T2

Probability Theory

Salvador Ruiz Correa

August 14, 2025

AGENDA:
1. σ-algebra generators.
2. Measureble space.
3. Product space.
4. Borel algebras.
5. Measures and probability measure.

A MEASURABLE SPACE is a foundational construct in measure theory, defined as an ordered pair $(X, \mathcal{F})$, where X is a non-empty set and $\mathcal{F}$ is a σ-algebra of subsets of X. The σ-algebra $\mathcal{F}$ specifies which subsets of X are considered measurable, thereby enabling the rigorous definition of measures, integration, and probability. Measurable spaces serve as the domain over which measures are defined, allowing for the formal treatment of size, probability, and convergence in both pure and applied contexts. This abstraction underpins much of modern analysis, probability theory, and statistical modeling, providing the structural framework for defining measurable functions, constructing product spaces, and formulating stochastic processes.

Borel Algebra

Informally speaking, you can think of a Borel algebra as a special collection of sets that helps you organize everyday numbers in a neat and organized way.

Imagine you have all the real numbers, like 1, 2, 3, and so on, and also the numbers with decimals like 1.5, 2.75, and so forth. These numbers can be really messy, and it's challenging to sort them neatly. But with a Borel algebra, you create a system where you can group these numbers together in a smart way.

Here's how it works: you start with simple sets, like all the numbers that are less than 3, or all the numbers between 1.5 and 2.5. Then, you can combine these sets in various ways to create more sets. For example, you can combine the set of numbers less than 3 with the set of numbers between 1.5 and 2.5 to create a new set: all the numbers between 1.5 and 3.

So, a Borel algebra is like a toolbox that helps you neatly organize numbers into sets that make sense, making it easier to study and work with these numbers in mathematics and statistics. It's a bit like putting numbers into labeled boxes to keep them tidy and manageable.

Consider the collection of all open intervals (a, b) of $\mathbb{R}$, where $a < b$. The minimum σ-algebra generated by this collection is called the σ-algebra of $\mathbb{R}$ and is denoted by $\mathcal{B}(\mathbb{R})$, A σ-algebra $\mathscr{F}$ on a set Ω is a family of subsets of Ω such that:

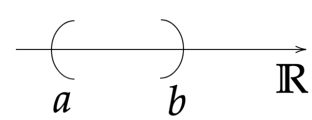

Figure 1: Open interval (a, b), $a \neq b$.

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{(a, b) \subseteq \mathbb{R} : a \leq b\}\right).$$

For any real numbers $a \leq b$, the following, are all elements of $\mathcal{B}(\mathbb{R})$.

- $(-\infty, a)$, (b, ∞), $(-\infty, a) \cup (b, \infty)$.
- $[a, b] = ((-\infty, a) \cup (b, \infty))^c$.
- $(-\infty, a] = \cup_{n=1}^{\infty}[a-n, a]$ and $[b, \infty) = \cup_{n=1}^{\infty}[b, b+n]$.
- $(a, b] = (a, \infty) \cap (-\infty, b]$
- $\{a\} = \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, a + \frac{1}{n})$ and $\{a_1, a_2, \ldots, a_n\} = \cup_{k=1}^{n}\{a_k\}$.

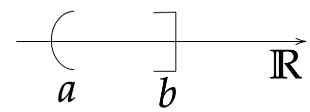

Figure 2: Semi-open interval $(a, b] = (a, \infty) \cap (-\infty, b]$, $a \neq b$.

*Borel Algebra in $\mathbb{R}^n$**

The Borel Algebra in $\mathbb{R}^n$ is defined as:

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{B}(\mathbb{R}) \times \cdots \times \mathcal{B}(\mathbb{R})).$$

The Borel σ-algebra is generated by many different systems of sets. The most interesting are:

- Open rectangles:

$$\mathscr{I}^{n,o} = \mathscr{I}^n(\mathbb{R}^n) = \{(a_1, b_1) \times \cdots \times (a_n, b_n) : a_i, b_i \in \mathbb{R}\}$$

.

- From the right half-open rectangles:

$$\mathscr{I} = \mathscr{I}(\mathbb{R}^n) = \{[a_1, b_1) \times \cdots \times [a_n, b_n) : a_i, b_i \in \mathbb{R}\}$$

.

- Specific examples
  - $\mathscr{I}^2(\mathbb{R}^2) = \sigma(\{(a, b) \times (c, d) \subseteq \mathbb{R}^2 : a \leq b, c \leq d\})$.
  - $\mathscr{I}(\mathbb{R}^2) = \sigma(\{[a, b) \times [c, d) \subseteq \mathbb{R}^2 : a \leq b, c \leq d\})$.
  - $\mathscr{I}^3(\mathbb{R}^3) = \sigma(\{(a, b) \times (c, d) \times (e, f) \subseteq \mathbb{R}^3 : a \leq b, c \leq d, e \leq f\})$.

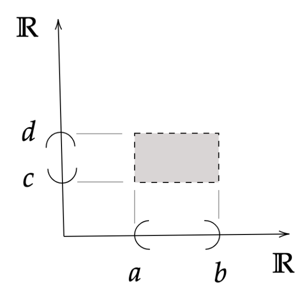

Figure 3: Rectangle $R \in \mathscr{I}(\mathbb{R}^2)$ for the real numbers a, b, c and d for which $a < b$ and $c < d$.

Measurable Space

Let $\mathscr{F}$ be a σ-algebra defined on the sample space Ω. The pair $(\Omega, \mathscr{F})$ is called *measurable space*.

Definition 1: Product Space

Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be two measurable spaces. The measurable spaces product is defined as

$$(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2),$$

where $\mathscr{F}_1 \otimes \mathscr{F}_2 = \sigma(\mathscr{F}_1 \times \mathscr{F}_2)$, and $\mathcal{G} = \mathscr{F}_1 \times \mathscr{F}_2$ is the σ-algebra generator. In general,

$$\mathscr{F}_1 \times \mathscr{F}_2 := \{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_1 \in \mathcal{F}_2\},$$

is not a σ-algebra.

Box 1: Product Space Example

Compute $(\Omega, \mathscr{F}) = (\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2)$ given:

$$\Omega_1 = \{\mathsf{H}, \mathsf{T}\},$$

$$\mathscr{F}_1 = \{\emptyset, \{\mathsf{H}\}, \{\mathsf{T}\}, X_1\},$$
$$\Omega_2 = \{\mathsf{h}, \mathsf{t}\},$$
$$F_2 = \{\emptyset, \{\mathsf{h}\}, \{\mathsf{t}\}, X_2\}.$$

- The cartesian product $\Omega_1 \times \Omega_2$ is:

$$\Omega = \Omega_1 \times \Omega_2 = \{(\mathsf{H}, \mathsf{h}), (\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{h}), (\mathsf{T}, \mathsf{t})\}.$$

- σ-algebra generator:

$$\begin{aligned} \mathscr{F}_1 \times \mathscr{F}_2 = \{ & (\emptyset, \emptyset), (\emptyset, \mathsf{h}), (\emptyset, \mathsf{t}), (\emptyset, \Omega_2), \\ & (\mathsf{H}, \emptyset), (\mathsf{H}, \mathsf{h}), (\mathsf{H}, \mathsf{t}), (\mathsf{H}, \Omega_2), \\ & (\mathsf{T}, \emptyset), (\mathsf{T}, \mathsf{h}), (\mathsf{T}, \mathsf{t}), (\mathsf{H}, \Omega_2), \\ & (\Omega_1, \emptyset), (\Omega_1, \mathsf{h}), (\Omega_1, \mathsf{t}), (\Omega_1, \Omega_2)\}. \end{aligned}$$

- σ-algebra:

$$\begin{aligned} \mathscr{F} = \mathscr{F}_1 \otimes \mathscr{F}_2 = \{ & \emptyset, \\ & \{(\mathsf{H}, \mathsf{h})\}, \{(\mathsf{H}, \mathsf{t})\}, \{(\mathsf{T}, \mathsf{h})\}, \{(\mathsf{T}, \mathsf{t})\} \\ & \{(\mathsf{H}, \mathsf{h}), (\mathsf{H}, \mathsf{t})\}, \{(\mathsf{H}, \mathsf{h}), (\mathsf{T}, \mathsf{h})\}, \{(\mathsf{H}, \mathsf{h}), (\mathsf{T}, \mathsf{t})\}, \\ & \{(\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{h})\}, \{(\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{t})\}, \\ & \{(\mathsf{T}, \mathsf{h}), (\mathsf{T}, \mathsf{t})\}, \\ & \{(\mathsf{H}, \mathsf{h}), (\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{h})\}, \{(\mathsf{H}, \mathsf{h}), (\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{t})\}, \\ & \{(\mathsf{H}, \mathsf{t}), (\mathsf{T}, \mathsf{h}), (\mathsf{T}, \mathsf{t})\}, \\ & \{(\mathsf{H}, \mathsf{h}), (\mathsf{T}, \mathsf{h}), (\mathsf{T}, \mathsf{t})\}, \\ & \Omega\}. \end{aligned}$$

Measures

Intuitively, a (measure-theoretic) measure is a way of giving value to things, but it's not as simple as measuring the length of a rope or the weight of an object. Instead, it's used for more complicated situations where you want to understand and compare things that aren't always straightforward to measure. Imagine you have a collection of different-sized buckets, and you want to know how much water each one can hold. Each bucket might have an irregular shape, and you can't just fill them to see how much they hold. A measure-theoretic measure helps you assign a value to each bucket's capacity, even when they're not simple shapes. This concept is used in advanced mathematics and statistics to tackle complex problems and understand things in a more abstract or generalized way.

Measure theory is indeed concerned with the problem of how to assign a size to certain sets. In daily life this is easy to do: we can count, take measurements, or calculate rates. In each case, we compare and express the result with respect to some base unit.

Definition 2: Measure Definition

A positive measure μ on Ω is a map $\mu : \mathscr{F} \to [0, \infty]$ satisfying:

- $\mathscr{F}$ is a σ-algebra in Ω.
- $\mu(\emptyset) = 0$
- If $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$ are pair-wise disjoint, then

$$\mu\left(\dot{\bigcup_{n\in\mathbb{N}}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

Examples of Measures

- The set function on λ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that assigns every half-open rectangle $[a, b)$ the value

$$\lambda([a,b)) = b - a.$$

  is called the *one-dimensional Lebesgue measure*.
- Let $\Omega = \{\omega_1, \omega_2, \dots\}$ be a countable and $(p_1, p_2, \dots)$ a sequence of numbers $p_n \in [0, 1]$, such that $\sum_{n\in\mathbb{N}} p_n = 1$. On $(\Omega, \mathcal{P}(\omega), P)$ the set function

$$P(A) = \sum_{n:\omega_n\in A}, \quad A \in \Omega,$$

  defines a *probability measure*. The triplet $(\Omega, \mathcal{P}(\omega), P)$ is called discrete probability space.
- Box 2 shows an example of a discrete probability measure. Here we reproduce one of the tables of Box 2.

Ω	$P(\omega)$
$\omega_1 = \mathsf{H^cW^cS^cR^c}$	α_1
$\omega_2 = \mathsf{H^cW^cS^cR}$	α_2
$\omega_3 = \mathsf{H^cW^cSR^c}$	α_3
$\omega_4 = \mathsf{H^cW^cSR}$	α_4
$\omega_5 = \mathsf{H^cWS^cR^c}$	α_5
$\omega_6 = \mathsf{H^cWS^cR}$	α_6
$\omega_7 = \mathsf{H^cWSR^c}$	α_7
$\omega_8 = \mathsf{H^cWSR}$	α_8
$\omega_9 = \mathsf{HW^cS^cR^c}$	α_9
$\omega_{10} = \mathsf{HW^cS^cR}$	α_{10}
$\omega_{11} = \mathsf{HW^cSR^c}$	α_{11}
$\omega_{12} = \mathsf{HW^cSR}$	α_{12}
$\omega_{13} = \mathsf{HWS^cR^c}$	α_{13}
$\omega_{14} = \mathsf{HWS^cR}$	α_{14}
$\omega_{15} = \mathsf{HWSR^c}$	α_{15}
$\omega_{16} = \mathsf{HWSR}$	α_{16}

$$P(\omega_k) = \alpha_k,$$
$$\sum_{k=1}^{16} P(\omega_k) = \sum_{k=1}^{16} \alpha_k = 1.$$

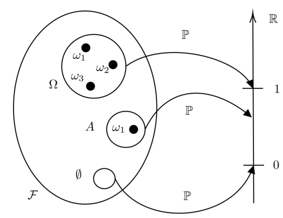

Figure 4: A probability measure is a special case of a measure. Suppose you have a measurable space. Then a function $\mu : \mathcal{F} \to [0, +\infty]$ is called a measure on $(\Omega, \mathcal{F})$ if it gives a size of zero, to the empty set and if it is countably additive. Notice how the measure takes elements of $\mathcal{F}$ and not elements of the sample space Ω; i.e., it maps events, not outcomes. A probability measure is just a measure, with the additional property that it maps to $[0, 1]$ rather than to $[0, +\infty]$. In order words, it assigns probabilities to events. We can see in the diagram above how the probability measure maps the empty set to zero, the sample space to 1, and any other event to some number in $[0, 1]$.

Probability Measures

In non-technical terms, a probability measure is a way of describing how likely or unlikely something is to happen. It's a bit like looking at weather forecasts that tell you the chances of rain. When we say there's a 50% probability of rain, it means that out of every two similar situations, it's expected to rain in one of them.

So, a probability measure assigns a number between 0 and 1 to events, where 0 means the event won't happen, 1 means it will definitely happen, and values in between tell us the likelihood of something occurring. It helps us understand and work with uncertainty, make predictions, and make informed decisions based on the chances of different events (Figure 19).

Definition 3: Probability Measure Definition

A probability measure is a a map $P = \mathbb{P} : \mathscr{F} \to [0, 1]$ satisfying:

- $\mathscr{F}$ is a σ-algebra in Ω.
- $P(\emptyset) = 0$
- $P(\Omega) = 1$
- If $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$ are pair-wise disjoint, then

$$P\left(\dot{\bigcup_{n\in\mathbb{N}}} A_n\right) = \sum_{n\in\mathbb{N}} P(A_n).$$

let $(\Omega, \mathscr{F}, P)$ be a probability space, $A, B, B_n, A_n \in \Omega$, $(A_n)_{n\in\mathbb{N}}$ pairwise disjoint.

1. $A \cap B = \emptyset \Longrightarrow P(A \cup B) = P(A) + P(B)$.

2. $A \subset B \implies P(A) \le P(B)$.
3. $A \subset B, \implies P(B \setminus A) = P(B) - P(A)$.
4. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.
5. $P(A \cup B) \le P(A) + P(B)$.
6. $P\left(\bigcup_{n\in\mathbb{N}} B_n\right) \le \sum_{n\in\mathbb{N}} P(B_n)$.
7. $\dot\bigcup_{n\in\mathbb{N}} A_n = \Omega \implies P(A) = \sum_{n\in\mathbb{N}} P(A \cap A_n)$.

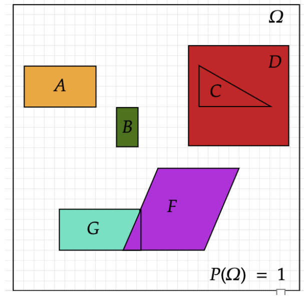

Figure 5: Events in the sample space Ω.

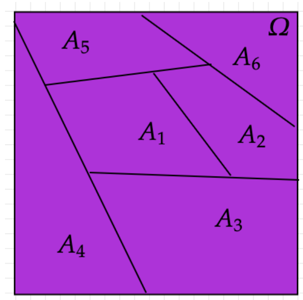

Figure 6: Property 7 example.

Probability Measure Examples

- We assume the outcome can be directly observed at the end of the experiment and thus $\mathscr{F}$ is chosen to be the largest possible σ-algebra, and we define P on it. The precise definition of P depends on the application:
  - For a finite sample space Ω where each outcome is equally likely, define P on $\mathscr{F} = \mathscr{P}(\Omega)$ via $P(A) = \frac{|A|}{|\Omega|}$ for any $A \in \mathscr{F}$.
  - To model the number of coin flip required to obtain the first head ($\Omega = \{1, 2, 3, ...\}$), define P on $\mathscr{F} = \mathscr{P}(\Omega)$ where P satisfies $P(\{\omega : \omega = \mathsf{k}\}) = (1 - \mathsf{p})^{\mathsf{k}-1}\mathsf{p}$. Here $\mathsf{p} \in (0, 1)$ represents the chance of getting a head in a single flip.
- To represent a uniform random number draw from $\Omega = [0, 1]$, define P on $\mathscr{F} = \mathcal{B}([0, 1])$ where P satisfies $P([a, b]) = b - a$ for $0 \le a \le b \le 1$. This is the Lebesgue measure on $[0, 1]$.