Probability theory is based on three fundamental **axioms**, known as **Kolmogorov's Axioms**. These axioms define a probability space and form the foundation of probability calculations.

**Kolmogorov’s Axioms of Probability**

Let SS be the **sample space**, and let P(A)P(A) represent the probability of an event AA. The three axioms are:

1. **Non-Negativity (Axiom 1)**

P(A)≥0for any event AP(A) \geq 0 \quad \text{for any event } A

This means that probabilities cannot be negative.

1. **Normalization (Axiom 2)**

P(S)=1P(S) = 1

The probability of the entire sample space (i.e., the occurrence of some event) is always **1**.

1. **Additivity (Axiom 3)**  
   If AA and BB are **mutually exclusive events** (i.e., they cannot happen at the same time, A∩B=∅A \cap B = \emptyset), then:

P(A∪B)=P(A)+P(B)P(A \cup B) = P(A) + P(B)

This property extends to any finite or countably infinite number of mutually exclusive events.

**Implications of the Axioms**

* The probability of an impossible event (empty set) is **zero**: P(∅)=0P(\emptyset) = 0
* The probability of any event is always between 0 and 1: 0≤P(A)≤10 \leq P(A) \leq 1
* If AcA^c is the **complement** of AA (i.e., the event that AA does not occur), then: P(Ac)=1−P(A)P(A^c) = 1 - P(A)
* If AA and BB are **not** mutually exclusive, then: P(A∪B)=P(A)+P(B)−P(A∩B)P(A \cup B) = P(A) + P(B) - P(A \cap B)

These axioms provide the foundation for probability theory and are widely used in statistics, machine learning, and data science.