Reminder

Gradescope enroll code: M8YYEZ

Homework due tonight by 11:59pm.
Conditional Probabilities

Often we want to know how likely something is conditioned on something else having happened.

**Example.** If we flip two fair coins, what is the probability that both outcomes are identical conditioned on the fact that at least one of them is heads?
\( \Omega = \{TT, TH, HT, HH\} \) \quad \forall \omega \in \Omega: \, \mathbb{P}(\omega) = \frac{1}{4} \\

“we get heads at least once” \quad \mathcal{A} = \{TH, HT, HH\} \\
“same outcome” \quad \mathcal{B} = \{TT, HH\} \\

If we know \( \mathcal{A} \) happened: (1) only three outcomes are possible, and (2) only one of them leads to \( \mathcal{B} \).

We expect: \( \mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{1}{3} \) \quad [Verbalized: Probability of \( \mathcal{B} \) conditioned on \( \mathcal{A} \).]
Conditional Probability – Formal Definition

**Definition.** The **conditional probability** of $\mathcal{B}$ given $\mathcal{A}$ is

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

Note: This is **only** defined if $P(A) \neq 0$.

If $P(A) = 0$, then $P(B|A)$ is **undefined**.
Example – Non-uniform Case

Pick a random ball

\[ \mathbb{P}(\text{green}) = \frac{1}{3} \]
\[ \mathbb{P}(\text{red}) = \frac{1}{6} \]
\[ \mathbb{P}(\text{blue}) = \frac{1}{3} \]
\[ \mathbb{P}(\text{black}) = \frac{1}{6} \]

“we do not get black” \[ \mathcal{A} = \{\text{red, blue, green}\} \]
“we get blue” \[ \mathcal{B} = \{\text{blue}\} \]

\[ \mathbb{P}(\mathcal{B} | \mathcal{A}) = \frac{\mathbb{P}(\{\text{blue}\} \cap \{\text{red, blue, green}\})}{\mathbb{P}(\{\text{red, blue, green}\})} = \frac{\mathbb{P}(\{\text{blue}\})}{\mathbb{P}(\{\text{red, blue, green}\})} = \frac{1/3}{5/6} = \frac{2}{5} \]
The Effects of Conditioning

\[ \mathbb{P}(B) < \mathbb{P}(B | A) \]

“\text{A-priori probability}” / prior

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

All three are possible!

\[ \mathbb{P}(B | A) > \mathbb{P}(B) \]

“A-posteriori probability” / posterior

“A-posteriori probability” / posterior
Prior Examples – A-posteriori vs a-priori

“heads at least once” \( \mathcal{A} = \{ \text{TH, HT, HH} \} \)

“same outcome” \( \mathcal{B} = \{ \text{TT, HH} \} \)

\[
\mathbb{P}(\mathcal{B}) = \frac{1}{2} > \mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{1}{3}
\]

“we do not get black” \( \mathcal{A} = \{ \text{red, blue, green} \} \)

“we get blue” \( \mathcal{B} = \{ \text{blue} \} \)

\[
\mathbb{P}(\mathcal{B}) = \frac{1}{3} < \mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{2}{5}
\]
Independence

**Definition.** Two events $\mathcal{A}$ and $\mathcal{B}$ are (statistically) **independent** if
\[
P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}) \cdot P(\mathcal{B}).
\]

**Note:** If $\mathcal{A}$, $\mathcal{B}$ independent, and $P(\mathcal{A}) \neq 0$, then:
\[
P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{A}\cap\mathcal{B})}{P(\mathcal{A})} = \frac{P(\mathcal{A})P(\mathcal{B})}{P(\mathcal{A})} = P(\mathcal{B})
\]
Reads as “The probability that $\mathcal{B}$ occurs is independent of $\mathcal{A}$.”
Independence - Example

Assume we toss two fair coins

“first coin is heads” \(\mathcal{A} = \{HH, HT\}\)

“second coin is heads” \(\mathcal{B} = \{HH, TH\}\)

\[
\mathbb{P}(\mathcal{A}) = 2 \times \frac{1}{4} = \frac{1}{2}
\]

\[
\mathbb{P}(\mathcal{B}) = 2 \times \frac{1}{4} = \frac{1}{2}
\]

\[
\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B})
\]

Note here we have defined the probability space **assuming** independence, so quite unsurprising – but this makes it all precise.
Gambler’s fallacy

Assume we toss 51 fair coins.
Assume we have seen 50 coins, and they are all “heads”.
What are the odds the 51st coin is also “heads”?

\[ A = \text{first 50 coins are heads} \]
\[ B = \text{51st coin is ”heads”} \]

\[ \mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{2^{-51}}{2^{-50}} = \frac{1}{2} \]

Gambler’s fallacy = Feels like it’s time for ”tails”!? 

51st coin is independent of outcomes of first 50 tosses!
Conditional Probability Define a Probability Space

The probability conditioned on $A$ follows the same properties as (unconditional) probability.

**Example.** $\mathbb{P}(B^c|A) = 1 - \mathbb{P}(B|A)$

**Formally.** $(\Omega, \mathbb{P})$ is a probability space + $\mathbb{P}(A) > 0$

$(\Omega, \mathbb{P}(\cdot|A))$ is a probability space
Recap

- \( \mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \).

- Independence: \( \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \).
Chain Rule

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

**Theorem. (Chain Rule)** For events $A_1, A_2, \ldots, A_n$,

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

(Proof: Apply above iteratively / formal proof requires induction)
Chain Rule – Applications

Often probability space \((\Omega, \mathbb{P})\) is given implicitly.

• Convenient: definition via a sequential process.
  – Use chain rule (implicitly) to define probability of outcomes in sample space.

• Allows for easy definition of experiments where \(|\Omega| = \infty\)
Sequential Process – Example

**Setting:** A fair die is thrown, and each time it is thrown, regardless of the history, it is equally likely to show any of the six numbers.

**Rules:** In each round
- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round
Sequential Process – Example

Events:
• $\mathcal{A}_i$ = Alice wins in round $i$
• $\mathcal{N}_i$ = nobody wins in round $i$

Rules: At each step:
• If outcome = 1,2 → Alice wins
• If outcome = 3 → Bob wins
• Else, play another round

\[
P(\mathcal{A})P(\mathcal{B}|\mathcal{A}) = P(\mathcal{A} \cap \mathcal{B})
\]

$P(\mathcal{A}_1) = \frac{1}{3}$

$P(\mathcal{A}_2) = P(\mathcal{A}_2 \cap \mathcal{N}_1)$

\[
= P(\mathcal{N}_1) \times P(\mathcal{A}_2 | \mathcal{N}_1)
\]

\[
= \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}
\]

[The event $\mathcal{A}_2$ implies $\mathcal{N}_1$, and this means that $\mathcal{A}_2 \cap \mathcal{N}_2 = \mathcal{A}_2$]
Sequential Process – Example

Events:
• $A_i = \text{Alice wins in round } i$
• $N_i = \text{nobody wins in round } i$

Rules:
- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round

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

$$P(A_i) = P(A_i \cap N_1 \cap N_2 \cap \cdots \cap N_{i-1})$$

$$= P(N_1) \times P(N_2|N_1) \times P(N_3|N_1 \cap N_2) \times \cdots \times P(N_{i-1}|N_1 \cap N_2 \cap \cdots \cap N_{i-2}) \times P(A_i|N_1 \cap N_2 \cap \cdots \cap N_{i-1})$$

$$= \left(\frac{1}{2}\right)^{i-1} \times \frac{1}{3}$$
Sequential Process – Example

Rules: At each step:
- If outcome = 1,2 → Alice wins
- If outcome = 3 → Bob wins
- Else, play another round
Sequential Process – Crazy Math?

\( \mathcal{A}_i = \text{Alice wins in round } i \rightarrow \mathbb{P}(\mathcal{A}_i) = \left( \frac{1}{2} \right)^{i-1} \times \frac{1}{3} \)

What is the probability that Alice wins?

\[
\mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots ) = \sum_{i=0}^{\infty} \left( \frac{1}{2} \right)^i \times \frac{1}{3} = \frac{1}{3} \times \sum_{i=0}^{\infty} \left( \frac{1}{2} \right)^i = \frac{1}{3} \times 2 = \frac{2}{3}
\]

**Fact.** If \(|x| < 1\), then \(\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}\).
Sequential Process – Another Example

Alice has two pockets:

- **Left pocket:** Two red balls, two green balls
- **Right pocket:** One red ball, two green balls.

Alice picks a random ball from a random pocket.
[Both pockets equally likely, each ball equally likely.]
Sequential Process – Another Example

\[ \mathbb{P}(R) = \mathbb{P}(R \cap \text{Left}) + \mathbb{P}(R \cap \text{Right}) \quad \text{(Law of total probability)} \]

\[ = \mathbb{P}(\text{Left}) \times \mathbb{P}(R|\text{Left}) + \mathbb{P}(\text{Right}) \times \mathbb{P}(R|\text{Right}) \]

\[ = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{2}{3} = \frac{1}{4} + \frac{1}{3} = \frac{7}{12} \]
Assume we observe high fever, what is the probability that the subject has Ebola? 

Posterior: $P(Ebola|High fever)$
**Bayes Rule**

**Theorem. (Bayes Rule)** For events $\mathcal{A}$ and $\mathcal{B}$, where $\mathbb{P}(\mathcal{A}), \mathbb{P}(\mathcal{B}) > 0$, 

$$
\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{B}) \cdot \mathbb{P}(\mathcal{A}|\mathcal{B})}{\mathbb{P}(\mathcal{A})}
$$

Rev. Thomas Bayes [1701-1761]

**Proof:** $\mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$
\[ P(\text{Ebola}|\text{High fever}) = \frac{P(\text{Ebola}) \cdot P(\text{High fever}|\text{Ebola})}{P(\text{High fever})} = \frac{10^{-8} \cdot 1}{0.15 \times 0.8 + 10^{-8} \times 1 + (0.85 - 10^{-8}) \times 0.1} \approx 7.4 \times 10^{-8} \]

\[ P(\text{Flu}|\text{High fever}) \approx 0.89 \]

\[ P(\text{Other}|\text{High fever}) \approx 0.11 \]

Most-likely a-posteriori outcome (MLA)
Bayes Rule – Example

Setting: An urn contains 6 balls:

• 3 red and 3 blue balls w/ probability $\frac{3}{4}$
• 6 red balls w/ probability $\frac{1}{4}$

We draw three balls at random from the urn.

All three balls are red. What is the probability that the remaining (undrawn) balls are all blue?
Sequential Process

Mixed

Not mixed

1/20

3R

2R1B

1R2B

3B

1/(6\choose 3)

Wanted: \( P(Mixed|3R) \)
Sequential Process

\[
\mathbb{P}(\text{Mixed}|3R) = \frac{\mathbb{P}(\text{Mixed}) \mathbb{P}(3R|\text{Mixed})}{\mathbb{P}(3R)} = \frac{\frac{3}{4} \times \frac{1}{20}}{\frac{3}{4} \times \frac{1}{20} + \frac{1}{4} \times 1} \approx 0.13
\]
Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

What would you do?
Monty Hall

Say you picked (without loss of generality) **Door 1**

```
Door 1
    /   \
  1/3   1/3
  |     |
Door 2 --- Door 3
          |   1
          |   1/2
          |   1
          |
          |
```

Car position

Open 2

Open 3
Monty Hall

\[
P(\text{Door 1} | \text{Open 3}) = \frac{P(\text{Door 1}) P(\text{Open 3} | \text{Door 1})}{P(\text{Open 3})} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times 1} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{2}} = \frac{1}{\frac{5}{6}} = \frac{6}{5} = \frac{1}{3}
\]

\[
P(\text{Door 2} | \text{Open 3}) = 1 - P(\text{Door 1} | \text{Open 3}) = \frac{2}{3}
\]
Monty Hall

Bottom line: Always swap!

Your choice