Quantitative Methods

Weeks 3 to 5: Probability

AY 2023-24

Department of Political Economy

Instructor: Daniele Girardi

Weeks 3 to 5 – Probability

1. Definitions & basic properties (week 3, Ch 4)

2. Random variables & their distribution (week 4, Ch 5-6)

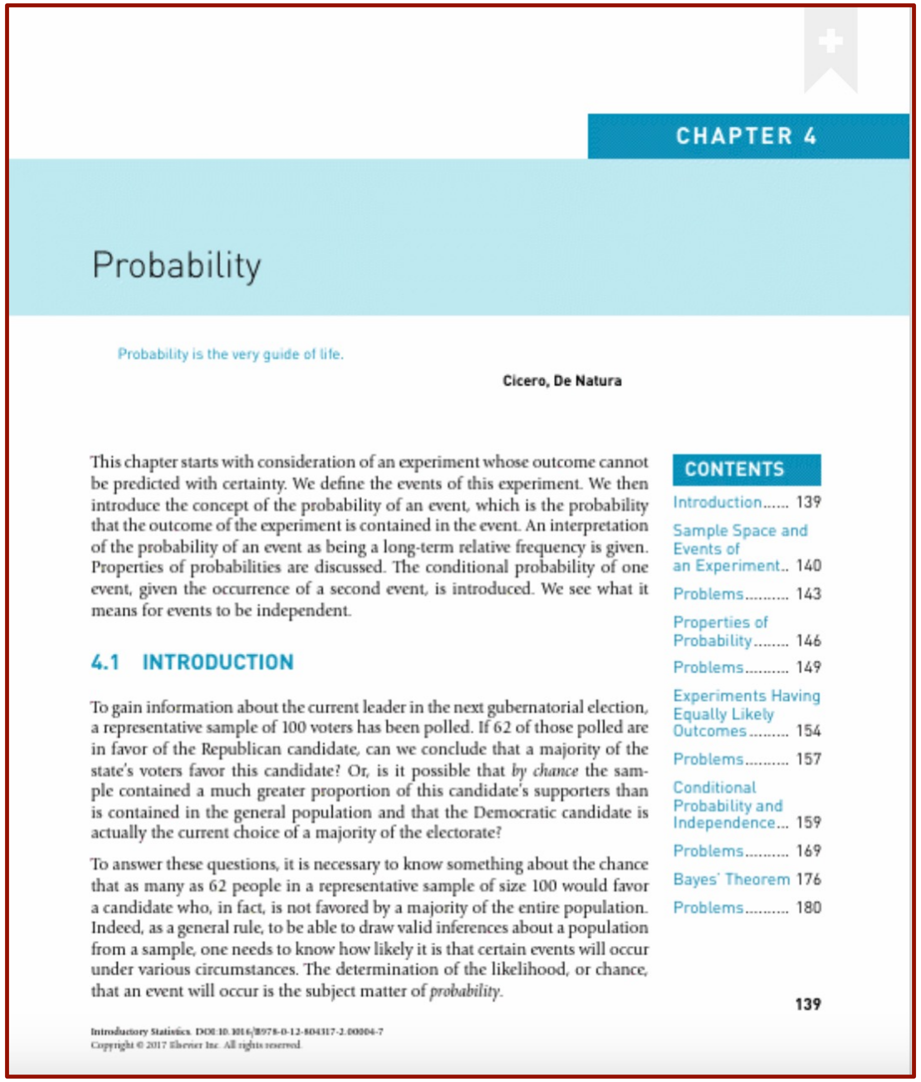
3. The distribution of sampling statistics (week 5, Ch 7)

Write down three things you learned from the reading

(textbook Chapter 4)

If you couldn't do the reading this week:

Write three things that come to your mind when you think about "probability".



CHAPTER 4

Probability

Probability is the very guide of life.

Cicero, De Natura

This chapter starts with consideration of an experiment whose outcome cannot be predicted with certainty. We define the events of this experiment. We then introduce the concept of the probability of an event, which is the probability that the outcome of the experiment is contained in the event. An interpretation of the probability of an event as being a long-term relative frequency is given. Properties of probabilities are discussed. The conditional probability of one event, given the occurrence of a second event, is introduced. We see what it means for events to be independent.

CONTENTS

4.1 INTRODUCTION

To gain information about the current leader in the next gubernatorial election, a representative sample of 100 voters has been polled. If 62 of those polled are in favor of the Republican candidate, can we conclude that a majority of the state's voters favor this candidate? Or, is it possible that *by chance* the sample contained a much greater proportion of this candidate's supporters than is contained in the general population and that the Democratic candidate is actually the current choice of a majority of the electorate?

To answer these questions, it is necessary to know something about the chance that as many as 62 people in a representative sample of size 100 would favor a candidate who, in fact, is not favored by a majority of the entire population. Indeed, as a general rule, to be able to draw valid inferences about a population from a sample, one needs to know how likely it is that certain events will occur under various circumstances. The determination of the likelihood, or chance, that an event will occur is the subject matter of *probability*.

139

Introductory Statistics. DOI:10.1016/B978-0-12-804317-2.00004-7

WHY PROBABILITY THEORY?

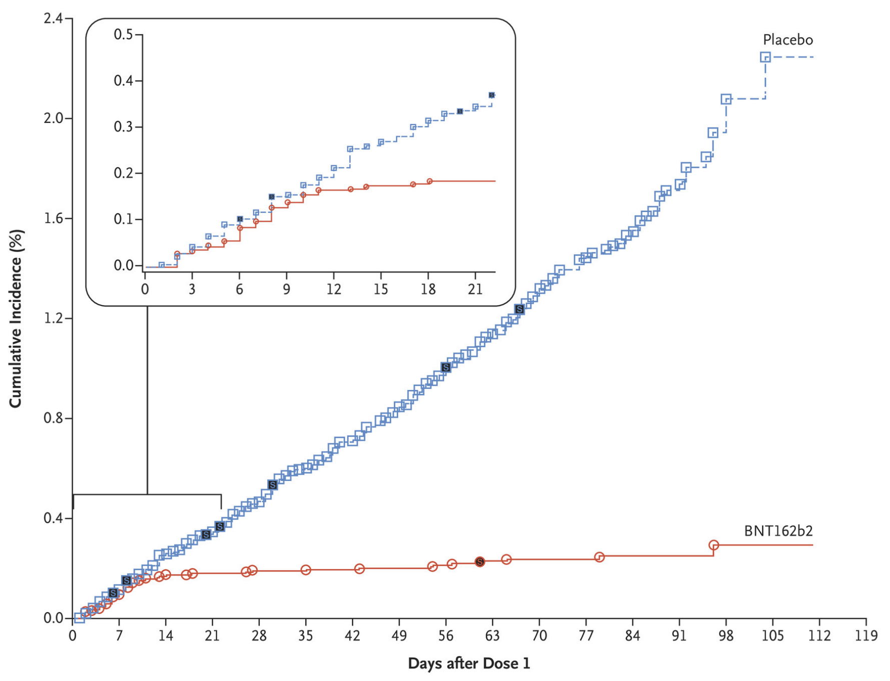
- **We live in an uncertain & risky world.**

- **Application to quantitative social science: *statistical inference***
  - *Ability to discern random fluctuations from changes caused by some intervention or influencing factor of interest.*
  - Allows to draw (probabilistic) conclusions from data

Example: Clinical trial for the Pfizer COVID vaccine

- 43,448 participants worldwide.
- ½ received vaccine, ½ a placebo.
- 9 infections among vaccinated.
- 172 infections in the placebo group.

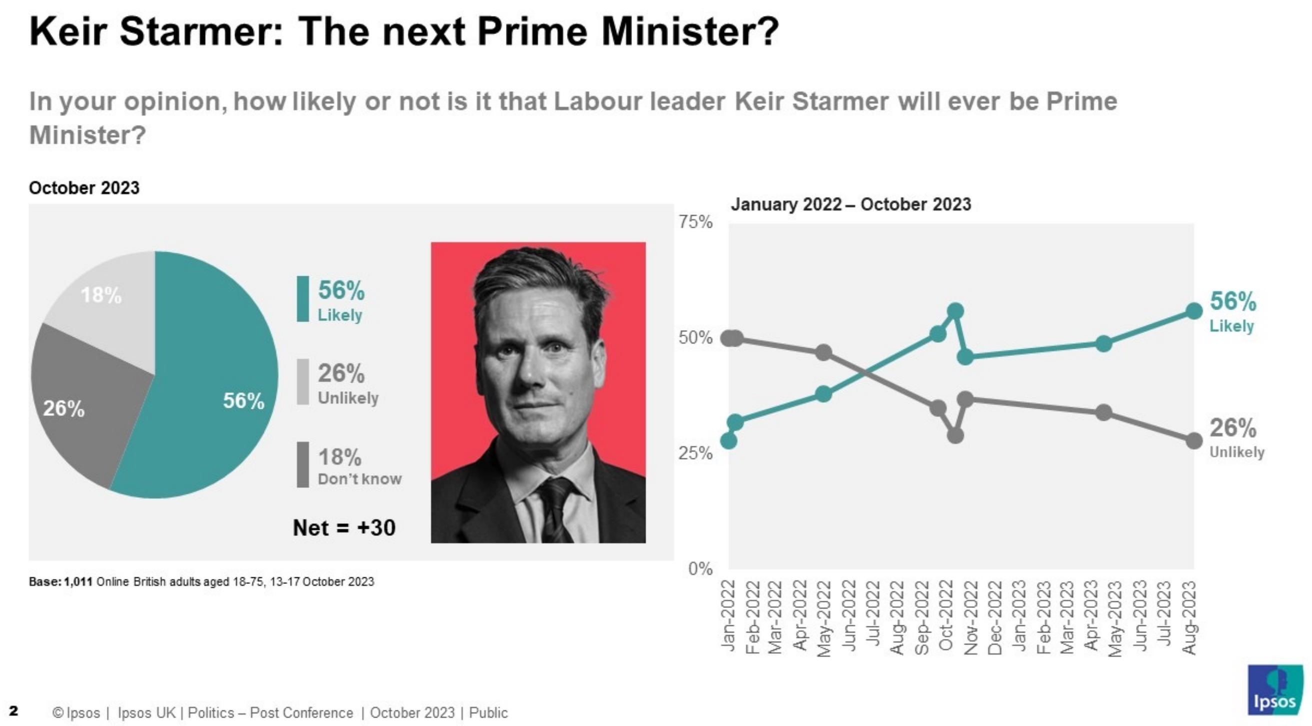
Do you remember how is this related to statistical inference?



Efficacy End-Point Subgroup	BNT162b2, 30 μg (N=21,669) No. of participants	Surveillance time person-yr (no. at risk)	Placebo (N=21,686) No. of participants	Surveillance time person-yr (no. at risk)	VE (95% CI) percent
Covid-19 occurrence					
After dose 1	50	4.015 (21,314)	275	3.982 (21,258)	82.0 (75.6–86.9)
After dose 1 to before dose 2	39		82		52.4 (29.5–68.4)
Dose 2 to 7 days after dose 2	2		21		90.5 (61.0–98.9)
≥7 Days after dose 2	9		172		94.8 (89.8–97.6)

Source: Polack et al (2020) "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine"

1. Probability: Definitions and basic properties



Key concepts

- **Random processes** (or *experiments*)
  - Flipping a coin.
  - Time it takes you to get on campus tomorrow.
  - Number of times your computer crashes while writing an essay.
  - *Influenced by something not known in advance, that will eventually be revealed.*
- **Outcomes**
- **Sample space**
- **Events**

Random process: flipping two coins

Possible outcomes:

HH or HT or TH or TT.

Sample space:

S = {(HH), (HT), (TH), (TT)}

Event A: Heads comes up in the first coin.

A = {(HH), (HT)}

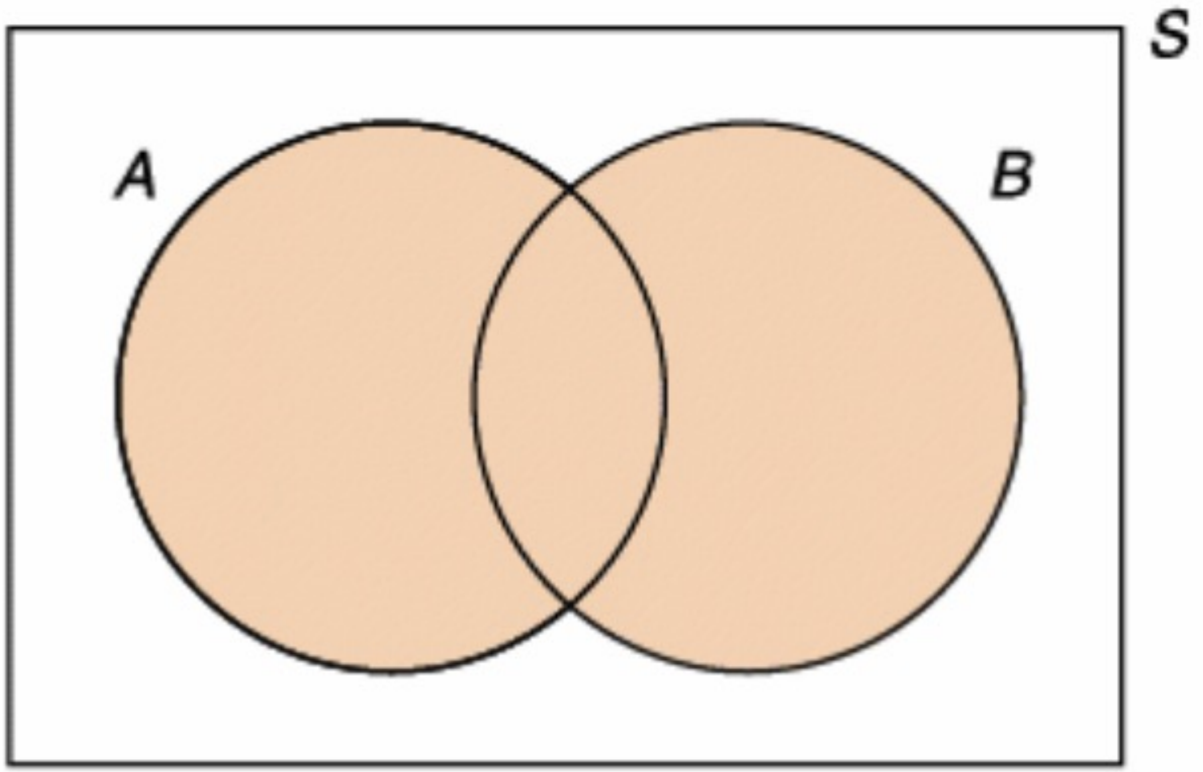
Event B: Tails comes up in at least one coin.

B = { (HT), (TH), (TT)}

Union of two events

$A \cup B$

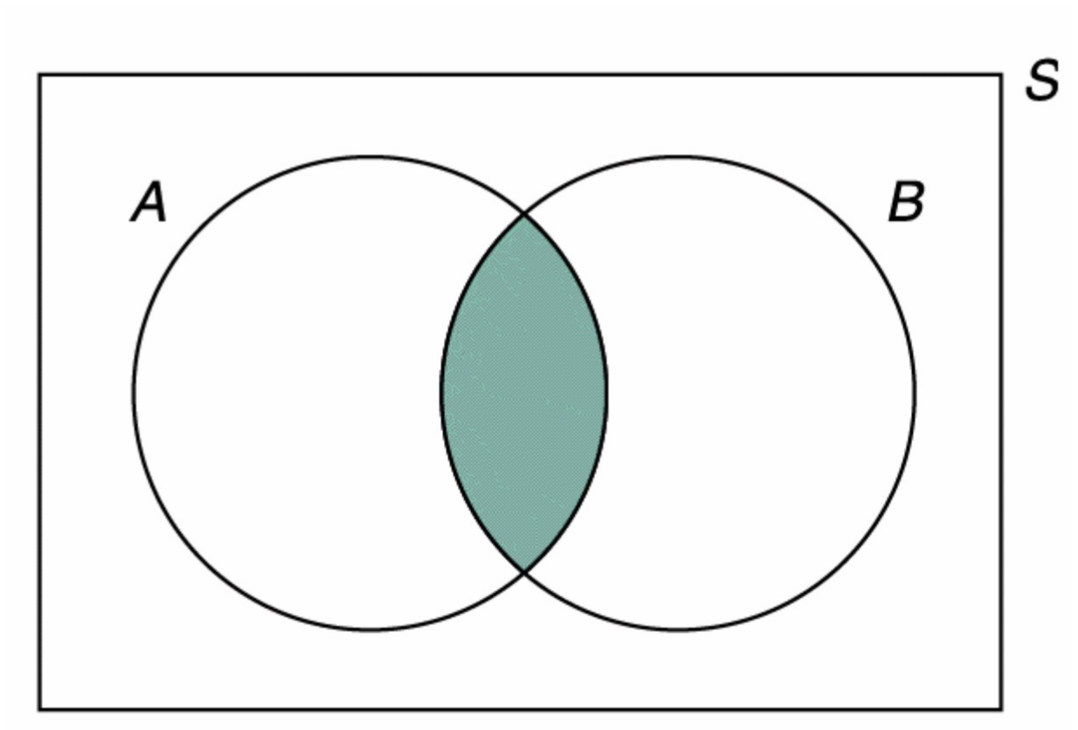
- All outcomes that are in A *or* in B.
- $A \cup B$ occurs if either A or B occurs.



Intersection of two events

$A \cap B$

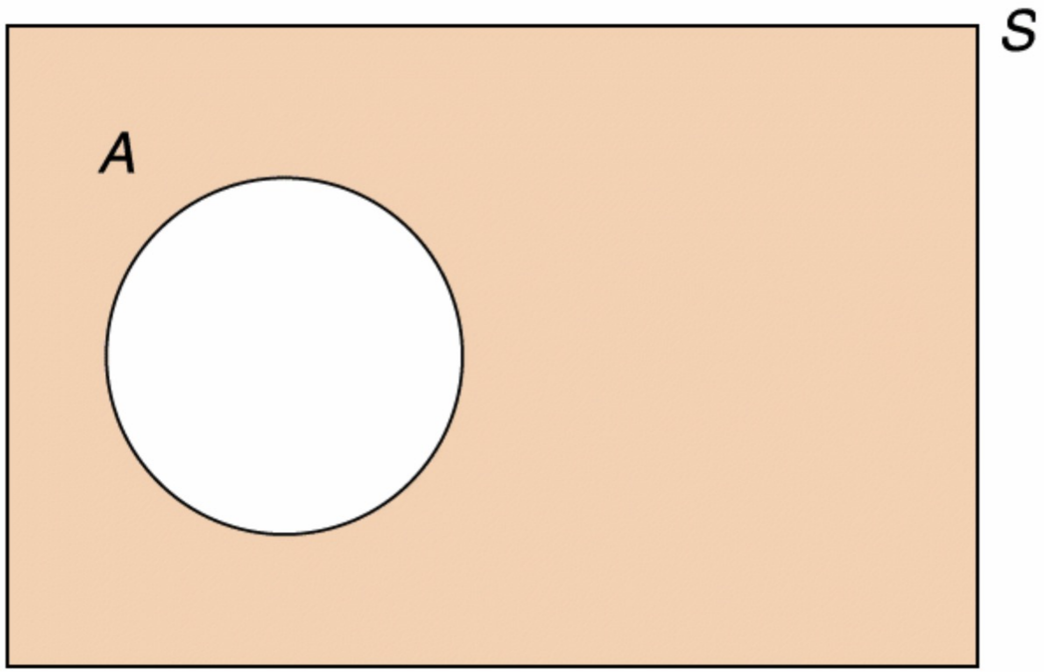
- All outcomes that are both in A *and* in B.
- $A \cap B$ occurs if both A *and* B occur.



Complement of an event

A^C

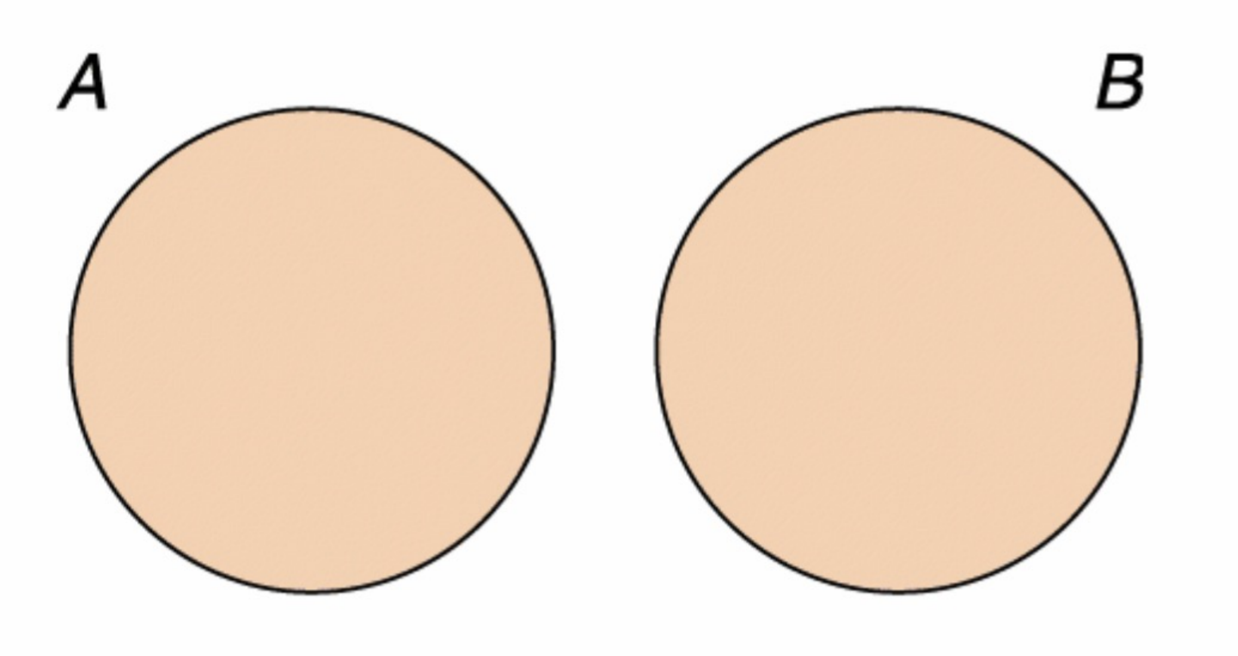
- All outcomes that are not in A.
- A^C occurs if A does not occur.



Disjoint events

Disjoint if $A \cap B = \emptyset$

- Mutually exclusive.
- No outcomes in common.
- A and B cannot both happen.



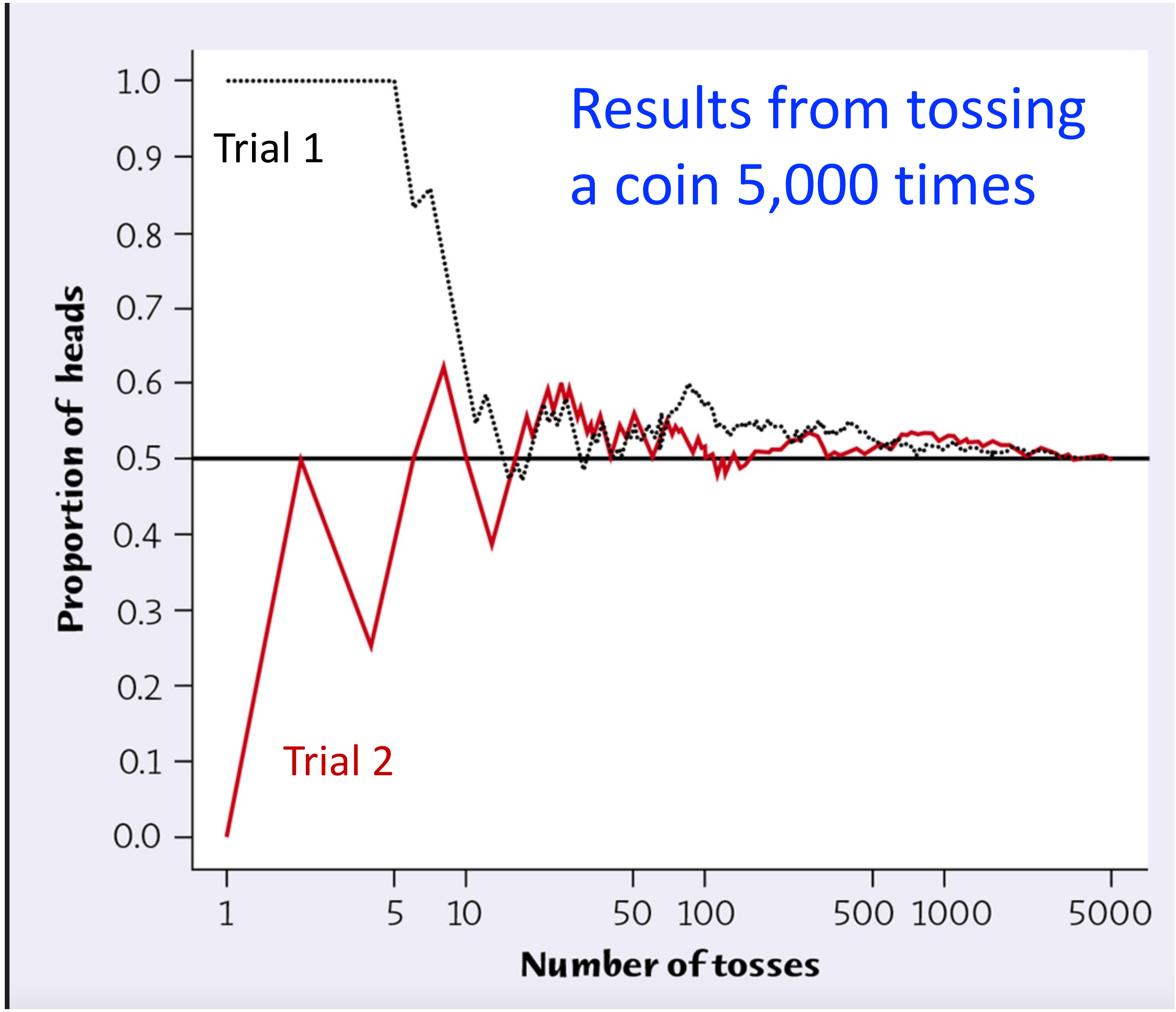
Recap of notation

- *Union* of A & B: $A \cup B$
- *Intersection* of A & B: $A \cap B$
- *Complement* of A: A^C
- *Null* event: $\emptyset$
- A & B *disjoint/mutually exclusive* if $A \cap B = \emptyset$

Probability

A measure of the likelihood of an outcome/event.

- *the proportion of times that the outcome would occur, if you repeated the process a very large number of times under identical conditions.*
- in short: *the long-run relative frequency of an outcome.*



Probability rules

1. $0 \leq P(A) \leq 1$

Probability is always between 0 and 1 (inclusive)

2. $P(Sample\ Space) = P(S) = 1$

All possible outcomes together have probability 1.

Probability rules

3. If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

If two events have no outcomes in common, the probability that one or the other occurs is the sum of the individual probabilities.

4. $P(A^c) = 1 - P(A)$

The probability that event A does not occur is 1 minus the probability that A occurs.

Q1: What is the probability that Rishi Sunak will still be PM on February 1st 2025?

Q2: *If Conservatives win more than 50% of House seats in the 2024 general election,* what is the probability that Rishi Sunak will still be PM in February 2025?

Conditional Probability

Unconditional probability:

P(*Sunak 2025 PM*)

Conditional probability:

P(*Sunak 2025 PM* | *Conservative victory*)

Conditional Probability of B given A

- The probability of event B, given that another event A occurs.
- $P(B|A)$.
- $P(B|A) = \frac{P(A \; and \; B)}{P(A)} = \frac{P(A \cap B)}{P(A)}$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

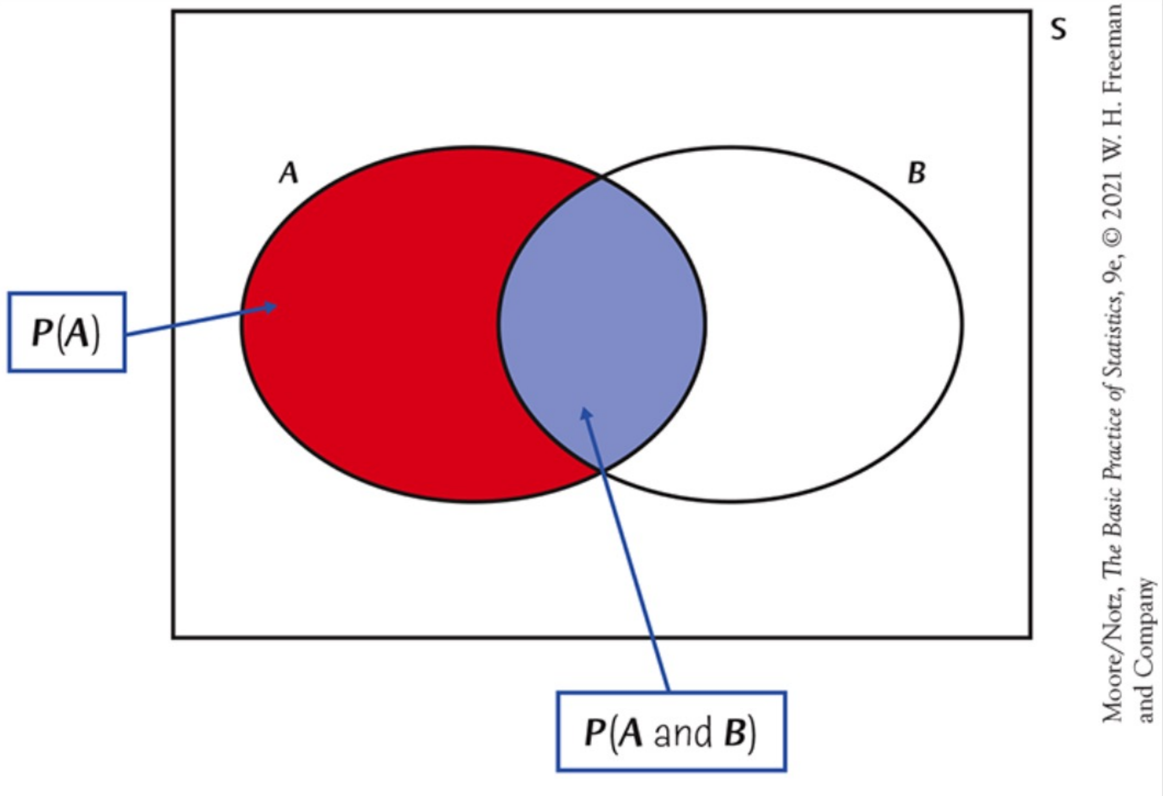


FIGURE 13.5

The probability of *B* given *A* is the proportion of the outcomes in the red- and blue-shaded oval (event *A*) that are in the blue-shaded area (event *A* and *B*).That is, $P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$\Downarrow$$

Multiplication Rule

$$P(A \cap B) = P(A)\, P(B|A)$$

For A and B to both happen… | A must happen | & | given that A has happened, B must happen.

Multiplication Rule: $P(A \cap B) = P(A)\,P(B|A)$

P(*Sunak '25 PM* | *Conservative win*)= 90%

P(*Conservative win*) = 30%

P(*Conservative win & Sunak '25 PM*) = ?

$$P(\textit{Conservative win \& Sunak '25 PM}) = 0.3 * 0.9 = 0.27 \text{ (or } 27\%)$$

Independence

- B is independent of A if

$$P(B|A) = P(B)$$

- If A and B are independent, then

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

- B independent of A $\rightarrow$ A independent of B

Thank you for your attention