Foundations of Computer Science
Lecture 16

Conditional Probability

Updating a Probability when New Information Arrives
Conditional Probability Traps
Law of Total Probability

Last Time

1. Outcome-tree method for computing probability.

2. Probability and sets.
   - Probability space.
   - Event is a subset of outcomes.
   - Can get complex events using set (logical) operations.

3. Uniform probability space
   - Toss 10 coins. Each sequence (e.g. HTHHHTTTHH) has equal probability.
   - Roll 3 dice. Each sequence (e.g. (2,4,5)) has equal probability.
   - Probability of event $\sim$ event size.

4. Infinite probability space.
   - Toss a coin until you get heads (possibly never ending).

Today: Conditional Probability

1. New information changes a probability.

2. Definition of conditional probability from regular probability.

3. Conditional probability traps
   - Sampling bias.
   - Transposed conditional.

4. Law of total probability.
   - Probabilistic case-by-case analysis.

Flu Season

1. Chances a random person has the flu is about 0.01 (or 1%) (*prior* probability).

$$\text{Probability of flu} : \mathbb{P}[\text{flu}] \approx 0.01.$$

2. You have a slight fever – *new information.* Chances of flu "increase".

$$\text{Probability of flu } \textit{given} \text{ fever} : \mathbb{P}[\text{flu} \mid \text{fever}] \approx 0.4.$$

   - New information changes the prior probability to the *posterior* probability.
   - Translate posterior as "*After* you get the new information."

$\mathbb{P}[A \mid B]$ is the (updated) *conditional* probability of A, *given* the new information B.

3. Roommie has flu (more new information). Flu for sure, take counter-measures.

$$\text{Probability of flu } \textit{given} \text{ fever and roommie flu} : \mathbb{P}[\text{flu} \mid \text{fever AND roommie flu}] \approx 1.$$

Pop Quiz. Estimate these probabilities:

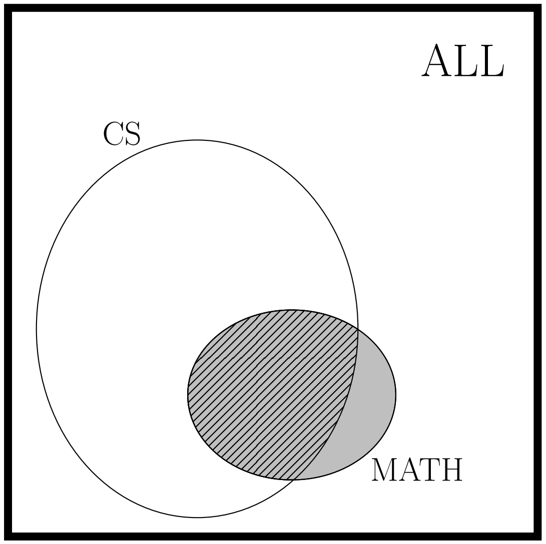
$\mathbb{P}[\text{Humans alive tomorrow}]$, $\mathbb{P}[\text{No Sun tomorrow}]$, $\mathbb{P}[\text{Humans alive tomorrow} \mid \text{No Sun tomorrow}]$.

CS, MATH and Dual CS-MATH Majors

5,000 students: **1,000** CS; **100** MATH; **80** dual MATH-CS.

Pick a random student:

$$\mathbb{P}[\text{CS}] = \frac{1000}{5000} = 0.2;$$
$$\mathbb{P}[\text{MATH}] = \frac{100}{5000} = 0.02;$$
$$\mathbb{P}[\text{CS AND MATH}] = \frac{80}{5000} = 0.016.$$



New information: student is MATH. What is $\mathbb{P}[\text{CS} \mid \text{MATH}]$?

- Effectively picking a random student from MATH.
- New probability of CS $\sim$ striped area $|\text{CS} \cap \text{MATH}|$.

$$\mathbb{P}[\text{CS} \mid \text{MATH}] = \frac{|\text{CS} \cap \text{MATH}|}{|\text{MATH}|} = \frac{80}{100} = 0.8.$$

MATH students are 4 times more likely to be CS majors than a random student.

Pop Quiz. What is $\mathbb{P}[\text{MATH} \mid \text{CS}]$? What is $\mathbb{P}[\text{CS} \mid \text{CS OR MATH}]$? **Exercise 16.2.**

Conditional Probability $\mathbb{P}[A \mid B]$

$\mathbb{P}[A \mid B]$ = frequency of outcomes known to be in B that are also in A.

n_B ooutcomes in event B when you repeat an experiment n times.

$$\mathbb{P}[B] = \frac{n_B}{n}.$$

Of the n_B outcomes in B, the number also in A is $n_{A\cap B}$,

$$\mathbb{P}[A \cap B] = \frac{n_{A\cap B}}{n}.$$

The frequency of outcomes in A among those outcomes in B is $n_{A\cap B}/n_B$,

$$\mathbb{P}[A \mid B] = \frac{n_{A\cap B}}{n_B} = \frac{n_{A\cap B}}{n} \times \frac{n}{n_B} = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

$$\mathbb{P}[A \mid B] = \frac{n_{A\cap B}}{n_B} = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A \text{ AND } B]}{\mathbb{P}[B]}$$

Chances of Rain Given Clouds

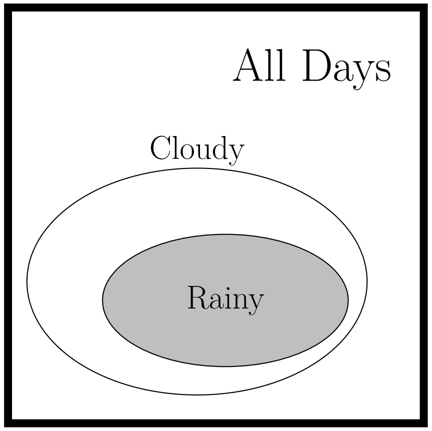
It is cloudy one in five days, $\mathbb{P}[\text{Clouds}] = \frac{1}{5}$. It rains one in seven days, $\mathbb{P}[\text{Rain}] = \frac{1}{7}$.

What are the chances of rain on a cloudy day?

$$\mathbb{P}[\text{Rain} \mid \text{Clouds}] = \frac{\mathbb{P}[\text{Rain} \cap \text{Clouds}]}{\mathbb{P}[\text{Clouds}]}.$$

$$\{\text{Rainy Days}\} \subseteq \{\text{Cloudy Days}\} \to \mathbb{P}[\text{Rain} \cap \text{Clouds}] = \mathbb{P}[\text{Rain}].$$

$$\mathbb{P}[\text{Rain} \mid \text{Clouds}] = \frac{\mathbb{P}[\text{Rain}]}{\mathbb{P}[\text{Clouds}]} = \frac{\frac{1}{7}}{\frac{1}{5}} = \frac{5}{7}.$$



5-times more likely to rain on a cloudy day than on a random day.

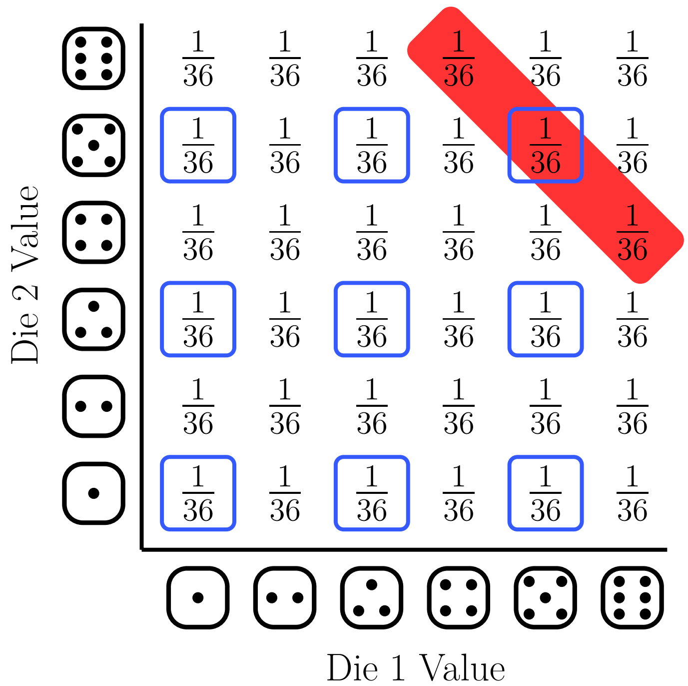
Crucial first step: identify the conditional probability. What is the "new information"?

$\mathbb{P}$[Sum of 2 Dice is 10 | Both are Odd]

Two dice have both rolled odd. What are the chances the sum is 10?

$$\mathbb{P}[\text{Sum is 10} \mid \text{Both are Odd}] = \frac{\mathbb{P}[(\text{Sum is 10}) \text{ AND } (\text{Both are Odd})]}{\mathbb{P}[\text{Both are Odd}]}$$

Probability Space



1. $\mathbb{P}[\text{Sum is 10}] = \frac{3}{36} = \frac{1}{12}$.
2. $\mathbb{P}[\text{Both are Odd}] = \frac{9}{36} = \frac{1}{4}$.
3. $\mathbb{P}[(\text{Sum is 10}) \text{ AND } (\text{Both are Odd})] = \frac{1}{36}$.
4. $\mathbb{P}[\text{Sum is 10} \mid \text{Both are Odd}] = \frac{1}{36} \div \frac{1}{4} = \frac{1}{9}$.

Pop Quiz. Compute $\mathbb{P}$[Both are Odd | Sum is 10]. Compare with $\mathbb{P}$[Sum is 10 | Both are Odd].

Computing a Conditional Probability

1: Identify that you need a conditional probability $\mathbb{P}[A \mid B]$.

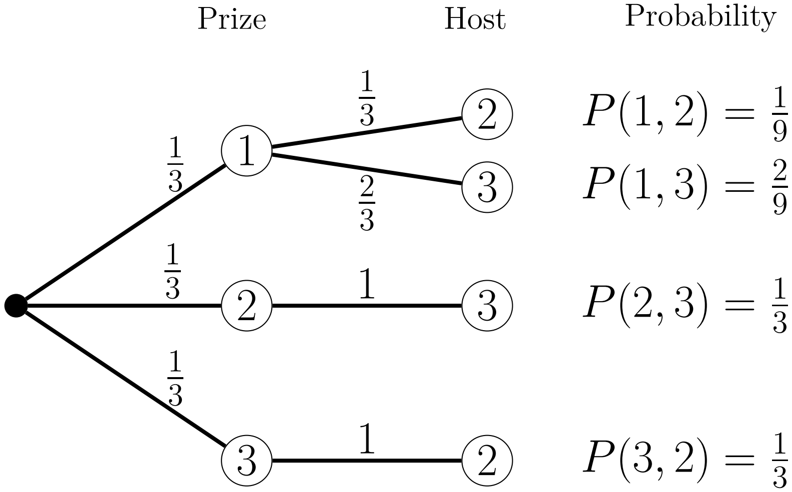
2: Determine the probability space $(\Omega, P(\cdot))$ using the outcome-tree method.

3: Identify the events A and B appearing in $\mathbb{P}[A \mid B]$ as subsets of Ω.

4: Compute $\mathbb{P}[A \cap B]$ and $\mathbb{P}[B]$.

5: Compute $\mathbb{P}[A \mid B] = \dfrac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.

Monty Prefers Door 3



Best strategy is always switch.
Winning outcomes: (2,3) or (3,2).

$$\mathbb{P}[\text{WinBySwitching}] = \tfrac{2}{3}.$$

Perk up if Monty opens door 2!

- Intuition: Why didn't Monty open door 3 if he prefers door 3?

$$\begin{aligned}
\mathbb{P}[\text{Win}|\text{Monty opens Door 3}] &= \frac{\mathbb{P}[\text{Win }\textsc{and}\text{ Monty opens Door 3}]}{\mathbb{P}[\text{Monty opens Door 3}]} \\
&= \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{9}} \\
&= \tfrac{3}{4}.
\end{aligned}$$

Your chances improved from $\frac{2}{3}$ to $\frac{3}{4}$!

A Pair of Boys

Your friends Ayfos, Ifar, Need and Niaz have two children each.
What is the probability of two boys? Answer: $\frac{1}{4}$.

New information:

1. Ayfos has at least one boy. (Answer: $\frac{1}{3}$.)
2. Ifar's older child is a boy. (Answer: $\frac{1}{2}$.)
3. One day you met Need on a walk with a boy. (Answer: $\frac{1}{2}$.)
4. Niaz is Clingon. Clingons always take a son on a walk if possible. One day, you met Niaz on a walk with a boy. (Answer: $\frac{1}{3}$.)

Now, what is the probability of two boys in each case?

It's the same question in each case, but with slightly different additional information.
You need conditional probabilities.

Conditional Probability Traps

These four probabilities are all different.

$$\mathbb{P}[A] \qquad \mathbb{P}\left[A \mid B\right] \qquad \mathbb{P}\left[B \mid A\right] \qquad \mathbb{P}\left[A \text{ AND } B\right]$$

Don't use one when you should use another.

Sampling Bias: Using $P[A]$ instead of $P[A \mid B]$

$$\mathbb{P}[\text{Voter will vote Republican}] \approx \tfrac{1}{2}.$$

Ask Apple™ to call up i-Phone™ users to see how they will vote.

$$\mathbb{P}[\text{Voter will vote Republican} \mid \text{Voter has an i-Phone}] \gg \tfrac{1}{2}. \qquad \text{(Why?)}$$

This has trapped many US election-pollers. For a famous example, Google™ "Dewey Defeats Truman."

Transposed Condtional: Using $P[B \mid A]$ instead of $P[A \mid B]$

Famous Lombard study on the riskiest profession: **Student!** Lombard confused:

$$\mathbb{P}[\text{Student} \mid \text{Die Young}] \qquad \text{with} \qquad \mathbb{P}\left[\text{Die Young} \mid \text{Student}\right]$$

The LAME Test and Transposed Conditionals

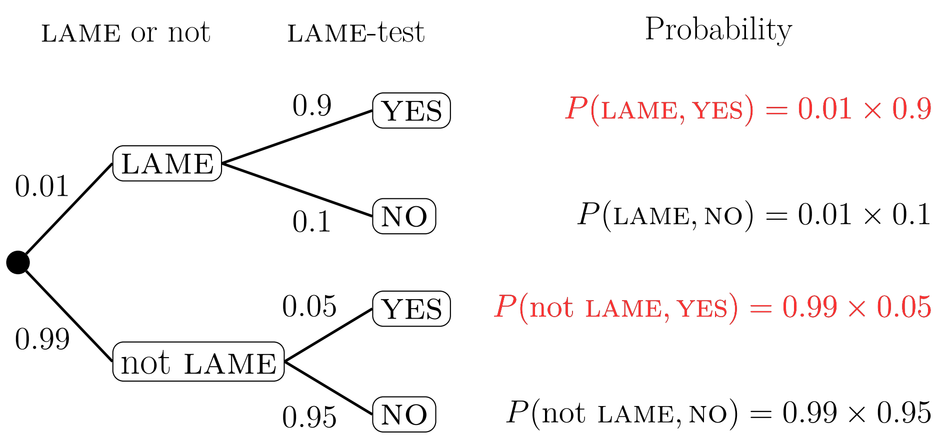
If you are LAME, the test makes a mistake in only 10% of cases.
If you are not LAME, the test makes a mistake in only 5% of cases.

You get tested positive. What are the chances you are LAME?

If you are not LAME, the test wouldn't make a mistake. So you are likely LAME.

It's *wrong* to look at $\mathbb{P}[\text{positive} \mid \text{not LAME}]$. We need $\mathbb{P}[\text{not LAME} \mid \text{positive}]$.

$$\mathbb{P}\left[\text{not LAME} \mid \text{YES}\right] = \frac{\mathbb{P}[\text{not LAME AND YES}]}{\mathbb{P}[\text{YES}]} = \frac{0.99 \times 0.05}{0.99 \times 0.05 + 0.9 \times 0.01} \approx 85\%.$$



The (accurate) test says YES but the chances are 85% that you are not LAME!

- You are LAME (rare) plus the test was right (likely)
- You are not LAME (very likely) plus the test got it wrong (rare). Wins!

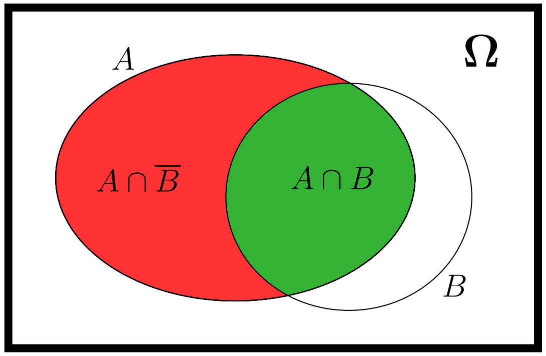
Total Probability: Case by Case Probability

Two types of outcomes in any event A:

- The outcomes in B (green);
- The outcomes not in B (red).

$$\mathbb{P}[A] = \mathbb{P}[A \cap B] + \mathbb{P}[A \cap \overline{B}]. \qquad (*)$$

(Similar to sum rule from counting.)



From the definition of conditional probability:

$$\begin{aligned}\mathbb{P}[A \cap B] &= \mathbb{P}[A \text{ AND } B] = \mathbb{P}[A \mid B] \times \mathbb{P}[B];\\ \mathbb{P}[A \cap \overline{B}] &= \mathbb{P}[A \text{ AND } \overline{B}] = \mathbb{P}[A \mid \overline{B}] \times \mathbb{P}[\overline{B}].\end{aligned}$$

Plugging these into $(*)$, we get a **FUNDAMENTAL** result for case by case analysis:

Law of Total Probability

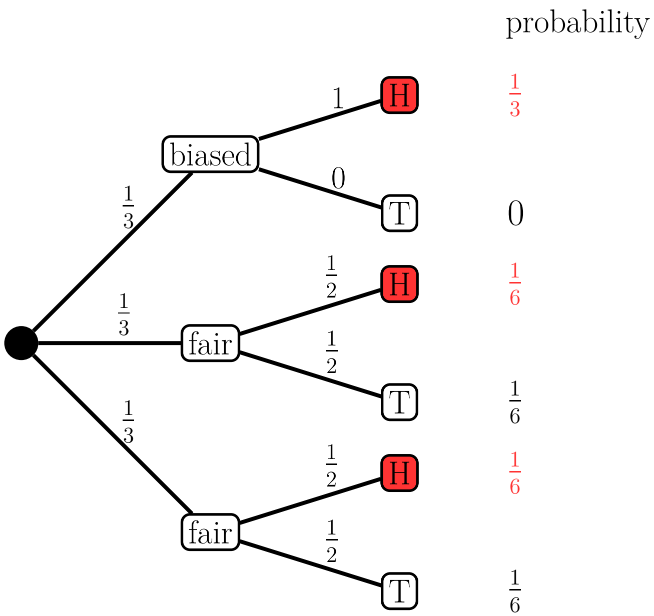
$$\mathbb{P}[A] = \mathbb{P}[A \mid B] \cdot \mathbb{P}[B] + \mathbb{P}[A \mid \overline{B}] \cdot \mathbb{P}[\overline{B}].$$

(Weight conditional probabilities for each case by probabilities of each case and add.)

Three Coins: Two Are Fair, One is 2-Headed

Pick a random coin and flip. What is the probability of H?

Outcome-Tree Method



$\mathbb{P}[\text{H}] = \frac{1}{6} + \frac{1}{6} + \frac{1}{3} = \frac{2}{3}$.

Total Probability

Case 1. B : You picked one of the fair coins
Case 2. $\overline{B}$: You picked the two-headed coin

$$\mathbb{P}[\text{H}] = \underset{\frac{1}{2}}{\mathbb{P}[\text{H} \mid B]} \cdot \underset{\frac{2}{3}}{\mathbb{P}[B]} + \underset{1}{\mathbb{P}[\text{H} \mid \overline{B}]} \cdot \underset{\frac{1}{3}}{\mathbb{P}[\overline{B}]}$$

$$= \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}$$

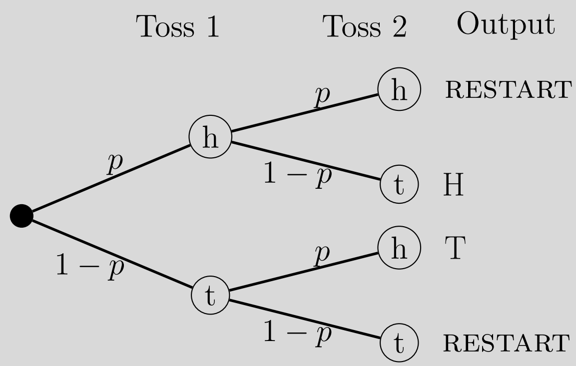
$$= \frac{2}{3}.$$

Exercise. A box has 10 coins: 6 fair and 4 biased (probability of heads $\frac{2}{3}$). What is $\mathbb{P}[\text{2 heads}]$ in each case?

(a) Pick a single random coin and flip it 3 times.
(b) Flip 3 times. For each flip, pick a random coin, flip it and then put the coin back.

Fair Toss from Biased Coin (*unknown* probability p of heads)?

- Make two tosses of the biased coin.
  (Lower case 'h' and 't' denote the outcomes of a toss.)
- If you get 'ht' output H; 'th' output T; otherwise RESTART.
- $P(\text{'ht'}) = P(\text{'th'}) = p(1-p)$.
- This suggests that an H is as likely as a T.



By the law of total probability (3 cases),

$$\mathbb{P}[\text{H}] = \underbrace{\mathbb{P}[\text{H} \mid \text{RESTART}]}_{\mathbb{P}[\text{H}]} \cdot \underbrace{\mathbb{P}[\text{RESTART}]}_{p^2+(1-p)^2} + \underbrace{\mathbb{P}[\text{H} \mid \text{'ht'}]}_{1} \cdot \underbrace{\mathbb{P}[\text{'ht'}]}_{p(1-p)} + \underbrace{\mathbb{P}[\text{H} \mid \text{'th'}]}_{0} \cdot \underbrace{\mathbb{P}[\text{'th'}]}_{p(1-p)}$$

$$= \mathbb{P}[\text{H}](p^2 + (1-p)^2) + p(1-p)$$

Solve for $\mathbb{P}[\text{H}]$

$$\mathbb{P}[\text{H}] = \frac{p(1-p)}{1-(p^2+(1-p)^2)} = \frac{p(1-p)}{2p-2p^2} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2}.$$

(You can also solve this problem using an infinite outcome tree and computing an infinite sum.)