

Lecture 1: Axioms of Probability

EECS 126 (UC Berkeley)

Spring 2026 — January 20, 2026

A probability space provides the mathematical framework for modeling random experiments. In this lecture we define the axioms of probability, introduce sample spaces and events, work through several foundational examples, and derive the birthday paradox bound.

1 Probability Spaces

Definition 1.1 (Probability space)

A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where:

- Ω is the *sample space* — the set of all possible outcomes of the experiment.
- $\mathcal{F} \subseteq 2^\Omega$ is a σ -algebra of subsets of Ω , called *events*.
- $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is a *probability measure* satisfying the Kolmogorov axioms.

The elements $\omega \in \Omega$ are called *base outcomes* (or *samples*). A σ -algebra $\mathcal{F}$ is a family of subsets of Ω that is closed under countable unions, countable intersections, and complements, and that contains Ω itself. In many discrete settings we simply take $\mathcal{F} = 2^\Omega$ (the power set), so that every subset of Ω is an event.

Example 1.1 (Tossing two fair coins)

Toss two fair coins. The sample space is

$$\Omega = \{HH, HT, TH, TT\}. \quad (1.1)$$

Each base outcome is equally likely, so we assign $\mathbb{P}(\{\omega\}) = \frac{1}{4}$ for each $\omega \in \Omega$. The event space is the power set $\mathcal{F} = 2^\Omega$, which has $|\mathcal{F}| = 2^{|\Omega|} = 2^4 = 16$ elements.

Example 1.2 (Toss until heads)

Toss a coin repeatedly until the first heads. Then

$$\Omega = \{H, TH, TTH, TTTH, \dots\} \quad (1.2)$$

is a *countably infinite* sample space.

Example 1.3 (Waiting for a bus)

Wait at a bus stop for the next bus, which arrives at some time in $[0, T]$. Now

$$\Omega = [0, T] \quad (1.3)$$

is an *uncountably infinite* (continuous) sample space. In this setting the choice of $\mathcal{F}$ becomes nontrivial; one typically takes $\mathcal{F}$ to be the Borel σ -algebra on $[0, T]$.

2 Kolmogorov Axioms

Definition 1.2 (Kolmogorov axioms)

A function $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}$ is a **probability measure** on $(\Omega, \mathcal{F})$ if it satisfies:

1. **Nonnegativity.** $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$.
2. **Normalization.** $\mathbb{P}(\Omega) = 1$.
3. **Countable additivity.** If $A_1, A_2, \dots \in \mathcal{F}$ are pairwise disjoint ($A_i \cap A_j = \emptyset$ for $i \neq j$), then

$$\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i). \quad (1.4)$$

Why countable additivity? In a continuous sample space such as $\Omega = [0, 1]$, every individual point has probability zero: $\mathbb{P}(\{x\}) = 0$. If we only had *finite* additivity, we could not recover $\mathbb{P}([0, 1]) = 1$ from the probabilities of individual points. Countable additivity bridges this gap and is essential for the theory of continuous distributions.

Example 1.4 (Biased coin)

Flip a biased coin: heads with probability p , tails with probability $1 - p$. Set

$$\Omega = \{H, T\}, \quad \mathcal{F} = 2^\Omega = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}. \quad (1.5)$$

Define

$$\mathbb{P}(H) = p, \quad \mathbb{P}(T) = 1 - p, \quad \mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\{H, T\}) = 1. \quad (1.6)$$

Then $(\Omega, \mathcal{F}, \mathbb{P})$ is a valid probability space.

3 Fundamental Probability Facts

The following basic identities are direct consequences of the Kolmogorov axioms.

Theorem 1.1 (Complement rule)

For any event $A \in \mathcal{F}$,

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A). \quad (1.7)$$

Proof. Since A and A^c are disjoint with $A \cup A^c = \Omega$, countable additivity gives $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$. $\square$

Theorem 1.2 (Inclusion–exclusion)

For any two events $A, B \in \mathcal{F}$,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \quad (1.8)$$

More generally, for events $A_1, A_2, \dots, A_n$,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_i \mathbb{P}(A_i) - \sum_{i < j} \mathbb{P}(A_i \cap A_j) + \sum_{i < j < k} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} \mathbb{P}(A_1 \cap \dots \cap A_n). \quad (1.9)$$

Theorem 1.3 (Union bound)

For any events $A_1, \dots, A_n$,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i). \quad (1.10)$$

Proof. Define $B_1 = A_1$ and $B_k = A_k \setminus \bigcup_{i < k} A_i$ for $k \geq 2$. The B_k are disjoint with $\bigcup B_k = \bigcup A_k$, and $\mathbb{P}(B_k) \leq \mathbb{P}(A_k)$, so $\mathbb{P}(\bigcup A_k) = \sum \mathbb{P}(B_k) \leq \sum \mathbb{P}(A_k)$. $\square$

4 Discrete Probability and Counting

In a discrete sample space, the probability of an event is just the sum of the probabilities of its base outcomes:

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}). \quad (1.11)$$

Definition 1.3 (Uniform sample space)

When all outcomes are equally likely, we say Ω is a **uniform sample space**, and for any event A ,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}. \quad (1.12)$$

5 The Birthday Paradox

Example 1.5 (Birthday paradox)

Let n people be in a room. What is the probability that at least two share a birthday? Let $K = 365$ and assume birthdays are uniform. Let $A =$ “at least two people share a birthday.” Compute via the complement:

$$\mathbb{P}(A^c) = \frac{K(K-1)(K-2)\cdots(K-n+1)}{K^n} = \prod_{j=0}^{n-1} \left(1 - \frac{j}{K}\right). \quad (1.13)$$

Using $1 - x \leq e^{-x}$:

$$\mathbb{P}(A^c) \leq \prod_{j=0}^{n-1} e^{-j/K} = e^{-\sum_{j=0}^{n-1} j/K} = e^{-n(n-1)/(2K)}. \quad (1.14)$$

Theorem 1.4 (Birthday bound)

$\mathbb{P}(A) \geq 1 - e^{-n(n-1)/(2K)}$. For $K = 365$ and $n = 23$, this gives $\mathbb{P}(A) \geq 50\%$.

Balls-and-bins interpretation. The birthday problem is equivalent to throwing n balls uniformly into K bins and asking for the probability of a collision. This connection arises frequently in hashing, load balancing, and networking.

End of Lecture 1.