Introduction to Probability

Grinnell College

March 3, 2025

Review

A couple of weeks ago we spent some time making tables and using them to answer questions.

- What percent of Titanic passengers survived?
- What percent of Florida voters supported citizenship pathways for immigrants who entered the US illegally.
- What percent of 30-50 year olds had low job satisfaction?

Today's Outline

- continue to use tables of data.
- introduce probabilities
- probability math

What is Probability?

We will be working with *events* of *random processes*

- random: cannot perfectly predict outcomes
- event: a specific outcome (or collection of outcomes)

Probabilities are numbers between 0 and 1 that represent how likely (or unlikely) an <u>event</u> is to happen.

- closer to zero = more unlikely
- closer to one = more likely

Probability of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times

What is Probability?

When multiple events are equally likely, probability can be thought of as a fraction

$$\frac{\#\text{ of ways an event can happen}}{\#\text{ of all possible outcomes}}$$

Examples:

Flipping a coin: 1 heads out of 2 possibilities $\rightarrow$ prob. heads $= 1/2 = 0.5$

Probability of drawing a red card from a deck of 52 cards? $26/52 = 0.5$

Empirical Examples

A report published in 1988 summarizes results of a Harvard Medical School clinical trial determining effectiveness of asprin in preventing heart attacks in middle-aged male physicians

	Heart Attack		
Treatment	Attack	No Attack	Total
Placebo	189	10,845	11,034
Aspirin	104	10,933	11,037
Total	293	21,778	22,071

Probability a randomly selected participant has a heart attack?

Probability that a randomly selected participant in the placebo group has a heart attack?

Notation

To save our selves some time, we often use some shorthand notation

P() is used to denote the probability of something, capital letters are quick ways to write down events

- P(patient having a heart attack) $\rightarrow$ P(heart attack) $\rightarrow$ P(H)
- read as "probability of patient having a heart attack"

Often times we may think of an event in terms of "success" (it happened) or "not success" (it did not happen)

Did patient have a heart attack?

- Yes = Success (unfortunate terminology)
- No = Failure

Notation

Sample Space: the set of all possible outcomes, denoted **S**

- uses brackets to denote a set, lists all the outcomes

Consider rolling a die where the set of possible outcomes is

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

We express probability of an outcome like so:

$$\text{P(rolling a 6)} = \frac{1}{6}$$

If context is clear, we can make it simpler:

$$\text{P}(6) = \frac{1}{6}$$

Marginal Probability

Marginal Probability – the probability of a single event

- P(H) = P(Heart attack)
- P(rolling a 6) with dice
- name comes from using the margins (totals) of a table
- simplest types of probability we can work with

Unions

Union – Scenario where one event happens **or** another event happens (or both)

- We will always use 'inclusive or' meaning both events happening is allowed
- denoted P(A or B), $P(A \cup B)$

Example: P(rolling a 1 *or* rolling a 2) $=$ P(1 or 2) $= P(1 \cup 2)$

Unions

How do we find union probabilities? When you define your events, add up all the individual outcomes before finding a probability.

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

Define the following events:

- A: roll a 2 $= \{2\}$
- B: roll an odd number $= \{1, 3, 5\}$

The *union* $A \cup B$ is the same as adding up all the individual events in both A and B

- $A \cup B = \{1, 2, 3, 5\}$
- P(A or B) $= \frac{4}{6} = \frac{2}{3}$

Intersection

Intersection – Scenario where two events happen at the same time

- denoted P(A and B), P(A $\cap$ B)

Example: define the following events

- A: roll a 1 or a 2 $= \{1, 2\}$.
- B: roll an odd number $= \{1, 3, 5\}$.

An **intersection** is looking at where these events 'meet' or 'overlap'

How could I satisfy both of these events? Look for the shared outcomes.

- $A \cap B = \{1\}$
- $P(A \cap B) = P(\text{A and B}) = \frac{1}{6}$

Disjoint

Two events are said to be **disjoint** or **mutually exclusive** if they cannot both happen at the same time.

Equivalenty, two events are **disjoint** if their intersection is empty

Example: Are the events 'rolling a 1' and 'rolling a 2' disjoint?

- Ask yourself. Can we roll both a 1 and 2 at the same time? No.
- 'rolling a 1' and 'rolling a 2' are *disjoint* events

We could also look at it like this:

- A: roll a 1 $= \{1\}$
- B: roll a 2 $= \{2\}$
- $A \cap B = \emptyset$
- P(A and B) $= 0$

Disjoint

Are the following events disjoint?

- Using a die to roll an even number or to roll a 3?
- Using a die to roll an odd number or a number greater than 4?
- 'taking a computer science class this semester' or 'taking a statistics class this semester'?

More on Unions

When we are working with more complicated event *unions* there is another (usually easier) way to calculate probabilities

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

Define the following events:

- A: roll a 1, 2, or 4 $= \{1, 2, 4\}$
- B: roll an even number $= \{2, 4, 6\}$

$A \cup B = \{1, 2, 4, 6\} \rightarrow \mathrm{P}(A \cup B) = \mathrm{P}(\text{A or B}) = 4/6$

This breaks down when we have events that each contain a lot of outcomes

- we'd have to combine a whole bunch of different outcomes together and keep track of everything in a new set $\rightarrow$ time consuming

Addition Rule

There is a connection between *unions* and *intersections* that lets us more easily calculate their probabilities

The **Addition Rule** states that for *any* events A and B, the probability that at least one will occur is:

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

Addition Rule – Practice

In a standard deck of 52 cards, there are 4 suits (diamonds, hearts, spades, and clubs), each containing 13 cards. Within each suit, there are 4 face cards (J, Q, K, and A)

What is the probability that we draw a card that is either a face card or a diamond?

Addition Rule – Practice

In a standard deck of 52 cards, there are 4 suits (diamonds, hearts, spades, and clubs), each containing 13 cards. Within each suit, there are 4 face cards (J, Q, K, and A)

What is the probability that we draw a card that is either a face card or a diamond?

$$\begin{aligned} P(\text{Face or diamond}) &= P(\text{Face}) + P(\text{diamond}) - P(\text{Face and diamond}) \\ &= \frac{12}{52} + \frac{13}{52} - \frac{4}{52} = \frac{21}{52} \approx 40\% \end{aligned}$$

Addition Rule – Special Case of Disjoint

The **Addition Rule** states that for *any* events A and B, the probability that at least one will occur is:

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

When events are **disjoint**, the Addition Rule simplifies.

$$\text{A and B are disjoint} \rightarrow \text{P(A and B)} = 0$$
$$\rightarrow \text{P(A or B)} = \text{P(A)} + \text{P(B)}$$

Example:

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

$$\text{P(1 or 2)} = \text{P(1)} + \text{P(2)} = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

Complements

The **Complement** of any event A is when the event *does not* occur

- denoted as A^C
- A^C represents all events in **S** that are not part of A

Together A and A^C must comprise all events that make up **S**, so from the Addition Rule we have P(A or A^C) = P(A) + P(A^C) = 1

This leads to the **Complement Rule**:

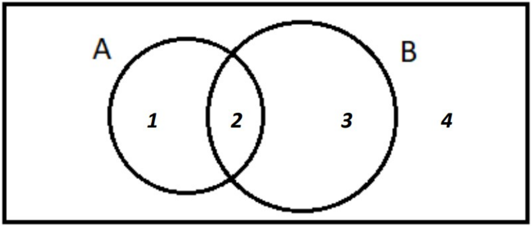
- P(A^C) = 1 - P(A)

Example: Let A: roll more than 2, so A = {3, 4, 5, 6}. P(A) = $\frac{4}{6} = \frac{2}{3}$

Then A^C is the event of *not* rolling more than 2, so A^C = {1, 2}, P(A^C) = 1 - P(A) = 1 - $\frac{2}{3} = \frac{1}{3}$

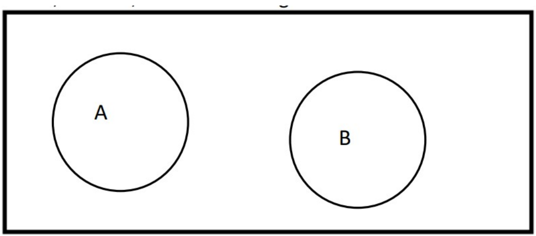
Venn Diagrams

Venn diagrams can be used as a way to help us think about these probabilities.



1. Which portion(s) of the Venn Diagram above is the *intersection* of A and B?
2. Which portion(s) of the Venn Diagram above are the *union* of A and B?
3. Which portion(s) of the Venn Diagram above is the *complement* of A?

Venn Diagrams



Disjoint Events can also be thought of as events that do not overlap

- $P(A \text{ and } B) = P(A \cap B) = 0$

Key Terms

Probability: number between 0 and 1 representing likelihood of an event
Union: when A or B can happen
Intersection: when A and B both happen
Disjoint: when events A and B *cannot* both happen

Additive Rule

- P(A or B) = P(A) + P(B) - P(A and B)
- Special: A and B are disjoint $\rightarrow$ P(A or B) = P(A) + P(B)

Complement Rule

- P(not A) = P(A^C) = 1 - P(A)