CHAPTER 12

Introduction to Probability

Goals

- Define random phenomena, outcomes, and sample spaces
- Identify and describe events
- Apply the three axioms of probability
- Use the addition rule (disjoint and general)
- Apply the complement rule
- Use Venn diagrams to visualize probabilities
- Distinguish finite and continuous probability models
- Define and interpret random variables

12.1. From Tables to Probability

In Chapter 6, we computed proportions from contingency tables. In this chapter, we formalise those ideas using the language of **probability**. The key insight is simple: proportions from contingency tables *are* probabilities.

Example 12.1 *From Chapter 6 to Chapter 12*
Recall the flu vaccine study (Example 6.7):

Group	**Developed Flu**	**No Flu**	**Total**
Vaccinated	26	3874	3900
Control	70	3830	3900
Total	96	7704	7800

In Chapter 6, you computed:

- Marginal proportion of flu: $\frac{96}{7800} = 0.0123$
- Conditional proportion of flu given vaccinated: $\frac{26}{3900} = 0.0067$

In Chapter 12 notation:

- $P(\text{Flu}) = 0.0123$ *(probability of flu in the whole study)*
- $P(\text{Flu} \mid \text{Vaccinated}) = 0.0067$ *(probability of flu given vaccinated)*

Key point: Probability theory gives us **general rules** that work in any situation, not just contingency tables. We need these rules for more complex situations: outcomes that happen in stages, combinations from multiple sources, and continuous measurements.

12.2. Foundations of Probability

Before we can state the rules of probability, we need to set up some vocabulary. This section introduces the building blocks: what counts as a random process, how we describe all possible results, and how we talk about collections of results we care about.

12.2.1 Random Phenomena and Outcomes

Definition 12.2 *Random Phenomenon and Outcome*
A **random phenomenon** is any process whose result is uncertain before it occurs.

An **outcome** is a single possible result of a random phenomenon.

Random phenomena include situations where:

- The outcome is not yet determined (flipping a coin, rolling a die, drawing a card)
- The outcome is determined but unknown to us (a patient's disease status before testing, tomorrow's weather)

Chapter 6 Connection
When you randomly selected a person from a contingency table, you were observing a random phenomenon. The uncertainty about their group membership is what made proportions meaningful.

12.2.2 Sample Space

Definition 12.3 *Sample Space*
The **sample space** is the set of **all possible outcomes** of a random phenomenon. We denote it by S.

We write sets using curly braces: $\{\ \}$. The braces hold all possible outcomes without regard to order or duplicates.

Quick examples:

- Flip a coin: $S = \{H, T\}$
- Roll a die: $S = \{1, 2, 3, 4, 5, 6\}$
- Pick a person from the flu study: $S = \{\text{all 7800 participants}\}$

Example 12.4 *Writing Sample Spaces*
Write the sample space for each random phenomenon.

1. Flipping a coin

2. Rolling a six-sided die

3. Flipping a coin twice

12.2.3 Events

Now that we can describe a full sample space, the next step is to focus on the parts of it we actually care about. We call these parts **events**.

> **Definition 12.5** *Event*
> An **event** is a set of outcomes from a sample space.

We rarely care about every outcome individually. Instead, we ask about **events**:

- "At least one head" (when flipping coins)
- "Rolling an even number" (when rolling a die)
- "The person has the flu" (when sampling from the flu study)

> **Chapter 6 Connection**
> In contingency tables, each row or column defined an event. "Vaccinated" was an event; "Developed Flu" was an event.

Exercise 12.6 *Identifying Events*
In every experiment, list the outcomes belonging to every event.

1. Flip a coin. Event A = "Get heads"

2. Roll a die. Event B = "Roll an even number"

3. Draw a card. Event C = "Draw a face card"

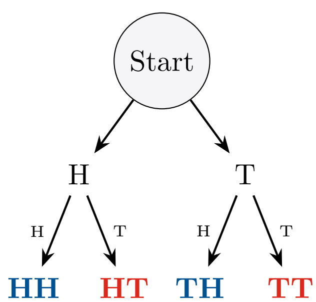
4. Flip two coins. Event D = "At least one tail"

12.2.4 Tree Diagrams

When a random process has multiple steps, listing all outcomes can get tricky. A tree diagram is a useful tool for keeping track.

> **Key point:** A **tree diagram** helps visualise all possible outcomes when a random process has multiple steps.

Example 12.7 *Flipping Two Coins*
When flipping a coin twice, the tree diagram shows:



Sample space: $S = \{HH, HT, TH, TT\}$: 4 possible outcomes.

Exercise 12.8 *Writing a Sample Space*
Flip a fair coin three times.

1. List all possible outcomes (sequences of flips)

2. Let $W =$ "At least two heads." List the outcomes in W

12.3. Rules of Probability

We now have the vocabulary: sample spaces, events, unions, intersections, and complements. The next question is: how do we actually *compute* probabilities? This section introduces the formal rules that let us assign numbers to events and combine them.

12.3.1 What Is Probability?

Definition 12.9 *Probability*
The **probability** of an event is a number between 0 and 1 that measures how likely the event is to occur.

There are three paradigms of probability:

1. **Classical:** Count equally likely outcomes.
Example: Fair die: each face has probability $\frac{1}{6}$.
2. **Empirical:** Use long-run relative frequencies.
Example: Track 1000 coin flips; proportion of heads ≈ 0.5.
3. **Subjective:** Based on personal judgment.
Example: "I think there's a 70% chance it rains tomorrow."

In our course, we will only use the **empirical** and **classical** paradigms.

Chapter 6 Connection
When you computed proportions from contingency tables, you were using the **empirical** approach: relative frequencies from observed data.

Convention
When no information suggests otherwise, assume equally likely outcomes.

12.3.2 Classical Probability

Key point: When outcomes are equally likely:

$$P(\text{event}) = \frac{\text{number of outcomes in event}}{\text{total number of outcomes}}$$

Example 12.10 *Classical Probability*
Roll a fair six-sided die. Calculate:

1. P(rolling a 4): ______________
2. P(rolling an even number): ______________
3. P(rolling greater than 4): ______________

Not every random process has equally likely outcomes. When one outcome is more likely than others, we need to figure out the individual probabilities before we can compute anything else.

Exercise 12.11 *When Outcomes Are Not Equally Likely*
A novelty die is weighted so that 6 appears twice as often as any other number. All other faces are equally likely.

1. Find $P(6)$

2. Find P(even number)

3. Find P(not a 6)

12.3.3 The Three Axioms of Probability

Whether we use the classical approach or the empirical approach, every valid probability assignment must follow the same three rules. These are the axioms of probability.

Definition 12.12 *Axioms of Probability*

A1: For any event A: $0 \leq P(A) \leq 1$.

A2: The probability of the entire sample space is 1: $P(S) = 1$.

A3: If A and B are **mutually exclusive** (cannot both occur), then

$$P(A \text{ or } B) = P(A) + P(B)$$

Exercise 12.13 *Applying the Axioms*
A lottery draws a three-digit number from 000 to 999. Each is equally likely.

- Total outcomes: __________
- Probability of each outcome: _______________

Find:

1. P(all three digits equal)

2. P(number < 500)

3. P(number ends in 7)

12.3.4 Combining Events: "Or" and "And"

Most interesting probability questions involve more than one event at a time. For example, we might want to know: did a person get the flu *and* were they vaccinated? Or did they get the flu *or* were they vaccinated? To handle questions like these, we need a way to combine events.

Definition 12.14 *Union and Intersection*

- **Union** "A or B": All outcomes in A, in B, or in both. Notation: $A \cup B$.
- **Intersection** "A and B": Only outcomes in *both* A and B. Notation: $A \cap B$.

Common Mistake: In mathematics, "or" is **inclusive**: it means at least one occurs, possibly both. This differs from everyday English where "or" often means "one or the other, but not both."

Example 12.15 *Union and Intersection with Cards*
Draw one card from a standard 52-card deck. Let A = "Draw a heart" (13 outcomes) and B = "Draw a face card (J, Q, K)" (12 outcomes).

Calculate the number of outcomes in each event:

1. $A \cap B$ ("heart **and** face card")

2. $A \cup B$ ("heart **or** face card")

Exercise 12.16 *Combining Events*
Pavlov has two dogs. Each dog is male (M) or female (F), all combinations equally likely.

1. Sample space S

2. A_1: "First dog is male"

3. A_2: "Second dog is male"

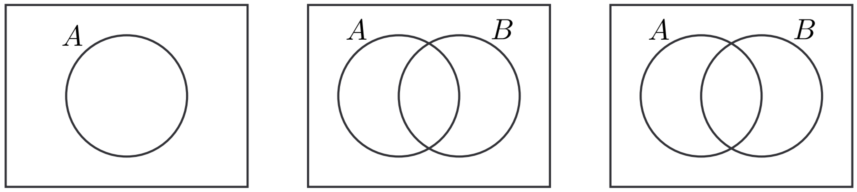
4. "At least one male" (union of A_1 and A_2)

5. "Both male" (intersection of A_1 and A_2)

12.3.5 Venn Diagrams

When we combine events, it helps to have a picture. Venn diagrams give us a way to see unions, intersections, and complements at a glance.

Definition 12.17 *Venn Diagram*
A **Venn diagram** represents events as regions inside a rectangle (the sample space). Overlapping regions show intersections.



Event A | A and B | A or B

Exercise 12.18 *Working with Venn Diagrams*
In a class of 30 students, 12 like apples (A), 7 like bananas (B), and 3 like both.

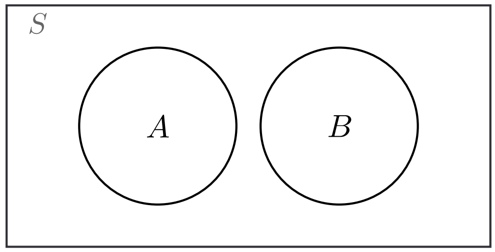
1. How many like only apples?

2. How many like only bananas?

3. How many like neither?

Draw a Venn diagram illustrating this situation:

Definition 12.19 *Mutually Exclusive (Disjoint)*
Two events are **mutually exclusive** if they cannot both occur. They share no outcomes.



Disjoint (mutually exclusive) events are important because they cannot happen at the same time. This means their probabilities can be added directly: $P(A \cup B) = P(A) + P(B)$. This rule makes calculations much simpler and avoids double-counting outcomes.

12.3.6 Mutually Exclusive Events

Axiom 3 tells us we can add probabilities when events are mutually exclusive. But what does that actually mean, and how do we check?

Exercise 12.20 *Checking Mutual Exclusivity*
Roll a fair six-sided die. Determine if each pair is mutually exclusive.

1. $A =$ "Roll a 2", $B =$ "Roll a 5"

2. $A =$ "Roll an even number", $B =$ "Roll greater than 3"

12.3.7 Addition Rule for Mutually Exclusive Events

Key point: If A and B are mutually exclusive:

$$P(A \text{ or } B) = P(A) + P(B)$$

Example 12.21 *Addition Rule for Disjoint Events*
For each of the following contexts, compute the probability.

1. Roll a die. Find $P(1 \text{ or } 6)$.

2. Draw a card. Find $P(\text{heart or spade})$.

Multiple Mutually Exclusive Events
If $A_1, A_2, \dots, A_n$ are all mutually exclusive, then

$$P(A_1 \text{ or } \cdots \text{ or } A_n) = P(A_1) + \cdots + P(A_n)$$

12.3.8 The General Addition Rule

What happens when events *can* overlap? If we just add $P(A) + P(B)$, we count the overlap twice. We need a correction.

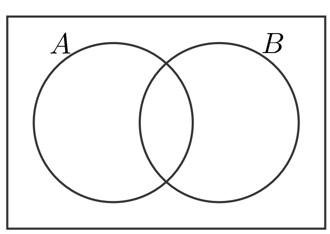
Exercise 12.22 *General Addition Rule Example*
A class has 100 students. 60 own a laptop, 40 own a tablet, and 20 own both.

	Tablet	No Tablet	Total
Laptop	20	40	60
No Laptop	20	20	40
Total	40	60	100

Find $P(\text{Laptop or Tablet})$

Definition 12.23 *Addition Rule (General)*
For **any** two events A and B:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$



Key point: We subtract $P(A \text{ and } B)$ because when A and B overlap, adding $P(A) + P(B)$ counts the overlap **twice**.

Exercise 12.24 *Movie Nights*
At a movie night with 300 attendees: 120 watched a comedy (A), 150 watched a drama (B), and 60 watched both ($A \cap B$).

What is the probability that a randomly selected attendee watched a comedy or a drama?

12.3.9 The Complement Rule

Sometimes the easiest way to find a probability is to figure out what *doesn't* happen. If we know the probability of the opposite event, we can subtract from 1.

Definition 12.25 *Complement*
The **complement** of event A, written $\overline{A}$ or A^c, is the event that A does **not** occur.

Example 12.26 *Finding Complements*
Find the complement for each event:

1. $A =$ "roll a 2" on a die: ____________________
2. $B =$ "get at least one head in 3 coin flips": ____________________
3. $C =$ "draw a heart from a deck": ____________________

Definition 12.27 *Complement Rule*
For any event A,
$$P(\overline{A}) = 1 - P(A)$$

Example 12.28 *Using the Complement Rule*
Tossing a coin 3 times. Event A = "At least one head in 3 flips."

Key point: Use the complement rule when it's easier to compute what you **don't** want.

Exercise 12.29 *Complement Rule Practice*
Suppose we roll a fair six-sided die. For each event, find the corresponding probabilities.

1. A = rolling a 2. Find $P(\overline{A})$

2. B = rolling less than 4. Find $P(\overline{B})$

3. Let C = "roll an even number" and D = "roll greater than 3". Find $P(C \cup D)$ and $P(\overline{C \cup D})$

12.4. Probability Models & Random Variables

So far we have been working with events described in words ("roll an even number," "draw a heart"). In practice, we often want to attach a *number* to each outcome so that we can do calculations with it. This section introduces the idea of a probability model and the concept of a random variable.

12.4.1 Two Types of Probability Models

Definition 12.30 *Finite Probability Model*
A model whose sample space has a **finite number** of outcomes. We can list every outcome and its probability.

Definition 12.31 *Continuous Probability Model*
A model whose sample space is **continuous**: outcomes can take any value in an interval. We describe probabilities using a density curve.

Quick Test
Ask: "Can I list every possible outcome?"

- **Yes** → finite (e.g., number of children, die roll)
- **No** → continuous (e.g., height, time, temperature)

Example 12.32 *Finite or Continuous?*
Classify each scenario:

1. Height of a person: ______________________
2. Number of cars sold: ______________________
3. Time to run a mile: ______________________
4. Number of siblings: ______________________
5. Goals in a hockey game: ______________________
6. Body temperature: ______________________
7. Eggs in a carton: ______________________
8. Gas remaining in tank: ______________________

12.4.2 Finite Probability Models in Practice

A finite probability model is just a table that lists every possible value and its probability. The probabilities must satisfy the axioms: each one is between 0 and 1, and they all add up to 1.

Exercise 12.33 *A Finite Probability Model*
"How many cups of coffee have you had today?"

Cups (X)	0	1	2	3	4	5+
Probability	0.40	0.22	0.15	0.10	0.07	0.06

1. Verify this is a valid probability model

2. $P(X < 4)$ = "fewer than 4 cups"

3. $P(X \geq 1)$ = "at least one cup"

Exercise 12.34 *Completing a Probability Model*
A discrete random variable Y has the following distribution. One probability is missing.

Value (Y)	10	20	30	40
Probability	0.15	0.30	**?**	0.10

Find $P(Y = 30)$,the probability that $Y = 30$.

12.4.3 Random Variables

Up to this point, our outcomes have been labels like H, T, or "Vaccinated." A random variable turns each outcome into a number, which lets us do arithmetic with probability.

> **Definition 12.35** *Random Variable*
> A **random variable** assigns a **numerical value** to each outcome of a random phenomenon.
>
> The **probability distribution** describes which values are possible and how likely each is.

Examples:

- X = number shown when rolling a die
- Y = number of heads in 10 coin flips
- Z = height (in cm) of a randomly selected student

> ✖ **Common Mistake:** Random variables must be **numerical**. "Favourite colour" is not a random variable (it's categorical). But we could define $X = 1$ if blue, $X = 0$ otherwise.

Exercise 12.36 *A Discrete Random Variable*
Let X = number of countries visited by a randomly selected DS1000 student.

Countries	0	1	2	3	4	5	···
Count	3	6	18	33	25	24	···
Proportion	0.015	0.029	0.088	0.162	0.123	0.118	···

Compute each of the following probabilities (or explain why it cannot be computed).

1. $P(X = 3)$

2. $P(X \leq 2)$

3. $P(X > 2)$

12.4.4 Continuous Random Variables

We will examine continuous random variables after conditional probability and before delving into Normal distributions. In the meantime, in case you are curious, the following are some facts about continuous random variables that make them different from discrete random variables:

- Continuous variables take values in an interval (e.g., height, time).
- Probabilities are represented by **area under a density curve**.
- The probability of any single exact outcome is 0.
- We calculate probabilities for intervals: $P(a < X < b)$.

We'll come back to continuous random variables after Chapter 13.

12.5. Chapter 12 Summary

Foundations

- **Random phenomenon:** uncertain outcome
- **Sample space S:** all possible outcomes
- **Event:** a set of outcomes
- **Union (or):** at least one occurs; **Intersection (and):** both occur
- **Complement:** event does not occur

Rules of Probability

- **A1:** $0 \leq P(A) \leq 1$
- **A2:** $P(S) = 1$
- **A3 (disjoint):** $P(A \cup B) = P(A) + P(B)$
- **General addition:** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Complement:** $P(\overline{A}) = 1 - P(A)$

Models

- **Finite:** it is possible to list all outcomes and probabilities in a table (e.g. a count)
- **Continuous:** outcomes can take any value in an interval (e.g. a measurement)
- **Random variable:** numerical outcome of a random phenomenon

Looking Ahead

Chapter 6 gave you **intuition** for probability through contingency tables. Chapter 12 gives you **formal rules** that work in any situation. Chapter 13 will focus on **conditional probability** in depth.

12.6. Additional Practice Problems

Exercise 12.37 *Languages*
At a company, $P(\text{speaks French}) = 0.30$, $P(\text{speaks Spanish}) = 0.25$, and $P(\text{speaks French or Spanish}) = 0.45$.

Find:

1. $P(\text{speaks both French and Spanish})$
2. $P(\text{speaks neither French nor Spanish})$
3. Are "speaks French" and "speaks Spanish" mutually exclusive? Justify.

Exercise 12.38 *Practice Problem 2*
A password consists of one letter (A, B, or C) followed by one digit (1, 2, 3, or 4). All passwords are equally likely.

Find:

1. How many outcomes are in the sample space?
2. $P(\text{password starts with A})$
3. $P(\text{password contains an even digit})$
4. $P(\text{starts with A } \textbf{or} \text{ contains an even digit})$
5. $P(\text{does not start with A } \textbf{and} \text{ contains an odd digit})$

Exercise 12.39 *Passwords*
For events A and B: $P(A) = 0.6$, $P(B) = 0.5$, and $P(\overline{A \cup B}) = 0.1$.

Find:

1. $P(A \cup B)$
2. $P(A \cap B)$
3. $P(\text{exactly one of } A \text{ or } B)$
4. Are A and B mutually exclusive? Justify.

Exercise 12.40 *More events*
A bag contains 3 marbles: Red (R), Green (G), and Blue (B). You draw two marbles one at a time **without replacement**.

Find:

1. List the sample space S. (Hint: use a tree diagram.)
2. Let A = "Red is drawn at some point." List the outcomes in A.
3. Let B = "Green is drawn first." List the outcomes in B.
4. Find $P(A)$, $P(B)$, and $P(A \cap B)$.
5. Are A and B mutually exclusive? Justify.

Exercise 12.41 *instagram and tiktok*
A survey of 200 university students found:

- 90 use Instagram (I)
- 110 use TikTok (T)
- 50 use both

Find:

1. Fill in all four regions of a Venn diagram with counts.
2. $P(I \cup T)$ (uses Instagram or TikTok)
3. $P(\overline{I \cup T})$ (uses neither)
4. P(uses exactly one of the two platforms)

Exercise 12.42 *Practice Problem 6: Finite Probability Model*
Let X = number of pets owned by a randomly selected student. The probability distribution is:

Pets (X)	0	1	2	3	4+
Probability	0.18	0.35	0.25	**?**	0.07

Compute each of the following probabilities: $P(X = 3), P(X \geq 2), P(X < 2)$.

Exercise 12.43 *Practice Problem 7: Non-Equally-Likely Outcomes*
A spinner has four sections: Red, Blue, Green, and Yellow. Red is twice as likely as Blue. Green and Yellow are each equally likely as Blue.

Find:

1. Find the probability of landing on each colour.
2. P(Red or Green). Are these events mutually exclusive?
3. P(not Yellow).