1

Elementary Probabilities and an Introduction to Stochastic Processes

This chapter reviews the basic concepts related to probability and random variables which will be useful for the rest of this text. For a more detailed explanation as well as demonstrations, the readers may refer to [BAR 07, DAC 82, FOA 03, OUV 08, OUV 09] in French and [BIL 12, CHU 01, DUR 10, KAL 02, SHI 00] in English. The readers who are already familiar with these concepts may proceed straight to section 1.3, which introduces the concept of stochastic processes.

This chapter begins with a brief summary of the concepts of a σ-algebra in section 1.1. These concepts will help in understanding the construction of the properties of conditional expectation in Chapter 2. We then study the chief definitions and properties of random variables and their distribution in section 1.2. There is an emphasis on discrete random variables as this entire book essentially studies discrete cases. Section 1.3 defines a stochastic process, which is the main subject studied in this book. Finally, there are exercises in handling these different concepts in section 1.4. The solutions are given in Chapter 8.

Throughout the rest of the text, Ω is a non-empty set and $\mathcal{P}(\Omega)$ denotes the set of the subsets of Ω :

$$\mathcal{P}(\Omega) = \{A; A \subset \Omega\}.$$

The set Ω is called the **universe** or the **fundamental set**. In practice, the set Ω contains all the possible outcomes of a random experiment.

1.1. Measures and σ-algebras

Let us start by reviewing the concept of a σ-algebra.

DEFINITION 1.1.– *A subset $\mathcal{F}$ of $\mathcal{P}(\Omega)$ is a σ-algebra over Ω if*

1) $\Omega \in \mathcal{F}$;

2) $\mathcal{F}$ is stable by complementarity: for any $A \in \mathcal{F}$, we have $A^c \in \mathcal{F}$, where A^c denotes the complement of A in Ω: $A^c = \Omega \backslash A$;

3) $\mathcal{F}$ is stable under a countable union: for any sequence of elements $(A_n)_{n\in\mathbb{N}}$ of $\mathcal{F}$, we have $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

*Elements of a σ-algebra are called **events**.*

EXAMPLE 1.1.– *The set $\mathcal{G} = \{\emptyset, \Omega\}$ is a σ-algebra and is also the smallest σ-algebra over Ω; it is called the **trivial** σ**-algebra**. Indeed, $\mathcal{G}$ is in fact a σ-algebra since $\Omega \in \mathcal{G}$, $\emptyset^c = \Omega \in \mathcal{G}$, $\Omega^c = \emptyset \in \mathcal{G}$, $\Omega \cup \emptyset = \Omega \in \mathcal{G}$ and by creating unions of $\emptyset$ and Ω we always obtain $\emptyset \in \mathcal{G}$ or $\Omega \in \mathcal{G}$. Further, for any other σ-algebra $\mathcal{F}$, we clearly have $\mathcal{G} \subset \mathcal{F}$.* ♢

EXAMPLE 1.2.– *The set $\mathcal{P}(\Omega)$ is the largest σ-algebra over Ω; it is called the **largest** σ**-algebra.** Indeed, by construction, $\mathcal{P}(\Omega)$ contains all the subsets of Ω, and thus it contains in particular Ω and it is stable by complementarity and under countable unions. In addition, any other σ-algebra $\mathcal{F}$ over Ω is clearly included in $\mathcal{P}(\Omega)$.* ♢

DEFINITION 1.2.– *Let Ω be a non-empty set and $\mathcal{F}$ be a σ-algebra over Ω. The couple $(\Omega, \mathcal{F})$ is called a **probabilizable space**.*

Among the elementary properties of σ-algebra, we can cite stability through any intersection (countable or not).

PROPOSITION 1.1.– *Any intersection of σ-algebras over a set Ω is a σ-algebra over Ω.*

PROOF.– Let $(\mathcal{F}_i)_{i\in I}$ be any family of σ-algebra indexed by a non-empty set I. Thus,

– first of all, for any i, $\Omega \in \mathcal{F}_i$, thus $\Omega \in \cap_{i\in I}\mathcal{F}_i$;

– secondly, if $A \in \cap_{i\in I}\mathcal{F}_i$, then for any i, $A \in \mathcal{F}_i$. As these are σ-algebras, we have that for any i, $A^c \in \mathcal{F}_i$, thus $A^c \in \cap_{i\in I}\mathcal{F}_i$;

– finally, if for any $n \in \mathbb{N}$, $A_n \in \cap_{i\in I}\mathcal{F}_i$, then for any i, n, $A_n \in \mathcal{F}_i$. As these are σ-algebras, we have that for any i, $\cup_{n\in\mathbb{N}}A_n \in \mathcal{F}_i$, thus

$$\cup_{n\in\mathbb{N}}A_n \in \cap_{i\in I}\mathcal{F}_i.$$ □

It is generally difficult to make explicit all the events in a σ-algebra. We often describe it using generating events.

DEFINITION 1.3.– *Let $\mathcal{E}$ be a subset of $\mathcal{P}(\Omega)$. The σ-algebra $\sigma(\mathcal{E})$* ***generated*** *by $\mathcal{E}$ is the intersection of all σ-algebras containing $\mathcal{E}$. It is the smallest σ-algebra containing $\mathcal{E}$. $\mathcal{E}$ is called the* ***generating system*** *of the σ-algebra $\sigma(\mathcal{E})$.*

It can be seen that $\sigma(\mathcal{E})$ is indeed a σ-algebra, being an intersection of σ-algebras.

EXAMPLE 1.3.– *If $A \subset \Omega$, then, $\sigma(A) = \{\emptyset, \Omega, A, A^c\}$ is the smallest σ-algebra Ω containing A.* ♢

EXAMPLE 1.4.– *If Ω is a topological space, the σ-algebra generated by the open sets of Ω is called the Borel* ***σ-algebra*** *of Ω. A* ***Borel set*** *is a set belonging to the Borel σ-algebra. On $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$ generally denotes the σ-algebra of Borel sets. It must be recalled that this is also the σ-algebra generated by the intervals, or by the intervals of the form $]-\infty, x]$, $x \in \mathbb{R}$. Thus, there is no unicity of the generating system.* ♢

We will now recall the concept of the product σ-algebra.

DEFINITION 1.4.– *Let $(E_i, \mathcal{F}_i)_{i \in \mathbb{N}}$ be a sequence of measurable spaces.*

– *Let $n \in \mathbb{N}$. The σ-algebra defined over $\prod_{i=0}^{n} E_i$ and generated by*

$$\{A_0 \times ... \times A_n; A_0 \in \mathcal{F}_0, ..., A_n \in \mathcal{F}_n\}$$

is denoted by $\mathcal{F}_0 \otimes ... \otimes \mathcal{F}_n$, and it is called the product σ-algebra over $\prod_{i=0}^{n} E_i$. We have, in particular,

$$\mathcal{F}_0 \otimes ... \otimes \mathcal{F}_n = \sigma\left(A_0 \times ... \times A_n; A_0 \in \mathcal{F}_0, ..., A_n \in \mathcal{F}_n\right).$$

In the specific case where $E_0 = ... = E_n = E$ and $\mathcal{F}_0 = ... = \mathcal{F}_n = \mathcal{F}$, we also write

$$\prod_{i=0}^{n} E_i = E^{n+1} \text{ and } \mathcal{F}_0 \otimes ... \otimes \mathcal{F}_n = \mathcal{F}^{\otimes n+1}.$$

– *We use $\otimes_{i \in \mathbb{N}} \mathcal{F}_i$ to denote the σ-algebra over the countable product space $\prod_{i \in \mathbb{N}} E_i$, generated by the sets of the form $\prod_{i \in \mathbb{N}} A_i$, where $A_i \in \mathcal{F}_i$ and $A_i = E_i$ except for a finite number of indices i. In the specific case where, for any $i \in \mathbb{N}$, $E_i = E$ and $\mathcal{F}_i = \mathcal{F}$, the product space $\prod_{i \in \mathbb{N}} E_i$ is denoted by $E^{\mathbb{N}}$, and the σ-algebra $\otimes_{i \in \mathbb{N}} \mathcal{F}_i$ is denoted by $\mathcal{F}^{\otimes \mathbb{N}}$.*

Finally, let us review the concepts of measurability and measure.

DEFINITION 1.5.– *Let* Ω *be non-empty set and* $\mathcal{F}$ *be a* σ*-algebra on* Ω*.*

*– A **measure** over a probabilizable space* $(\Omega, \mathcal{F})$ *is defined as any mapping* μ *defined over* $\mathcal{F}$*, with values in* $[0, +\infty] = \mathbb{R}_+ \cup \{+\infty\}$*, such that* $\mu(\emptyset) = 0$ *and for any family* $(A_i)_{i\in\mathbb{N}}$ *of pairwise disjoint elements of* $\mathcal{F}$*, we have the property of* σ*-additivity:*

$$\mu\left(\bigcup_{i=0}^{+\infty} A_i\right) = \sum_{i=0}^{+\infty} \mu(A_i).$$

– A measure μ *over a probabilizable space* $(\Omega, \mathcal{F})$ *is said to be **finite**, or have **finite total mass**, if* $\mu(\Omega) < \infty$*.*

– If μ *is a **measure** over a probabilizable space* $(\Omega, \mathcal{F})$*, then the triplet* $(\Omega, \mathcal{F}, \mu)$ *is called a **measured space**.*

DEFINITION 1.6.– *Let* $(\Omega, \mathcal{F})$ *and* $(E, \mathcal{E})$ *be two probabilizable spaces. A mapping* X*, defined over* Ω *taking values in* E*, is said to be* $(\mathcal{F}, \mathcal{E})$*-**measurable**, or just measurable, if there is no ambiguity regarding the reference* σ*-algebras, if*

$$\forall B \in \mathcal{E}, \quad X^{-1}(B) \in \mathcal{F}.$$

In practice, when $E \subset \mathbb{R}$, we set $\mathcal{E} = \mathcal{B}(E)$ the set of Borel subsets of E, that is, the set of subsets of E. We can simply say that X is $\mathcal{F}$-measurable. When, in addition, we manipulate a single σ-algebra $\mathcal{F}$ over Ω, it can be simply said that X is measurable. If we work with several σ-algebras over Ω, the concerned σ-algebra must always be specified: X is $\mathcal{F}$-measurable.

EXAMPLE 1.5.– *If* $(\Omega, \mathcal{F})$ *is a measurable space and* $A \in \mathcal{F}$*, then the indicator function*

$$\mathbb{1}_A : x \longmapsto \begin{cases} 1 \; if \; x \in A \\ 0 \; if \; x \notin A \end{cases}$$

is $\mathcal{F}$*-measurable. Indeed, for any Borel set* B *in* $\mathbb{R}$*, we have*

$$(\mathbb{1}_A)^{-1}(B) = \begin{cases} \Omega \; if \; \{0,1\} \subset B \\ A \; if \; 1 \in B \; and \; 0 \notin B \\ A^c \; if \; 0 \in B \; and \; 1 \notin B \\ \emptyset \; if \; 0 \notin B \; and \; 1 \notin B \end{cases}$$

Thus, in all cases, we do have $(\mathbb{1}_A \in B) \in \mathcal{F}$*.* ♢

EXAMPLE 1.6.– *The composition of two measurable functions is measurable. Indeed, if* $(\Omega, \mathcal{F})$*,* $(E, \mathcal{E})$ *and* $(G, \mathcal{G})$ *are three probabilizable spaces,* $f : \Omega \longmapsto E$ *and*

$g : E \longmapsto G$ are two $(\mathcal{F}, \mathcal{E})$ and $(\mathcal{E}, \mathcal{G})$-measurable mappings, respectively, then for any $B \in \mathcal{G}$, $g^{-1}(B) \in \mathcal{E}$ and consequently,

$$(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B)) \in \mathcal{F}.$$

Thus, the composition $g \circ f$ is indeed measurable on $(\Omega, \mathcal{F})$ in $(G, \mathcal{G})$. ♢

1.2. Probability elements

We will now review the concept of a probability measure or probability distribution, and the concept of random variable, as well as the chief properties of these concepts.

1.2.1. *Probabilities*

A probability measure or probability distribution is a finite measure whose total mass is equal to 1.

DEFINITION 1.7.– *A* ***probability*** *or* ***probability measure****, or* ***law of probability*** *or* ***distribution*** *over a probability space $(\Omega, \mathcal{F})$ is a measure with a total mass equal to* 1. *In other words, a probability over $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow \mathbb{R}$ such that*

– *for any $A \in \mathcal{F}$, $\mathbb{P}(A) \geq 0$,*

– $\mathbb{P}(\Omega) = 1$,

– *for any sequence of pairwise disjoint events in $\mathcal{F}$, denoted by $(A_n)_{n \in \mathbb{N}}$, we have*

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is then called a ***probability space****.*

EXAMPLE 1.7.– *Ω is endowed with the coarse σ-algebra $\mathcal{F} = \{\emptyset, \Omega\}$. Thus, the single probability measure on $(\Omega, \mathcal{F})$ is given by:*

$$\mathbb{P} : A \in \mathcal{F} \longmapsto \mathbb{P}(A) = \begin{cases} 1 \; if \; A = \Omega \\ 0 \; if \; A = \emptyset. \end{cases}$$ ♢

EXAMPLE 1.8.– *Let $\Omega = [0, 1]$ and $\mathcal{F} = \mathcal{B}([0, 1])$ be the Borel σ-algebra of $[0, 1]$. If λ denotes the Lebesgue measure, then the mapping:*

$$\mathbb{P} : A \in \mathcal{F} \longmapsto \lambda(A)$$

is a probability measure on $(\Omega, \mathcal{F})$. ♢

EXAMPLE 1.9.– *Let Ω be non-empty set such that $card(\Omega) < \infty$, where $card(\Omega)$ denotes the cardinal of Ω, that is, the number of elements in Ω. Consider the mapping $\mathbb{P}$ from $\mathcal{P}(\Omega)$ onto $[0,1]$ such that for every $A \in \mathcal{P}(\Omega)$, $\mathbb{P}(A) = \dfrac{card(A)}{card(\Omega)}$.*

The mapping $\mathbb{P}$ is then a probability on $(\Omega, \mathcal{P}(\Omega))$, said to be the uniform probability on Ω. ♢

We will only review those properties of a probability that will be useful for this book.

PROPOSITION 1.2.– *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(A_n)_{n\in\mathbb{N}}$ be a sequence of events in $\mathcal{F}$.*

– If $(A_n)_{n\in\mathbb{N}}$ is increasing (for the inclusion), then,

$$\mathbb{P}\left(\bigcup_{i=0}^{\infty} A_i\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

– If $(A_n)_{n\in\mathbb{N}}$ is decreasing (for the inclusion), then,

$$\mathbb{P}\left(\bigcap_{i=0}^{\infty} A_i\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

We will now review the concept of independent events and σ-algebras.

DEFINITION 1.8.– *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.*

*– Two events, A and B, are **independent** if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$.*

*– A family of events $(A_i \in \mathcal{F}_i, i \in I)$ is said to be **mutually independent** if for any finite family $J \subset I$, we have*

$$\mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}(A_j).$$

– Two σ-algebras $\mathcal{F}$ and $\mathcal{G}$ are independent if for any $A \in \mathcal{F}$ and $B \in \mathcal{G}$, A and B are independent.

*– A family of sub-σ-algebra $\mathcal{F}_i \subset \mathcal{F}, i \in I$ is **mutually independent** if any family of events $(A_i \in \mathcal{F}_i, i \in I)$ is mutually independent.*

EXAMPLE 1.10.– *We roll a six-faced die and write*

– A_1 *the event "the number obtained is even"; and*

– A_2 *the event "the number obtained is a multiple of* 3*" .*

The universe of possible outcomes is $\Omega = \{1, 2, 3, 4, 5, 6\}$ *which has a finite number of elements and as all its elements have the same chance of occurring, we can endow it with the uniform probability* $\mathbb{P} : A \in \mathcal{P}(\Omega) \longmapsto \dfrac{card(A)}{card(\Omega)}$. *Since*

$$A_1 = \{2, 4, 6\},\ A_2 = \{3, 6\} \text{ and } A_1 \cap A_2 = \{6\},$$

we have

$$\mathbb{P}(A_1) = \frac{1}{2},\ \mathbb{P}(A_2) = \frac{1}{3} \text{ and } \mathbb{P}(A_1 \cap A_2) = \frac{1}{6} = \mathbb{P}(A_1) \times \mathbb{P}(A_2).$$

Therefore, A_1 *and* A_2 *are two independent events.* ♢

EXAMPLE 1.11.– *A coin is tossed twice. The following events are considered:*

– A_1 *"Obtaining tails (T) on the first toss";*

– A_2 *"Obtaining heads (H) on the second toss"; and*

– A_3 *"Obtaining the same face on both tosses".*

The universe of possible outcomes is

$$\Omega = \{(T, H), (T, T), (H, T), (H, H)\}$$

which has four elements, and as all elements have the same chance of occurring, it can be endowed with uniform probability. Since

$$A_1 = \{(T, H), (T, T)\},$$
$$A_2 = \{(T, H), (H, H)\},$$
$$A_3 = \{(T, T), (H, H)\},$$

we have

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \frac{1}{2},$$
$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2 \cap A_3) = \mathbb{P}(A_1 \cap A_3) = \frac{1}{4},$$

and $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 0 \neq \mathbb{P}(A_1) \times \mathbb{P}(A_2) \times \mathbb{P}(A_3)$. *Thus, the events* A_1, A_2 *and* A_3 *are pairwise independent, but are not mutually independent. Unless specified, the notion of independence by default always signifies mutual independence and not pairwise independence.* ♢

1.2.2. *Random variables*

Let us now recall the definition of a generic random variable, and then the specific case of discrete random variables.

DEFINITION 1.9.– *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probabilizable space and $(E, \mathcal{E})$ be a measurable space. A* ***random variable*** *on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in the measurable space $(E, \mathcal{E})$, is any mapping $X : \Omega \longrightarrow E$ such that, for any B in $\mathcal{E}$, $X^{-1}(B) \in \mathcal{F}$; in other words, $X : \Omega \longrightarrow E$ is a random variable if it is an $(\mathcal{F}, \mathcal{E})$-measurable mapping. We then write the event "*X belongs to B*" by*

$$X^{-1}(B) = \{\omega \in \Omega; X(\omega) \in B\} = (X \in B).$$

In the specific case where $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$, the mapping X is called a ***real random variable****. If $E = \mathbb{R}^d$ with $d \geq 2$, and $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$, the mapping X is said to be a* ***real random vector****.*

EXAMPLE 1.12.– *Let us return to the experiment where a six-sided die is rolled, where the set of possible outcomes is $\Omega = \{1, 2, 3, 4, 5, 6\}$, which is endowed with the uniform probability. Consider the following game:*

– *if the result is even, you win 10 €;*

– *if the result is odd, you win 20 €.*

This game can be modeled using the random variable $X : \Omega \longmapsto \{10, 20\}$, defined by:

$$X(\omega) = \begin{cases} 10 \text{ if } \omega \in \{2, 4, 6\} \\ 20 \text{ if } \omega \in \{1, 3, 5\}. \end{cases}$$

This mapping is a random variable, since for any $B \in \mathcal{P}(\{10, 20\})$, we have

$$(X \in B) = X^{-1}(B) = \begin{cases} \{2, 4, 6\} \text{ if } B = \{10\} \\ \{1, 3, 5\} \text{ if } B = \{20\} \\ \Omega \text{ if } B = \{10, 20\} \\ \emptyset \text{ if } B = \emptyset. \end{cases}$$

and all these events are in $\mathcal{P}(\Omega)$. ♢

DEFINITION 1.10.– *The* ***distribution*** *of a random variable X defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E, \mathcal{E})$ is the mapping $\mathbb{P}_X : \mathcal{E} \longrightarrow [0, 1]$ such that, for any$B \in \mathcal{E}$,*

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B).$$

The distribution of X is a probability distribution on $(E, \mathcal{E})$; it is also called the ***image distribution*** *of $\mathbb{P}$ by X.*

DEFINITION 1.11.– *A random real variable is* ***discrete*** *if* $X(\Omega)$ *is at most countable. In other words, if* $X(\Omega) = \{x_i, i \in I\}$*, where* $I \subset \mathbb{N}$*. In this case, the probability distribution of* X *is characterized by the family*

$$(\mathbb{P}(X = x_i))_{i \in I} \, .$$

EXAMPLE 1.13.– ***Uniform distribution****: Let* $N \in \mathbb{N}^{\star}$ *and* $\{x_1, ..., x_N\} \subset \mathbb{R}$*. Let* X *be a random variable on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that* $X(\Omega) = \{x_1, ..., x_N\}$ *and for any* $i \in \{1, ..., N\}$*,*

$$\mathbb{P}(X = x_i) = \frac{1}{N}.$$

It is then said that X *follows a uniform distribution on* $\{x_1, ..., x_N\}$*.* ♢

EXAMPLE 1.14.– ***The Bernoulli distribution****: Let* $p \in [0, 1]$*. Let* X *be a random variable on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that* $X(\Omega) = \{0, 1\}$ *and*

$$\mathbb{P}(X = 1) = p \;\; and \;\; \mathbb{P}(X = 0) = 1 - p.$$

It is then said that X *follows a Bernoulli distribution with parameter* p*, and we write* $X \sim \mathcal{B}(p)$*.*

The Bernoulli distribution models random experiments with two possible outcomes: success, with probability p*, and failure, with probability* $1 - p$*. This is the case in the following game. A coin is tossed* N *times. This experiment is modeled by* $\Omega = \{T, H\}^N$*, endowed with the* σ*-algebra of its subsets and the uniform distribution. For* $1 \leq n \leq N$*, the mappings* X_n *from* Ω *onto* $\mathbb{R}$ *are considered, defined by*

$$X_n(\omega_1, \omega_2, \ldots, \omega_n, \ldots, \omega_N) = \mathbb{1}_{\{T\}}(\omega_n),$$

the number of tails at the n*th toss. Thus,* X_n*,* $1 \leq n \leq N$*, are random real variables in the Bernoulli distribution with parameter* $1/2$ *if the coin is balanced.* ♢

EXAMPLE 1.15.– ***Binomial distribution****: Let* $p \in [0, 1]$*,* $N \in \mathbb{N}^{\star}$ *and* X *be a random variable on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that* $X(\Omega) = \{0, 1, ..., N\}$ *and for any* $k \in \{0, 1, ..., N\}$*,*

$$\mathbb{P}(X = k) = C_N^k p^k (1 - p)^{N-k}.$$

It is then said that X *follows a binomial distribution with parameters* N *and* p*, and we write* $X \sim \mathcal{B}(N, p)$*.*

If the Bernoulli experiment with probability of success p *is repeated* N *times, independently, then the binomial distribution is the distribution of the random* *variable containing the number of successes at the end of the N repetitions of the experiment.* ♢

EXAMPLE 1.16.– ***Hypergeometric distribution***: *Let n and N be two integers such that $n \leq N$, $p \in]0,1[$ such that $pN \in \mathbb{N}$, and let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$X(\Omega) = [\max(0; n - N(1-p)), \min(n, Np)] \cap \mathbb{N}$$

and for any $k \in X(\Omega)$,

$$\mathbb{P}(X = k) = \frac{C_{Np}^{k} C_{N(1-p)}^{n-k}}{C_N^n}.$$

X is then said to follow a hypergeometric distribution with parameters N, n and p, and we write $X \sim \mathcal{H}(N, n, p)$.

If we consider an urn containing N indistinguishable balls, k red balls and $N - k$ white balls, with $k \in \{1, ...N-1\}$, and if we simultaneously draw n balls, then the random variable X, equal to the number of red balls obtained, follows a hypergeometric distribution with parameters N, n and $p = \frac{k}{N}$. ♢

EXAMPLE 1.17.– ***Poisson distribution***: *Let $\lambda > 0$ and X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$X(\Omega) = \mathbb{N}$$

and for any $k \in X(\Omega)$,

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

It is then said that X follows a Poisson distribution with parameter λ, and we write $X \sim \mathcal{P}(\lambda)$. ♢

DEFINITION 1.12.– *Let X be a discrete random variable such that $X(\Omega) = \{x_i, i \in I\}$, where $I \subset \mathbb{N}$.*

*– X or the distribution of X is said to be **integrable** (or **summable**) if*

$$\sum_{i \in I} |x_i| \mathbb{P}(X = x_i) < \infty.$$

*– If X is integrable, then the **expectation** of X is the real number defined by*

$$\mathbb{E}[X] = \sum_{i \in I} x_i \mathbb{P}(X = x_i).$$

EXAMPLE 1.18.– *The random variable X defined in Example 1.12 admits an expectation equal to*

$$\begin{aligned}\mathbb{E}[X] &= 10 \times \mathbb{P}(X = 10) + 20 \times \mathbb{P}(X = 20) \\ &= 10 \times \mathbb{P}(\{2, 4, 6\}) + 20 \times \mathbb{P}(\{1, 3, 5\}) \\ &= 10 \times \frac{1}{2} + 20 \times \frac{1}{2} = 15.\end{aligned}$$

The average winnings in the die-rolling game is therefore equal to 15 *€.* ♢

The following proposition establishes a link between the expectation of a discrete, random variable and measure theory.

PROPOSITION 1.3.– *Let X be a discrete random variable such that $X(\Omega) = \{x_i, i \in I\}$, where $I \subset \mathbb{N}$. It is assumed that $\sum_{i \in I} |x_i| \mathbb{P}(X = x_i) < \infty$. Then,*

$$\mathbb{E}[X] = \int_\Omega X(w) d\mathbb{P}(w).$$

The above proposition also justifies the concept of integrability introduced in Definition 1.12. Further, in this case (i.e. when X is integrable: $\sum_{i \in I} |x_i| \mathbb{P}(X = x_i) < \infty$), we write $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

When X^p is integrable for a certain real number $p \geq 1$ (i.e. when $\sum_{i \in I} |x_i|^p \mathbb{P}(X = x_i) < \infty$), we write

$$X \in L^p(\Omega, \mathcal{F}, \mathbb{P}) \ \ and \ \ \|X\|_p = \left(\mathbb{E}\left[|X|^p\right]\right)^{1/p}.$$

Let us look at some of the properties of expectations.

PROPOSITION 1.4.– *Let X and Y be two integrable, discrete random variables, $a, b \in \mathbb{R}$. Then,*

1) Linearity: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

2) Transfer theorem: *if g is a measurable function such that $g(X)$ is integrable, then*

$$\mathbb{E}[g(X)] = \sum_{i \in I} g(x_i) \mathbb{P}(X = x_i).$$

3) Monotonicity: *if $X \leq Y$ almost surely (a.s.), then* $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

4) Cauchy–Schwartz inequality: *If X^2 and Y^2 are integrable, then XY is integrable and*

$$(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2].$$

5) Jensen inequality: *if g is a convex function such that $g(X)$ is integrable, then,*

$$g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)].$$

DEFINITION 1.13.– *Let X be a discrete random variable, such that $X(\Omega) = \{x_i, i \in I\}$, $I \subset \mathbb{N}$ and X^2 is integrable. The* ***variance*** *of X is the real number:*

$$\mathbb{V}(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \sum_{i \in I}(x_i - \mathbb{E}[X])^2\mathbb{P}(X = x_i).$$

Variance satisfies the following properties.

PROPOSITION 1.5.– *If a discrete random variable X admits variance, then,*

1) $\mathbb{V}(X) \geq 0$.

2) $\mathbb{V}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

3) For any $(a, b) \in \mathbb{R}^2$, $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$.

1.2.3. *σ-algebra generated by a random variable*

We now define the σ-algebra generated by a random variable. This concept is important for several reasons. For instance, it can make it possible to define the independence of random variables. It is also at the heart of the definition of conditional expectations; see Chapter 2.

PROPOSITION 1.6.– *Let X be a real random variable, defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E, \mathcal{E})$. Then, $\mathcal{F}_X = X^{-1}(\mathcal{E}) = \{X^{-1}(A); A \in \mathcal{E}\}$ is a sub-σ-algebra of $\mathcal{F}$ on Ω. This is called the* ***σ-algebra generated by the random variable*** *X. It is written as $\sigma(X)$. It is the smallest σ-algebra on Ω that makes X measurable:*

$$\sigma(X) = X^{-1}\left(\mathcal{B}(\mathbb{R})\right) = \{X^{-1}(B); B \in \mathcal{B}(\mathbb{R})\} = \{(X \in B); B \in \mathcal{B}(\mathbb{R})\}.$$

EXAMPLE 1.19.– *Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $X = c \in \mathbb{R}$ be a constant. Then, for any Borel set B in $\mathbb{R}$, $(X \in B)$ has the value $\emptyset$ if $c \notin B$ and Ω if $c \in B$. Thus, the σ-algebra generated by X is $\mathcal{F}_0$. Reciprocally, it can be demonstrated that the only $\mathcal{F}_0$-measurable random variables are the constants. Indeed, let X be a $\mathcal{F}_0$-measurable random variable. Assume that it takes at least two different values, x and y. It may be assumed that $y \geq x$ without loss of generality. Therefore, let $B = [x, \frac{x+y}{2}]$. We have that $(X \in B)$ is non-empty because $x \in B$ but is not Ω since $y \notin B$. Therefore, X is not $\mathcal{F}_0$-measurable.* $\diamondsuit$

PROPOSITION 1.7.– *Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E, \mathcal{E})$ and let $\sigma(X)$ be the σ-algebra generated by X. Thus, a random variable Y is $\sigma(X)$-measurable if and only if there exists a measurable function f such that $Y = f(X)$.*

This technical result will be useful in certain demonstrations further on in the text. In general, if it is known that Y is $\sigma(X)$-measurable, we cannot (and do not need to) make explicit the function f. Reciprocally, if Y can be written as a measurable function of X, it automatically follows that Y is $\sigma(X)$-measurable.

EXAMPLE 1.20.– *A die is rolled 2 times. This experiment is modeled by $\Omega = \{1, 2, 3, 4, 5, 6\}^2$ endowed with the σ-algebra of its subsets and the uniform distribution. Consider the mappings X_1, X_2 and Y from Ω onto $\mathbb{R}$ defined by*

$$X_1(\omega_1, \omega_2) = \omega_1,$$
$$X_2(\omega_1, \omega_2) = \omega_2,$$
$$Y(\omega_1, \omega_2) = \mathbb{1}_{\{2,4,6\}}(\omega_1),$$

thus, X_i is the result of the ith roll and Y is the parity indicator of the first roll. Therefore, $Y = \mathbb{1}_{\{2,4,6\}}(X_1)$; thus, Y is $\sigma(X_1)$-measurable. On the other hand, Y cannot be written as a function of X_2. ♢

The σ-algebra generated by X represents all the events that can be observed by drawing X. It represents the information revealed by X.

DEFINITION 1.14.– *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.*

– Let X and Y be two random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$. Then, X and Y are said to be independent if the σ-algebras $\sigma(X)$ and $\sigma(Y)$ are independent.

– Any family $(X_i)_{i \in I}$ of random variables is independent if the σ-algebras $\sigma(X_i)$ are independent.

– Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let X be a random variable. Then, X is said to be independent of $\mathcal{G}$ if $\sigma(X)$ is independent of $\mathcal{G}$ or, in other words, $\forall A \in \mathcal{G}$, X and $\mathbb{1}_A$ are independent.

PROPOSITION 1.8.– *If X and Y are two integrable and independent random variables, then their product XY is integrable and $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.*

1.2.4. *Random vectors*

We will now more closely study random variables taking values in $\mathbb{R}^d$, with $d \geq 2$. This concept has already been defined in Definition 1.9. We will now look at the relations between the random vector and its coordinates. When $d = 2$, we then speak of a random couple.

PROPOSITION 1.9.– *Let X be a real random vector on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathbb{R}^d$. Then,*

$$X(w) = \begin{pmatrix} X_1(w) \\ \vdots \\ X_n(w) \end{pmatrix}$$

is such that for any $i \in \{1, ..., d\}$, X_i is a real random variable.

DEFINITION 1.15.– *A random vector is said to be discrete if each of its components, X_i, is a discrete random variable.*

DEFINITION 1.16.– *Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ be a discrete random couple such that*

$$X_1(\Omega) = \{x_{1j}, j \in I_1\} \textit{ et } X_2(\Omega) = \{x_{2k}, k \in I_2\}.$$

The ***conjoint distribution*** *(or* ***joint distribution*** *or, simply, the distribution) of X is given by the family*

$$\{\mathbb{P}(X_1 = x_{1j}, X_2 = x_{2k});\ (j,k) \in I_1 \times I_2\}.$$

The ***marginal distributions*** *of X are the distributions of X_1 and X_2. These distributions may be derived from the conjoint distribution of X through:*

$$\forall j \in I_1,\ \ \mathbb{P}(X_1 = x_{1j}) = \sum_{k \in I_2} \mathbb{P}(X_1 = x_{1j}, X_2 = x_{2k})$$

and

$$\forall k \in I_2,\ \ \mathbb{P}(X_2 = x_{2k}) = \sum_{j \in I_1} \mathbb{P}(X_1 = x_{1j}, X_2 = x_{2k}).$$

The concept of joint distributions and marginal distributions can naturally be extended to vectors with dimension larger than 2.

EXAMPLE 1.21.– *A coin is tossed 3 times, and the result is noted. The universe of possible outcomes is $\Omega = \{T, H\}^3$. Let X denote the total number of tails obtained and Y denote the number of tails obtained at the first toss. Then,*

$$X(\Omega) = \{0, 1, 2, 3\} \textit{ and } Y(\Omega) = \{0, 1\}.$$

The couple (X, Y) is, therefore, a random vector (referred to here as a "random couple"), with joint distribution defined by

$$\mathbb{P}((X,Y) = (i,j)) = \begin{cases} \dfrac{1}{2^3} \; if \; (i,j) \in \{(0,0),(1,1),(2,0),(3,1)\} \\ 2 \times \dfrac{1}{2^3} \; if \; (i,j) \in \{(1,0),(2,1)\} \\ 0 \; if \; (i,j) \in \{(3,0),(0,1)\}, \end{cases}$$

for any $(i, j) \in X(\Omega) \times Y(\Omega)$, which makes it possible to derive the distributions of X and Y (called the marginal distributions of the couple (X, Y)):

Distribution of X:

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 0, Y = 0) + \mathbb{P}(X = 0, Y = 1) = \frac{1}{2^3}$$

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 1, Y = 0) + \mathbb{P}(X = 1, Y = 1) = \frac{3}{2^3}$$

$$\mathbb{P}(X = 2) = \mathbb{P}(X = 2, Y = 0) + \mathbb{P}(X = 2, Y = 1) = \frac{3}{2^3}$$

$$\mathbb{P}(X = 3) = \mathbb{P}(X = 3, Y = 0) + \mathbb{P}(X = 3, Y = 1) = \frac{1}{2^3}.$$

Distribution of Y:

$$\mathbb{P}(Y = 0) = \sum_{i=0}^{3} \mathbb{P}(X = i, Y = 0) = 4 \times \frac{1}{2^3} = \frac{1}{2}$$

$$\mathbb{P}(Y = 1) = \sum_{i=0}^{3} \mathbb{P}(X = i, Y = 1) = 4 \times \frac{1}{2^3} = \frac{1}{2}. \qquad \diamondsuit$$

1.2.5. *Convergence of sequences of random variables*

To conclude this section on random variables, we will review some classic results of convergence for sequences of random variables. Throughout the rest of this book, the abbreviation *r.v.* signifies *random variable*.

DEFINITION 1.17.– *Let $(X_n)_{n\geq 1}$ and X be r.v.s defined on $(\Omega, \mathcal{F}, \mathbb{P})$.*

1) It is assumed that there exists $p > 0$ such that, for any $n \geq 0$, $\mathbb{E}[|X_n|^p] < \infty$, and $\mathbb{E}[|X|^p] < \infty$. It is said that the sequence of random variables $(X_n)_{n\geq 1}$ ***converges on the average of the order p*** *or* ***converges in L^p*** *towards X, if*

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0.$$

We then write $X_n \xrightarrow{L^p} X$. In the specific case where $p = 2$, we say there is a ***convergence in quadratic mean****.*

2) The sequence of r.v. $(X_n)_{n\geq 1}$ is called ***almost surely (a.s.) convergent*** *towards X, if*

$$\lim_{n\to\infty} \mathbb{P}\left(w \in \Omega; \lim_{n\to\infty} X_n(w) = X(w)\right) = 1.$$

We then write $X_n \xrightarrow{a.s.} X$.

THEOREM 1.1 (Monotone convergence theorem).– *Let $(X_n)_{n\geq 1}$ be a sequence of positive and non-decreasing random variables and let X be an integrable random variable, all of these defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If (X_n) converges almost surely to X, then*

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

THEOREM 1.2 (Dominated convergence theorem).– *Let $(X_n)_{n\geq 1}$ be a sequence of random variables and let X be another random variable, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If the sequence (X_n) converges to X a.s., and for any $n \geq 1$, $|X_n| \leq Z$, where Z is an integrable random variable, then $X_n \xrightarrow{L^1} X$ and, in particular,*

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

THEOREM 1.3 (Strong law of large numbers).– *Let $(X_n)_{n\geq 1}$ be a sequence of integrable, independent random variables from the same distribution. Then,*

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}[X_1].$$

1.3. Stochastic processes

The main objective of this book is to study certain families of stochastic (or random) processes in discrete time. There are two ways of seeing such objects:

– as a sequence $(X_n)_{n\in\mathbb{N}}$ of real random variables;

– as a single random variable X taking values in the set of real sequences.

The index n represents time. Since $n \in \mathbb{N}$, we speak of processes in discrete time. In the rest of this book, unless indicated otherwise, we will only consider processes taking discrete real values. The notation E thus denotes a finite or countable subset of $\mathbb{R}$ and $\mathcal{E} = \mathcal{P}(E)$, the set of subsets of E.

DEFINITION 1.18.– *A* ***stochastic process*** *is a sequence $X = (X_n)_{n\in\mathbb{N}}$ of random variables taking values in $(E, \mathcal{E})$. The process X is then a random variable taking values in $(E^{\mathbb{N}}, \mathcal{E}^{\otimes\mathbb{N}})$.*

EXAMPLE 1.22.– *A coin is tossed an infinite number of times. This experiment is modeled by $\Omega = \{T, H\}^{\mathbb{N}^*}$. For $n \in \mathbb{N}^*$, consider the mappings X_n to Ω in $\mathbb{R}$ defined by*

$$X_n(\omega_1, \omega_2, \ldots, \omega_n, \ldots) = \mathbb{1}_{\{T\}}(\omega_n),$$

the number of tails at the nth toss. Therefore, X_n, $n \in \mathbb{N}^*$ are discrete, real random variables and the sequence $X = (X_n)_{n\in\mathbb{N}}$ is a stochastic process. $\diamondsuit$

DEFINITION 1.19.– *Let $X = (X_n)_{n\in\mathbb{N}}$ be a stochastic process. For all $n \in \mathbb{N}$, the distribution of the vector $(X_0, X_1, \ldots, X_n)$ is denoted by μ_n. The probability distributions $(\mu_n)_{n\in\mathbb{N}}$ are called **finite-dimensional distributions** or **finite-dimensional marginal distributions** of the process $X = (X_n)_{n\in\mathbb{N}}$.*

PROPOSITION 1.10.– *Let $X = (X_n)_{n\in\mathbb{N}}$ be a stochastic process and let $(\mu_n)_{n\in\mathbb{N}}$ be its finite-dimensional distributions. Then, for all $n \in \mathbb{N}^*$ and $(A_0, \ldots, A_{n-1}) \in \mathcal{E}^n$, we have*

$$\mu_{n-1}(A_0 \times \ldots \times A_{n-1}) = \mu_n(A_0 \times \ldots \times A_{n-1} \times E).$$

In other words, the restriction of the marginal distribution of the vector $(X_0, \ldots, X_n)$ to its first n coordinates is exactly the distribution of the vector $(X_0, \ldots, X_{n-1})$.

PROOF.– This proof directly follows from the definition of the objects. We have

$$\begin{aligned}\mu_{n-1}(A_0 \times \ldots \times A_{n-1}) &= \mathbb{P}(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}) \\ &= \mathbb{P}(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}, X_n \in E) \\ &= \mu_n(A_0 \times \ldots \times A_{n-1} \times E),\end{aligned}$$

and hence, the desired equality. $\square$

Indeed, this property completely characterizes the distribution of the process X according to the following theorem.

THEOREM 1.4 (Kolmogorov).– *The **canonical space** $(\Omega, \mathcal{F})$ is defined in the following manner. Let $\Omega = E^{\mathbb{N}}$. The* coordinate mappings *$(X_n)_{n\in\mathbb{N}}$ are defined by $X_n(\omega) = \omega_n$ for any $\omega = (\omega_n)_{n\in\mathbb{N}} \in \Omega$ and we write $\mathcal{F} = \sigma(X_n, n \in \mathbb{N})$. Let $(\mu_n)_{n\in\mathbb{N}}$ be a family of probability distributions such that*

1) for any $n \in \mathbb{N}$, μ_n is defined on $(E^{n+1}, \mathcal{E}^{\otimes(n+1)})$,

2) for any $n \in \mathbb{N}^*$ and $(A_0, \ldots, A_{n-1}) \in \mathcal{E}^n$, we have $\mu_{n-1}(A_0 \times \ldots \times A_{n-1}) = \mu_n(A_0 \times \ldots \times A_{n-1} \times E)$.

Therefore, there exists a unique probability distribution μ *over the canonical space* $(\Omega, \mathcal{F})$ *such that the process* $X = (X_n)_{n\in\mathbb{N}}$ *for the coordinate mapping has the distribution* μ *and for the finite-dimensional distributions has the sequence* $(\mu_n)_{n\in\mathbb{N}}$.

This result is very important for the theory of processes as it signifies that it is sufficient to specify (all) the finite-dimensional distributions and for them to be compatible with each other, to uniquely define a process distribution over the space of infinite random sequences. In practice, this makes it possible to justify the construction of processes (existence property) as well as showing that two processes have the same distribution (unicity property).

To study the random variables taking values in the set of sequences, we need new definitions for σ-algebras and measurability.

DEFINITION 1.20.– *In a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, a* ***filtration*** *is a sequence* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *of sub-*σ*-algebras of* $\mathcal{F}$ *such that, for any* $n \in \mathbb{N}$*,* $\mathcal{F}_n \subset \mathcal{F}_{n+1}$*. This is, thus, a non-decreasing sequence (for inclusion) of sub-*σ*-algebras of* $\mathcal{F}$.

When $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *is a filtration defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, the quadruplet* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n\in\mathbb{N}})$ *is said to be a* ***filtered probability space****.*

EXAMPLE 1.23.– *Let* $(X_n)_{n\in\mathbb{N}}$ *be a sequence of random variables and we consider, for any* $n \in \mathbb{N}$*,* $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$*, the* σ*-algebra generated by* $\{X_0, ..., X_n\}$*. The sequence* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *is, therefore, a filtration, called a* ***natural filtration*** *of* $(X_n)_{n\in\mathbb{N}}$ *or filtration generated by* $(X_n)_{n\in\mathbb{N}}$*. This filtration represents the information revealed over time, by the observation of the drawings of the sequence* $X = (X_n)_{n\in\mathbb{N}}$. $\diamondsuit$

DEFINITION 1.21.– *Let* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n\in\mathbb{N}})$ *be a filtered probability space, and let* $X = (X_n)_{n\in\mathbb{N}}$ *be a stochastic process.*

– X *is said to be* ***adapted*** *to the filtration* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *(or again* $(\mathcal{F}_n)_{n\in\mathbb{N}}$*−adapted), if* X_n *is* $\mathcal{F}_n$*-measurable for any* $n \in \mathbb{N}$;

– X *is said to be* ***predictable*** *with respect to the filtration* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *(or again* $(\mathcal{F}_n)_{n\in\mathbb{N}}$*−predictable), if* X_n *is* $\mathcal{F}_{n-1}$*-measurable for any* $n \in \mathbb{N}^*$.

EXAMPLE 1.24.– *A process is always adapted with respect to its natural filtration.* $\diamondsuit$

As its name indicates, for a predictable process, we know its value X_n from the instant $n-1$.

1.4. Exercises

EXERCISE 1.1.– Let $\Omega = \{a, b, c\}$.

1) Completely describe all the σ-algebras of Ω.

2) State which are the sub-σ-algebras of which.

EXERCISE 1.2.– Let $\Omega = \{a, b, c, d\}$. Among the following sets, which are σ-algebras?

1) $\mathcal{A}_1 = \big\{\emptyset, \Omega, \{b\}, \{a, c, d\}\big\}$;

2) $\mathcal{A}_2 = \big\{\emptyset, \Omega, \{b, c, d\}, \{c, d\}\big\}$;

3) $\mathcal{A}_3 = \big\{\{b\}\big\}$;

4) $\mathcal{A}_4 = \big\{\emptyset, \Omega, \{a, c\}, \{b, d\}\big\}$.

For those which are not σ-algebras, completely describe the σ-algebras they generate.

EXERCISE 1.3.– Let X be a random variable on $(\Omega, \mathcal{F})$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Show that X is $\mathcal{G}$-measurable if and only if $\sigma(X) \subset \mathcal{G}$.

EXERCISE 1.4.– Let $A \in \mathcal{F}$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Show that $\mathbb{1}_A$ is $\mathcal{G}$-measurable if and only if $A \in \mathcal{G}$.

EXERCISE 1.5.– Let $\Omega = \{P, F\} \times \{P, F\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$, corresponding to two successive coin tosses. Let

– X_1 be the random variable *number of T on the first toss*;

– X_2 be *the number of T on the second toss*;

– Y be *the number of T obtained on the two tosses*;

– and $Z = 1$ if the two tosses yielded an identical result; otherwise, it is 0.

1) Describe $\mathcal{F}_1 = \sigma(X_1)$ and $\mathcal{F}_2 = \sigma(X_2)$. Is X_1 $\mathcal{F}_2$-measurable?

2) Describe $\mathcal{G} = \sigma(Y)$. Is Y $\mathcal{F}_1$-measurable? Is X_1 $\mathcal{G}$-measurable?

3) Describe $\mathcal{H} = \sigma(Z)$. Is Z $\mathcal{F}_1$-measurable, $\mathcal{G}$-measurable? Is X_1 $\mathcal{H}$-measurable?

4) Give the inclusions between $\mathcal{F}$, $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{G}$ and $\mathcal{H}$.

EXERCISE 1.6.– Let $(X_n)_{n\geq 1}$ be a sequence of independent random variables with the same Rademacher distribution with parameter $1/2$:

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2.$$

Let $S_0 = 0$ and $S_n = \sum_{k=1}^{n} X_k$ for $n \geq 1$. The process $S = (S_n)_{n\in\mathbb{N}}$ is called the simple symmetric random walk on $\mathbb{Z}$. It will be studied in detail in Chapter 3. We write $X_0 = 0$. Show that the filtration generated by the sequence $(X_n)_{n\in\mathbb{N}}$ is the same as that generated by the sequence $(S_n)_{n\in\mathbb{N}}$.

EXERCISE 1.7.– Consider the following game of chance. A player begins by choosing a number between 6 and 8 (inclusive), which we call the principal. The player then rolls 2 uncut, six-sided, non-rigged dice and sums the result. The wins are as follows:

– If the sum is 2 or 3, the player loses 1 DT (Tunisian dinar).

– If the sum is 11, the player wins 1 DT if the principal is 7; otherwise, they lose 1 DT.

– If the sum is 12, the player wins 1 DT if the principal is 6 or 8; otherwise, they lose 1 DT.

– Finally, in all other cases, nothing happens (no win, no loss).

1) Determine Ω, the universe of all outcomes of the experiment.

2) S is the random variable giving the sum of the two dice. Determine the distribution of S.

3) X_7 is the random variable giving the player's winnings (negative in the case of a loss) when the principal is 7. Determine the distribution of X_7.

4) Calculate $\mathbb{E}[X_7]$ and $\mathbb{V}(X_7)$.

5) We consider X_6 to be the winnings of the player when the principal is 6. Determine the distribution of X_6 and then calculate $\mathbb{E}[X_6]$.

EXERCISE 1.8.– Let $p \in]0, 1[$. We have one coin that leads to tails with the probability p. We toss this coin until we obtain tails for the second time. Let X be the number of heads obtained during this experiment.

1) Determine the distribution of X.

2) Justify the existence of the expectation of X.

3) Calculate $\mathbb{E}[X]$.