Advanced Probability: WEEK 1

1 σ-fields

In probability, the fundamental set \( \Omega \) describes the set of all possible outcomes (or realizations) of a given experiment. It might be any set, without any particular structure, such as for example \( \Omega = \{1, \ldots, 6\} \) representing the outcomes of a die roll, or \( \Omega = [0, 1] \) representing e.g. the outcomes of a concentration measurement of some chemical product. Notice moreover that the set \( \Omega \) need not be composed of numbers exclusively. It is e.g. perfectly valid to consider the set \( \Omega = \{\text{banana, apple, orange}\} \).

Given a fundamental set \( \Omega \), it is important to describe what information does one have on the system, namely on the outcomes of the experiment. This notion of information is well captured by the mathematical notion of \( \sigma \)-field, which is defined below. Notice that in elementary probability courses, it is generally assumed that the information one has about a system is complete, so that it becomes useless to introduce the concept below.

Definition 1.1. Let \( \Omega \) be a set. A \( \sigma \)-field (or \( \sigma \)-algebra) on \( \Omega \) is a collection \( \mathcal{F} \) of subsets of \( \Omega \) (or events) satisfying the following three properties or axioms:

(i) \( \emptyset \in \mathcal{F} \).

(ii) If \( A \in \mathcal{F} \), then \( A^c \in \mathcal{F} \).

(iii) If \( (A_n)_{n=1}^\infty \subset \mathcal{F} \), then \( \bigcup_{n=1}^\infty A_n \in \mathcal{F} \). In particular, if \( A, B \in \mathcal{F} \), then \( A \cup B \in \mathcal{F} \).

The following properties can be further deduced from the above axioms (this is left as an exercise):

(iv) \( \Omega \in \mathcal{F} \).

(v) If \( (A_n)_{n=1}^\infty \subset \mathcal{F} \), then \( \bigcap_{n=1}^\infty A_n \in \mathcal{F} \). In particular, if \( A, B \in \mathcal{F} \), then \( A \cap B \in \mathcal{F} \).

(vi) If \( A, B \in \mathcal{F} \) and \( A \subset B \), then \( B \setminus A \in \mathcal{F} \).

Terminology. The pair \((\Omega, \mathcal{F})\) is called a measurable space and the events belonging to \( \mathcal{F} \) are said to be \( \mathcal{F} \)-measurable, that is, they are the events that one can decide on whether they happened or not, given the information \( \mathcal{F} \). In other words, if one knows the information \( \mathcal{F} \), then one is able to tell to which events of \( \mathcal{F} \) (= subsets of \( \Omega \)) does the realization of the experiment \( \omega \) belong.

Example. For a generic set \( \Omega \), the following are always \( \sigma \)-fields:

\[ \mathcal{F}_0 = \{\emptyset, \Omega\} \] (= trivial \( \sigma \)-field).

\[ \mathcal{P}(\Omega) = \{\text{all subsets of } \Omega\} \] (= complete \( \sigma \)-field).

Example 1.2. Let \( \Omega = \{1, \ldots, 6\} \). The following are \( \sigma \)-fields on \( \Omega \):

\[ \mathcal{F}_1 = \{\emptyset, \{1\}, \{2, \ldots, 6\}, \Omega\} \]

\[ \mathcal{F}_2 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\} \]

Example 1.3. Let \( \Omega = [0, 1] \) and \( I_1, \ldots, I_n \) be a family of disjoint intervals in \( \Omega \) such that \( I_1 \cup \ldots \cup I_n = \Omega \) \((\{I_1, \ldots, I_n\} \text{ is also called a partition of } \Omega)\). The following is a \( \sigma \)-field on \( \Omega \):

\[ \mathcal{F}_3 = \{\emptyset, I_1, \ldots, I_n, I_1 \cup I_2, \ldots, I_1 \cup I_2 \cup I_3, \ldots, \Omega\} \] (NB: there are \( 2^n \) events in total in \( \mathcal{F}_3 \)).

1.1 \( \sigma \)-field generated by a collection of events

An event carries in general more information than itself. As an example, if one knows whether the result of a die roll is odd (corresponding to the event \( \{1, 3, 5\} \)), then one also knows of course whether the result is even (corresponding to the event \( \{2, 4, 6\} \)). It is therefore convenient to have a mathematical description of the information generated by a single event, or more generally by a family of events.
Definition 1.4. Let $\mathcal{A} = \{A_i, i \in I\}$ be a collection of events, where $I$ need not be a countable set. The \textit{\sigma-field generated by} $\mathcal{A}$ is the smallest \sigma-field on $\Omega$ containing all the events $A_i$. It is denoted as $\sigma(\mathcal{A})$.

Remark. A natural question is whether such a vague definition makes sense. Observe first that there is always at least one \sigma-field containing $\mathcal{A}$: it is $\mathcal{P}(\mathcal{A})$. Then, one can show that an arbitrary intersection of \sigma-fields is still a \sigma-field. One can therefore provide the following alternative definition of $\sigma(\mathcal{A})$: it is the intersection of all \sigma-fields containing the collection $\mathcal{A}$, which is certainly a well-defined object.

Example. Let $\Omega = \{1, \ldots, 6\}$ (cf. Example 1.2).

Let $\mathcal{A}_1 = \{\{1\}\}$. Then $\sigma(\mathcal{A}_1) = \mathcal{F}_1$.
Let $\mathcal{A}_2 = \{\{1, 3, 5\}\}$. Then $\sigma(\mathcal{A}_2) = \mathcal{F}_2$.
Let $\mathcal{A} = \{\{1\}, \ldots, \{6\}\}$. Then $\sigma(\mathcal{A}) = \mathcal{P}(\Omega)$.

Exercise. Let $\mathcal{A} = \{\{1, 2, 3\}, \{1, 3, 5\}\}$. Compute $\sigma(\mathcal{A})$.

Example. Let $\Omega = [0, 1]$ and let $\mathcal{A}_3 = \{I_1, \ldots, I_n\}$ (cf. Example 1.3). Then $\sigma(\mathcal{A}_3) = \mathcal{F}_3$.

Borel \sigma-field. Another important example of generated \sigma-field on $\Omega = [0, 1]$ is the following:

$$B([0, 1]) = \sigma(\{\{0\}, \{1\}, [a, b] : a, b \in [0, 1], a < b\}),$$

is the \textit{Borel} \sigma-field on $[0, 1]$ and elements of $B([0, 1])$ are called the Borel subsets of $[0, 1]$. As surprising as it may be, $B([0, 1]) \neq \mathcal{P}([0, 1])$, which generates some difficulties from the theoretical point of view. Nevertheless, it is quite difficult to construct explicit examples of subsets of $[0, 1]$ which are \textit{not} in $B([0, 1])$.

Notice indeed that

a) All singletons belong to $B([0, 1])$. Indeed, for any $0 < x < 1$, $\{x\} = \bigcap_{n \geq 1}[x - \frac{1}{n}, x + \frac{1}{n}]$ belongs to $B([0, 1])$, by the property (v) above and the fact that the Borel \sigma-field is by definition the smallest \sigma-field containing all open intervals.

b) Therefore, all closed intervals, being unions of open intervals and singletons, also belong to $B([0, 1])$.

c) Likewise, all countable intersections of open intervals $B([0, 1])$, as well as all countable unions of closed intervals belong to $B([0, 1])$.

d) The story goes on with countable unions of countable intersections of open intervals, etc. Even though the list is quite long, not all the subsets of $[0, 1]$ are part of $B([0, 1])$, as mentioned above.

Remark. In general, the \sigma-field generated by a collection of events contains many more elements than the collection itself! The Borel \sigma-field is a good example (where “many more” has to be interpreted here in the sense that the cardinality of $B([0, 1])$ is the same as that of $\mathcal{P}([0, 1])$, while the cardinality of $\{\{0\}, \{1\}, [a, b] : a, b \in [0, 1], a < b\}$ is the same as that of $[0, 1]$). In the finite case, you will observe the same phenomenon while computing $\sigma(\{1, 2, 3\}, \{1, 3, 5\})$ on $\Omega = \{1, \ldots, 6\}$.

1.2 Sub-\sigma-field

One may have more or less information about a system. In mathematical terms, this translates into the fact that a \sigma-field has more or less elements. It is therefore convenient to introduce a (partial) ordering on the ensemble of existing \sigma-fields, in order to establish a \textit{hierarchy} of information. This notion of hierarchy is important and will come back when we will be studying stochastic processes that evolve in time.

Definition 1.5. Let $\Omega$ be a set and $\mathcal{F}$ be a \sigma-field on $\Omega$. A \textit{sub-\sigma-field of} $\mathcal{F}$ is a collection $\mathcal{G}$ of events such that:

(i) If $A \in \mathcal{G}$, then $A \in \mathcal{F}$.

(ii) $\mathcal{G}$ is itself a \sigma-field.

Notation. $\mathcal{G} \subset \mathcal{F}$. 
**Remark.** Let $\Omega$ be a generic set. The trivial $\sigma$-field $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is a sub-$\sigma$-field of any other $\sigma$-field on $\Omega$. Likewise, any $\sigma$-field on $\Omega$ is a sub-$\sigma$-field of the complete $\sigma$-field $\mathcal{P}(\Omega)$.

**Example.** Let $\Omega = \{1, \ldots, 6\}$ (cf. Example 1.2). Notice that $\mathcal{F}_1$ is not a sub-$\sigma$-field of $\mathcal{F}_2$ (even though $\{1\} \subset \{1,3,5\}$), nor is $\mathcal{F}_2$ a sub-$\sigma$-field of $\mathcal{F}_1$. In general, notice that

1) If $A \in \mathcal{G}$ and $\mathcal{G} \subset \mathcal{F}$, then it is true that $A \in \mathcal{F}$.

but

2) $A \subset B$ and $B \in \mathcal{G}$ together do not imply that $A \in \mathcal{G}$.

**Example.** Let $\Omega = [0,1]$ (cf. Example 1.3). Then $\mathcal{F}_3$ is a sub-$\sigma$-field of $\mathcal{B}([0,1])$. Also, if $\mathcal{F}_4 = \sigma(J_1, \ldots, J_m)$, where $\{J_1, \ldots, J_m\}$ represents a finer partition of the interval $[0,1]$ (i.e. each interval $I$ of $\mathcal{F}_3$ is a disjoint union of intervals $J$), then $\mathcal{F}_3 \subset \mathcal{F}_4$.

## 2 Random variables

The notion of random variable is usually introduced in elementary probability courses as a vague concept, essentially characterized by its distribution. In mathematical terms however, random variables do exist prior to their distribution: they are functions from the fundamental set $\Omega$ to $\mathbb{R}$. Here is a preliminary definition.

**Definition 2.1.** On the set $\mathbb{R}$, one defines the Borel $\sigma$-field as

$$\mathcal{B}(\mathbb{R}) = \sigma([a,b[ : a, b \in \mathbb{R}, a < b]).$$

The elements of $\mathcal{B}(\mathbb{R})$ are called Borel sets. Again, notice that $\mathcal{B}(\mathbb{R})$ is strictly included in $\mathcal{P}(\mathbb{R})$.

**Definition 2.2.** Let $(\Omega, \mathcal{F})$ be a measurable space. A random variable on $(\Omega, \mathcal{F})$ is a map $X : \Omega \to \mathbb{R}$ satisfying

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (1)$$

**Notation.** One often simply denotes the set $\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\} = X^{-1}(B)$: it is called the inverse image of the set $B$ through the map $X$ (watch out that $X$ need not be a bijective function in order for this set to be well defined).

**Terminology.** The above random variable $X$ is sometimes called $\mathcal{F}$-measurable, in order to emphasize that if one knows the information $\mathcal{F}$, then one knows the value of $X$.

**Example.** If $\mathcal{F} = \mathcal{P}(\Omega)$, then condition (1) is always satisfied, so every map $X : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable random variable. On the contrary, if $\mathcal{F} = \{\emptyset, \Omega\}$, then the only random variables which are $\mathcal{F}$-measurable are the maps $X : \Omega \to \mathbb{R}$ which are constant.

**Remark.** Condition (1) can be shown to be equivalent to the following condition:

$$\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{F}, \quad \forall t \in \mathbb{R},$$

which is significantly easier to check.

**Definition 2.3.** Let $(\Omega, \mathcal{F})$ be a measurable space and $A \in \mathcal{F}$ be an event. Then the map $\Omega \to \mathbb{R}$ defined as

$$\omega \mapsto 1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise,} \end{cases}$$

is a random variable on $(\Omega, \mathcal{F})$. It is called the indicator function of the event $A$.

**Example.** Let $\Omega = \{1, \ldots, 6\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$ (cf. Example 1.2). Then $X_1(\omega) = \omega$ and $X_2(\omega) = 1_{\{1,3,5\}}(\omega)$ are both random variables on $(\Omega, \mathcal{F})$. Moreover, $X_2$ is $\mathcal{F}_2$-measurable, but notice that $X_1$ is neither $\mathcal{F}_1$- nor $\mathcal{F}_2$-measurable.
Example. Let $\Omega = [0, 1]$ and $\mathcal{F} = \mathcal{B}([0, 1])$ (cf. Example 1.3). Then $X_3(\omega) = \sum_{i=1}^{n} x_i 1_{I_i}(\omega)$ and $X_4(\omega) = \omega$ are both random variables on $(\Omega, \mathcal{F})$. Notice however that only $X_3$ is $\mathcal{F}_3$-measurable.

We will need to consider not only random variables, but also functions of random variables. This is why we introduce the following definition.

**Definition 2.4.** A map $g : \mathbb{R} \to \mathbb{R}$ such that

$$\{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R}), \quad \forall B \in \mathcal{B}(\mathbb{R})$$

is called a Borel-measurable function on $\mathbb{R}$.

**Remark.** A Borel-measurable function on $\mathbb{R}$ is therefore nothing but a random variable on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

As it is difficult to construct explicitly sets which are not Borel sets, it is equally difficult to construct functions which are not Borel-measurable. Nevertheless, one often needs to check that a given function is Borel-measurable. A useful criterion for this is the following (given here without proof).

**Proposition 2.5.** If $g : \mathbb{R} \to \mathbb{R}$ is continuous, then it is Borel-measurable.

Finally, let us mention this useful property of functions of random variables.

**Proposition 2.6.** If $X$ is an $\mathcal{F}$-measurable random variable and $g : \mathbb{R} \to \mathbb{R}$ is Borel-measurable, then $Y = g(X)$ is also an $\mathcal{F}$-measurable random variable.

**Proof.** Let $B \in \mathcal{B}(\mathbb{R})$. Then

$$\{Y \in B\} = \{g(X) \in B\} = \{X \in g^{-1}(B)\} \in \mathcal{F},$$

since $X$ is an $\mathcal{F}$-measurable random variable and $g^{-1}(B) \in \mathcal{B}(\mathbb{R})$ by assumption. \qed

### 2.1 $\mathcal{\sigma}$-field generated by a collection of random variables

The amount of information contained in a random variable, or more generally in a collection of random variables, is given by the definition below.

**Definition 2.7.** Let $(\Omega, \mathcal{F})$ be a measurable space and $\{X_i, i \in I\}$ be a collection of random variables on $(\Omega, \mathcal{F})$. The $\mathcal{\sigma}$-field generated by $X_i$, $i \in I$, denoted as $\sigma(X_i, i \in I)$, is the smallest $\mathcal{\sigma}$-field $\mathcal{G}$ on $\Omega$ such that all the random variables $X_i$ are $\mathcal{G}$-measurable.

**Remark.** Notice that

$$\sigma(X_i, i \in I) = \sigma(\{\{X_i \in B\}, i \in I, B \in \mathcal{B}(\mathbb{R})\}),$$

where the right-hand side expression refers to Definition 1.4. It turns out that one also has

$$\sigma(X_i, i \in I) = \sigma(\{\{X_i \leq t\}, i \in I, t \in \mathbb{R}\}).$$

**Example.** Let $(\Omega, \mathcal{F})$ be a measurable space. If $X_0$ is a constant random variable (i.e. $X_0(\omega) = c \in \mathbb{R}$), then $\sigma(X_0) = \{\emptyset, \Omega\}$.

**Example.** Let $\Omega = \{1, \ldots, 6\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$ (cf. Example 1.2). Then $\sigma(X_1) = \mathcal{P}(\Omega)$ and $\sigma(X_2) = \mathcal{F}_2$.

**Example.** Let $\Omega = [0, 1]$ and $\mathcal{F} = \mathcal{B}([0, 1])$ (cf. Example 1.3). Then $\sigma(X_3) = \mathcal{F}_3$ and $\sigma(X_4) = \mathcal{B}([0, 1])$.

Following the proof of Proposition 2.6, the proposition below can be easily shown.

**Proposition 2.8.** If $X$ is a random variable on a measurable space $(\Omega, \mathcal{F})$ and $g : \mathbb{R} \to \mathbb{R}$ is Borel-measurable, then $Y = g(X)$ is a $\sigma(X)$-measurable random variable (this applies in particular to $Y = X$).

As a matter of fact, it turns out that the reciprocal statement is also true: if $Y$ is a $\sigma(X)$-measurable random variable, then there exists a Borel-measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $Y = g(X)$. 

4