AI 5090: Stochastic Processes

Lecture 02

Scribe: Yashmitha G., Rahul Madhogarhia

07 January 2025

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Contents

1 Sample Space 1

2 σ-Algebra 2

3 Borel σ-Algebra 3
3.1 Need for Borel σ-Algebra . . . 3
3.2 Defining Borel σ-Algebra . . . 4

4 Probability Measure 4
4.1 Computing Probability in Practice . . . 5

5 Random Variable 5

This lecture is meant to be a primer on probability theory, covering concepts essential for the study of stochastic processes. The lecture covers the concepts of sample space, σ-algebra, Borel σ-algebra, probability measure, and random variables succinctly. The following lectures will touch upon the concepts of cumulative distribution function (CDF), random vectors, and sequences of random variables.

1 Sample Space

We begin this section with some basic assumptions. The theory of probability assumes two important quantities:

1. **Random Experiment** - An experiment or trial is any procedure that has a well-defined set of possible outcomes. An experiment is said to be random if the results of the experiment are not always predictable.

2. **Outcome** - It is the result of conducting the experiment once.

Once an experiment is decided, the next step would be to quantify the set of all possible outcomes or the sample space as it is popularly known. Below, we formally define the sample space of an experiment.

Definition 1 (Sample Space of a Random Experiment)**.** *The set of all possible outcomes associated with a random experiment is called the* **sample space** *of the experiment, and is denoted by the symbol* Ω.

Example 1. *Consider an experiment in which a coin with two faces labelled H (for heads) and T (for tails) is tossed n times in succession. The outcome of this experiment depends on the quantity of interest, as detailed below.*

- *The face of the coin that shows up: if this is the quantity of interest, then* $\Omega = \{H, T\}$.
- *The speed with which the coin lands: if this is the quantity of interest, then* $\Omega = [0, \infty) = \mathbb{R}_+$.
- *The number of times the coin flips in the air: if this is the quantity of interest, then* $\Omega = \{0, 1, 2, 3, \dots\} = \mathbb{N} \cup \{0\}$.

Remark 1. *The sample space can be a combination of one or more quantities of interest. For instance, consider the set of all tuples of the form (x, y), where:*

- *x represents the number of coin flips before landing.*

- y *represents the face of the coin that shows up.*

In this case, we have

$$\Omega = \{(x, y) : \ (x, y) \in \mathbb{N} \times \{H, T\}\}$$

2 σ-Algebra

Often, the object of interest to us is not merely a single outcome of the experiment, but rather a collection of outcomes that forms a subset of the sample space Ω. The overarching goal here is to be able to assign probabilities to one or more subsets of Ω of interest to us. The notion of a σ-algebra, a formal definition of which is provided below, provides a framework to collect subsets of Ω whose probabilities we would like to speak of.

Definition 2 (σ-Algebra)**.** *Let Ω be a sample space. A σ-**algebra** (denoted as $\mathscr{F}$) on Ω is a collection of subsets of Ω that satisfies the following properties:*

1. $\Omega \in \mathscr{F}$.
2. *If $A \in \mathscr{F}$, then $A^c \in \mathscr{F}$, where $A^c = \Omega \setminus A$ is the complement of A in Ω (*closure under set complement).*
3. *If $A_1, A_2, A_3, \dots \in \mathscr{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$, or $A_1 \cup A_2 \cup A_3 \cup \dots \in \mathscr{F}$ (closure under *countable union).*

Aside:

- **Closure:** A set is said to be closed under an operation if applying that operation to elements of the set always results in an element that is within the set.
- **Countable Union** (denoted as σ): A countable union refers to the union of countably infinitely many sets, indexed by a countable index set. Without loss of generality, we may assume the index set to be $\mathbb{N}$. Allowing for repetition of sets participating in the countable union, this operation also covers finite unions (or unions of finitely many sets).
- **Countable Intersection** (denoted as δ): A countable intersection refers to the intersection of a sequence of sets, indexed by natural numbers. In other words, it is the intersection of an infinite number of sets, $A_1, A_2, A_3, \dots$, and it is denoted by $\bigcap_{i=1}^{\infty} A_i$.
- **Algebra on a Set:** An algebra $\mathscr{A}$ on a set Ω is similar to a σ-algebra but is closed only under finite unions, and not necessarily under countable unions. That is, for all $n \in \mathbb{N}$ and $A_1, \dots, A_n \in \mathscr{A}$, we have $A_1 \cup \dots \cup A_n \in \mathscr{A}$. Every σ-algebra is an algebra, but the converse is not true in general.

Remark 2. *Because we have*

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c \right)^c,$$

courtesy of de Morgan's law, it follows that a σ-algebra is closed under countable intersections.

Remark 3. *A σ-algebra is also referred to as σ-field.*

σ-Algebra as a Set of Sets: A σ-algebra is essentially a set of subsets of the sample space Ω. In essence, it is a "set of sets" because it is a collection of sets (subsets of Ω) that is closed under certain operations like complements, countable unions, and countable intersections.

Example 2. *Let $\Omega = \{1, 2, 3, 4, 5, 6\}$. Following σ-algebras can be defined on Ω:*

- $\mathscr{F}_1 = \{\emptyset, \Omega\}$. *This is the "smallest" σ-algebra possible.*
- $\mathscr{F}_2 = \{A \mid A \subseteq \Omega\}$, *where $\mathscr{F}_2$ contains all subsets of Ω. Notice that $\mathscr{F}_2$ is essentially the power set of Ω (denoted as $\mathcal{P}(\Omega)$ or 2^{Ω}), and is the "biggest" σ-algebra possible.*
- $\mathscr{F}_3 = \{\Omega, \emptyset, \{1\}, \{2, 3, 4, 5, 6\}\}$
- $\mathscr{F}_4 = \{\emptyset, \{1, 2, 3\}, \{4, 5, 6\}, \Omega\}$

Example 3 (Building a σ-Algebra)**.** *Let $\Omega = \{1, 2, 3, 4, 5, 6\}$. Start with the following collection of subsets of Ω:*

$$\mathscr{F}_0 = \{\emptyset, \Omega, \{1\}, \{2, 3\} \{4, 5\}\}.$$

Notice that $\{1\}^c = \{2, 3, 4, 5, 6\}$ is not present in $\mathscr{F}_0$. Clearly,$\mathscr{F}_0$ is NOT a σ-algebra on Ω. To make it a σ-algebra, we undertake the following procedure. First, we add the complement of each set in $\mathscr{F}_0$. Doing so gives us the collection

$$\mathscr{F}_1 = \{\emptyset, \Omega, \{1\}, \{2, 3\}, \{4, 5\}, \{2, 3, 4, 5, 6\}, \{1, 4, 5, 6\} \{1, 2, 3, 6\}\}$$

We then include all possible unions and intersections sets in $\mathscr{F}_1$ to arrive at

$$\mathscr{F} = \Big\{\emptyset, \Omega, \{1\}, \{6\}, \{2, 3\}, \{4, 5\}, \{1, 6\}, \{4, 5, 6\},$$
$$\{1, 2, 3\}, \{2, 3, 6\}, \{1, 4, 5\}, \{1, 4, 5, 6\}, \{1, 2, 3, 6\}, \{2, 3, 4, 5\}, \{2, 3, 4, 5, 6\}, \{1, 2, 3, 4, 5\}\Big\}.$$

Note that $\mathscr{F}$ satisfies all the conditions of Definition 2, and is therefore a valid σ-algebra on Ω.

Remark 4. *In the previous example, observe that $|\mathscr{F}| = 16 = 2^4$.*
It is a non-trivial fact that the cardinality of any σ-algebra, if finite, is always a power of 2. For more details, the reader is invited to look at the solution to question 2 in this document.

Remark 5 (Different levels of infinity)**.** *If a set Ω is countably infinite (i.e., there exists a bijection between Ω and $\mathbb{N}$), then its power set 2^Ω has the same cardinality as the set of real numbers $\mathbb{R}$. Therefore:*

$$|\Omega| < |2^\Omega|.$$

More generally, by Cantor's theorem, for any set S, the following hierarchy holds:

$$|S| < |2^S| < |2^{2^S}| < |2^{2^{2^S}}| < \dots$$

This illustrates that the cardinality of power sets increases hierarchically, with each power set representing a larger level of infinity. Traditionally, the cardinality of $\mathbb{N}$ is referred to as aleph$_0$ (represented by $\aleph_0$), that of $2^\mathbb{N}$ as aleph$_1$ (represented by $\aleph_1$), that of $2^\mathbb{R}$ as aleph$_2$ (represented by $\aleph_2$), and so on.

The purpose of defining a σ-algebra $\mathscr{F}$ on a sample space Ω is to **assign probabilities** to each set in $\mathscr{F}$. A natural question that arises is: why not simply consider $\mathscr{F} = 2^\Omega$ always, and assign probabilities to every subset of Ω? It turns out that when Ω is an uncountably infinite set (of cardinality greater than or equal to $\aleph_1$), the cardinality of 2^Ω is too large to assign probabilities to every set in a meaningful way. To address this issue, **Émile Borel**, a French mathematician and pioneer in measure theory, introduced a special construct of σ-algebra called the **Borel σ-algebra**.
In the next section, we will study the Borel σ-algebra in more detail.

3 Borel σ-Algebra

3.1 Need for Borel σ-Algebra

Suppose that Ω is an uncountably infinite set (i.e., has cardinality greater than or equal to $\aleph_1$); for instance, $\Omega = \mathbb{R}$. A natural question is: why not simply consider $\mathscr{F} = 2^\Omega$ as the corresponding σ-algebra on Ω? This question was studied in depth by Émile Borel, who realised that it is impossible to assign probabilities to every set in $2^\mathbb{R}$ in a meaningful way. Simply put, Émile Borel showed that one or more axioms of probability (to be outlined in the next section) will be violated in the process of assigning probabilities to every set in 2^Ω whenever Ω is uncountably infinite). To circumvent this problem, Borel found the necessity to define a framework for constructing a σ-algebra of subsets of Ω that is non-trivial (i.e., contains all sets of interest to us), yet not equal to the power set 2^Ω. This led him to the construction of a novel σ-algebra, now popularly known as the Borel σ-algebra, defined formally below considering $\Omega = \mathbb{R}$.

3.2 Defining Borel σ-Algebra

Let $\Omega = \mathbb{R}$. Consider the collection

$$\mathscr{D} = \{(-\infty, x] : x \in \mathbb{R}\}.$$

The collection $\mathscr{D}$ is closed under finite intersections, meaning the intersection of any finite number of sets in $\mathscr{D}$ is also in $\mathscr{D}$. Such a collection of sets is called a π**-system**. The smallest σ-algebra that can be constructed from the sets in $\mathscr{D}$ is called the Borel σ-algebra of subsets of $\mathbb{R}$, denoted by $\mathscr{B}(\mathbb{R})$. Importantly, Borel demonstrated that $|\mathscr{B}(\mathbb{R})| = \aleph_1 = |\mathbb{R}|$. That is, the cardinality of $\mathscr{B}(\mathbb{R})$ is same as that of $\mathbb{R}$.

A word of caution: Try not to visualize what $\mathscr{B}(\mathbb{R})$ looks like, as it contains infinitely many elements.

Remark 6. *The concept of Borel σ-algebra is not limited to $\mathbb{R}$.*

- *For a finite or countably infinite set Ω, the Borel σ-algebra on Ω is identical to the power set 2^{Ω}.*
- *When Ω is uncountably infinite, the cardinality of the Borel σ-algebra on Ω is one level of infinity smaller than that of the power set.*

The pair $(\Omega, \mathscr{F})$, where $\mathscr{F}$ is a σ-algebra of subsets of Ω, is called a measurable space.

4 Probability Measure

As alluded to in the previous section, the purpose of constructing a σ-algebra is to assign probabilities to every set in the σ-algebra. Such an assignment of probabilities is formally defined via a probability measure as outlined below.

Definition 3 (Probability Measure)**.** *Fix a measurable space $(\Omega, \mathscr{F})$. A* **probability measure** $\mathbb{P} : \mathscr{F} \to [0, 1]$ *is a function on $\mathscr{F}$ satisfying the following axioms:*

1. $\mathbb{P}(\Omega) = 1.$
2. $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$ *for all $A \in \mathscr{F}$.*
3. *For any countable collection of mutually disjoint sets $A_1, A_2, A_3, \ldots \in \mathscr{F}$,*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Here, the sets $A_1, A_2, \ldots$ are mutually disjoint, meaning that $A_i \cap A_j = \varnothing$ for all $i \neq j$.

The above conditions are known as the **axioms of probability**.

Remark 7. *From axioms 1 and 2 above, we deduce that $\mathbb{P}(\varnothing) = 0$, where $\varnothing$ is the empty set.*

The tuple $(\Omega, \mathscr{F}, \mathbb{P})$ is referred to as a **probability space**.

Example 4. *Consider rolling a single dice with sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathscr{F} = \{\varnothing, \Omega, \{1\}, \{2, 3, 4, 5, 6\}\}$. From the first axiom, $\mathbb{P}(\Omega) = 1$, and $\mathbb{P}(\varnothing) = 0$. Assuming the die is fair (i.e., each outcome is equally likely), we have*

$$\mathbb{P}(\{1\}) = \frac{1}{6}, \qquad \mathbb{P}(\{2, 3, 4, 5, 6\}) = 5 \cdot \frac{1}{6} = \frac{5}{6}.$$

Note: For any subset $S \subset \Omega$, if $S \notin \mathscr{F}$, we cannot define $\mathbb{P}(S)$. To define $\mathbb{P}(S)$, S must belong to $\mathscr{F}$. Thus, defining $\mathscr{F}$ appropriately is crucial.

4.1 Computing Probability in Practice

Hereafter, we shall refer to any set in the σ-algebra $\mathscr{F}$ as an **event**. Suppose that $A \in \mathscr{F}$. Upon performing the random experiment, suppose that $\omega \in \Omega$ results as the outcome of the experiment. If $\omega \in A$, we say that the event A **occurs**.

One practical method to compute probability of any event $A \in \mathscr{F}$ is to repeat the random experiment indefinitely, and to track the limiting proportion of the number of times the event A occurs. Mathematically:

$$\mathbb{P}(A) = \lim_{N\to\infty} \frac{N(A)}{N} = \lim_{N\to\infty} \frac{\text{\#times } A \text{ occurs}}{N}.$$

Some remarks are in order.

1. Ω is called the **sure event**
2. Any set $A \in \mathscr{F}$ such that $\mathbb{P}(A) = 1$ is called an **almost-sure event**
3. If $\mathbb{P}(A) = 0$, it does **not** imply that $A = \varnothing$
4. If $\mathbb{P}(A) = 1$, it does **not** imply that $A = \Omega$

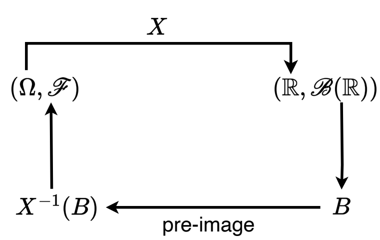
Example 5. *Consider N trials of tossing a coin in which heads appears only once and the remaining trials yield tails. Using an earlier definition for computing probabilities, we have*

$$\mathbb{P}(\{\textit{Head}\}) = \lim_{N\to\infty} \frac{1}{N} = 0, \qquad \mathbb{P}(\{\textit{Tail}\}) = \lim_{N\to\infty} \frac{N-1}{N} = 1.$$

Here, even though $\mathbb{P}(\{\textit{Head}\}) = 0$, the event $\{\textit{Head}\}$ itself is **not** *the empty set. The empty set occurs when no other outcome occurs at all. Similarly, $\mathbb{P}(\{\textit{Tail}\}) = 1$ does* **not** *imply that the event $\{\textit{Tail}\}$ is itself the complete sample space. In this case, the event $\{\textit{Tail}\}$ is an almost-sure event.*

5 Random Variable

Having understood the probability measure $\mathbb{P}$ defined on a measurable space $(\Omega, \mathscr{F})$, we will now introduce the concept of a random variable.



$$\forall B \in \mathscr{B}(\mathbb{R}), \quad X^{-1}(B) \in \mathscr{F}$$

Figure 1: Pictorial representation of a random variable.

Definition 4 (Random Variable)**.** *Fix $(\Omega, \mathscr{F})$. A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if*

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

Equivalently, $X : \Omega \to \mathbb{R}$ is a random variable with respect to $\mathscr{F}$ if

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \in (-\infty, x]\} = \{\omega \in \Omega : X(\omega) \le x\} \in \mathscr{F}, \qquad \forall x \in \mathbb{R}.$$

The above equation is known as the condition of measurability.

The equivalent definition is a remarkable statement as it implies that we only need to check if the pre-images of *generating sets* (in this case, sets of the form $(-\infty, x]$ for $x \in \mathbb{R}$, which collectively generate every set in $\mathscr{B}(\mathbb{R})$ through the set operations of countable unions, complements, and countable intersections) belong to $\mathscr{F}$, in contrast to checking the preceding condition for the pre-image of every set in $\mathscr{B}(\mathbb{R})$.

Note: $X^{-1}(B)$ is not an inverse function. It is simply a notation for the set of all $\omega \in \Omega$ that map to the elements of B, i.e., the set of elements in the sample space Ω whose images under the random variable X lie in the set B. Mathematically,

$$X^{-1}(B) \coloneqq \{\omega \in \Omega : X(\omega) \in B\}.$$

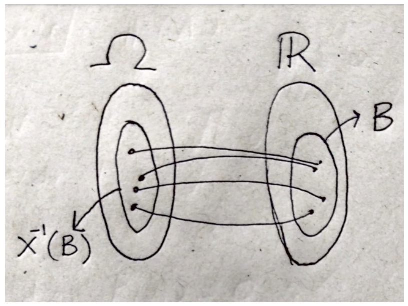


Figure 2: Set representation of $X^{-1}(B)$.

In the coming lecture, we will touch upon some examples of a valid random variable X as defined, among other topics.