CS 70 Su23: Lecture 16

Conditional probability

Conditional probability

How do we “add knowledge” to our calculations?

- Some examples. What’s the probability of:
    - finding someone who is taking CS 70?
        - in Dwinelle?
    - finding your keys in your room?
        - if you know they’re in your closet?
    - getting the card you want?
        - if a bunch of cards have already been drawn?

Given some information, how does that affect the probability of an event?

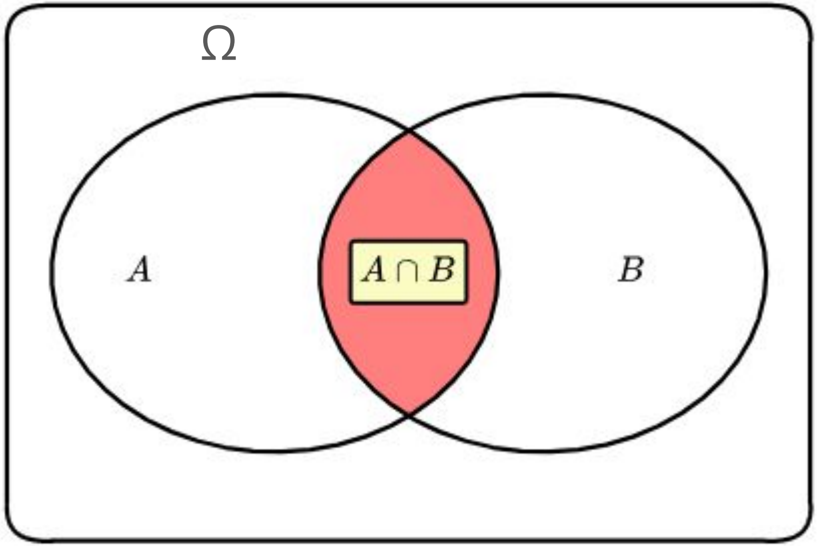
It’s like probability, but… conditional

Ω = our universe

A, B = events

$$\Pr[A] = \frac{|A|}{|\Omega|} \quad \Pr[B] = \frac{|B|}{|\Omega|}$$

$$\Pr[A \cap B] = \frac{|A \cap B|}{|\Omega|}$$



What is the probability of A if we know B is going to happen?

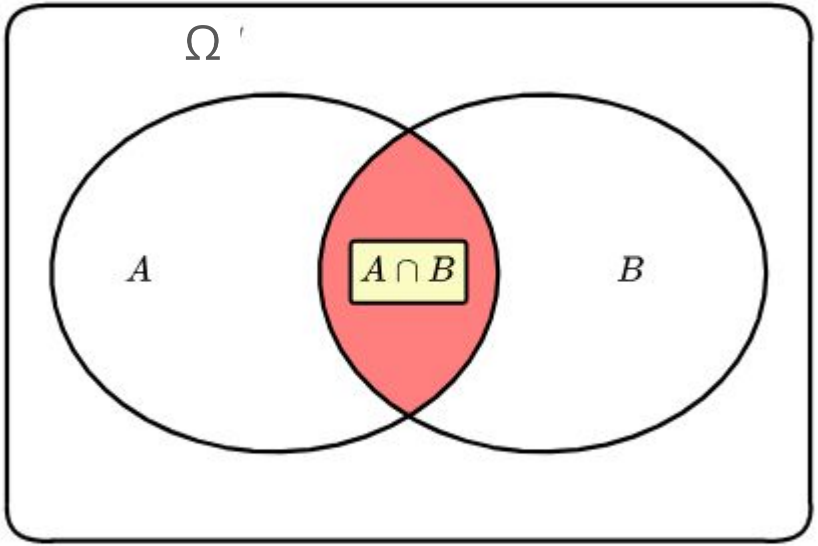
It’s like probability, but… conditional

Notation:

- $\Pr[A|B]$. “The probability of A given B”

$$\Pr[A|B] = \frac{|A \cap B|}{|B|}$$

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$



Berkeley
UNIVERSITY OF CALIFORNIA

Conditional probability

[link to visualization]

Bayes's Theorem

It's just rearranging the conditional probability formula:

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} \qquad \Pr[B|A] = \frac{\Pr[A \cap B]}{\Pr[A]}$$

$$\Pr[A \cap B] = \Pr[A|B] \times \Pr[B] = \Pr[B|A] \times \Pr[A]$$

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

If you know Pr[B|A], Pr[A], and Pr[B], you can compute Pr[A|B]

- this may not be immediately obvious!
- lets us do some interesting things

Inference

- Structured as follows:
  - We have a **prior distribution** Pr[A]
  - We have a **likelihood estimation** Pr[B|A]
  - We **observe** some event in B
  - Use Bayes's Theorem to update our **posterior distribution** Pr[A|B]

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

Bayes theorem and Bayes rule are the same thing (also note sometimes the apostrophe gets omitted for no real good reason)

Example 1: Counterfeits

We want to detect counterfeit bills.

- We have a test to see if a bill is fake (result is “fake” or “not fake”):
    - When used on a fake bill, **90%** of the time, it says “fake”
        - 10% of the time, it says “not fake”
    - When used on a real bill, **80%** of the time, it says “not fake”
        - 20% of the time, it says “fake”
- We know 5% of the bills out there are fake

If we test a random bill and our test says “fake”, what is the probability the bill is fake?

Example 1: Counterfeits

Let's define some events:

- A = a bill is fake
  - $\bar{A}$ = a bill is real (this is the complement of A)
- B = the test results in "fake"

We are trying to find $\Pr[A|B]$.

What we know:

- $\Pr[A] = 0.05$
- $\Pr[B|A] = 0.9$
- $\Pr[B|\bar{A}] = 0.2$

Berkeley
UNIVERSITY OF CALIFORNIA

Example 1: Counterfeits

Time for Bayes's rule!

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

... but what is Pr[B]?

Example 1: Counterfeits

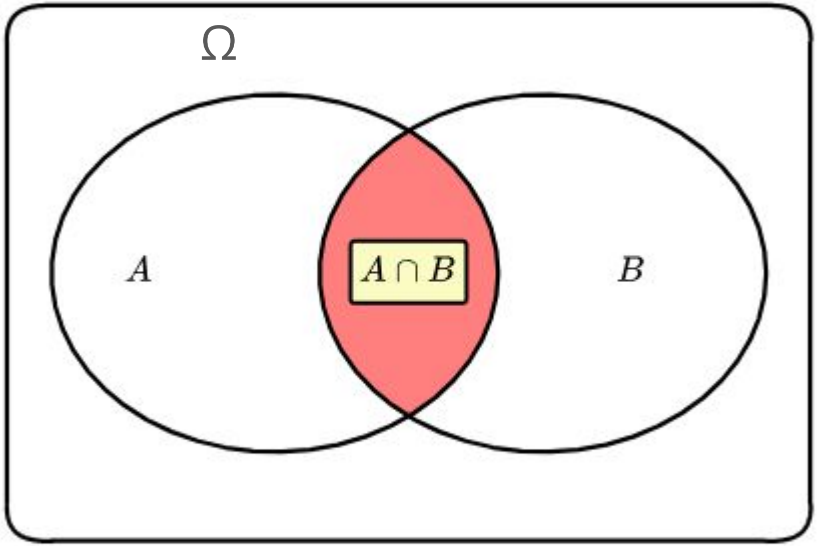
Observation:

- $\Pr[B] = \Pr[A \cap B] + \Pr[\bar{A} \cap B]$
  - $\bar{A}$ is the same as $\Omega \setminus A$
- Can we take advantage of this?

More Bayes rule:

$\Pr[A \cap B] = \Pr[B|A] \times \Pr[A]$

$\Pr[\bar{A} \cap B] = \Pr[B|\bar{A}] \times \Pr[\bar{A}]$



Example 1: Counterfeits

Time for Bayes's rule!

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

$$= \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B|A] \times \Pr[A] + \Pr[B|\bar{A}] \times \Pr[\bar{A}]} = \frac{0.9 \times 0.05}{0.9 \times 0.05 + 0.2 \times 0.95}$$

≈ 0.19

If a random bill tests as fake, there's less than a 1 in 5 chance it's actually fake!

Berkeley
UNIVERSITY OF CALIFORNIA

Example 2: Biased coins

Your friend has two indistinguishable coins. One comes up heads 80% of the time, while the other comes up heads 30% of the time.

You pick a coin at random and flip it, and it comes up heads. Which coin did you (probably) pick?

Example 2: Biased coins

Let’s define some events:

- A = you picked the 80% coin
  - $\bar{A}$ = you picked the 30% coin (this is the complement of A)
- B = the coin you flipped came up heads

We are trying to find $\Pr[A|B]$.

What we know:

- $\Pr[A] = 0.5$
- $\Pr[B|A] = 0.8$
- $\Pr[B|\bar{A}] = 0.3$

Example 2: Biased coins

Time for Bayes's rule!

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

… but what is Pr[B]?

Example 2: Biased coins

Time for Bayes's rule!

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B]}$$

$$= \frac{\Pr[B|A] \times \Pr[A]}{\Pr[B|A] \times \Pr[A] + \Pr[B|\bar{A}] \times \Pr[\bar{A}]} = \frac{0.8 \times 0.5}{0.8 \times 0.5 + 0.3 \times 0.5}$$

$$= \frac{8}{11}$$

Roughly a 73% chance you picked the 80% coin, 27% chance you picked the 30% coin

Aside: meta point

Manipulating the equations is not the hard part, setting up the events is!

When you see complicated word problems, always define your events carefully, and define their probabilities.

Total Probability Rule

$\Pr[B] = \Pr[A \cap B] + \Pr[\bar{A} \cap B]$

$= \Pr[B|A] \times \Pr[A] + \Pr[B|\bar{A}] \times \Pr[\bar{A}]$

Let's generalize this!

Let $A_1, A_2, \ldots, A_n$ **partition** the sample space Ω:

- $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$
- $\forall\ i \neq j, A_i \cap A_j = \varnothing$ (they are mutually exclusive; the sets are disjoint)

$\Pr[B] = \sum_i \Pr[B|A_i] \times \Pr[A_i]$

Berkeley
UNIVERSITY OF CALIFORNIA

Inference

- Structured as follows:
  - We have a **prior distribution** Pr[A]
  - We have a **likelihood estimation** Pr[B|A]
  - We **observe** some event in B
  - Use **Bayes's Theorem** and **Total Probability Rule** to compute our **posterior distribution** Pr[A|B]

$$\Pr[A|B] = \frac{\Pr[B|A] \times \Pr[A]}{\sum_i \Pr[B|A_i] \times \Pr[A_i]}$$

Example 3: Rock paper scissors

You are playing against one of three opponents. There is a 20% chance you'll face the first one, a 50% you'll face the second one, and a 30% chance you'll face the third one. The first opponent throws Rock 60% of the time, the second throws Rock 40% of the time, and the third throws Rock 33% of the time.

In the first round, your opponent throws Rock. What's the likelihood that you will win the next round if you throw Paper?

Example 3: Rock paper scissors

Defining events:

- Prior distribution
  - Likelihood of opponent
- Likelihood estimation
  - How likely each opponent is to throw Rock
- Observation
  - they threw Rock
- Posterior distribution
  - how likely it is we're facing each opponent

Example 3: Rock paper scissors

Let's define some events:

- A_i = you're against opponent i
- B = your opponent threw Rock

We are trying to find Pr[A|B].

What we know:

- $Pr[A_1] = 0.2$, $Pr[A_2] = 0.5$, $Pr[A_3] = 0.3$,
- $Pr[B|A_1] = 0.6$
- $Pr[B|A_2] = 0.4$
- $Pr[B|A_3] = 0.33$

Berkeley
UNIVERSITY OF CALIFORNIA

Example 3: Rock paper scissors

Time for Bayes's rule!

$$\Pr[A_1|B] = \frac{\Pr[B|A_1] \times \Pr[A_1]}{\Pr[B]}$$

$$= \frac{\Pr[B|A_1] \times \Pr[A_1]}{\sum_i \Pr[B|A_i] \times \Pr[A_i]} = \frac{0.6 \times 0.2}{0.6 \times 0.2 + 0.4 \times 0.5 + 0.33 \times 0.3}$$

$$= \frac{6}{21}$$

$\Pr[A_2|B] = 10/21$

$\Pr[A_3|B] = 5/21$

Berkeley UNIVERSITY OF CALIFORNIA

Example 3: Rock paper scissors

What's the likelihood that you'll win the next round if you throw Paper?

- C = your opponent throws Rock on round 2

We are trying to find Pr[C|B].

New "priors":

- $Pr[A_1|B] = 6/21$, $Pr[A_2|B] = 10/21$, $Pr[A_3|B] = 5/21$

Likelihood estimates remain the same (opponents don't switch it up):

- $Pr[C|A_1, B] = 0.6$
- $Pr[C|A_2, B] = 0.4$
- $Pr[C|A_3, B] = 0.33$

Berkeley
UNIVERSITY OF CALIFORNIA

Example 3: Rock paper scissors

$$\Pr[C|B] = \textstyle\sum_i \Pr[C|A_i, B] \times \Pr[A_i|B]$$

$$= 0.6 \times 6/21 + 0.4 \times 10/21 + 0.33 \times 5/21$$

$$= 6/35 + 4/21 + 5/63$$

$$= 139/315$$

There is roughly a 44% chance you'll win the second round by throwing Paper

Berkeley
UNIVERSITY OF CALIFORNIA

Example 4: Nikki likes chess

You're about to play chess against a randomly-chosen instructor. If Victor is chosen, your chances of winning are 90%. If Nate is chosen, your chances of winning are 50%. If Nikki is chosen, your chances of winning are 10%.

What is your overall likelihood of winning?

- A = you win
- B_1 = you face Victor, B_2 = you face Nate, B_3 = you face Nikki

Total probability:

$\Pr[A] = \sum_i \Pr[A|B_i] \times \Pr[B_i] = ⅓ (0.9 + 0.5 + 0.1) = 0.5$

Example 5: Biased coins again

You have two coins, one that flips heads 80% of the time, and one that flips heads 30% of the time. You randomly pick one of the coins and flip it, then you flip the other coin.

What is the probability that the first flip is heads?

Let A_i = the i^{th} flip is heads, B_i = you picked the i^{th} coin

$Pr[A_1] = Pr[A_1|B_1] \times Pr[B_1] + Pr[A_1|B_2] \times Pr[B_2]$

$= 0.8 \times 0.5 + 0.3 \times 0.5$

$= 0.55$

Berkeley
UNIVERSITY OF CALIFORNIA

Example 5: Biased coins again

What is the probability that the second flip is heads?

$Pr[A_2] = Pr[A_1 \cap A_2] + Pr[\bar{A}_1 \cap A_2]$

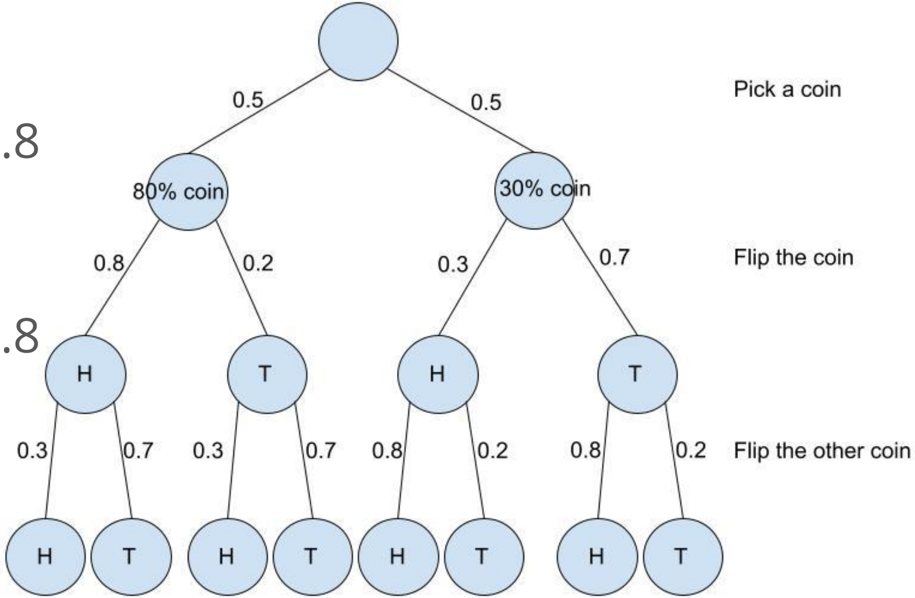
$Pr[A_1 \cap A_2] = 0.5 \times 0.8 \times 0.3 + 0.5 \times 0.3 \times 0.8$

$= 0.24$

$Pr[\bar{A}_1 \cap A_2] = 0.5 \times 0.2 \times 0.3 + 0.5 \times 0.7 \times 0.8$

$= 0.31$

$Pr[A_2] = 0.55$



Example 5: Biased coins again

$\Pr[A_1] = \Pr[A_2]$. Why?

Knowing whether the first flip is heads affects our guess about the second flip.
That is, $\Pr[A_2 | A_1] \neq \Pr[A_2]$.

But if we're looking only at $\Pr[A_2]$, it's **unconditional**. In other words, here, we *don't* have knowledge of the previous flip.

In fact, it ends up being a *symmetric* argument for why the unconditional probabilities are equal.

Think about the question rephrased: "You flip both coins at the same time. What's the probability that the second coin is heads?"

If you don't get to observe the result of the first flip, the order in which you flip them doesn't really matter!

Berkeley
UNIVERSITY OF CALIFORNIA

Next class: more conditional probability

There's a special name for when $\Pr[A_2 | A_1] \neq \Pr[A_2]$. We'll talk about this in more depth (independence).

Nikki will be back!