San José State University

Math 161A: Applied Probability & Statistics

Lecture 1: Probability Basics

Prof. Guangliang Chen

Section 2.1 Sample Spaces and Events

Section 2.2 Axioms, Interpretations, and Properties of Probability

Introduction

To model a random phenomenon (such as flipping a coin, rolling a die), we need to specify the following components:

- **Sample space**
- **Events**
- **Probability**

Collectively, they define a **probability space**.

We'll go through the above concepts one by one.

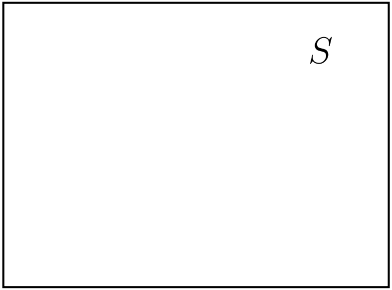
A little bit set terminology and notation:

- A **set** is an unordered collection of objects: $A = \{1, 2, 3\} = \{3, 1, 2\}$, $B = (1, 3) = \{x \mid 1 < x < 3\}$, $C = [0, \infty) = \{x \mid x \geq 0\}$
- **Notation**: To indicate an object is in or not in the set: $1 \in A$, $4 \notin A$
- A **subset** is a subcollection of the objects: $\{1\} \subset A \subset C$
- The **size** (cardinality) of a set is the number of objects in it: $|A| = 3$
- **Special sets**: $\emptyset$ (empty set), $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$, $\mathbb{R} = (-\infty, \infty)$

Sample space

Def 0.1. The set of all possible outcomes of a random phenomenon is called the sample space for that experiment.

- We denote a sample space by S (some books use Ω instead).
- A sample space is typically represented by a rectangle, and the outcomes of the sample space are denoted by points within the rectangle.



Example 0.1. Write down the sample space of each of the following experiments:

- Tossing a coin: $S = \{H,\ T\}$.
- Rolling a die: $S = \{1, 2, 3, 4, 5, 6\}$.
- Drawing a card from an ordinary deck of 52: $S = \{\text{All 52 cards}\}$.

Example 0.2. Write down the sample space of each of the following experiments:

- Throw a coin twice:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

- Roll two dice:

$$\begin{aligned} S &= \{(1,1), \ldots, (1,6), (2,1), (2,2), \ldots, (6,6)\} \longleftarrow \text{by enumeration} \\ &= \{(i,j) \mid 1 \le i \le 6, 1 \le j \le 6\} \quad \longleftarrow \text{by formula} \end{aligned}$$

- Throw a coin repeatedly until a head first appears:

$$S = \{H,\ TH,\ TTH,\ TTTH, \ldots\}$$

The sample spaces in the previous example are **discrete** sets, which are countable. That is, the number of objects in the set must be

- finite (e.g., $\{1, 2, \ldots, 6\}$), or
- countably infinite: There is a 1-to-1 correspondence to the set of natural numbers, $\mathbb{N} = \{1, 2, 3, \ldots\}$.

For example, the set of all integers $\mathbb{Z}$ is countable:

1	2	3	4	5	6	7	8	9	10	$\cdots$
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\cdots$
0	1	-1	2	-2	3	-3	4	-4	5	$\cdots$

In contrast, the set of all real numbers $\mathbb{R}$ is uncountable.

In the following example, the sample spaces are **continuous intervals** (which are uncountable).

Example 0.3.

- Life time of a new light bulb. The sample space is an interval $S = (0, \infty)$.
- Waiting time (in minutes) to talk to a customer service representative: $S = (0, \infty)$
- Throw a dart to a unit disk and measure its distance to center: $S = [0, 1]$

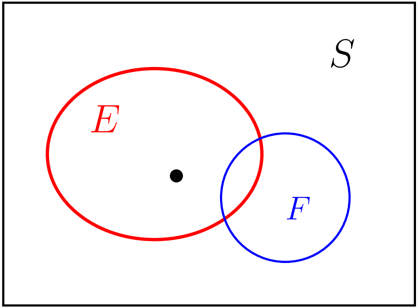
Events

Consider the following probability questions about events:

- (**Toss two fair dice**) What is the probability of getting a sum of 8?
- (**Toss two fair dice**) What is the probability of getting two even numbers?
- (**Toss two fair dice**) What is the probability of getting two identical numbers?
- (**Toss a fair coin repeatedly until a head first appears**) What is the probability that at most 3 tails are observed?

Def 0.2. Mathematically, an event is just a subset E of outcomes in the sample space S.

- In particular, $S, \emptyset$ are events.
- We say an event E **occurred** if the actual outcome of the experiment lies in E.
- It is called a **simple** event if it contains only one outcome. Otherwise, it is called a **compound** event.



(Event E occurred, while F did not)

Example 0.4 (Roll a single die). The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. The following are events:

- $A = \{1\} = \{\text{The smallest number}\}$ $\longleftarrow$ simple event
- $B = \{6\} = \{\text{The largest number}\}$ $\longleftarrow$ simple event
- $C = \{2, 4, 6\} = \{\text{An even number}\}$ $\longleftarrow$ compound event
- $D = \{1, 3, 5\} = \{\text{An odd number}\}$ $\longleftarrow$ compound event

If an outcome of 1 was observed when performing the experiment, then which of the above events occurred (and which of them did not occur)?

Example 0.5 (Throw two dice)**.** The sample space is

$$S = \{(1,1),(1,2),\ldots,(6,6)\} = \{(i,j) \mid 1 \le i,j \le 6\}.$$

The following are events:

$$\begin{aligned} A &= \{\text{Sum equals } 6\} \\ &= \{(1,5),(2,4),(3,3),(4,2),(5,1)\} \\ B &= \{\text{Two identical numbers}\} \\ &= \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\} \\ C &= \{\text{Two even numbers}\} \\ &= \{(2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2),(6,4),(6,6)\}. \end{aligned}$$

Example 0.6. Consider the experiment where you repeatedly toss a coin until you see the first head. The following is an event:

$$\begin{aligned} E &= \{\text{At most 4 tails are obtained}\} \\ &= \{H, TH, TTH, TTTH, TTTTH\} \end{aligned}$$

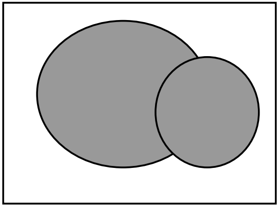
Event operations

Def 0.3. Let $A, B \subseteq S$ be two events. We define

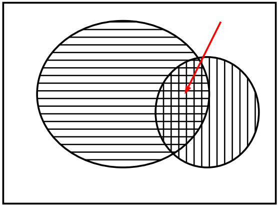
- **Complement** A^c: set of all outcomes not in A
- **Union** $A \cup B$: set of all outcomes in A or B (or both)
- **Intersection** $A \cap B$: set of all outcomes in both A and B
- **Difference** $A - B = A \cap B^c$: set of all outcomes in A and not in B

They can be represented by the so-called Venn diagrams (see next slide).

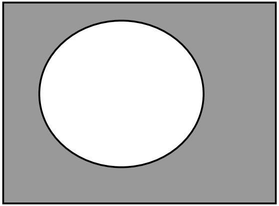
$A \cup B$



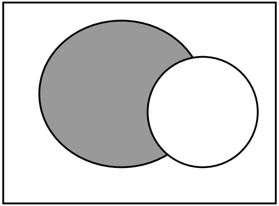
$A \cap B$



A^c



$A - B$



Example 0.7 (Throw two dice)**.** Let

- $A = \{\text{Sum equals 6}\}$
- $B = \{\text{Two identical numbers}\}$
- $C = \{\text{Two even numbers}\}$

Compute $|C|, A \cap B, A \cup B, B^c, A - C$

Two useful set laws:

- **Distributive law**:

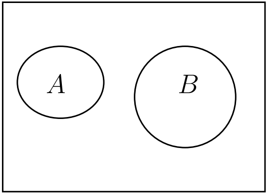
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

- **De Morgan's Laws**

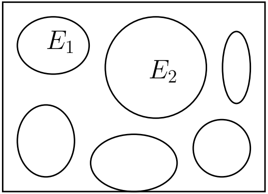
$$(A \cup B)^c = A^c \cap B^c,$$
$$(A \cap B)^c = A^c \cup B^c$$

Disjoint events

Def 0.4. Two events A, B are said to be **disjoint**, or **mutually exclusive**, if their intersection is empty: $A \cap B = \emptyset$.



A sequence of events $E_1, E_2, \ldots$ are said to be **pairwise disjoint**, or **mutually exclusive**, if $E_i \cap E_j = \emptyset$ for all $i \neq j$.



Example 0.8 (Toss two fair dice)**.** Are the following two events disjoint?

- $A = \{\text{Sum equals 7}\}$.
- $B = \{\text{Two identical numbers}\}$.

Probability

Intuitively, probability is a number $P(E)$ describing the chance of an event E occurring.

The larger the probability, the more likely for the event to occur.

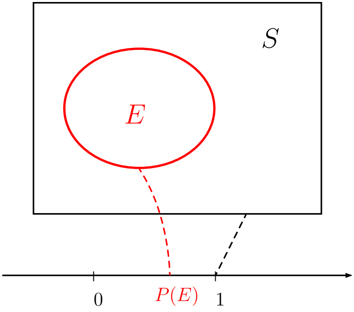
And it needs to satisfy certain conditions in order to be valid/meaningful.

Below is the formal definition of probability.

Def 0.5. Probability is a function defined on the space of events that satisfies the following Kolmogorov Axioms of Probability:

1. $P(E) \geq 0$ for any $E \subseteq S$.
2. $P(S) = 1$.
3. For any infinite sequence of pairwise disjoint events $E_1, E_2, \ldots$,

$$P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$



Theorem 0.1. The Axioms of Probability imply the following are true:[1]

- $P(\emptyset) = 0$.
- If $E_1, E_2, \ldots, E_k \subset S$ are pairwise disjoint, then

$$P\left(\cup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} P(E_i)$$

- $P(E^c) = 1 - P(E)$, from which we obtain that $P(E) \leq 1$.
- $P(B - A) = P(B) - P(B \cap A)$: If $A \subseteq B$, then it simplifies to $P(B - A) = P(B) - P(A)$.

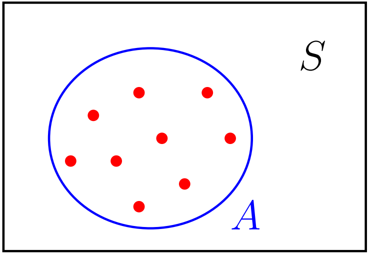
[1] This is why we did not include these properties in the definition of probability.

Countable sample spaces

The following property implies that, to define the probability function over a **countable** sample space, it suffices to specify only the probabilities of simple events.

Theorem 0.2. If the sample space S is countable, then for any event $A \subseteq S$,

$$P(A) = \sum_{a \in A} P(\{a\}).$$



Example 0.9 (Fair coin model)**.** Let $S = \{H, T\}$ with $P(\{H\}) = P(\{T\}) = .5$.

Example 0.10 (Biased coin model)**.** Let $S = \{H, T\}$ with $P(\{H\}) = .55, P(\{T\}) = .45$.

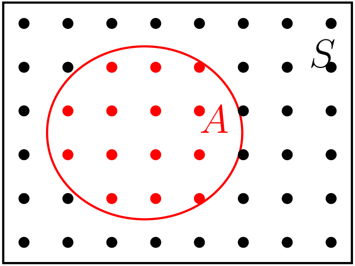
Example 0.11 (Fair die model)**.** Let $S = \{1, 2, \ldots, 6\}$ with $P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = \frac{1}{6}$. The probability of getting an even number is

$$P(\{\text{An even number}\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

Finite sample spaces with equally likely outcomes

Theorem 0.3. If $|S| < \infty$ (i.e., S is a finite set) and all the outcomes are equally likely to occur, then for any event $A \subseteq S$,

$$P(A) = \frac{|A|}{|S|} = \frac{\#\text{ outcomes in } A}{\#\text{ outcomes in } S}.$$



Proof. By the preceding theorem,

$$P(A) = \sum_{a \in A} P(\{a\}) = \sum_{a \in A} \frac{1}{|S|} = \frac{1}{|S|} \cdot |A| = \frac{|A|}{|S|}.$$

Joke: What is a probability to meet a dinosaur?

A: What is a probability to meet a dinosaur on the street?

B: Well, 50x50!

A: How, why???

B: You either meet it or not!

So, i met it!

Example 0.12 (Throw a fair die). Find the following probabilities:

$P(\{\text{An even number})\}) =$

$P(\{\text{At least } 5\}) =$

$P(\{\text{Not a } 3\}) =$

Example 0.13 (Throw two fair dice)**.** Find the following probabilities:

$P(\{\text{Sum equals } 6\}) =$

$P(\{\text{Two identical numbers}\}) =$

$P(\{\text{Both even}\}) =$

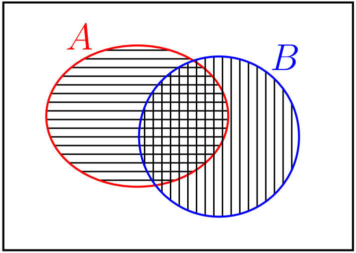
Example 0.14 (Toss a fair coin 5 times). What is the probability of getting at least one head?

Inclusive-exclusive formula (2 events)

Theorem 0.4. For any $A, B \subseteq S$,

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

In particular, if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.



Proof. By additivity for mutually exclusive events,

$$\begin{aligned} P(A \cup B) &= P(A - B) + P(A \cap B) + P(B - A) \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(B \cap A) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Example 0.15. In a large discrete math class, 55% of the students have a major in math, and 35% of the class have a major in CS. Among the two groups of students combined, 5% of them are dual majors (in math and CS). What is the probability that a randomly selected student from the class majors in

(a) at least one of math and CS,

(b) one and only one of math and CS,

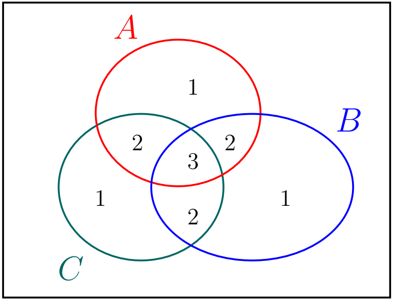
(c) neither math nor CS?

(blank slide)

Inclusive-exclusive formula (3 events)

Theorem 0.5. For any three events $A, B, C \subseteq S$, we have

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



Example 0.16 (Select an integer from $\{1, \ldots, 100\}$ at random)**.** What is the probability that it is divisible by at least one of the three prime numbers 2, 3, 5? (Answer: .74)

Summary

We first introduced the concept of a **probability space** associated to a random phenomenon, which consists of the following:

- **Sample space** S (set of all possible outcomes)
- **Events** $E \subseteq S$ (subsets of outcomes, often with a common trait)
- **Probability** (chance that an event occurs): a mapping from events to numbers, $P : E \subseteq S \mapsto P(E) \in \mathbb{R}$, that satisfies the three Axioms of Probability

  1. $P(E) \geq 0$ for any $E \subseteq S$.

2. $P(S) = 1$.

3. If an infinite sequence of events $E_1, E_2, \ldots$ are pairwise disjoint, then

$$P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

The Axioms imply more properties for the probability function:

- $P(\emptyset) = 0$.
- If $E_1, E_2, \ldots, E_k$ are pairwise disjoint, then

$$P\left(\cup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} P(E_i)$$

- $P(E^c) = 1 - P(E)$, from which we obtain that $P(E) \leq 1$.
- If $A \subseteq B$, then $P(A) \leq P(B) \longleftarrow$ This is due to the property $P(B - A) = P(B) - P(A \cap B)$
- Inclusive-exclusive formula for any two events $A, B \subseteq S$:

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

- Inclusive-exclusive formula for any three events $A, B, C \subseteq S$:

$$\begin{aligned} P(A \cup B \cup C) = & \, P(A) + P(B) + P(C) \\ & - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ & + P(A \cap B \cap C). \end{aligned}$$

Lastly, there are two special settings:

- If the sample space S is countable, then for any event $A \subseteq S$,

$$P(A) = \sum_{a \in A} P(\{a\}).$$

- If the sample space is finite and all the outcomes are equally likely to occur, then for any event $A \subseteq S$,

$$P(A) = \frac{|A|}{|S|} = \frac{\#\text{ outcomes in } A}{\#\text{ outcomes in } S}.$$