Monday is a holiday (MLK day)

- No lecture
  - CCG due Wed 10:30.
- No office hours
  - a few T/W will have an extra TA.

Bayes' Rule

CSE 312 Winter 25
Lecture 6

Last Time

Conditional Probability

For an event B, with $\mathbb{P}(B) > 0$,
the "Probability of A conditioned on B" is

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

Roll a fair 6-sided die. Let A be "the die roll is a 4", B be "the die roll is an even number."

$P(A|B) = 1/6 / 1/2 = 1/3$

$P(B|A) = 1/6 / 1/6 = 1$

Conditioning...

Let A be "the red die is 5"
Let B be "the sum is 4"
Let C be "the blue die is 3"

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning…

Let A be "the red die is 5"
Let B be "the sum is 4"
Let C be "the blue die is 3"

$\mathbb{P}(A|B)$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning…

Let A be "the red die is 5"
Let B be "the sum is 4"
Let C be "the blue die is 3"

$\mathbb{P}(A|B)$

$\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$
$\mathbb{P}(B) = 3/36$

$$P(A|B) = \frac{0}{3/36}$$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	$\boxed{(1,3)}$	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	$\boxed{(2,2)}$	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	$\boxed{(3,1)}$	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning…

Let A be "the red die is 5"
Let B be "the sum is 4"
Let C be "the blue die is 3"

$\mathbb{P}(A|C)$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning…

Let A be "the red die is 5"
Let B be "the sum is 4"
Let C be "the blue die is 3"

$\mathbb{P}(A|C)$

$\mathbb{P}(A \cap C) = 1/36$
$\mathbb{P}(C) = 6/36$

$$P(A|C) = \frac{1/36}{6/36}$$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning Practice

- Red die 6 conditioned on sum 7
- Red die 6 conditioned on sum 9
- Sum 7 conditioned on red die 6

Take a few minutes to work on this with the people around you! (also on your handout)

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning Practice

A ~ Red die 6

B ~ Sum is 7

$\frac{1}{36}$ $\frac{6}{36}$

$\mathbb{P}(A|B)$

$= \mathbb{P}(A \cap B)/P(B)$

$=$

$= 1/6$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning Practice

A ~ Red die 6

C ~ Sum is 9

$\mathbb{P}(A|C)$ $\frac{1}{36}$ $\frac{4}{36}$

$= \mathbb{P}(A \cap C)/P(C)$

$=$

$= 1/4$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning Practice

B ~ Sum is 7

A ~ Red die is 6

1/36   5/36

$\mathbb{P}(B|A)$

$= \mathbb{P}(B \cap A)/P(A)$

$=$

$= 1/6$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditioning Practice

Red die 6 conditioned on sum 7 1/6

Red die 6 conditioned on sum 9 1/4

Sum 7 conditioned on red die 6 1/6

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Direction Matters

Are $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$ the same?

Direction Matters

No! $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$ are different quantities.

$\mathbb{P}$("traffic on the highway" | "it's snowing") is close to 1

$\mathbb{P}$("it's snowing" | "traffic on the highway") is much smaller; there many other times when there is traffic on the highway

It's a lot like implications – order can matter a lot!

(but there are some A, B where the conditioning doesn't make a difference)

Bayes’ Rule

Wonka Bars

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

You want to get a golden ticket. You could buy a 1000-or-so of the bars until you find one, but that's expensive…you've got a better idea!

You have a test – a very precise scale you've bought.

- If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.
- If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

If you pick up a bar and it alerts, what is the probability you have a golden ticket?

Willy Wonka

Fill out the poll everywhere so
Robbie knows how long to explain
Go to pollev.com/robbie

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.

If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

You pick up a bar and it alerts, what is the probability you have a golden ticket?

Which of these is closest to the right answer?

A. 0.1%
B. 10%
C. 50%
D. 90%
E. 99%
F. 99.9%

46

43

Conditioning

Let S be the event that the Scale alerts you

Let G be the event your bar has a Golden ticket.

What conditional probabilities are each of these?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.

If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

You pick up a bar and it alerts, what is the probability you have a golden ticket?

Conditioning

Estimate $\mathbb{P}(S)$.

Let S be the event that the **S**cale alerts you

Let G be the event your bar has a **G**olden ticket.

What conditional probabilities are each of these?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.	$\mathbb{P}(G)$
If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.	$\mathbb{P}(S\|G)$
If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.	$\mathbb{P}(S\|\bar{G})$
You pick up a bar and it alerts, what is the probability you have a golden ticket?	$\mathbb{P}(G\|S)$

Reversing the Conditioning

All of our information conditions on whether G happens or not – does your bar have a golden ticket or not?

But we're interested in the "reverse" conditioning. We know the scale alerted us – we know the test is positive – but do we have a golden ticket?

Bayes' Rule

Bayes' Rule

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

Bayes' Rule

Bayes' Rule

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

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{\mathbb{P}(S|G) \cdot \mathbb{P}(G)}{\mathbb{P}(S)}$$

Bayes' Rule

Bayes' Rule

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

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{.999 \cdot .001}{\mathbb{P}(S)}$$

Filling In

What's $\mathbb{P}(S)$?

We'll use a trick called "the law of total probability":

$$\mathbb{P}(S) = \mathbb{P}(S|G) \cdot \mathbb{P}(G) + \mathbb{P}(S|\bar{G}) \cdot \mathbb{P}(\bar{G})$$

$$= 0.999 \cdot .001 + .01 \cdot .999$$

$$= .010989$$

Law of Total Probability

Let $A_1, A_2, \ldots, A_k$ be a **partition** of Ω.

A partition of a set S is a family of subsets $S_1, S_2, \ldots, S_k$ such that:

$S_i \cap S_j = \emptyset$ for all i, j and

$S_1 \cup S_2 \cup \cdots \cup S_k = S$.

i.e. every element of Ω is in exactly one of the A_i.

Law of Total Probability

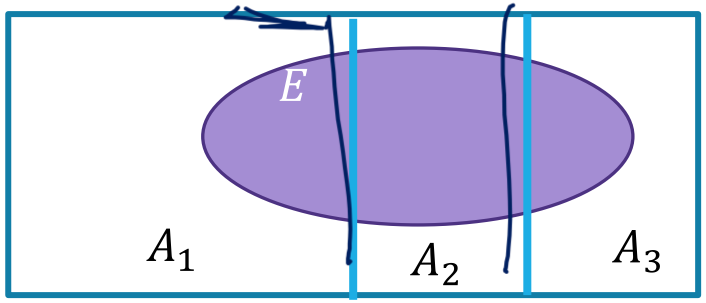
Law of Total Probability

Let $A_1, A_2, \ldots, A_k$ be a **partition** of Ω.

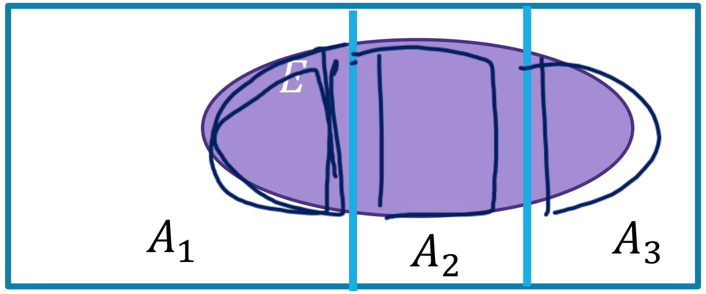
For any event E,

$$\mathbb{P}(E) = \sum_{\text{all } i} \mathbb{P}(E|A_i)\mathbb{P}(A_i)$$

Ω, split into partition A_1, A_2, A_3 with event E inside.



Why?



The Proof is actually pretty informative on what's going on.

$\sum_{\text{all } i} \mathbb{P}(E|A_i)\mathbb{P}(A_i)$

$= \sum_{\text{all } i} \frac{\mathbb{P}(E \cap A_i)}{\mathbb{P}(A_i)} \cdot \mathbb{P}(A_i)$ (definition of conditional probability)

$= \sum_{\text{all } i} \mathbb{P}(E \cap A_i)$

$= \mathbb{P}(E)$

The A_i partition Ω, so $E \cap A_i$ partition E. Then we just add up those probabilities.

Ability to add follows from the "countable additivity" axiom.

Bayes Rule

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{.999 \cdot .001}{.010989}$$

Solving $\mathbb{P}(G|S) = \frac{1}{11}$, i.e. about 0.0909.

Only about a 10% chance that the bar has the golden ticket!

Wait a minute…

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.
If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.
If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

That doesn't fit with many of our guesses. What's going on?

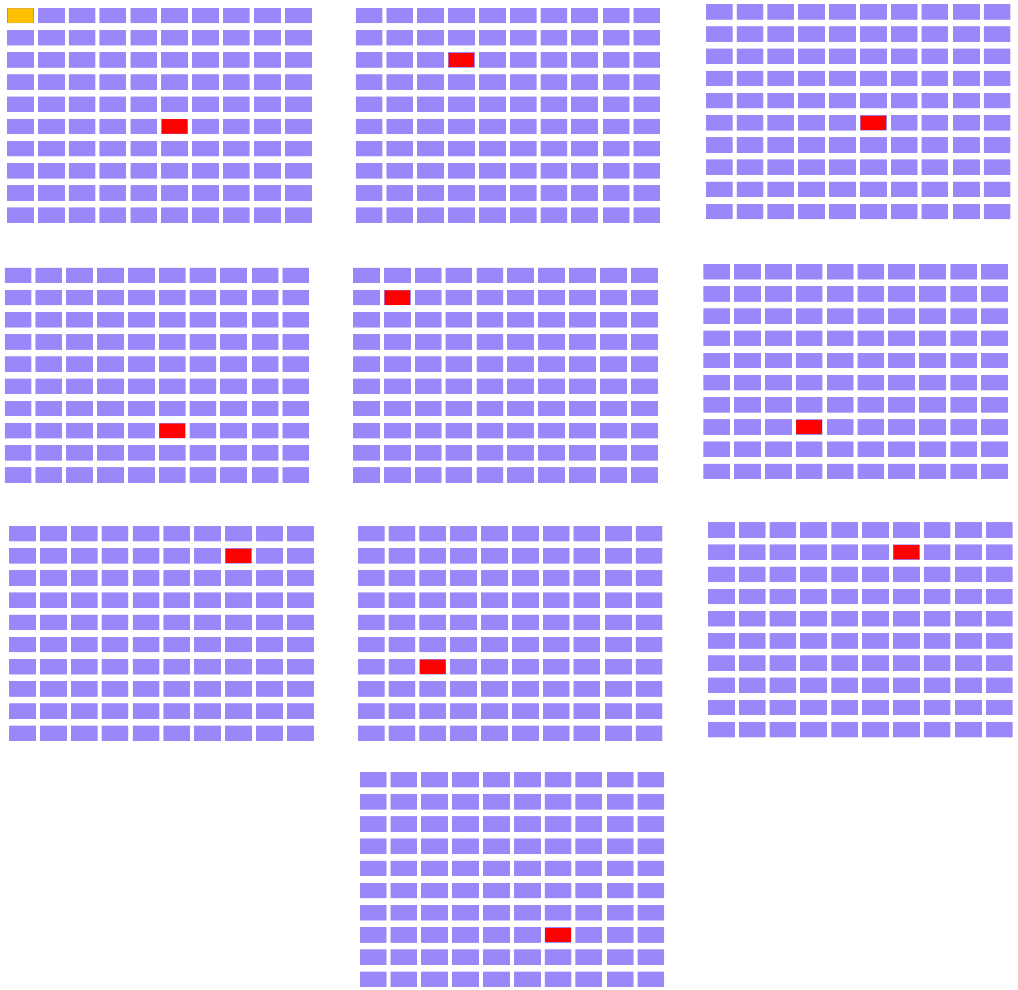
Instead of saying "we tested one and got a positive" imagine we tested 1000. **ABOUT** how many bars of each type are there?

(about) 1 with a golden ticket 999 without. Let's say those are exactly right.

Let's just say that one golden is truly found.

(about) 1% of the 999 without would be a positive. Let's say it's exactly 10.

Visually



Gold bar is the one (true) golden ticket bar.
Purple bars don't have a ticket and tested negative.
Red bars don't have a ticket, but tested positive.

The test is, in a sense, doing really well.
It's almost always right.

The problem is it's also the case that the correct answer is almost always "no."

Updating Your Intuition

🔥 Take 1: The test is **actually good** and has VASTLY increased our belief that there IS a golden ticket when you get a positive result.

If we told you "your job is to find a Wonka Bar with a golden ticket" without the test, you have 1/1000 chance, with the test, you have (about) a 1/11 chance. That's (almost) 100 times better!

This is actually a huge improvement!

Updating Your Intuition

🔥 Take 2: Humans are really bad at intuitively understanding very large or very small numbers.

When I hear "99% chance", "99.9% chance", "99.99% chance" they all go into my brain as "well that's basically guaranteed" And then I forget how many 9's there actually were.

But the number of 9s matters because they end up "cancelling" with the "number of 9's" in the population that's truly negative. We'll talk about this a little more on Friday in the applications.

Updating Your Intuition

🔥 Take 3: View tests as updating your beliefs, not as revealing the truth.

Bayes' Rule says that $\mathbb{P}(B|A)$ has a factor of $\mathbb{P}(B)$ in it. You have to translate "The test says there's a golden ticket" to "the test says you should increase your estimate of the chances that you have a golden ticket."

A test takes you from your "prior" beliefs of the probability to your "posterior" beliefs.

More Bayes Practice

A contrived example

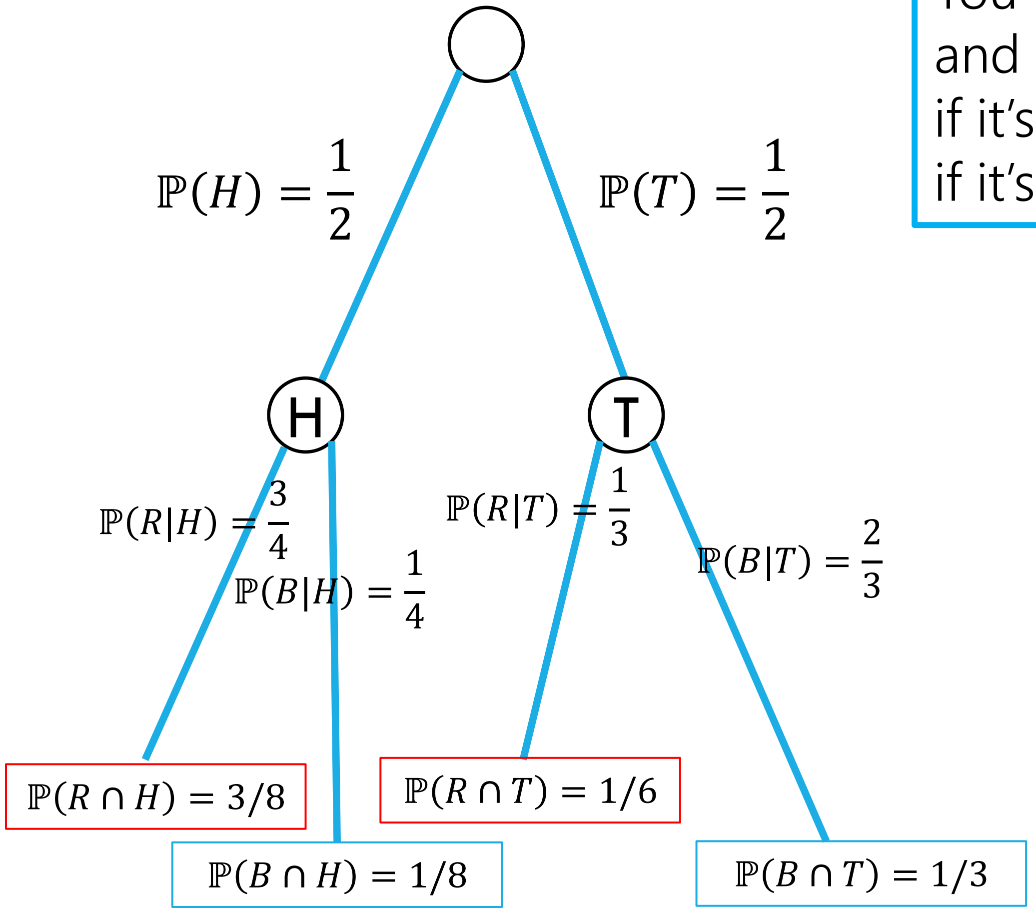
You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.

You will flip a fair coin; if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.

Let B be you draw a blue marble. Let T be the coin is tails.

What is $\mathbb{P}(B|T)$ what is $\mathbb{P}(T|B)$?

Updated Sequential Processes

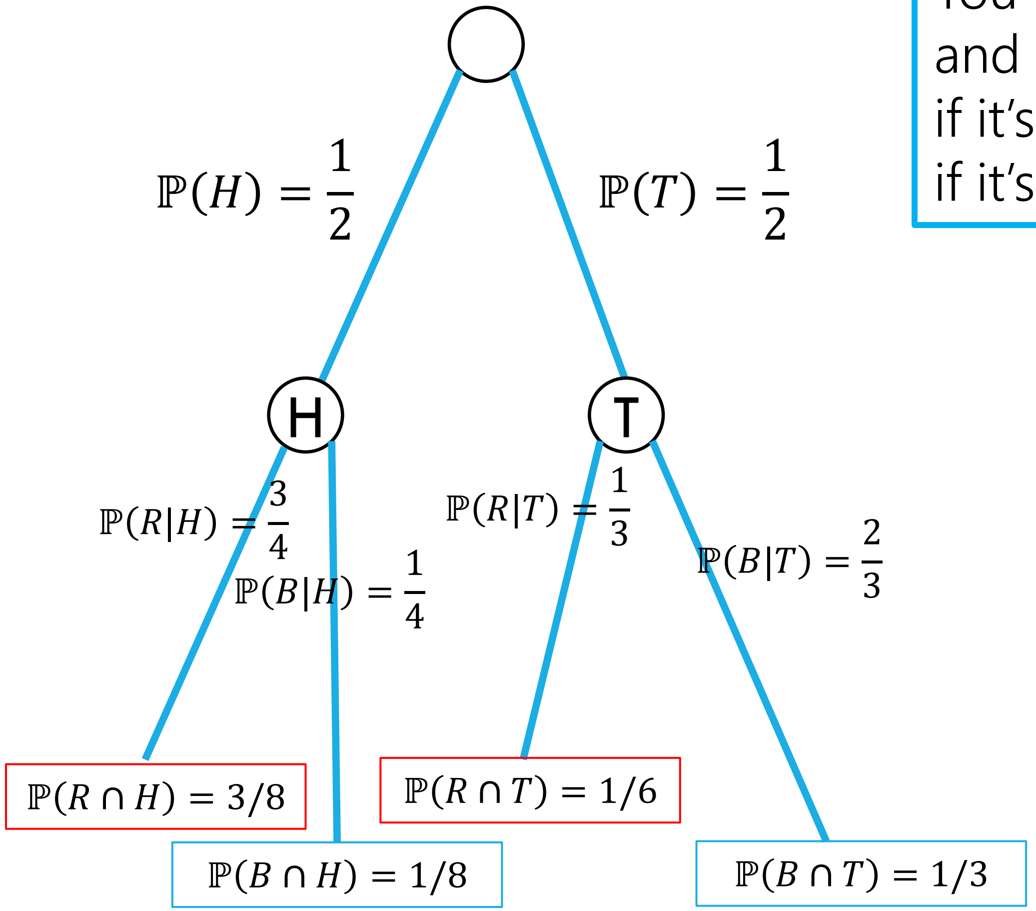


You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.

For sequential processes with probability, at each step multiply by
$\mathbb{P}(\text{next step} \mid \text{all} \cap \text{prior} \cap \text{steps})$

Updated Sequential Processes

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.



For sequential processes with probability, at each step multiply by
$\mathbb{P}(\text{next step } | \text{all} \cap \text{prior} \cap \text{steps})$

$$\mathbb{P}(B|T) = 2/3;\ \mathbb{P}(B) = \frac{1}{8} + \frac{1}{3} = \frac{11}{24}$$

Flipping the conditioning

What about $\mathbb{P}(T|B)$?

You have three red marbles and one blue marble in your left pocket,
and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.

Pause, what's your intuition?

Is this probability

A. less than ½

B. equal to ½

C. greater than ½

Flipping the conditioning

What about $\mathbb{P}(T|B)$?

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.

Pause, what's your intuition?

Is this probability

A. less than ½

B. equal to ½

C. greater than ½

The right (tails) pocket is far more likely to produce a blue marble if picked than the left (heads) pocket is. Seems like $\mathbb{P}(T|B)$ should be greater than ½.

Flipping the conditioning

What about $\mathbb{P}(T|B)$?

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.

Bayes' Rule says:

$$\mathbb{P}(T|B) = \frac{\mathbb{P}(B|T)\mathbb{P}(T)}{\mathbb{P}(B)}$$

$$= \frac{\frac{2}{3} \cdot \frac{1}{2}}{11/24} = 8/11$$

The Technical Stuff

Proof of Bayes' Rule

$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ by definition of conditional probability

Now, imagining we get $A \cap B$ by conditioning on A, we should get a numerator of $\mathbb{P}(B|A) \cdot \mathbb{P}(A)$

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

As required.

A Technical Note

After you condition on an event, what remains is a probability space.

With B playing the role of the sample space,

$\mathbb{P}(\omega|B)$ playing the role of the probability measure.

All the axioms are satisfied (it's a good exercise to check)

That means any theorem we write down has a version where you condition everything on B.

An Example

Bayes Theorem still works in a probability space where we've already conditioned on S.

$$\mathbb{P}(A|[B \cap S]) = \frac{\mathbb{P}(B|[A \cap S]) \cdot \mathbb{P}(A|S)}{\mathbb{P}(B|S)}$$

A Quick Technical Remark

I often see students write things like

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

This is not a thing.

You probably want $\mathbb{P}(A|[B \cap C])$

$A|B$ isn't an event – it's describing an event **and** telling you to restrict the sample space. So you can't ask for the probability of that conditioned on something else.