n), \quad \text{for } n \in \mathbb{N}$$

a) Determine the value of $\alpha > 0$ such that M is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

Let now a be a strictly positive integer and let

$$T = \inf\{n \geq 1 \,:\, |S_n| \geq a\}$$

As seen in class, T is a stopping time with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

b) Apply the one of the versions of the optional stopping theorem (*fully justifying its use*) to compute $\mathbb{E}(\exp(-\alpha T))$.

Hint: For this part, you can take for granted the fact that S_T and T are independent.

c) Let now T' be the stopping time defined as

$$T' = \inf\{n \geq 1 \,:\, S_n \geq a\}$$

Can you apply the same procedure as above in order to compute the value of $\mathbb{E}(\exp(-\alpha T'))$?

Exercise 14.6. Let $(U_n,\, n \geq 1)$ be a sequence of i.i.d. random variables, all uniform on $[0,1]$, and let $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(U_1, \ldots, U_n)$, $n \geq 1$. Let us also define the three processes

$$X_0 = 1/2, \quad X_{n+1} = \begin{cases} \frac{1+X_n}{2}, & \text{if } U_{n+1} > X_n \\ \frac{X_n}{2}, & \text{if } U_{n+1} \leq X_n \end{cases}$$

$$Y_0 = 1/2, \quad Y_{n+1} = \begin{cases} \frac{1+Y_n}{2}, & \text{if } U_{n+1} \leq Y_n \\ \frac{Y_n}{2}, & \text{if } U_{n+1} > Y_n \end{cases}$$

and

$$Z_0 = 1/2, \quad Z_{n+1} = \begin{cases} \frac{1+Z_n}{2}, & \text{if } U_{n+1} \leq 1/2 \\ \frac{Z_n}{2}, & \text{if } U_{n+1} > 1/2 \end{cases}$$

a) Are these three processes confined to some interval?

b) Compute $\mathbb{E}(X_{n+1}|\mathcal{F}_n)$, $\mathbb{E}(Y_{n+1}|\mathcal{F}_n)$ and $\mathbb{E}(Z_{n+1}|\mathcal{F}_n)$.

c) Which of the three processes is a martingale with respect to $(\mathcal{F}_n,\, n \geq 1)$?

d) Is this martingale converging a.s. as n goes to infinity? To what limiting random variable? Run a simulation in order to answer this question, then justify theoretically your answer.

e) Run a simulation to see what are the *other two* processes doing! Plot in particular (what you will see is quite interesting...):

e1) a trajectory of each process over 100 time slots;

e2) the histogram of all possible values taken by each process over 10'000 time slots.

Exercise 14.7. Let $Y = (Y_n,\, n \in \mathbb{N})$ be the process defined recursively as

$$Y_0 = 1, \quad Y_{n+1} = \begin{cases} \dfrac{3Y_n}{2}, & \text{with probability } 1/2 \\ \dfrac{Y_n}{2}, & \text{with probability } 1/2 \end{cases}$$

a) Is the process Y a submartingale, supermartingale or martingale with respect to its natural filtration $(\mathcal{F}_n,\, n \in \mathbb{N})$? Justify your answer.

b) Compute $\mathbb{E}(Y_n)$ and $\mathrm{Var}(Y_n)$ recursively, for all $n \geq 1$.

c) Is the process Y confined to some interval?

d) Does there exist a random variable Y_∞ such that $Y_n \underset{n\to\infty}{\to} Y_\infty$ almost surely?

e) If it exists, what is the random variable Y_∞?

Hint: In order to answer this question rigorously, consider the process Z defined as $Z_n = \log(Y_n)$.

f) If Y_∞ exists, does it also hold that $Y_n = \mathbb{E}(Y_\infty|\mathcal{F}_n)$?

Exercise 14.8. Let $(M_n,\, n \in \mathbb{N})$ be a square-integrable martingale with respect to some filtration $(\mathcal{F}_n,\, n \in \mathbb{N})$ and $(H_n,\, n \in \mathbb{N})$ be a predictable process with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$, such that $|H_n(\omega)| \leq K_n$ for every $\omega \in \Omega$ and $n \in \mathbb{N}$.

Let also $(G_n,\, n \in \mathbb{N})$ be the process defined as $G_0 = 0$, $G_n = \sum_{j=1}^n H_j\,(M_j - M_{j-1})$, $n \geq 1$. By the proposition seen in class, we know that G is a martingale.

a) Show that $\mathbb{E}(G_n^2) = \sum_{j=1}^n \mathbb{E}\left(H_j^2\,(A_j - A_{j-1})\right)$, for every $n \geq 1$, where $(A_n,\, n \in \mathbb{N})$ is the (unique) predictable and increasing process such that $(M_n^2 - A_n,\, n \in \mathbb{N})$ is a martingale.

b) Consider $M = S$, the simple symmetric random walk. Find a sufficient condition on the process H (other than $H \equiv 0$:) such that there exists a random variable G_∞ with $\mathbb{E}(G_\infty|\mathcal{F}_n) = G_n$, for every $n \in \mathbb{N}$.

c) Numerical application: still with $M = S$ (i.e., $M_n = S_n = \sum_{j=1}^n X_j$ with X_j i.i.d. ± 1 with equal probability), observe numerically how does the process G behave when $n \to +\infty$ with the following H's (which are all equal to 0 at time 0, by convention)

$$H_n^{(1)} = \frac{1}{n} \qquad H_n^{(2)} = \frac{X_{n-1}}{n} \qquad H_n^{(3)} = \frac{X_{n-1}}{\sqrt{n}} \qquad H_n^{(4)} = \frac{X_n}{\sqrt{n}} \qquad H_n^{(5)} = \frac{\sum_{j=1}^{n-1} X_j}{n}$$

NB: One of these H's is problematic!

Exercise 14.9. Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables such that $\mathbb{P}(\{X_1 = 1\}) = 1/2$ and $\mathbb{P}(\{X_1 = -1\}) = 1/2$. Among the following martingales, which are the ones which are guaranteed to converge a.s. to some limiting random variable as n goes to infinity? For each one, explain why it is or it is not the case. When the martingale converges a.s., try describing the limiting random variable.

a) $S_0 = 0$, $S_n = X_1 + \ldots + X_n$, $n \geq 1$.

b) $M_n = S_n^2 - n$, $n \geq 0$.

c) Let $\alpha > 0$ be some fixed parameter; $M_n = \exp(\alpha\, S_n - c(\alpha)\, n)$, $n \geq 0$. In this example, you should first compute what the function $c(\alpha)$ is, in order for M to be a martingale.

d) Let H be a predictable process such that $|H_n(\omega)| \leq K_n$ for all $n \geq 1$ and $\omega \in \Omega$ and $\sum_{n\geq 0} \mathbb{E}(H_n^2) < \infty$. Let $M = (H \cdot S)$ be the martingale transform defined as

$$M_0 = 0, \quad M_n = (H \cdot S)_n = \sum_{i=1}^{n} H_i\, (S_i - S_{i-1}), \quad n \geq 1.$$

Hint: In this example, you should first prove that

$$\mathbb{E}((H \cdot S)_n^2) = \sum_{i=1}^{n} \mathbb{E}(H_i^2), \quad n \geq 1.$$

For the following processes, first verify that they are indeed martingales.

e) Let $a < 0 < b$ and $T' = \inf\{n \geq 0 \,:\, S_n \leq a \quad \text{or} \quad S_n \geq b\}$. Define then

$$M_n = S_{\min(T,n)} = \begin{cases} S_n, & \text{if } n < T, \\ S_T, & \text{if } n \geq T. \end{cases}$$

f) Let now $T' = \inf\{n \geq 0 \,:\, S_n \geq 1\}$ and $M_n = S_{\min(T',n)}$.

g) Let $(\mathcal{F}_n,\, n \geq 0)$ be a filtration and let X be an $\mathcal{F}_N$-measurable random variable, for some $N > 0$. Let

$$M_n = \mathbb{E}(X|\mathcal{F}_n), \quad n \geq 0.$$

How do the trajectories of this process look like?

15 Five exercises sampled from former exams

Exercise 15.1. Let $S = (S_n,\, n \in \mathbb{N})$ be the simple symmetric random walk. Let also $(X_n,\, n \in \mathbb{N})$ be the process defined as

$$X_0 = S_0 = 0, \quad X_n = S_1 + \ldots + S_n,\; n \geq 1.$$

a) Compute $\mathbb{E}(X_n)$ and $\mathrm{Var}(X_n)$. How does $\mathrm{Var}(X_n)$ grow with n?

b) Fix $s > 0$. Compute an upper bound on $\mathbb{E}(\exp(sX_n))$, for $n \geq 1$.

Hint: You may use the inequality $\log(\cosh(x)) \leq x^2/2$.

c) Fix now $t > 0$ and deduce an upper bound on $\mathbb{P}(X_n \geq n^2 t)$ which is tight for n large.

d) What can you deduce on the convergence of the sequence of random variables

$$\left(Y_n = \frac{X_n}{n^2},\, n \geq 1\right)?$$

Exercise 15.2. Let $0 < a < b < +\infty$ and $(X_n,\, n \geq 1)$ be a sequence of i.i.d. random variables all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $a \leq X_n(\omega) \leq b$ for all $n \geq 1$ and $\omega \in \Omega$.

a) Let $\mu = \mathbb{E}(X_1)$. Show that there exists $\nu \in \mathbb{R}$ such that

$$\frac{1}{n}\sum_{j=1}^{n}(X_j - \mu)^3 \underset{n\to\infty}{\to} \nu \quad \text{almost surely}$$

b) Does it always hold that $\nu \geq 0$? If yes, prove it; if no, provide a counter-example.

c) Compute the values of μ and ν in the particular case where $\mathbb{P}(\{X_1 = a\}) = \mathbb{P}(\{X_1 = b\}) = \frac{1}{2}$.

d) Back to the general case now. Let $(Z_n,\, n \geq 1)$ be the sequence of random variables defined as

$$Z_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j - \mu) \quad \text{for } n \geq 1$$

Show that there exists a random variable Z such that $Z_n \underset{n\to\infty}{\overset{d}{\to}} Z$ and compute the distribution of Z in the particular case mentioned in part c).

e) In the general case, compute the value of $\mathbb{E}(Z_n^3)$ for a fixed (arbitrary) value of n.

Exercise 15.3. a) Let X, Y be two i.i.d.$\sim \mathcal{N}(0,1)$ random variables. Compute $\mathbb{E}(\exp(sXY))$ for $s \geq 0$, specifying for which values of s the expectation is well-defined and finite / well-defined but infinite / ill-defined.

b) Let $(X_n,\, n \geq 1)$, $(Y_n,\, n \geq 1)$ be two independent sequences of i.i.d.$\sim \mathcal{N}(0,1)$ random variables, and let $Z_n = \sum_{j=1}^{n} X_j Y_j$ for $n \geq 1$. Let also $t > 0$ be fixed. Show that there exists $c > 0$ (possibly depending on t) such that $\mathbb{P}(\{Z_n > nt\}) \leq \exp(-cn)$ for all $n \geq 1$.

Hints:

- If $X \sim \mathcal{N}(0, \sigma^2)$, then $\mathbb{E}(\exp(sX)) = \exp(\sigma^2 s^2/2)$ for every $s > 0$ and $\mathbb{E}(\exp(sX^2)) = \frac{1}{\sqrt{1-2\sigma^2 s}}$ for every $0 < s < \frac{1}{2\sigma^2}$.

- You do not need to compute the best value for the constant $c > 0$ in part b). Any value of $c > 0$ will do, or even just providing a *convincing* argument that such a $c > 0$ exists is enough.

Exercise 15.4. Let $(X_n,\, n \geq 1)$ be a sequence of random variables, each taking values in the set $\{-1,+1\}$. Let also $S_0 = 0$, $S_n = X_1 + \ldots + X_n$, $n \geq 1$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$. Assume now that for $n \geq 0$,

$$\mathbb{P}(\{X_{n+1} = +1\} \,|\, \mathcal{F}_n) = \frac{1}{2} - \frac{S_n}{2n}, \quad \mathbb{P}(\{X_{n+1} = -1\} \,|\, \mathcal{F}_n) = \frac{1}{2} + \frac{S_n}{2n}.$$

a) Describe how the process $(S_n,\, n \in \mathbb{N})$ behaves up to time $n = 4$.

b) Is the process $(S_n,\, n \in \mathbb{N})$ a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$? If yes, then prove it. If no, then can you find a process closely related to $(S_n,\, n \in \mathbb{N})$ which is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

c) Find a recursion formula for $\mathrm{Var}(S_n)$, and deduce from there what the value of $\mathrm{Var}(S_n)$ is for all values of $n \geq 0$.

d) Show that $\dfrac{S_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$.

Hint: If you have not found the exact value of $\mathrm{Var}(S_n)$ in part c), you might still be able to deduce from the recursion an upper bound on $\mathrm{Var}(S_n)$ that can help you solving part d).

Exercise 15.5. Let $(M_n,\, n \in \mathbb{N})$ be a martingale with respect to a filtration $(\mathcal{F}_n,\, n \in \mathbb{N})$, such that

$$\sup_{n\in\mathbb{N}} \sup_{\omega\in\Omega} |M_n(\omega)| \leq K, \quad \text{for some } 0 < K < +\infty$$

Let $(A_n,\, n \in \mathbb{N})$ be the process defined recursively as

$$A_0 = 0, \quad A_{n+1} = A_n + \log \mathbb{E}(\exp(M_{n+1} - M_n)|\mathcal{F}_n)$$

a) Show that the process A is increasing.

b) Show that the process $(X_n = \exp(M_n - A_n),\, n \in \mathbb{N})$ is a martingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

Let now M be the martingale defined recursively as $M_0 = x \in\,]0,1[$, $M_{n+1} = \begin{cases} M_n^2, & \text{with prob. } \frac{1}{2} \\ 2M_n - M_n^2, & \text{with prob. } \frac{1}{2} \end{cases}$
and A, X be the processes defined above in this particular case.

c) Compute what the process A is in this particular case.

d) Does there exist a random variable X_∞ such that $\mathbb{E}(X_\infty|\mathcal{F}_n) = X_n$ for all $n \in \mathbb{N}$? Explain!