Measure Theory

1 σ-Algebras

Definition 1.1. A σ-algebra $\mathcal{A}$ on a set X is a family of subsets of X such that:

1. $X \in \mathcal{A}$ $(\mathbf{\Sigma}_1)$
2. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ $(\mathbf{\Sigma}_2)$
3. If $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ $(\mathbf{\Sigma}_3)$

A set $A \in \mathcal{A}$ is said to be *measurable* or $\mathcal{A}$-*measurable*.

Example 1.1.

1. $\mathcal{P}(X)$ is a σ-algebra (the maximal σ-algebra on X).
2. $\{\emptyset, X\}$ is a σ-algebra (the minimal σ-algebra on X).
3. $A := \{A \subseteq X : \#A < \infty \text{ or } \#A^c < \infty\}$ is a σ-algebra.
4. (Trace σ-algebra) Let $E \subseteq X$ be any set and let $\mathcal{A}$ be a σ-algebra on X. Then
$$\mathcal{A}_E := \{E \cap A : A \in \mathcal{A}\}$$
is a σ-algebra on E.

Proof. We verify the three defining properties of a σ-algebra on E:

(a) Since $X \in \mathcal{A}$, we have $E = E \cap X \in \mathcal{A}_E$.
(b) If $E \cap A \in \mathcal{A}_E$, then $E \setminus (E \cap A) = E \cap A^c$, and since $A^c \in \mathcal{A}$, it follows that $E \cap A^c \in \mathcal{A}_E$.
(c) If $(E \cap A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}_E$, then $\bigcup_n (E \cap A_n) = E \cap \bigcup_n A_n$, and since $\bigcup_n A_n \in \mathcal{A}$, we conclude that $\bigcup_n (E \cap A_n) \in \mathcal{A}_E$.

Hence, $\mathcal{A}_E$ is a σ-algebra on E. $\square$

5. (Pre-image σ-algebra) Let $f : X \to X'$ be a function and let $\mathcal{A}'$ be a σ-algebra on X'. Then

$$\mathcal{A} := \{f^{-1}(A') : A' \in \mathcal{A}'\}$$

is a σ-algebra on X.

Theorem 1.1. Let X be a set and let $\{\mathcal{A}_i : i \in I\}$ be a family of σ-algebras on X. Define

$$\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i = \{A \subseteq X : A \in \mathcal{A}_i \text{ for all } i \in I\}.$$

Then, $\mathcal{A}$ is a σ-algebra on X.

Proof. We verify the σ-algebra properties for $\mathcal{A}$:

1. Since $X \in \mathcal{A}_i$ for all $i \in I$, we have $X \in \mathcal{A}$.
2. If $A \in \mathcal{A}$, then $A \in \mathcal{A}_i$ for all $i \in I$, so $A^c = X \setminus A \in \mathcal{A}_i$ for all $i \in I$, hence $A^c \in \mathcal{A}$.
3. If $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$, then $A_n \in \mathcal{A}_i$ for all n and i, so $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_i$ for all $i \in I$, and thus $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Therefore, $\mathcal{A}$ is a σ-algebra on X. □

Definition 1.2. Let X be a set and let $\mathcal{E} \subseteq \mathcal{P}(X)$ be a collection of subsets of X. The *σ-algebra generated by* $\mathcal{E}$, denoted by $\sigma(\mathcal{E})$, is the smallest σ-algebra on X containing all sets in $\mathcal{E}$. That is,

$$\sigma(\mathcal{E}) := \bigcap \big\{\mathcal{A} \subseteq \mathcal{P}(X) : \mathcal{A} \text{ is a } \sigma\text{-algebra on } X,\ \mathcal{E} \subseteq \mathcal{A}\big\}.$$

Remark 1.1 (Generated σ-algebras)**.**

1. If $\mathcal{A}$ is a σ-algebra, then $\sigma(\mathcal{A}) = \mathcal{A}$.
2. For $A \subseteq X$, we have $\sigma(\{A\}) = \{\emptyset, A, A^c, X\}$.
3. If $\mathcal{F} \subseteq \mathcal{G} \subseteq \mathcal{A}$, then $\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{G}) \subseteq \sigma(\mathcal{A})$.

Definition 1.3 (Topological Space)**.** A *topological space* is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T} \subseteq \mathcal{P}(X)$ is a collection of subsets of X, called *open sets*, satisfying the following properties:

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,

2. If $\{U_\alpha \in \mathcal{T} : \alpha \in I\}$ is an arbitrary collection of open sets, then the union $\bigcup_{\alpha \in I} U_\alpha \in \mathcal{T}$,
3. If $\{U_i \in \mathcal{T} : i = 1, \ldots, n\}$ is a finite collection of open sets, then the intersection $\bigcap_{i=1}^n U_i \in \mathcal{T}$.

The collection $\mathcal{T}$ is called a *topology* on X. The complement of an open set is called a *closed set*.

Remark 1.2 (Standard Topology on $\mathbb{R}^n$)**.** A subset $U \subseteq \mathbb{R}^n$ is called *open* if for every point $x \in U$, there exists an $\varepsilon > 0$ such that the open ball

$$B_\varepsilon(x) := \{y \in \mathbb{R}^n : \|x - y\| < \varepsilon\},$$

where $\|\cdot\|$ denotes the Euclidean norm, is contained in U; that is, $B_\varepsilon(x) \subseteq U$.

The collection of all such open sets is denoted by $\mathcal{O} = \mathcal{O}_{\mathbb{R}^n}$ and forms the *standard topology* on $\mathbb{R}^n$.

Definition 1.4 (Borel σ-algebra)**.** The σ-algebra $\sigma(\mathcal{O})$ generated by the collection of open sets $\mathcal{O} = \mathcal{O}_{\mathbb{R}^n}$ of $\mathbb{R}^n$ is called the *Borel σ-algebra* on $\mathbb{R}^n$.

Its elements are called *Borel sets* or *Borel measurable sets*. We denote the Borel σ-algebra by $\mathcal{B}(\mathbb{R}^n)$.

Definition 1.5. Let X be a topological space and let $A \subseteq X$. A collection $\{U_\alpha\}_{\alpha \in A} \subseteq \mathcal{T}$ of open sets is called an *open cover* of A if

$$A \subseteq \bigcup_{\alpha \in A} U_\alpha.$$

A *subcover* is a subcollection that still covers A. The set A is called *compact* if every open cover of A admits a finite subcover.

Remark 1.3. In $\mathbb{R}^n$, a set is compact if and only if it is closed and bounded (Heine–Borel Theorem).

Theorem 1.2 (Borel σ-algebra from Different Generators)**.** Let $\mathcal{O}, \mathcal{C}, \mathcal{K} \subseteq \mathcal{P}(\mathbb{R}^n)$ denote the collections of open, closed, and compact subsets of $\mathbb{R}^n$, respectively. Then,

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{O}) = \sigma(\mathcal{C}) = \sigma(\mathcal{K}).$$

Proof. Since compact sets are closed, we have $\mathcal{K} \subseteq \mathcal{C}$, and by Remark 1.1(3), $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{C})$. Conversely, for any $C \in \mathcal{C}$, define $C_k := C \cap B_k(0)$, where $B_k(0)$ is the closed ball of radius k centered at the origin. Each C_k is closed and bounded, hence compact, so $C_k \in \mathcal{K}$. Since $C = \bigcup_{k\in\mathbb{N}} C_k$, it follows that $C \in \sigma(\mathcal{K})$, and thus $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{K})$.

Next, since $\mathcal{C} = \mathcal{O}^c := \{U^c : U \in \mathcal{O}\}$, and complements of sets in a σ-algebra are again in the σ-algebra, it follows that $\mathcal{C} \subseteq \sigma(\mathcal{O})$, hence $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{O})$. The reverse inclusion follows similarly from $\mathcal{O} = \mathcal{C}^c$. Therefore, we have:

$$\sigma(\mathcal{K}) = \sigma(\mathcal{C}) = \sigma(\mathcal{O}).$$

□

Generating Sets of the Borel Algebra. The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ can be generated by various systems of sets. Of particular importance are:

- The family of open rectangles:

$$\mathcal{J}_{o,n} := \{(a_1, b_1) \times \cdots \times (a_n, b_n) : a_i, b_i \in \mathbb{R}\},$$

- The family of half-open rectangles:

$$\mathcal{J}_n := \{[a_1, b_1) \times \cdots \times [a_n, b_n) : a_i, b_i \in \mathbb{R}\}.$$

We denote by $\mathcal{J}_n^{\mathrm{rat}}, \mathcal{J}_{o,n}^{\mathrm{rat}}$ the subsets with rational endpoints. These sets represent intervals in $\mathbb{R}$, rectangles in $\mathbb{R}^2$, cuboids in $\mathbb{R}^3$, and hypercubes in higher dimensions.

Theorem 1.3. We have the following equality of Borel σ-algebras on $\mathbb{R}^n$:

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}_n^{\mathrm{rat}}) = \sigma(\mathcal{J}_{o,n}^{\mathrm{rat}}) = \sigma(\mathcal{J}_n) = \sigma(\mathcal{J}_{o,n}),$$

Remark 1.4. Let $D \subseteq \mathbb{R}$ be a dense subset, for example $D = \mathbb{Q}$ or $D = \mathbb{R}$. Then the Borel sets on $\mathbb{R}$ can also be generated by any of the following families of intervals:

$$\{(-\infty, a) : a \in D\}, \quad \{(-\infty, a] : a \in D\}, \quad \{(a, \infty) : a \in D\}, \quad \{[a, \infty) : a \in D\}.$$

2 Measure Spaces

Definition 2.1. A *(positive) measure* μ on X is a map $\mu : \mathcal{A} \to [0, \infty]$, where $\mathcal{A}$ is a σ-algebra on X, satisfying:

$$\mu(\emptyset) = 0, \tag{M1}$$

and for any pairwise disjoint sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$,

$$\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n). \tag{M2}$$

Property (M2) is also called *countable additivity.*

If μ satisfies (M1), (M2), but $\mathcal{A}$ is not a σ-algebra, then μ is called a *pre-measure.*

Remark 2.1. (M2) requires implicitly that $\bigsqcup_n A_n$ is again in $\mathcal{A}$ this is clearly the case for σ-algebras, but needs special attention when dealing with pre-measures.

Definition 2.2 (Monotone sequences of sets)**.** Let $(A_n)_{n\in\mathbb{N}}$ and $(B_n)_{n\in\mathbb{N}}$ be sequences of subsets of X.

We say (A_n) is *increasing* if

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$$

and write $A_n \uparrow A$ where

$$A := \bigcup_{n\in\mathbb{N}} A_n$$

Similarly, (B_n) is *decreasing* if

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$$

and write $B_n \downarrow B$ where

$$B := \bigcap_{n\in\mathbb{N}} B_n$$

Definition 2.3. Let X be a set and $\mathcal{A}$ a σ-algebra on X. The pair $(X, \mathcal{A})$ is called a *measurable space.* If μ is a measure on $(X, \mathcal{A})$, then $(X, \mathcal{A}, \mu)$ is called a *measure space.*

A measure μ is called:

- *finite* if $\mu(X) < \infty$,
- a *probability measure* if $\mu(X) = 1$.

Accordingly, we speak of a *finite measure space* and a *probability space.*

Definition 2.4. A measure μ on a measurable space $(X, \mathcal{A})$ is called σ*-finite* if there exists a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that:

$$A_n \uparrow X \quad \text{and} \quad \mu(A_n) < \infty \quad \text{for all } n \in \mathbb{N}.$$

In this case, the measure space $(X, \mathcal{A}, \mu)$ is called σ*-finite.*

Lemma 2.1 (Basic properties of measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Then:

(i) If $A_0, \ldots, A_k \in \mathcal{A}$ are pairwise disjoint, then $\mu\big(\bigcup_{n=0}^{k} A_n\big) = \sum_{n=0}^{k} \mu(A_n)$.

(ii) If $A, B \in \mathcal{A}$ with $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

(iii) If $A, B \in \mathcal{A}$, $A \subseteq B$, and $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.

Proof. (i) Extend (A_n) by $A_n = \varnothing$ for $n > k$. Then by countable additivity,

$$\mu\big(\bigcup_{n=0}^{k} A_n\big) = \mu\big(\bigcup_{n=0}^{\infty} A_n\big) = \sum_{n=0}^{\infty} \mu(A_n) = \sum_{n=0}^{k} \mu(A_n).$$

(ii) Since $B = A \cup (B \setminus A)$ with disjoint union,

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(iii) Rearranging gives

$$\mu(B \setminus A) = \mu(B) - \mu(A),$$

which is well-defined if $\mu(A) < \infty$. □

Lemma 2.2 (Main properties of measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Then:

(i) **Countable subadditivity:** For any countable family $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$,

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mu(A_i).$$

(ii) **Continuity from below (increasing sequence):** If $A_1 \subseteq A_2 \subseteq \cdots$ (i.e., $A_n \uparrow A$), then

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

(iii) **Continuity from above (decreasing sequence):** If $B_1 \supseteq B_2 \supseteq \cdots$ (i.e., $B_n \downarrow B$), then

$$\mu\left(\bigcap_{n \in \mathbb{N}} B_n\right) = \lim_{n \to \infty} \mu(B_n).$$

Proof. (i) For countable subadditivity, set $B_k := A_k \setminus \bigcup_{i=1}^{k-1} A_i$, so that (B_k) are disjoint with $B_k \subseteq A_k$. Then,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\bigsqcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

(ii) Let $A_n \uparrow A$, i.e., $A_n \subseteq A_{n+1}$ and $A = \bigcup_{n\in\mathbb{N}} A_n$. Define $B_n := A_n \setminus A_{n-1}$ with $A_0 := \emptyset$. Then (B_n) is disjoint and $\bigsqcup_n B_n = A$. By countable additivity,

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n)$$

(iii) Assume $B_n \downarrow B$, i.e., $B_n \supseteq B_{n+1}$ and $B = \bigcap_n B_n$, with $\mu(B_1) < \infty$. Set $A_n := B_1 \setminus B_n$, so $A_n \uparrow A := B_1 \setminus B$. Then

$$\mu(B) = \mu(B_1) - \mu(A) = \mu(B_1) - \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu(B_n)$$

$\square$

Remark 2.2. With appropriate modifications, these properties also hold for pre-measures, i.e., when $\mathcal{A}$ is not necessarily a σ-algebra.

Example 2.1 (Dirac measure)**.** Let $(X, \mathcal{A})$ be a measurable space and let $x \in X$. Define $\delta_x : \mathcal{A} \to \{0, 1\}$ by

$$\delta_x(A) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then δ_x is a measure on $(X, \mathcal{A})$, called the *Dirac measure* (or unit mass) at the point x.

Example 2.2 (Counting measure)**.** Let $(X, \mathcal{A})$ be a measurable space. Define $\#A : \mathcal{A} \to [0, \infty]$ by

$$\#A := \begin{cases} \text{number of elements in } A & \text{if } A \text{ is finite}, \\ \infty & \text{if } A \text{ is infinite}. \end{cases}$$

Then $\#$ is a measure on $(X, \mathcal{A})$, called the *counting measure*.

Example 2.3 (Discrete probability measure)**.** Let $\Omega = \{\omega_1, \omega_2, \dots\}$ be a countable set, and let $(p_n)_{n\in\mathbb{N}} \subseteq [0,1]$ be a sequence such that $\sum_{n\in\mathbb{N}} p_n = 1$. Define the set function $P : \mathcal{P}(\Omega) \to [0,1]$ by

$$P(A) := \sum_{\{n\in\mathbb{N}:\omega_n\in A\}} p_n = \sum_{n\in\mathbb{N}} p_n\, \delta_{\omega_n}(A), \quad A \subseteq \Omega,$$

where δ_{ω_n} denotes the Dirac measure at ω_n. Then P is a probability measure on $(\Omega, \mathcal{P}(\Omega))$, and the triplet $(\Omega, \mathcal{P}(\Omega), P)$ is called a *discrete probability space.*

Example 2.4 (Linear combination of measures)**.** Let $(X, \mathcal{A})$ be a measurable space, and let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of measures on $(X, \mathcal{A})$. Let $(x_n)_{n\in\mathbb{N}} \subseteq [0, \infty]$. Then the set function

$$\mu := \sum_{n\in\mathbb{N}} x_n \mu_n$$

defined by

$$\mu(A) := \sum_{n\in\mathbb{N}} x_n \mu_n(A), \quad \text{for all } A \in \mathcal{A},$$

is a measure on $(X, \mathcal{A})$

Proof. We verify the axioms of a measure:

(M1) *(Null empty set):* For all $n \in \mathbb{N}$, $\mu_n(\emptyset) = 0$, so

$$\mu(\emptyset) = \sum_{n\in\mathbb{N}} x_n \mu_n(\emptyset) = \sum_{n\in\mathbb{N}} x_n \cdot 0 = 0.$$

(M2) *(Countable additivity):* Let $(A_k)_{k\in\mathbb{N}} \subseteq \mathcal{A}$ be pairwise disjoint. Since each μ_n is a measure, we have

$$\mu_n\left(\bigsqcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mu_n(A_k), \quad \text{for all } n \in \mathbb{N}.$$

Then,

$$\mu\left(\bigsqcup_{k\in\mathbb{N}} A_k\right) = \sum_{n\in\mathbb{N}} x_n \mu_n\left(\bigsqcup_{k\in\mathbb{N}} A_k\right) = \sum_{n\in\mathbb{N}} x_n \sum_{k\in\mathbb{N}} \mu_n(A_k).$$

Since all terms are non-negative, we may exchange the order of summation:

$$\sum_{n\in\mathbb{N}} x_n \sum_{k\in\mathbb{N}} \mu_n(A_k) = \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} x_n \mu_n(A_k) = \sum_{k\in\mathbb{N}} \mu(A_k).$$

Therefore, μ is countably additive. $\square$

Example 2.5 (Restriction of a measure)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and let $A \in \mathcal{A}$. Define the set function $\mu_A : \mathcal{A} \to [0, \infty]$ by

$$\mu_A(B) := \mu(A \cap B), \quad \text{for all } B \in \mathcal{A}.$$

Then μ_A is a measure on $(X, \mathcal{A})$, called the *restriction of* μ *to* A.

Proof. We verify the two defining properties of a measure:

(M1): $\mu_A(\emptyset) = \mu(A \cap \emptyset) = \mu(\emptyset) = 0$.

(M2): Let $(B_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ be pairwise disjoint. Then $(A \cap B_n)_{n\in\mathbb{N}}$ are also pairwise disjoint, and

$$\mu_A\left(\bigsqcup_{n\in\mathbb{N}} B_n\right) = \mu\left(A \cap \bigsqcup_{n\in\mathbb{N}} B_n\right) = \mu\left(\bigsqcup_{n\in\mathbb{N}} (A \cap B_n)\right) = \sum_{n\in\mathbb{N}} \mu(A \cap B_n) = \sum_{n\in\mathbb{N}} \mu_A(B_n).$$

Hence, μ_A is a measure. $\square$

Definition 2.5 (Lebesgue measure on $\mathbb{R}^n$)**.** Define the set function λ_n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ by

$$\lambda_n\left([\![a, b)\!)\right) := \prod_{i=1}^{n} (b_i - a_i),$$

for all $[\![a, b)\!) := [a_1, b_1) \times \cdots \times [a_n, b_n) \in \mathcal{J}_n$. This is called the n-dimensional Lebesgue measure.

Remark 2.3. The set function λ_n is defined only on the family $\mathcal{J}_n$ of half-open rectangles and hence is not yet a measure. Extending λ_n to a measure on $\mathcal{B}(\mathbb{R}^n)$ requires the Carathéodory extension theorem, which will be developed later.