

# Set Theory and Probability I

Mahbubur Rahman Khan

October 24, 2025

## Abstract

This note presents fundamental concepts in Set Theory and Probability Theory. Material from CSE422 class (taught by Md. Shahriar Karim, Assistant Professor, North South University) lectures is enhanced with definitions, theorems, examples, and proofs from faculty notes [2,3] and textbooks by Ross [4] and Chan [1]. Topics include set operations, cardinality, probability spaces, axioms, corollaries, probability assignment (pmf/pdf), measure zero sets, and conditional probability.

## Contents

<b>1</b>	<b>Set Theory</b>	<b>1</b>
1.1	Sets and Elements . . . . .	1
1.2	Describing Sets . . . . .	1
1.3	Fundamental Set Concepts . . . . .	1
1.4	Set Operations . . . . .	2
1.5	Disjoint Sets and Partitions . . . . .	4
1.6	Set Algebra Properties . . . . .	4
<b>2</b>	<b>Cardinality and Countability</b>	<b>4</b>
2.1	Equicardinality and Bijective Functions . . . . .	5
2.2	Types of Sets by Cardinality . . . . .	5
<b>3</b>	<b>The Probability Space <math>(\Omega, \mathcal{F}, \mathbb{P})</math></b>	<b>6</b>
3.1	Component 1: The Sample Space $(\Omega)$ . . . . .	7
3.2	Component 2: The Event Space $(\mathcal{F})$ . . . . .	7
3.3	Component 3: The Probability Law $(\mathbb{P})$ . . . . .	7
3.4	Measurable Sets and the Need for a $\sigma$ -Field . . . . .	8

<b>4</b>	<b>The Axioms of Probability</b>	<b>9</b>
<b>5</b>	<b>Corollaries from the Axioms</b>	<b>9</b>
<b>6</b>	<b>Conditional Probability</b>	<b>10</b>
6.1	Definition . . . . .	10
6.2	Conditional Probability as a Valid Probability Measure . . . . .	10

# 1 Set Theory

Set theory forms the fundamental language and mathematical basis for understanding probability. We begin by defining the core concepts.

## 1.1 Sets and Elements

A **set** is a collection of distinct objects, considered as an entity in itself. The objects within the set are called its **elements** or members. Sets are typically denoted by capital letters (e.g.,  $A, B, S$ ) and their elements by lowercase letters or symbols.

**Definition 1.1** (Set). A **set** is a collection of distinct elements. We often denote a set by listing its elements within curly braces  $[1, 3]$ .

$$A = \{\xi_1, \xi_2, \dots, \xi_n\}$$

where  $\xi_i$  is the  $i$ -th element in the set.

*Remark 1.1.* The elements in a set are **distinct** (no repetitions) and **unordered** (the order in which elements are listed does not matter).

**Example 1.1.** •  $A = \{2, 7, 9\}$  is a finite set of numbers.

- $B = \{\text{apple, orange, pear}\}$  is a finite set of objects.
- $C = \{2, 4, 6, 8, \dots\}$  is an infinite set (the set of positive even integers).

**Definition 1.2** (Element Membership). If an object  $\xi$  is an element of a set  $A$ , we write  $\xi \in A$ . If  $\xi$  is not an element of  $A$ , we write  $\xi \notin A$ .

## 1.2 Describing Sets

Sets can be described explicitly by listing elements or implicitly using properties.

**Definition 1.3** (Set Builder Notation). A set can be defined by specifying a property that its elements must satisfy.

$$A = \{x \mid P(x)\}$$

This reads as "A is the set of all elements  $x$  such that property  $P(x)$  is true."

**Example 1.2.** •  $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$  can be written as  $A = \{x \mid x \in \mathbb{Z}^+ \text{ and } 1 \leq x \leq 9\}$ .

- The set of even integers can be written as  $E = \{x \mid x = 2k \text{ for some integer } k\}$ .

## 1.3 Fundamental Set Concepts

**Definition 1.4** (Set Equality). Two sets  $A$  and  $B$  are **equal** ( $A = B$ ) if and only if they contain exactly the same elements [3, 4].

**Definition 1.5** (Cardinality). The **cardinality** of a finite set  $A$ , denoted  $|A|$ , is the number of distinct elements in the set [3].

**Example 1.3.** For  $A = \{2, 7, 9\}$ ,  $|A| = 3$ .

**Definition 1.6** (Universal Set ( $\Omega$ )). In a particular context, the **Universal Set** ( $\Omega$  or  $S$ ) is the set containing all possible elements under consideration [1, 3].

**Definition 1.7** (Empty Set ( $\emptyset$ )). The **Empty Set** ( $\emptyset$  or  $\{\}$ ) is the unique set containing no elements [3, 4].

**Definition 1.8** (Subset ( $A \subseteq B$ )). Set  $A$  is a **subset** of set  $B$  if every element of  $A$  is also an element of  $B$  [1]. If  $A \subseteq B$  and  $A \neq B$ , then  $A$  is a **proper subset** of  $B$ , denoted  $A \subset B$ .

**Theorem 1.1.** Two sets  $A$  and  $B$  are equal if and only if  $A \subseteq B$  and  $B \subseteq A$  [1].

## 1.4 Set Operations

These operations allow us to combine or modify sets.

**Definition 1.9** (Union ( $A \cup B$ )). The **union** of sets  $A$  and  $B$  is the set containing all elements that are in  $A$ , or in  $B$ , or in both [4].

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

**Example 1.4.** If  $A = \{1, 2, 3\}$  and  $B = \{3, 4, 5\}$ , then  $A \cup B = \{1, 2, 3, 4, 5\}$ .

Union of Two Sets ( $A \cup B$ )

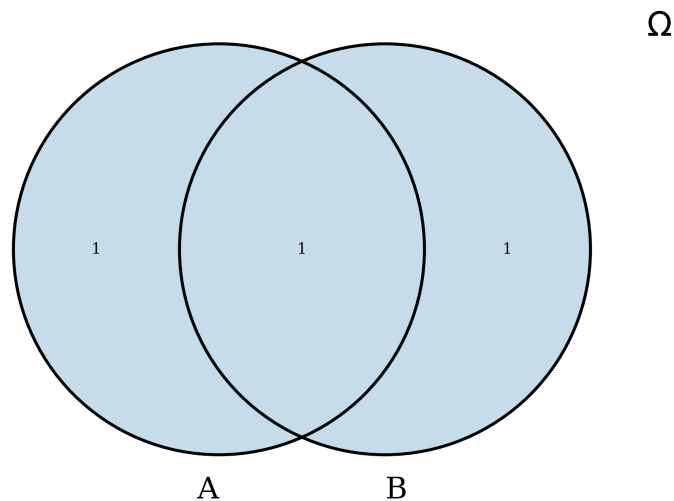


Figure 1: Venn Diagram for Union  $A \cup B$ .

**Definition 1.10** (Intersection ( $A \cap B$ )). The **intersection** of sets  $A$  and  $B$  is the set containing all elements that are common to both  $A$  and  $B$  [4].

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

Intersection of Two Sets ( $A \cap B$ )

$\Omega$

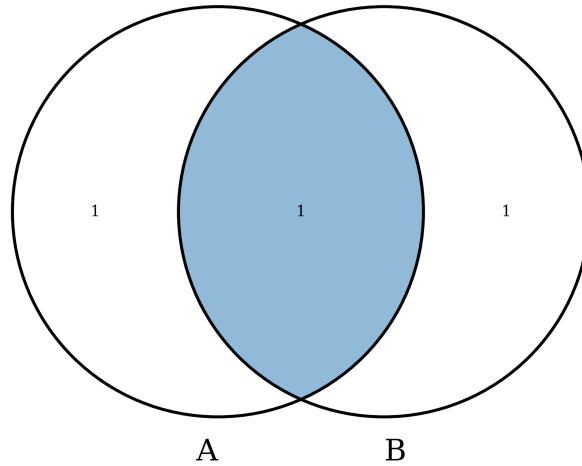


Figure 2: Venn Diagram for Intersection  $A \cap B$ .

**Example 1.5.** If  $A = \{1, 2, 3\}$  and  $B = \{3, 4, 5\}$ , then  $A \cap B = \{3\}$ . If  $C = \{1, 2\}$  and  $D = \{3, 4\}$ , then  $C \cap D = \emptyset$ .

**Definition 1.11** (Complement ( $A^c$  or  $\bar{A}$ )). Relative to a universal set  $\Omega$ , the **complement** of  $A$  is the set of all elements in  $\Omega$  that are *not* in  $A$  [3, 4].

$$A^c = \{x \mid x \in \Omega \text{ and } x \notin A\}$$

Complement of a Set ( $A^c$ )

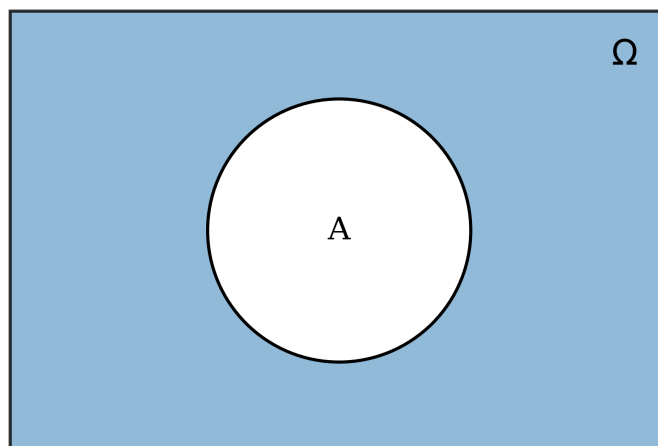


Figure 3: Venn Diagram for Complement  $A^c$ .

**Definition 1.12** (Set Difference ( $A \setminus B$ )). The **difference** of sets  $A$  and  $B$  is the set containing all elements that are in  $A$  but *not* in  $B$  [1, 3].

$$A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$$

Note that  $A \setminus B = A \cap B^c$ .

### Set Difference ( $A \setminus B$ )

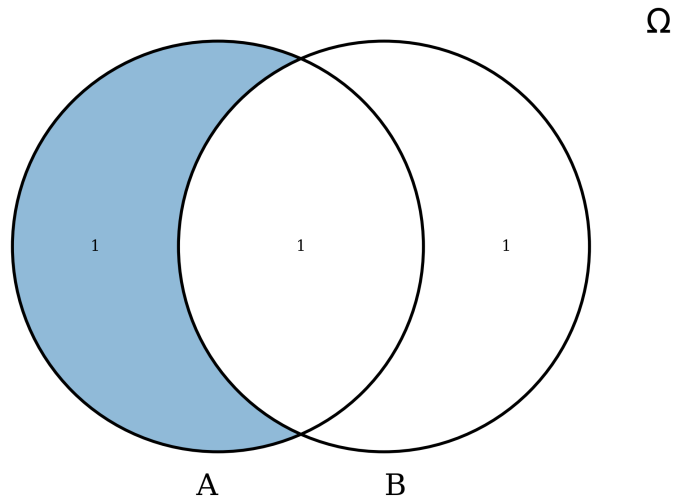


Figure 4: Venn Diagram for Set Difference  $A \setminus B$ .

## 1.5 Disjoint Sets and Partitions

**Definition 1.13** (Disjoint Sets). Two sets  $A$  and  $B$  are **disjoint** (or mutually exclusive) if they have no elements in common, i.e.,  $A \cap B = \emptyset$  [1, 3].

**Definition 1.14** (Partition). A collection of non-empty sets  $\{A_1, \dots, A_n\}$  forms a **partition** of a set  $\Omega$  if [1, 3]:

1. The sets are pairwise disjoint:  $A_i \cap A_j = \emptyset$  for all  $i \neq j$ .
2. The sets are collectively exhaustive:  $\bigcup_{i=1}^n A_i = \Omega$ .

## 1.6 Set Algebra Properties

These properties are essential for manipulating sets.

- **Commutative Laws:**  $A \cup B = B \cup A$ ;  $A \cap B = B \cap A$  [1, 3].
- **Associative Laws:**  $A \cup (B \cup C) = (A \cup B) \cup C$ ;  $A \cap (B \cap C) = (A \cap B) \cap C$  [1, 3].
- **Distributive Laws:**  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ ;  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$  [1, 3].
- **De Morgan's Laws:**  $(A \cup B)^c = A^c \cap B^c$ ;  $(A \cap B)^c = A^c \cup B^c$  [1, 3].

## 2 Cardinality and Countability

Understanding the "size" of sets, especially infinite sets, is important for defining sample spaces.

## Partition of a Set $\Omega$

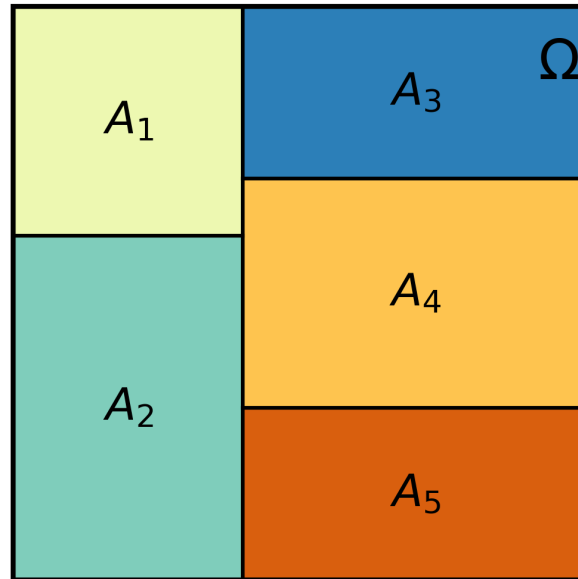


Figure 5: A partition of the set  $\Omega$  [1].

### 2.1 Equicardinality and Bijective Functions

**Definition 2.1** (Equicardinal Sets). Two sets  $A$  and  $B$  are **equicardinal** (have the same cardinality) if there exists a **bijective function**  $f : A \rightarrow B$  between them [1, 3].

**Definition 2.2** (Bijective Function). A function  $f : A \rightarrow B$  is **bijective** (or a one-to-one correspondence) if it is both:

- **Injective (One-to-One):** For every  $a_1, a_2 \in A$ , if  $f(a_1) = f(a_2)$ , then  $a_1 = a_2$ . (Distinct inputs map to distinct outputs) [3].
- **Surjective (Onto):** For every  $b \in B$ , there exists at least one  $a \in A$  such that  $f(a) = b$ . (Every element in the codomain is mapped to) [3].

### 2.2 Types of Sets by Cardinality

- **Finite Set:** A set that is equicardinal with  $\{1, 2, \dots, n\}$  for some non-negative integer  $n$ . Its cardinality is  $n$  [3].
- **Infinite Set:** A set that is not finite [3].
  - **Countably Infinite Set:** An infinite set that is equicardinal with the set of natural numbers  $\mathbb{N} = \{1, 2, 3, \dots\}$  [1, 3].
  - **Uncountable Set:** An infinite set that is not countably infinite [1, 3].

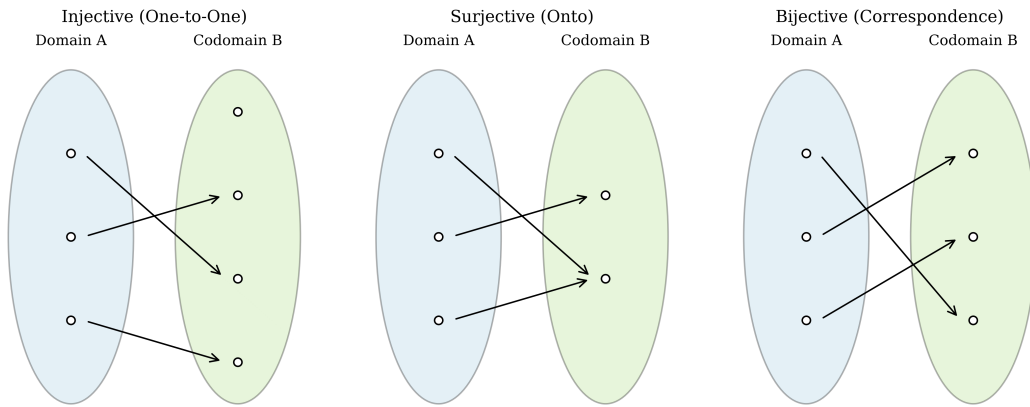


Figure 6: Illustrations of injective, surjective, and bijective functions.

**Definition 2.3** (Countable Set). A set is **countable** if it is either finite or countably infinite [1].

**Example 2.1.** •  $\mathbb{N}$  (natural numbers),  $\mathbb{Z}$  (integers),  $\mathbb{Q}$  (rational numbers) are countable [1,3].

•  $\mathbb{R}$  (real numbers), intervals like  $[0, 1]$  or  $(0, 1)$  are uncountable [1,3].

*Remark 2.1.* Every countably infinite set can be listed as a sequence.

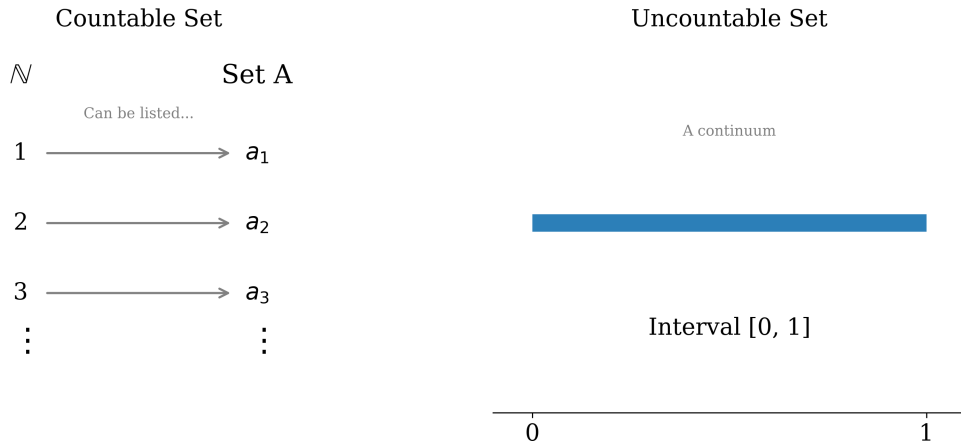


Figure 7: Visualization of countable and uncountable sets.

### 3 The Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$

Probability theory provides a mathematical framework for modeling experiments with random outcomes. This framework is the **probability space**, a triplet  $(\Omega, \mathcal{F}, \mathbb{P})$  [1,2].

### 3.1 Component 1: The Sample Space ( $\Omega$ )

**Definition 3.1** (Sample Space). The **Sample Space** ( $\Omega$ ) is the set of *all possible outcomes* of a random experiment [1, 4]. Each outcome  $w$  is an element of  $\Omega$  ( $w \in \Omega$ ). A sample space must be **exhaustive** (includes all possibilities) and its elements must be **mutually exclusive** (distinct outcomes) [1].

**Example 3.1.** • Experiment: Flip a coin.  $\Omega = \{H, T\}$  [4].

- Experiment: Roll a die.  $\Omega = \{1, 2, 3, 4, 5, 6\}$  [4].
- Experiment: Measure lifetime of a device.  $\Omega = [0, \infty) = \{t \mid t \geq 0\}$  [4].

### 3.2 Component 2: The Event Space ( $\mathcal{F}$ )

**Definition 3.2** (Event). An **Event** ( $A$ ) is a subset of the sample space ( $A \subseteq \Omega$ ) [1, 4]. An event is said to **occur** if the outcome of the experiment is an element of that event.

**Example 3.2.** Experiment: Roll a die,  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .

- Event  $A$ : "An even number is rolled."  $A = \{2, 4, 6\}$ .
- Event  $B$ : "A number less than 3 is rolled."  $B = \{1, 2\}$ .
- Event  $C$ : "The number 5 is rolled."  $C = \{5\}$ .

**Definition 3.3** (Event Space ( $\sigma$ -field)). The **Event Space** ( $\mathcal{F}$ ) is a collection of subsets of  $\Omega$  (i.e., a set of events) that satisfies the properties of a  **$\sigma$ -field** (or  $\sigma$ -algebra). This ensures that we can consistently assign probabilities and perform operations like complements and unions [1]. A collection  $\mathcal{F}$  of subsets of  $\Omega$  is a  $\sigma$ -field if:

1.  $\Omega \in \mathcal{F}$  (The sample space itself is an event).
2. If  $A \in \mathcal{F}$ , then  $A^c \in \mathcal{F}$  (Closed under complementation).
3. If  $A_1, A_2, \dots \in \mathcal{F}$  (a countable sequence), then  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$  (Closed under countable unions).

*Remark 3.1.* For discrete sample spaces,  $\mathcal{F}$  is usually the **power set**  $P(\Omega)$  (the set of all subsets) [1, 3]. For continuous spaces like  $\mathbb{R}$ ,  $\mathcal{F}$  is the **Borel  $\sigma$ -field**  $B(\mathbb{R})$ , generated by intervals [1, 2].

### 3.3 Component 3: The Probability Law ( $\mathbb{P}$ )

**Definition 3.4** (Probability Law/Measure). The **Probability Law** or **Probability Measure**  $\mathbb{P}$  is a function that assigns a real number  $\mathbb{P}[A]$  to each event  $A \in \mathcal{F}$ , satisfying the Axioms of Probability [1, 2].

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

It quantifies the likelihood of each measurable event in  $\mathcal{F}$ , serving as a measure of its relative size within the sample space [1].

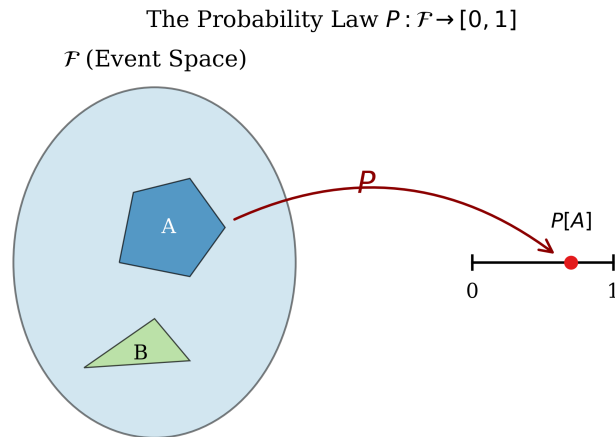


Figure 8: The Probability Law  $\mathbb{P}$  maps events from  $\mathcal{F}$  to values in  $[0, 1]$  [1].

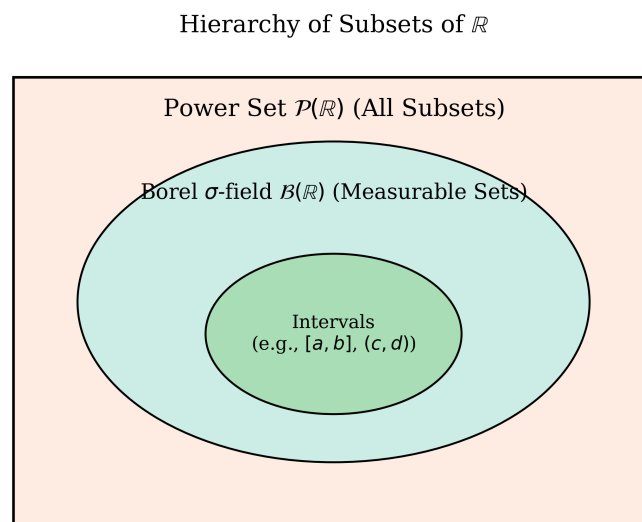
### 3.4 Measurable Sets and the Need for a $\sigma$ -Field

Not every subset of a sample space can be consistently assigned a probability. To avoid contradictions, probability is defined only on a specific collection of subsets  $F$  called a  $\sigma$ -field.

The sets in  $F$  are called **measurable sets**. A probability measure  $P$  assigns values only to these measurable events, not to arbitrary subsets [1].

*Remark 3.2.* In continuous sample spaces such as  $\mathbb{R}$ , the event space is usually the Borel  $\sigma$ -field  $\mathcal{B}(\mathbb{R})$ , which contains intervals and any set formed from them using countable unions, intersections, and complements.

**Example 3.3.** Every interval in  $\mathbb{R}$  such as  $(0, 1)$  or  $[2, \infty)$  is measurable. However, there exist subsets of  $\mathbb{R}$  that are not measurable, and probability cannot be meaningfully assigned to them.



Intuitively, probability measures the likelihood of measurable events. A  $\sigma$ -field ensures that this measurement behaves consistently.

## 4 The Axioms of Probability

Any function  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  must satisfy the following three axioms to be considered a valid probability measure [1, 4].

**Axiom I (Non-negativity):** For any event  $A \in \mathcal{F}$ ,

$$\mathbb{P}[A] \geq 0$$

**Axiom II (Normalization):** The probability assigned to the entire sample space is 1.

$$\mathbb{P}[\Omega] = 1$$

**Axiom III (Countable Additivity):** For any countable sequence of **pairwise disjoint** events  $\{A_1, A_2, \dots\} \subseteq \mathcal{F}$  (i.e.,  $A_i \cap A_j = \emptyset$  for all  $i \neq j$ ), the probability of their union is the sum of their individual probabilities.

$$\mathbb{P} \left[ \bigcup_{i=1}^{\infty} A_i \right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$$

*Remark 4.1.* Axiom III implies **finite additivity**: For any finite collection of pairwise disjoint events  $A_1, \dots, A_n$ ,

$$\mathbb{P} \left[ \bigcup_{i=1}^n A_i \right] = \sum_{i=1}^n \mathbb{P}[A_i]$$

In particular, for two disjoint events  $A$  and  $B$ ,  $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$  [1].

## 5 Corollaries from the Axioms

These fundamental properties are direct consequences of the axioms.

**Corollary 5.1** (Probability of the Complement). *For any event  $A$ ,  $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$  [1, 2, 4].*

*Proof.* Since  $A$  and  $A^c$  form a partition of  $\Omega$  (they are disjoint and  $A \cup A^c = \Omega$ ), Axioms II and III (finite additivity) give:  $1 = \mathbb{P}[\Omega] = \mathbb{P}[A \cup A^c] = \mathbb{P}[A] + \mathbb{P}[A^c]$ . Rearranging yields the result.  $\square$

**Corollary 5.2** (Probability of the Empty Set).  $\mathbb{P}[\emptyset] = 0$  [1, 2, 4].

*Proof.* Since  $\emptyset = \Omega^c$ , apply Corollary 1:  $\mathbb{P}[\emptyset] = \mathbb{P}[\Omega^c] = 1 - \mathbb{P}[\Omega] = 1 - 1 = 0$ .  $\square$

**Corollary 5.3** (Monotonicity). *If  $A \subseteq B$ , then  $\mathbb{P}[A] \leq \mathbb{P}[B]$  [1].*

*Proof.* We can write  $B = A \cup (B \setminus A)$ . The sets  $A$  and  $B \setminus A$  are disjoint. By Axiom III,  $\mathbb{P}[B] = \mathbb{P}[A] + \mathbb{P}[B \setminus A]$ . Since  $\mathbb{P}[B \setminus A] \geq 0$  by Axiom I, it follows that  $\mathbb{P}[B] \geq \mathbb{P}[A]$ .  $\square$

**Corollary 5.4** (Inclusion-Exclusion Principle (for 2 events)). *For any two events  $A$  and  $B$ :*

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$$

[1, 2, 4]

*Proof.* We can write  $A \cup B$  as the union of three disjoint sets:  $(A \setminus B)$ ,  $(B \setminus A)$ , and  $(A \cap B)$ . Thus,  $\mathbb{P}[A \cup B] = \mathbb{P}[A \setminus B] + \mathbb{P}[B \setminus A] + \mathbb{P}[A \cap B]$ . Also,  $A = (A \setminus B) \cup (A \cap B)$  (disjoint union), so  $\mathbb{P}[A] = \mathbb{P}[A \setminus B] + \mathbb{P}[A \cap B]$ , which implies  $\mathbb{P}[A \setminus B] = \mathbb{P}[A] - \mathbb{P}[A \cap B]$ . Similarly,  $\mathbb{P}[B \setminus A] = \mathbb{P}[B] - \mathbb{P}[A \cap B]$ . Substituting these into the first equation gives:  $\mathbb{P}[A \cup B] = (\mathbb{P}[A] - \mathbb{P}[A \cap B]) + (\mathbb{P}[B] - \mathbb{P}[A \cap B]) + \mathbb{P}[A \cap B]$ , which simplifies to  $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$ .  $\square$

## 6 Conditional Probability

Conditional probability allows us to update the probability of an event based on the knowledge that another event has occurred.

### 6.1 Definition

**Definition 6.1** (Conditional Probability). Given two events  $A$  and  $B$ , with  $\mathbb{P}[B] > 0$ , the **conditional probability of  $A$  given  $B$**  is defined as:

$$\mathbb{P}[A | B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

[2]

*Remark 6.1.* Intuitively, knowing that  $B$  has occurred restricts the relevant sample space to just the outcomes within  $B$ .  $\mathbb{P}[A | B]$  measures the proportion of the probability mass within  $B$  that also corresponds to outcomes in  $A$ .

### 6.2 Conditional Probability as a Valid Probability Measure

For a fixed event  $B$  with  $\mathbb{P}[B] > 0$ , the function  $P_B(\cdot) = \mathbb{P}[\cdot | B]$  defines a new, valid probability measure on the original sample space  $\Omega$  and event space  $\mathcal{F}$ . It satisfies the three axioms [2].

*Proof.* (Verifying the axioms for  $P_B(\cdot)$ )

1. **Non-negativity:** For any event  $A$ ,  $\mathbb{P}[A \cap B] \geq 0$  (Axiom I for  $\mathbb{P}$ ) and  $\mathbb{P}[B] > 0$ . Therefore,  $P_B(A) = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \geq 0$ .
2. **Normalization:** We need to show  $P_B(\Omega) = 1$ .  $P_B(\Omega) = \mathbb{P}[\Omega | B] = \frac{\mathbb{P}[\Omega \cap B]}{\mathbb{P}[B]}$ . Since  $B \subseteq \Omega$ ,  $\Omega \cap B = B$ . Thus,  $P_B(\Omega) = \frac{\mathbb{P}[B]}{\mathbb{P}[B]} = 1$ .

3. **Countable Additivity:** Let  $A_1, A_2, \dots$  be a sequence of pairwise disjoint events. We need to show  $P_B(\bigcup A_i) = \sum P_B(A_i)$ .

$$\begin{aligned} P_B\left(\bigcup_{i=1}^{\infty} A_i\right) &= \mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i \mid B\right] \\ &= \frac{\mathbb{P}[(\bigcup_{i=1}^{\infty} A_i) \cap B]}{\mathbb{P}[B]} \\ &= \frac{\mathbb{P}[\bigcup_{i=1}^{\infty} (A_i \cap B)]}{\mathbb{P}[B]} \quad (\text{Distributive Law}) \end{aligned}$$

Since  $A_i$  are disjoint, the events  $(A_i \cap B)$  are also disjoint. Applying Axiom III for  $\mathbb{P}$ :

$$\begin{aligned} P_B\left(\bigcup_{i=1}^{\infty} A_i\right) &= \frac{\sum_{i=1}^{\infty} \mathbb{P}[A_i \cap B]}{\mathbb{P}[B]} \\ &= \sum_{i=1}^{\infty} \frac{\mathbb{P}[A_i \cap B]}{\mathbb{P}[B]} \\ &= \sum_{i=1}^{\infty} \mathbb{P}[A_i \mid B] = \sum_{i=1}^{\infty} P_B(A_i) \end{aligned}$$

Thus,  $P_B(\cdot)$  is a valid probability measure. □

## References

- [1] Stanley H Chan. *Introduction to probability for data science*. Michigan Publishing, Ann Arbor, 2021.
- [2] Md. Shahriar Karim. Cse 422: Probability measures, 2025. Faculty course material. Available at: [https://ece.northsouth.edu/~shahriar.karim/CSE\\_422.html](https://ece.northsouth.edu/~shahriar.karim/CSE_422.html).
- [3] Md. Shahriar Karim. Cse 422: Set theory, 2025. Faculty course material. Available at: [https://ece.northsouth.edu/~shahriar.karim/CSE\\_422.html](https://ece.northsouth.edu/~shahriar.karim/CSE_422.html).
- [4] Sheldon M Ross. *Introduction to Probability Models*. Academic Press, Dec 2006.