When asked to define a miracle—

“Anything with a probability of less than 10%”
-Enrico Fermi
1 Introduction

2 Chevalier de Méré

3 Probability foundations

4 Perspectives: Bayesian & Frequentist

5 Bayes Theorem

6 Probability: X-men
Meet Linda¹

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in the 2017 Women’s March.

Assign probabilities for each statement that describes Linda:

1. Linda is a teacher.
2. Linda is a bookstore cashier and enjoys yoga.
3. Linda is a bank teller.
4. Linda describes herself as a feminist.
5. Linda sells insurance.
6. Linda is a bank teller and describes herself as a feminist.

¹Tversky & Kahneman experiment
What is the probability 2 people share the same birthday (month and day)?

The probability is 99.99%!

If there are only 30 people in a room the probability at least 2 people share a birthday is 70.63%!
Chevalier de Méré questioned Pascal which dice game he would be more likely to win:

Game 1: Throw one die, up to 4 times, winning once you roll a six.

Game 2: Throw 2 dice, up to 24 times, winning once you roll a double-six.
A simulation of the Chevalier playing his two games

- Game 1: one die
- Game 2: two dice

Proportion of wins

0.0

0.4

0.6

0.8

1.0

0

20

40

60

80

100

0

2000

4000

6000

8000

10000

Proportion of wins

0.40

0.45

0.50

0.55

0.60
Short-term v. long-term: repeated!
Kolmogorov’s Probability Axioms

1. Probabilities cannot be negative
2. Out of all possible outcomes, at least one occurring is 100%
3. Events that are independent (disjoint/mutually exclusive) have a probability equal to adding both event probabilities together.

1. $0 \leq P(A)$
2. $P(S) = 1$
3. $P(A \cup B) = P(A) + P(B)$
The sample space, $S$ or $\Omega$, is defined as all possible outcomes of a single trial random experiment. On a single dice roll, the outcome can be a value of 1 to 6.

The circle $A$ includes event $A$: one subset of outcomes. This could be that you roll a 3 on one single dice roll.

$A^c$ is the complement of event $A$, all other possible outcomes. If you rolled a 3 then the complement is 1, 2, 4, 5, 6.
The event A or B or both A and B happen is the union of A and B.

Set notation is $A \cup B$
Principles: AND

The event $A$ AND $B$ is the intersection of $A$ and $B$.

Set notation is $A \cap B$
Principles: disjoint

Events $A$ and $B$ do not both happen together is disjoint or mutually exclusive events.

Set notation is $A \cap B = \emptyset$
What is the probability of drawing a Jack or a spade?

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

$$P(\text{Jack} \cup \spadesuit) = P(\text{Jack}) + P(\spadesuit) - P(\text{Jack} \cap \spadesuit)$$

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52}$$

$$= \frac{16}{52} = 30.77\%$$
Consider independent events as event A will not provide information about event B’s occurrence.

\[ P(A \cap B) = P(A) \times P(B) \]

Rolling dice will not inform you about the next roll.

Rolling consecutive sixes on dice rolls is calculated as sample space of 6 events, and the event a 6 occurs twice:

\[ \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} = 2.78\% \]
Monotonicity

\[ A \subseteq B = P(A) \leq P(B) \]

✔️ If you’re the president (A) then you are a U.S. citizen (B).
❌ If you’re a U.S. citizen then you’re president.
Méré’s game probabilities

Game 1

\[ P(\text{rolling at least one } 6) = 1 - \text{probability(not rolling a } 6) = 1 - \left( \frac{5}{6} \right)^4 = 0.5177 \]

Game 2

6 outcomes x 6 outcomes = 36 total outcomes; only 1 is double-six;

\[ P(\text{rolling at least one } 6) = 1 - (\text{probability not rolling double-6}) = 1 - \left( \frac{35}{36} \right)^4 = 0.491 \]
Frequentist: Conditioned on theory

**Frequentist**

- Parameters are fixed and unknown.
- Sample data is random from a population distribution.
- Neyman-Pearson’s Hypothesis test with Fisher’s $p$-value approach
- Probability is interpreted as long-run repeated sampling, under almost identical conditions, from a population distribution.
The $p$-value is the probability of obtaining a test statistic equal to, or more extreme than, the observed value under the assumption that the null hypothesis is true.
Bayesian: Conditioned on data

- Parameters are random and unknown.
- Sample data is fixed.
- Probability is interpreted as our beliefs about random variables; a measure of uncertainty.
- Distributions are updated with the arrival of new information—use of priors with the likelihood.
Conditional Probability

Confusing $P(A \mid B)$, $P(B \mid A)$, and $P(A)$...

$P(\text{cancer} \mid \text{positive test})$ is the probability of having cancer when your test results are positive. Is this a false positive or do you have cancer?

$P(\text{positive test} \mid \text{cancer})$ is a test's sensitivity. If you have cancer, what is the chance the test will accurately yield a positive result?

$P(\text{cancer})$ is probability of having cancer; the prevalence.
Bayes Theorem

\[ P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} \]

\[ P(B|A) = \frac{P(A|B) \times P(B)}{P(A|B) \times P(B) + P(A|B^c) \times P(B^c)} \]

Divide a positive result by all the ways it is possible to get a positive result—true positive plus a false alarm.
The prevalence of breast cancer is 0.8% for women in their forties. The mammogram tests have a 7% false positive rate. The test sensitivity is 90% which means 10% of people with cancer won’t be detected by the test.

What is the probability a woman has cancer given a positive test result?

Hint: don’t confuse $P(\text{cancer} | \text{positive result})$ with $P(\text{positive result} | \text{cancer})$
Example: Breast cancer

\[ P(\text{cancer}|\text{positive result}) \]

\[ P(BC|t+) = \frac{\text{true positive}}{\text{true positive} + \text{false alarm}} \]

\[ P(BC|t+) = \frac{(0.9)(0.008)}{(0.9)(0.008) + (0.07)(0.992)} \]

Probability of having cancer given a positive test result \( \approx 9.4\% \)
X-men: Sentinel project
Statement: “More mortal people are killed by Sentinels; the estimated numbers in 2019: there were 370 mortal people killed by Sentinels and 235 mutant people killed by the Sentinels”...use probability to check assertion:

Estimated population: 191 million mortal and 42 million mutant.

This means \( \frac{191}{42} = 4.6 \) times more mortal people.

Find expected mutant deaths by Sentinels: \( \frac{370}{4.6} = 80.43 \) mutant people killed by Sentinels if everything was equally proportionate.

Calculate discrepancy: \( \frac{235}{80.43} = 2.9. \)

Which can be interpreted as 3 times more mutant people are killed than mortal people than proportionally expected.
If someone asks “Are you a Bayesian or are you a Frequentist?” the answer should be yes.