Lecture 2

Probability

Grace Tompkins

Last modified — 06 May 2026

Last Class:

- Set Theory

Today's Learning Outcomes:

By the end of this lecture, students are anticipated to be able to

- Define probability as a mathematical object
- Use set-theory to compute probabilities and prove probabilistic properties

1 Probability

The basics of probability

- A probability is the formal treatment of randomness
- The key to this are models
- Defining a probability requires the following:
  - Random experiment
  - Sample space
  - Event
  - Rules to combine events (set operations)

Experiments

📖 DEFINITION

Experiment: an action undertaken to make a discovery, test a hypothesis, or confirm a known fact.

Example: Release your pen from 4.9 meters above the ground

Predicted outcomes:

- The pen will fall to the ground.
- It will take about 1 sec to reach the ground.

Actual observations:

- The pen hit the ground
- Unsure if it took exactly one second….

Uncertainty

The outcome of some experiments cannot be determined beforehand. They are uncertain.

- If I roll a die, which number will show?
- How many times will the R4 fly pass me with the "SORRY, BUS FULL" sign on?
- If I walk to school without my rain jacket, will it rain?
- If I randomly choose one sweater and one pair of pants from my closet, will they match?

Probability theory

Even though die rolls are random, patterns emerge when we repeat the experiment many times. We can study these using probability theory.

📖 DEFINITION

Probability Theory: The study of uncertainty, random phenomena, and patterns via mathematical models

- Based on a set of **axioms** (statements or propositions accepted to hold true) and **theorems** (propositions which are established to hold true using sound logical reasoning)
- This is what we will study in STAT 302!

Sample Space

In order to think about probabilities, we need to consider the possible outcomes of an experiment.

📖 DEFINITION

Sample space: the set of all possible outcomes of a random experiment.

Denoted by Ω (or S in the textbook)

We also can consider a generic outcome, also called sample point, by ω (i.e. $\omega \in \Omega$). Note that the textbook uses $s \in S$.

Sample Space

💡 EXAMPLE

- Roll a die:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Draw a card from a poker deck:

$$\Omega = \{A\clubsuit, 2\clubsuit, \ldots, K\heartsuit\}$$

- Wind speed at YVR (km/h):

$$\Omega = \{\omega : 0 \leq \omega < 200\} \quad (\text{kmph}) \quad \subset \mathbb{R}$$

- Wait time for R4 at UBC (min):

$$\Omega = \{\omega : 0 \leq \omega < \infty\} \quad \subset \mathbb{R}$$

Events

📖 DEFINITION

Event: a subset of the sample space Ω

Notation: We commonly use upper case letters $(A, B, C, \ldots)$ for events.

Events are sets!

- $\omega \in A$ means "ω is an element of A".
- $C \subset D$ means "C is a subset of D".

Examples

- Events are often formed by outcomes sharing some property.
- It's a good idea to practice listing explicitly the sample points of events described with words.

💡 **EXAMPLE**

- Roll a dice: $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - $A =$ "roll an even number" $= \{2, 4, 6\}$
  - $B =$ "roll a 3 or less" $= \{1, 2, 3\}$
  - $F =$ "roll an even number no higher than 3" $= \{2\}$
- Bus wait time: $H =$ "wait is less than half an hour" $= \{w : 0 \leq w < 30\}$
- Max-wind-speed: $G =$ "wind is over 80 km/hour" $= \{w : 80 < w < \infty\}$

$\{w : w > 80\}$

Examples

💡 EXAMPLE

Consider flipping two coins in a row:

a. Write out all elements in the sample space for this experiment.

b. How many elements are in Ω (denoted $|\Omega|$)?

c. Let $A = \{\text{first roll is heads}\}$. What is $|A|$?

a) $\Omega = \{HH, HT, TT, TH\}$

H= head
T= tail

b) $|\Omega| = 4$

c) $A = \{HH, HT\}$

$|A| = 2$

Exercises

✏️ **EXERCISE: WORK OR FAIL**

A system has **3 components**, which can either **work** or **fail**.

The experiment consists of observing the status (W/F) of the 3 components.

a. Describe the sample space for this experiment without listing all of the possible outcomes.

b. What is $|\Omega|$?

c. Let $A = \{\text{component 3 fails}\}$. What is $|A|$?

a) The possible outcomes are sequences (x_1, x_2, x_3) where each $x_i = \begin{cases} W & \text{if working} \\ F & \text{if failing} \end{cases}$ for $i = 1, 2, 3$.

$\Omega = \{ WWW, WWF, WFW, \ldots, FFF \}$

Exercises

b) $|\Omega| = 8$

(x_1, x_2, x_3) $\rightarrow$ $2 \cdot 2 \cdot 2 = 2^3 = 8$

W|F W|F W|F

c) $A = \{$component 3 fails$\}$. What is $|A|$?

(x_1, x_2, F) $\rightarrow$ $2 \cdot 2 = 2^2 = 4$

W|F W|F

$A = \{WWF, WFF, FWF, FFF\}$

2 Properties of Probability

The Probability of an Event

- Even though random outcomes cannot be predicted, in some cases we have an idea about the chance that an outcome occurs.
  - If you toss a fair coin, the chance of observing a head is the same as that of observing a tail.
  - If you buy a lottery ticket, the chance of winning is very small.
- A probability function $\mathbb{P}$ quantifies these chances.
- Probability functions are computed on events $A \in \mathcal{B}$. We calculate $\mathbb{P}(A)$. Mathematically / formally, we have:

$$\mathbb{P} : \mathcal{B} \to [0, 1]$$

where $\mathcal{B}$ is a collection of possible events.

Probability Axioms

Let Ω be a sample space and $\mathcal{B}$ be a collection of events (i.e. subsets of Ω).

📖 DEFINITION

Probability measure (or probability function): Any function $\mathbb{P}$ with domain $\mathcal{B}$ that satisfies:

1. **Axiom 1:** $\mathbb{P}(\Omega) = 1$;
2. **Axiom 2:** $\mathbb{P}(A) \geq 0$ for any $A \in \mathcal{B}$; (non-negative)
3. **Axiom 3: Additivity** If $\{A_n\}_{n\geq 1}$ is a sequence of **disjoint** events, then

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Note: $\{A_n\}_{n\geq 1}$ is a sequence of **disjoint** events when $A_i \cap A_j = \varnothing$ if $i \neq j$

We also assume $\mathbb{P}(\varnothing) = 0$.

Probability Axiom: Additivity

In a nutshell, additivity says that so long as events A, B, C are disjoint, the probability of the union $\mathbb{P}(A \cup B \cup B) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$.

✏️ **EXERCISE: CANDIES**

Suppose you encounter a bowl of candies. There are 13 red candies, 20 blue candies, and 17 orange candies. What is the probability that you randomly select a red or blue candy from the bowl?

$$P(\text{Red or Blue}) = P(\text{Red} \cup \text{Blue})$$

$$= P(\text{Red}) + P(\text{Blue})$$

$$= \frac{13}{50} + \frac{20}{50}$$

$$= \frac{33}{50}$$

→ disjoint b/c a candy cannot be red and blue at the same time.
(additivity)

Properties of the Probability Function

Let A and B denote arbitrary events, where Ω is the sample space.

- **Probability of the complement:** $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

- **Monotonicity:** $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

  subset

- **Probability of the union:** $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

(not necessarily disjoint)

- **Boole's inequality:** $\mathbb{P}(\bigcup_{i=1}^{m} A_i) \leq \sum_{i=1}^{m} \mathbb{P}(A_i)$

Properties of the Probability Function

✏️ **EXERCISE: PROOF OF PROBABILITY OF THE COMPLEMENT**

Prove the probability of the complement: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

To do this, show that if $\mathbb{P}$ satisfies Axioms 1, 2, and 3, and A is an arbitrary event, then necessarily $\mathbb{P}(A^c) \; = \; 1 - \mathbb{P}(A)$.

Hint: What is $A \cup A^c$? Use Axiom 1: $P(\Omega) = 1$

Recall: $A \cup A^c = \Omega$

$P(\Omega) = 1$ Axiom 1

$P(A \cup A^c) = 1$ $A \cup A^c = \Omega$

$P(A) + P(A^c) = 1$ since disjoint

$\hookrightarrow$ $P(A) = 1 - P(A^c)$ rearrange

Properties of the Probability Function

Properties of the Probability Function

💡 **EXAMPLE**

Prove monotonicity: $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

1. first show $B = (B \cap A) \cup (B \cap A^c)$

$$B = B \cap \Omega$$
$$= B \cap (A \cup A^c)$$
$$= (B \cap A) \cup (B \cap A^c)$$ dist. property

Properties of the Probability Function

Step 2: show $A \subset B \Rightarrow A \cap B = A$

① let $\omega \in A \cap B$

$\Rightarrow \omega \in A$ and $\omega \in B$

$\Rightarrow \omega \in A$

$\therefore \omega \in A \cap B \Rightarrow \omega \in A$. This means $A \cap B \subseteq A$

② let $\omega \in A$

$\Rightarrow \omega \in B$ since $A \subset B$. (subset def: if $\omega \in A \Rightarrow \omega \in B$ when $A \subset B$)

$\Rightarrow \omega \in A$ and $\omega \in B$

$\Rightarrow \omega \in A \cap B$

This means $A \subseteq A \cap B$ $\qquad \therefore A = A \cap B$

Properties of the Probability Function

We've shown

- $B = (B \cap A) \cup (B \cap A^c)$
- $A \subset B \Rightarrow A \cap B = A$

Therefore, as $A \subset B$:

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

$$P(B) = P[A \cup (B \cap A^c)]$$
$$= P(A) + P(B \cap A^c]$$

since A disjoint from. $B \cap A^c$

Properties of the Probability Function

$$P(B) = P(A) + \underbrace{P(B \cap A^c]}_{\geq 0 \text{ by Axiom}}$$

$$\therefore P(B) \geq P(A)$$

Properties of the Probability Function

✏️ **EXERCISE: PROOF OF PROBABILITY OF THE UNION**

Prove the probability of the union: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Hint: First prove that $A \cup B = A \cup (B \cap A^c)$

1

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

$= (A \cup B) \cap (A \cup A^c)$ dist. prop

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

$= A \cup B$

Axiom 3: $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ if A_1, A_2 disjoint

Properties of the Probability Function

2

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

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

$= P(A) + P(B \cap A^c)$ since A, $B \cap A^c$ are disjoint.

3

We can expand $P(B)$

$P(B) = P(B \cap \Omega)$

$= P[B \cap (A \cup A^c)]$

$= P[(B \cap A) \cup (B \cap A^c)]$ dist. law

$P(B) = P(B \cap A) + P(B \cap A^c)$ disjoint

$\Rightarrow P(B \cap A^c) = P(B) - P(B \cap A)$

Properties of the Probability Function

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

sub ③ into ②

$$= P(A) + \underline{P(B \cap A^c)}$$

$$= P(A) + \underline{P(B) - P(A \cap B)}$$

✓

Properties of the Probability Function

💡 **EXAMPLE**

Prove Boole's inequality: $\mathbb{P}(\bigcup_{i=1}^{m} A_i) \leq \sum_{i=1}^{m} \mathbb{P}(A_i)$

Proof by induction

$\boxed{n=0}$ We have empty set A. This case is trivial

$$P(\{\emptyset\}) \leq \sum P(\{0\})$$

$$0 \leq 0 \quad ✓$$

$\boxed{n=1}$ $P(\bigcup_{i=1}^{1} A_i) \leq \sum_{i=1}^{1} P(A_i) \Rightarrow P(A_1) \leq P(A_1)$ ✓

Properties of the Probability Function

$\boxed{n=2}$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - \underbrace{P(A_1 \cap A_2)}_{\geq 0}$$

$$\leq P(A_1) + P(A_2)$$

$\therefore$ holds for $n=2$

Now: assume this holds for $m \leq n$

Properties of the Probability Function

$$P\left[\bigcup_{i=1}^{m+1} A_i\right] = P\left[\left(\bigcup_{i=1}^{m} A_i\right) \cup A_{m+1}\right]$$ assoc property

$$\leq P\left(\bigcup_{i=1}^{m} A_i\right) + P(A_{m+1})$$ from n=2 case

$$\leq \sum_{i=1}^{m} P(A_i) + P(A_{m+1})$$ induction hyp.

$$= \sum_{i=1}^{m+1} P(A_i)$$

//

To Do:

- No reading for next class: most of Section 1.4 we are de-emphasizing.
- Start working on your Assignment due May 20th, 11:59pm

Next Class: More Applied Problems + Permutations and Combinations!

$$\bigcup_{i=1}^{3} A_i = A_1 \cup A_2 \cup A_3$$
$$= \left(\bigcup_{i=1}^{2} A_i\right) \cup A_3$$