CECS-381: Stochastic Computing

Instructor: Yutong Zhao
Semester: Fall 2025
Class-2: Axioms of Probability

Recap from Week 1: Counting Principles

- Last week, we established the groundwork for quantifying possibilities. We learned:
  - **Product Rule:** When multiple independent tasks are performed in sequence, the total number of ways is the product of the ways for each task. E.g., choosing a shirt AND pants.
  - **Sum Rule:** When a task can be performed in one of several mutually exclusive ways, the total number of ways is the sum of the ways for each alternative. E.g., choosing a math course OR a CS course.
  - **Permutations:** Arrangements of objects where the order of selection is crucial. E.g., awarding gold, silver, bronze medals. Formula: $P(n,k) = \frac{n!}{(n-k)!}$.
  - **Combinations:** Selections of objects where the order does NOT matter. E.g., forming a committee. Formula: $C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Recap from Week 1: Counting Principles

- These counting principles are indispensable for determining the size of the sample space ($|\Omega|$) and the number of outcomes in an event ($|E|$), which are the building blocks for classical probability.
- Today, we transition from "how many ways" to "how likely" – introducing the formal framework of probability.

What is Probability? (Intuitive Understanding)

- Probability is a numerical measure of the likelihood or chance that a specific event will occur.
- It is always a value between 0 and 1, inclusive.
- A probability of 0 ($P(E) = 0$) indicates that the event E is impossible and will never occur.
- A probability of 1 ($P(E) = 1$) indicates that the event E is certain to occur.
- Values between 0 and 1 represent varying degrees of likelihood.
- Probabilities are often expressed as fractions (e.g., $1/2$), decimals (e.g., 0.5), or percentages (e.g., 50%).

Classical Definition of Probability

- The **Classical Definition** of probability applies when all outcomes in the sample space are equally likely.
- If an experiment has a finite sample space Ω with equally likely outcomes, and an event $E \subseteq \Omega$, then the probability of event E is:

$$P(E) = \frac{\text{Number of favorable outcomes for } E}{\text{Total number of possible outcomes}} = \frac{|E|}{|\Omega|}$$

- **Example:** Rolling a fair six-sided die.
  - Sample Space $\Omega = \{1, 2, 3, 4, 5, 6\}$, so $|\Omega| = 6$.
  - Event $E = \{\text{rolling an even number}\} = \{2, 4, 6\}$, so $|E| = 3$.
  - $P(E) = \frac{|E|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$.
- This definition is intuitive for simple, symmetric experiments.

Limitations of the Classical Definition

- The classical definition relies heavily on the assumption of "equally likely outcomes".
  - **Problem:** What if outcomes are not equally likely? (e.g., a biased coin where heads is more likely than tails). The formula $|E|/|\Omega|$ would not apply directly.
- It struggles with situations involving infinite sample spaces.
  - **Problem:** What is the probability of choosing exactly 0.5 from all real numbers between 0 and 1? The "number of outcomes" becomes ill-defined.
- It doesn't provide a rigorous mathematical foundation for deriving more complex probability rules and theorems.
- To overcome these limitations and build a universal framework, we turn to the axiomatic approach.

Axiomatic Approach to Probability (Kolmogorov, 1933)

- Andrey Kolmogorov introduced an axiomatic framework for probability.
- Probability is defined through three fundamental axioms.
- These axioms guarantee mathematical consistency and provide the basis for all probability theorems.

Axiom 1: Non-negativity

- **Statement:** For any event $E \in \mathcal{F}$,

$$P(E) \geq 0$$

- **Meaning:** Probabilities cannot be negative.
- **Example:** Rolling a 7 on a standard die: $P(\{7\}) = 0$, never negative.

Axiom 2: Normalization

- **Statement:** The probability of the entire sample space is 1:

$$P(\Omega) = 1$$

- **Meaning:** When an experiment is performed, some outcome in Ω must occur.
- **Example:** Rolling a die always gives some face $\{1, 2, 3, 4, 5, 6\}$, so the total probability is 1.

Axiom 3: Countable Additivity

- **Statement:** For any sequence of pairwise disjoint events $E_1, E_2, \dots \in \mathcal{F}$,

$$P\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} P(E_i)$$

- **Meaning:** If events cannot happen together, the probability of "any of them happening" is the sum of their probabilities.
- **Example:** On a die, $P(\{2\} \cup \{4\}) = P(\{2\}) + P(\{4\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Definition 1.1: Probability Space

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

- Ω: the sample space, the set of all possible outcomes.
- $\mathcal{F}$: a collection of events (subsets of Ω).
- P: the probability measure, a function assigning probabilities to events in $\mathcal{F}$.

Definition 1.2: Probability Function

A **probability function** is a map $P : \mathcal{F} \to \mathbb{R}$ satisfying:

- Non-negativity: $P(E) \geq 0$ for all $E \in \mathcal{F}$.
- Normalization: $P(\Omega) = 1$.
- Countable additivity: For any disjoint $E_1, E_2, \ldots,$

$$P\Big(\bigcup_i E_i\Big) = \sum_i P(E_i).$$

Remarks (Discrete Probability Spaces)

- In many cases we use **discrete** spaces: Ω is finite or countably infinite, and $\mathcal{F}$ is often all subsets of Ω.
- In such cases, P is uniquely determined by the probabilities of the simple (elementary) events in Ω.

Lemma 1.1: Two-Event Union

Statement. For any $E_1, E_2 \in \mathcal{F}$,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

Idea. When we sum $P(E_1)$ and $P(E_2)$, we are adding the probabilities of all outcomes in E_1 and all outcomes in E_2. Any outcome in the intersection $(E_1 \cap E_2)$ gets its probability counted twice. To correct this, we must subtract the probability of the intersection once.
Use. Corrects the double-counting of the intersection.

Lemma 1.2: Union Bound

Statement. For any finite or countable collection of events $E_1, E_2, \dots \in \mathcal{F}$,

$$P\Big(\bigcup_i E_i\Big) \;\le\; \sum_i P(E_i).$$

Use. Provides a simple upper bound on the probability of a union, even when events are not disjoint.

Example. For events $E_1 = \{\text{rain}\}$, $E_2 = \{\text{snow}\}$:

$$P(\text{rain or snow}) \;\le\; P(\text{rain}) + P(\text{snow}).$$

Lemma 1.3: Inclusion–Exclusion Principle

Statement. For any n events $E_1, \ldots, E_n \in \mathcal{F}$,

$$P\Big(\bigcup_{i=1}^{n} E_i\Big) = \sum_{i=1}^{n} P(E_i) - \sum_{1 \le i < j \le n} P(E_i \cap E_j) + \sum_{1 \le i < j < k \le n} P(E_i \cap E_j \cap E_k) - \cdots + (-1)^{n+1} P(E_1 \cap \cdots \cap E_n).$$

Use. Gives the exact probability of a union by alternating signs to correct for over-counting.

Definition 1.3: Independence of Two Events

Definition. Two events $E, F \in \mathcal{F}$ are independent if and only if

$$P(E \cap F) = P(E) \cdot P(F).$$

Intuition. The occurrence of E does not affect the probability of F, and vice versa. Knowing that one event happened gives no extra information about the other.

Example.

- Tossing two fair coins:
  $E = \{\text{first coin is heads}\}$, $F = \{\text{second coin is heads}\}$. Then $P(E) = P(F) = \frac{1}{2}$ and $P(E \cap F) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$. $\Rightarrow$ E and F are independent.

Definition 1.4: Mutual Independence

Definition. A collection of events $E_1, E_2, \ldots, E_n \in \mathcal{F}$ are mutually independent if for every subset $\{E_{i_1}, E_{i_2}, \ldots, E_{i_k}\}$ (with $1 \leq k \leq n$),

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i_1}) \cdot P(E_{i_2}) \cdots P(E_{i_k}).$$

Intuition. Not only are events independent in pairs, but every group of them behaves independently.

Example.

- Tossing three fair coins. Let $E_1 = \{\text{coin 1 is heads}\}$, $E_2 = \{\text{coin 2 is heads}\}$, $E_3 = \{\text{coin 3 is heads}\}$. Then for any subset (pairs or all three together), the independence condition holds.

Axioms of Probability: Example 1 - Rolling a Die

Rolling a Fair Six-Sided Die

Experiment: roll a fair die. Sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$, each outcome has probability $1/6$.

Check Axiom 1 (Non-negativity): For any $x \in \Omega$, $P(\{x\}) = 1/6 \geq 0$. Thus, for any event $E \subseteq \Omega$, $P(E) \geq 0$.

Axioms of Probability: Example 1 - Rolling a Die (Cont.)

Checking Axiom 2 (Normalization)

$$P(\Omega) = \sum_{x=1}^{6} P(\{x\}) = 6 \times \frac{1}{6} = 1.$$

Hence, the total probability is 1 (100%).

Axioms of Probability: Example 1 - Rolling a Die (Cont.)

Checking Axiom 3 (Additivity)

Let $E_1 = \{2, 4, 6\}$ (even), $E_2 = \{1, 3, 5\}$ (odd). They are disjoint: $E_1 \cap E_2 = \emptyset$, and

$$P(E_1) = 3/6 = 1/2, \quad P(E_2) = 3/6 = 1/2.$$

Then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) = 1/2 + 1/2 = 1 = P(\Omega).$$

Additivity holds.

Axioms of Probability: Example 2 - Drawing a Card

Drawing a Single Card from a Deck

Experiment: draw one card from a standard 52-card deck. Sample space: $\Omega = \{\text{all 52 cards}\}$, each card equally likely.

Define events:

$$A = \{\text{draw an Ace}\}, \quad B = \{\text{draw a Heart}\}.$$

$$P(A) = \frac{4}{52} = \frac{1}{13}, \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

Axioms of Probability: Example 2 - Drawing a Card (Cont.)

Intersection of A and B

Are A and B disjoint? No. The Ace of Hearts belongs to both events:

$$A \cap B = \{\text{Ace of Hearts}\}, \quad P(A \cap B) = \tfrac{1}{52}.$$

Axioms of Probability: Example 2 - Drawing a Card (Cont.)

Applying Addition Rule

Since A and B are not disjoint, use Lemma 1.1:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Substitute values:

$$P(A \cup B) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}.$$

Correctly accounts for the overlap.

Axioms of Probability: Exercise Questions

1. A bag contains 10 red, 8 blue, and 2 green marbles. If one marble is drawn at random:
    1. What is the probability of drawing a red marble?
    2. What is the probability of drawing a blue marble?
    3. What is the probability of drawing a green marble?

Axioms of Probability: Solutions to Exercise Questions

1. Total number of marbles $= 10(\text{red}) + 8(\text{blue}) + 2(\text{green}) = 20$ marbles.
    1. $P(\text{Red}) = \frac{10}{20} = \frac{1}{2} = 0.5.$
    2. $P(\text{Blue}) = \frac{8}{20} = \frac{2}{5} = 0.4.$
    3. $P(\text{Green}) = \frac{2}{20} = \frac{1}{10} = 0.1.$

Axioms of Probability: Exercise Questions (Cont.)

1. A bag contains 10 red, 8 blue, and 2 green marbles. If one marble is drawn at random:
   4. What is the probability of drawing a red or a blue marble?
   5. What is the probability of not drawing a green marble?
2. Let $P(A) = 0.6$, $P(B) = 0.4$, and $P(A \cap B) = 0.2$. Using the appropriate rule, find $P(A \cup B)$.
3. If the probability of event E occurring is $P(E) = 0.7$, what is the probability of its complement, $P(E^c)$?

Axioms of Probability: Solutions to Exercise Questions (Cont.)

1. Total number of marbles $= 20$.
   4. By Axiom 3 (disjoint events):

   $$P(\text{Red or Blue}) = P(\text{Red}) + P(\text{Blue}) = 0.5 + 0.4 = 0.9$$

   5. By Complement Rule:

   $$P(\text{Not Green}) = 1 - P(\text{Green}) = 1 - 0.1 = 0.9$$

Axioms of Probability: Solutions to Exercise Questions (Cont.)

1. Given $P(A) = 0.6$, $P(B) = 0.4$, $P(A \cap B) = 0.2$. Use Addition Rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B) = 0.6 + 0.4 - 0.2 = 0.8$$

2. Given $P(E) = 0.7$. By Complement Rule:

$$P(E^c) = 1 - P(E) = 1 - 0.7 = 0.3$$

Axioms of Probability: Exercise Questions 2

4. A coin is flipped twice.
   1. List the sample space Ω.
   2. What is the probability of getting exactly one Head?
   3. What is the probability of getting at least one Tail?
   4. Let $A = \{\text{first flip is Head}\}$, $B = \{\text{second flip is Head}\}$. Are A and B independent?

Axioms of Probability: Solutions to Exercise 2

4. Two flips of a fair coin.
   1. $\Omega = \{HH, HT, TH, TT\}$.
   2. Exactly one Head: outcomes $\{HT, TH\}$.

      $$P(\text{exactly 1 H}) = \frac{2}{4} = 0.5.$$

   3. At least one Tail: outcomes $\{HT, TH, TT\}$.

      $$P(\text{at least 1 T}) = \frac{3}{4} = 0.75.$$

   4. $P(A) = \frac{2}{4} = 0.5$, $P(B) = \frac{2}{4} = 0.5$, $P(A \cap B) = P(\{HH\}) = \frac{1}{4}$. Since $P(A \cap B) = P(A) \cdot P(B) = 0.25$, events A and B are **independent**.