***Conditional* Probability**

CSE 312 23Su

Lecture 5

Uniform Probability Spaces

Another Example

Suppose you shuffle a deck of cards so any arrangement is equally likely. What is the probability that the top two cards have the same value?

Sample Space: $\Omega = \{(x, y): x \text{ and } y \text{ are different cards } \}$

Probability Measure: uniform measure $\mathbb{P}(\omega) = \frac{1}{52 \cdot 51}$

Event: all pairs with equal values

Probability: $\frac{13 \cdot P(4,2)}{52 \cdot 51}$

Another Example

Suppose you shuffle a deck of cards so any arrangement is equally likely. What is the probability that the top two cards have the same value?

Sample Space: Set of all orderings of all 52 cards

Probability Measure: uniform measure $\mathbb{P}(\omega) = \frac{1}{52!}$

Event: all lists that start with two cards of the same value

Probability: $\frac{13 \cdot P(4,2) \cdot 50!}{52!}$

Another Example

Suppose you shuffle a deck of cards so any arrangement is equally likely. What is the probability that the top two cards have the same value?

Sample Space: Set of all orderings of all 52 cards

Probability Measure: uniform measure $\mathbb{P}(\omega) = \frac{1}{52!}$

Event: all lists that start with two cards of the same value

Probability: $\frac{13 \cdot P(4,2) \cdot 50 * 49 * 48 * \cdots . * 2 * 1}{52 * 51 * 50 * 49 * 48 * \ldots * 2 * 1}$

Takeaway

There’s often information you “don’t need” in your sample space.

It won’t give you the wrong answer.

But it sometimes makes for extra work/a harder counting problem,

Good indication: you cancelled A LOT of stuff that was common in the numerator and denominator.

Few notes about events and samples spaces

- If you're dealing with a situation where you may be able to use a uniform probability space, make sure to set up the sample space in a way that every outcome is equally likely.

- Try not overcomplicate the sample space – only include the information that you need in it.

- When you define an event, make sure it is a <u>subset</u> of the sample space!
  e.g., if order matters in the sample space, it should also matter in the event space

Some Quick Observations

For discrete probability spaces (the kind we've seen so far)

$\mathbb{P}(E) = 0$ if and only if ?

$\mathbb{P}(E) = 1$ if and only if ?

Some Quick Observations

For discrete probability spaces (the kind we've seen so far)

$\mathbb{P}(E) = 0$ if and only if an event can't happen.

$\mathbb{P}(E) = 1$ if and only if an event is guaranteed (every outcome outside E has probability 0).

Birthday Paradox

Sharing Birthdays 🎂

There are about 50 people in this class. What is the probability that at least 2 of us share a birthday? Assume that there are exactly 365 possible birthdays, and each possibility is equally likely.

We know from the pigeonhole principle that if there are >365 people in the group, there will certainly be at least 2 people that share the same birthday. But what's the *probability* of this happening if there are only 50 people?

Sharing Birthdays 🎂

There are about 50 people in this class. What is the probability that at least 2 of us share a birthday? Assume that there are exactly 365 possible birthdays, and each possibility is equally likely.

What do you think this probability is closest to?

A) 0.001

B) 0.5

C) 0.99

D) 1

Fill out the poll everywhere **pollev.com/cse312** and login with your UW identity

Sharing Birthdays 🎂

There are about 50 people in this class. What is the probability that at least 2 of us share a birthday? Assume that there are exactly 365 possible birthdays, and each possibility is equally likely.

Sample Space:

Probability Measure:

Event:

Probability:

There are about 50 people in this class. What is the probability that at least 2 of us share a birthday? Assume that there are exactly 365 possible birthdays, and each possibility is equally likely.

Sample Space: Set of assignments of birthdays to people. $|\Omega| = 365^{50}$

Probability Measure: Uniform probability measure. $\mathbb{P}(\omega) = \frac{1}{365^{50}}$ for $\omega \in \Omega$

Event: Let E be the event that at least 2 people share a birthday.

Probability: $\mathbb{P}(E) = 1 - \mathbb{P}(\bar{E})$. $\bar{E}$ is the event that **no one** shares a birthday.

$$\mathbb{P}(\bar{E}) = \frac{|\bar{E}|}{|\Omega|} = \frac{P(365,50)}{365^{50}}.$$

We use a permutation for $|\bar{E}|$ because birthdays are "selected" without replacement, (all have different birthdays) and order matters (my birthday is different from your birthday, etc.)

$$\mathbb{P}(E) = 1 - \frac{P(365,50)}{365^{50}} \approx 0.97$$

Sharing Birthdays 🎂

There are about 50 people in this class. What is the probability that at least 2 of us share a birthday? Assume that there are exactly 365 possible birthdays, and each possibility is equally likely.

$$\mathbb{P}(E) = 1 - \frac{P(365,50)}{365^{50}} \approx 0.97$$

This is pretty high! So almost definitely, two of us here share the same birthday 🥳

That’s very likely! Why?

It turns out that human brains find thinking about probabilities difficult!

Our brains are a bit selfish! When it comes to the probability that someone shares our birthday, that would be $\frac{1}{365}$ - not quite so likely.

But if we’re looking at any pair’s birthday in a group of n people, there are $\binom{n}{2}$ pairs of people, which grows quadratically with n. So the probability of at least one pair of people sharing a birthday approaches 1 pretty fast!

Summary

- Probability allows us to assign a value between 0 and 1 to outcomes
- A ***random experiment*** is any process where the outcome is not known for certain
- The ***sample space*** of an experiment is the set of all possible outcomes
- An ***event*** is a subset of the sample space (some set of outcomes)
- The ***probability space*** is the pair $(\Omega, \mathbb{P})$ where Ω is the sample space and $\mathbb{P}$ is the probability measure (a *function* that assigns probabilities to every outcome ω in the sample space)
  - A ***uniform probability space*** is a common type of probability space where every outcome is equally likely. To find the probability of an event in a uniform probability space, we find the size of the event divided by the size of the sample space

Outline

Last time we introduced **probability**

Terms like sample space, event, probability space

We talked about *uniform probability spaces*, a common type of probability space

Today, we’re talking about *conditional* probability

How do we compute probabilities when we’re given some partial/extra information?

- What is conditioning, and what does conditional probability mean?
- Rules to work with conditional probability (Bayes’ rule, law of total probability)
- Intuition about conditional probability

Conditional Probabilities

Conditioning

You roll a fair red die and a fair blue die (without letting the dice affect each other).

But they fell off the table and you can't see the results.

I can see the results – I tell you the sum of the two dice is 4.

What's the probability that the red die shows a 5, **conditioned** on knowing the sum is 4?

Conditioning

You roll a fair red die and a fair blue die (without letting the dice affect each other).

But they fell off the table and you can't see the results.

I can see the results – I tell you the sum of the two dice is 4.

What's the probability that the red die shows a 5, **conditioned** on knowing the sum is 4?

It's 0.

Without the conditioning it was $1/6$.

Conditioning

When I told you “the sum of the dice is 4” we **restricted the sample space.**

The only remaining outcomes are $\{(1,3),(2,2),(3,1)\}$ out of $\{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$.

Outside the (restricted) sample space, the probability is going to become 0. What about the probabilities inside?

Conditional Probability

Conditional Probability

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

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

Just like with the formal definition of probability, this is pretty abstract.

It does accurately reflect what happens in the real world.

If $\mathbb{P}(B) = 0$, we can't condition on it (it can't happen! There's no point in defining probabilities where we know B has not happened) – $\mathbb{P}(A|B)$ is **undefined** when $\mathbb{P}(B) = 0$.

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)	$\boxed{(1,3)}$	(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"

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

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

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

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

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

	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"

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

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

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

	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

A ~ Red die 6

B ~ Sum is 7

C ~ Sum is 9

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

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

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

	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)

Take a few minutes to work on this with the people around you!
(also on your handout)
PollEv.com/cse312

Conditioning Practice

A ~ Red die 6

B ~ Sum is 7

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

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

$=$

$= 1/6$

	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

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

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

$=$

$= 1/4$

	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

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

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

$=$

$= 1/6$

	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)

Wonka Bars 🍫

Example working with conditional probabilities

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

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.

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

Which do you think is closest to the right answer?

A. 0.1%
B. 10%
C. 50%
D. 99%

PollEv.com/cse312

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

What probability are we looking for?

*"<u>If you pick up a bar and it alerts</u>, what is the **probability you have a golden ticket**?"*

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

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

What probability are we looking for?

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

What probabilities are each of these from the problem?

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.

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

What probability are we looking for?

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

What probabilities are each of these from the problem?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.	$\mathbb{P}(B) = 0.001$
If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.	$\mathbb{P}(A|B) = 0.999$
If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.	$\mathbb{P}(A|\bar{B}) = 0.01$

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

To summarize…

$\mathbb{P}(B) = 0.001$

$\mathbb{P}(A|B) = 0.999$

$\mathbb{P}(A|\bar{B}) = 0.01$

We're looking for $\mathbb{P}(B|A) = ???$

Reversing the Conditioning

All of our information conditions on whether B happens or not

$\mathbb{P}(A|B)$, $\mathbb{P}(A|\bar{B})$ -- "is the test positive if *we know there is/there is not a golden ticket*?"

But we're interested in the "*reverse*" conditioning.

We know the scale alerted us/the test is positive, but do we have a golden ticket?

$$\mathbb{P}(B) = 0.1\%$$
$$\mathbb{P}(A|B) = 99.9\%$$
$$\mathbb{P}(A|\bar{B}) = 1\%$$
$$\mathbb{P}(B|A) = ???$$

Is there a relationship between these probabilities?

Bayes' Rule: relates $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$

Bayes' Rule

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

Bayes' Rule: relates $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$

Bayes' Rule

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

$\mathbb{P}(B) = 0.1\%$
$\mathbb{P}(A|B) = 99.9\%$
$\mathbb{P}(A|\bar{B}) = 1\%$
$\mathbb{P}(B|A) = ???$

What do we know about Wonka Bars?

$$0.999 = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{.001}$$

Bayes' Rule: relates $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$

Bayes' Rule

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

$\mathbb{P}(B) = 0.1\%$
$\mathbb{P}(A|B) = 99.9\%$
$\mathbb{P}(A|\bar{B}) = 1\%$
$\mathbb{P}(B|A) = ???$

What do we know about Wonka Bars?

$$0.999 = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{.001}$$

We're looking for $\mathbb{P}(B|A)$, so to solve for that, all we need left is $\mathbb{P}(A)$

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

We know the probability of A *conditioned on* whether the bar has the ticket (B and $\bar{B}$)

How can we use those conditional probabilities to find the probability of A when we're not conditioning on anything?

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

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

We know the probability of A *conditioned on* whether the bar has the ticket (B and $\bar{B}$)

How can we use those conditional probabilities to find the probability of A when we're not conditioning on anything?

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

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

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

$$= .010989$$

(detour) Law of Total Probability

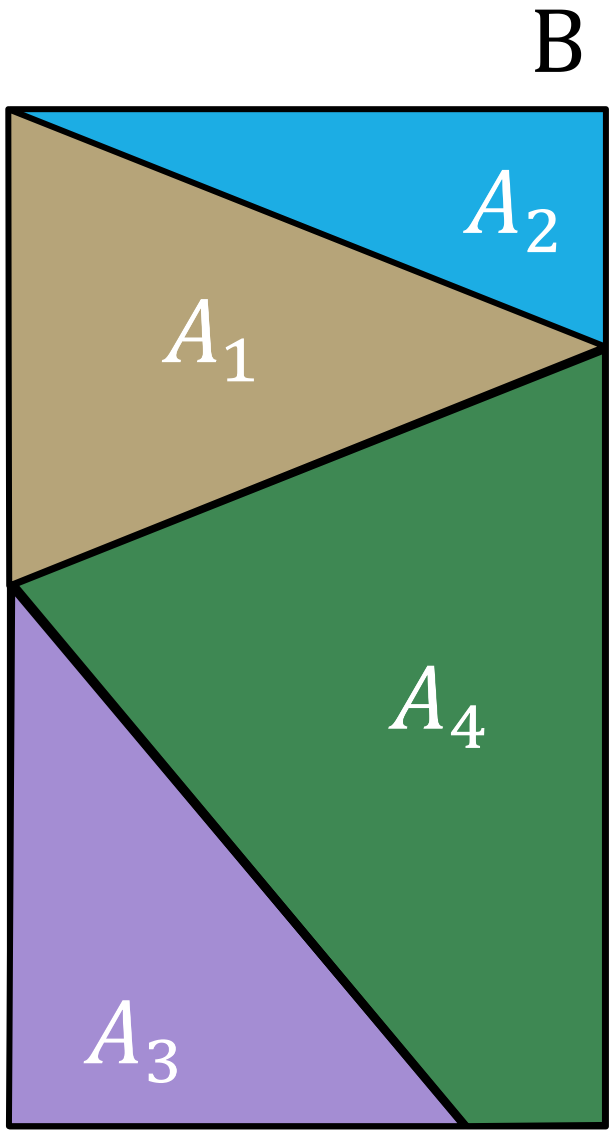
Partitions

Sets A_1, A_2, …., A_n <u>partition</u> a set B if:

1. They **do not overlap** (they are *mutually exclusive*)
   $A_i \cap A_j = \emptyset$ for all pairs of i and j

2. They cover/**exhaust the entire set** $\boldsymbol{B}$
   $A_1 \cup A_2 \cup \cdots \cup A_n = B$

e.g., the sets $A = \{1,4\}, B = \{2,5,6\}, C = \{3\}$ *partition* the set $B = \{1,2,3,4,5,6\}$

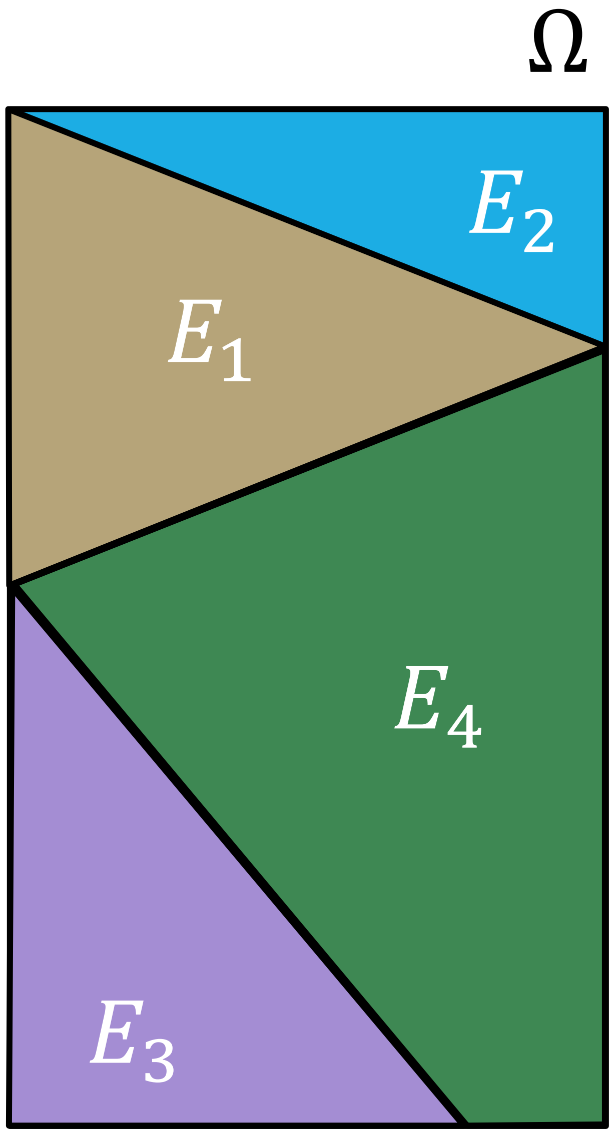


Partitions *of a sample space*

Events E_1, E_2, E_3, E_4 partition the sample space Ω if:

1. They **do not overlap** (they are *mutually exclusive*)
   $E_i \cap E_j = \emptyset$ for all pairs of i and j

2. They cover/**exhaust the entire sample space Ω**
   $E_1 \cup E_2 \cup \cdots \cup E_n = \Omega$



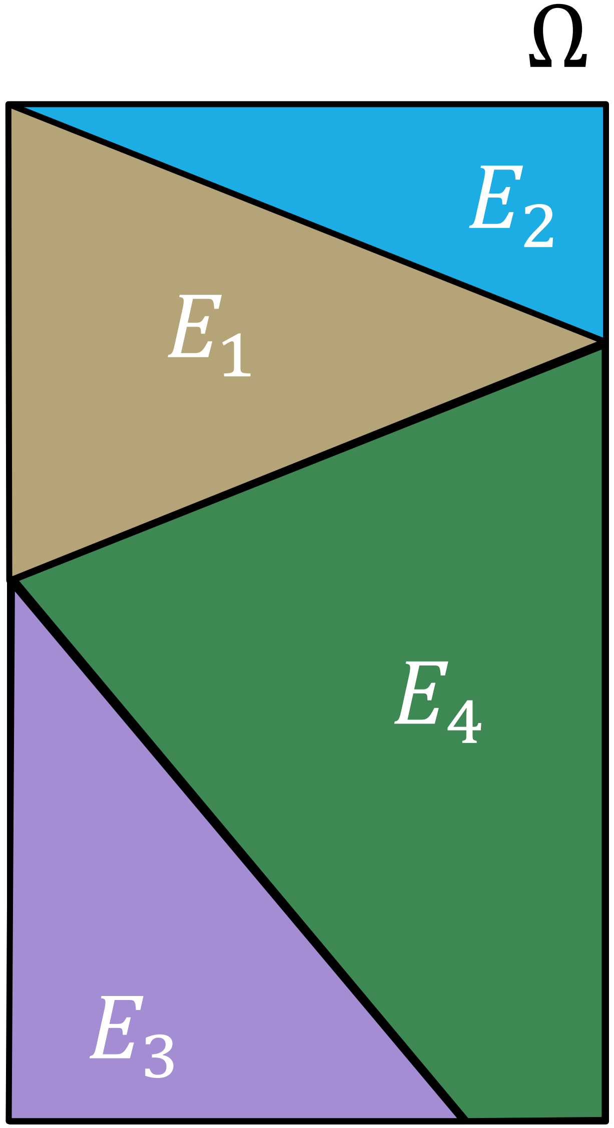
An event and it's complement (A and $\bar{A}$) partition Ω

Because (1) they are mutually exclusive

(2) $A \cup \bar{A} = \Omega$, every outcome is either in the A or not

Law of Total Probability

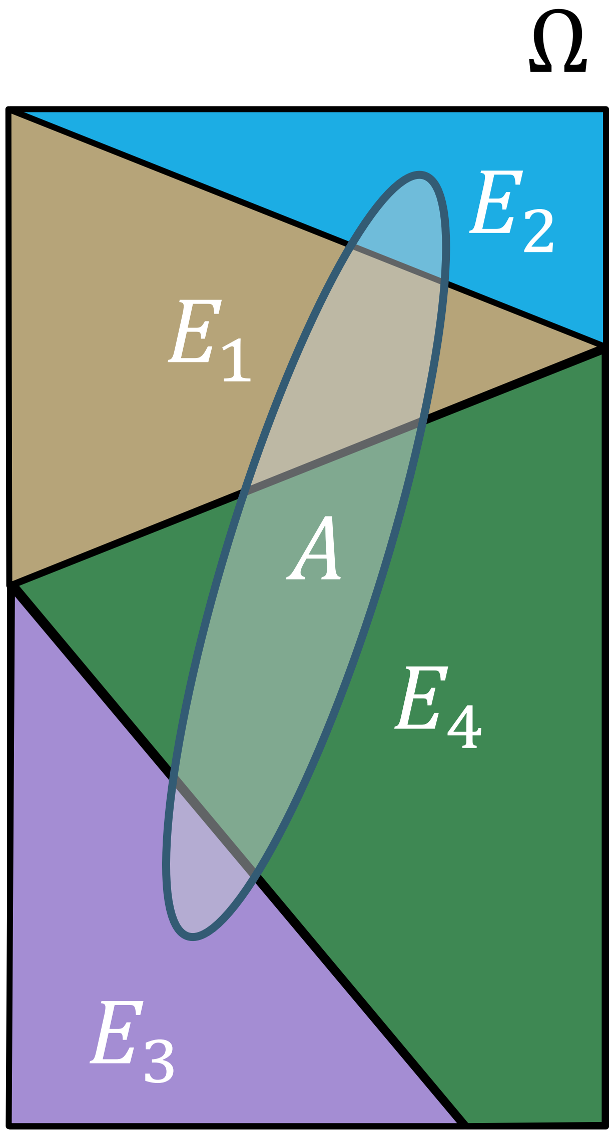
E_1, E_2, E_3, E_4 partition the sample space Ω



Law of Total Probability

E_1, E_2, E_3, E_4 partition the sample space Ω

Now, we want to find the probability of some event A in this sample space - $\mathbb{P}(A)$



Law of Total Probability

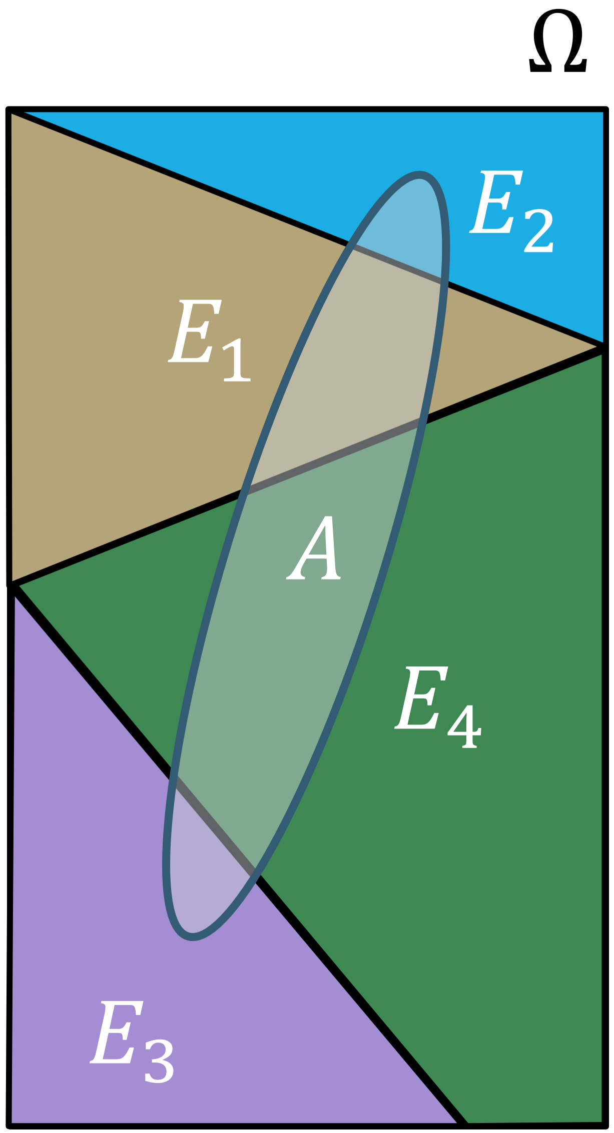
E_1, E_2, E_3, E_4 partition the sample space Ω

Now, we want to find the probability of some event A in this sample space - $\mathbb{P}(A)$

Because it might overlap with the partitions:

$\mathbb{P}(A) =$

$\mathbb{P}(A \cap E_1) \quad + \mathbb{P}(A \cap E_2) \quad + \mathbb{P}(A \cap E_3) \quad + \mathbb{P}(A \cap E_4) =$



Law of Total Probability

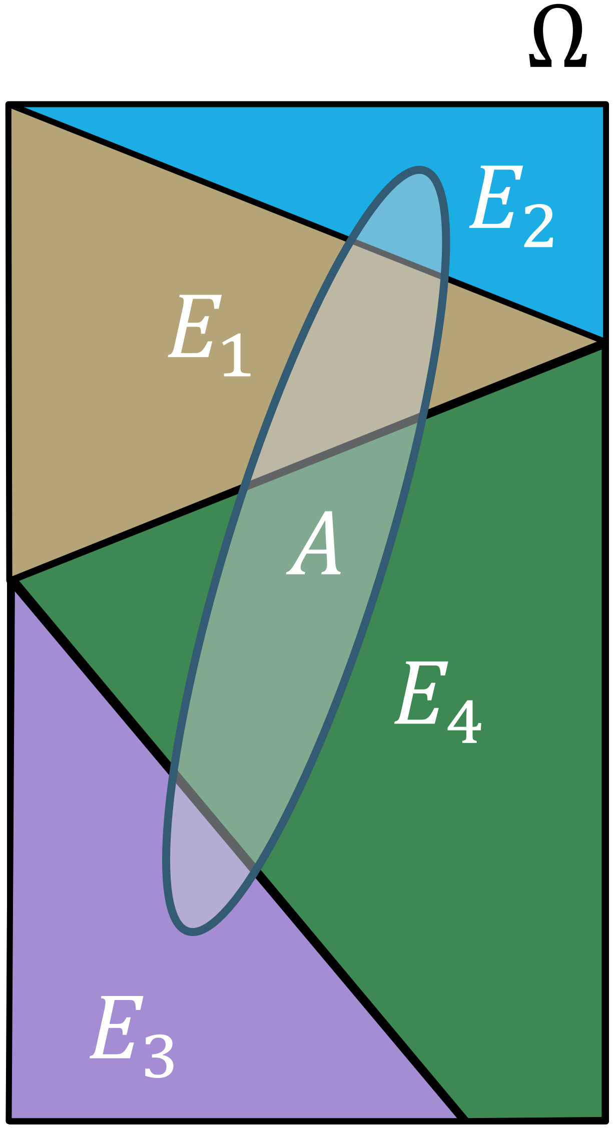
E_1, E_2, E_3, E_4 partition the sample space Ω

Now, we want to find the probability of some event A in this sample space - $\mathbb{P}(A)$

Because it might overlap with the partitions:

$\mathbb{P}(A) =$

$\mathbb{P}(A \cap E_1) \quad + \mathbb{P}(A \cap E_2) \quad + \mathbb{P}(A \cap E_3) \quad + \mathbb{P}(A \cap E_4) =$



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

Law of Total Probability

E_1, E_2, E_3, E_4 partition the sample space Ω

Now, we want to find the probability of some event A in this sample space - $\mathbb{P}(A)$

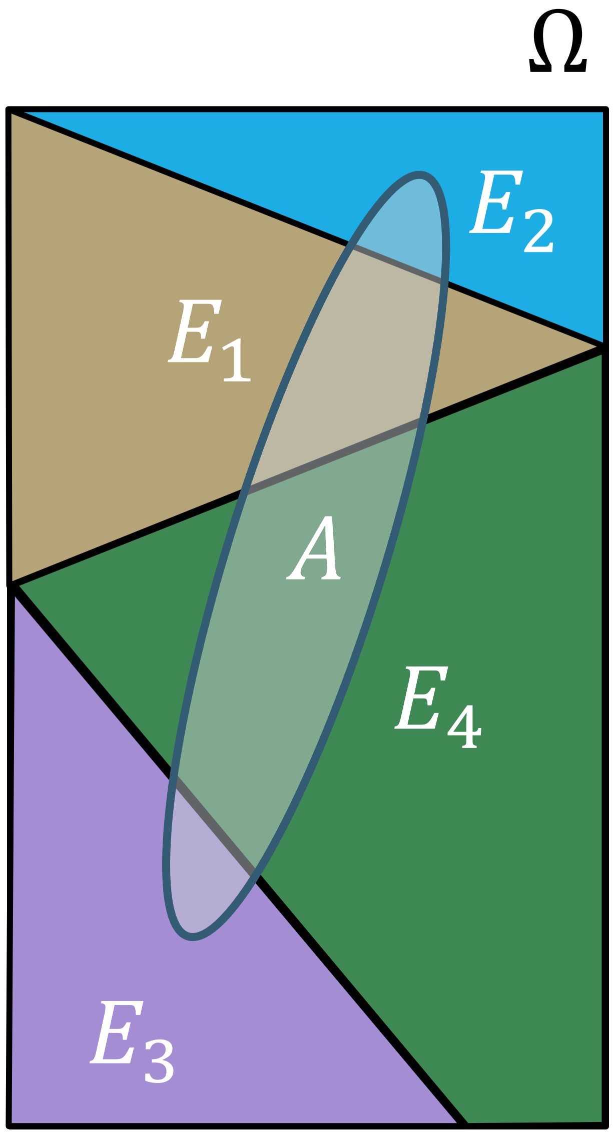
Because it might overlap with the partitions:

$\mathbb{P}(A) =$

$\mathbb{P}(A \cap E_1) + \mathbb{P}(A \cap E_2) + \mathbb{P}(A \cap E_3) + \mathbb{P}(A \cap E_4) =$

$\mathbb{P}(A|E_1)\mathbb{P}(E_1) + \mathbb{P}(A|E_2)\mathbb{P}(E_2) + \mathbb{P}(A|E_3)\mathbb{P}(E_3) + \mathbb{P}(A|E_4)\mathbb{P}(E_4)$

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



Law of Total Probability (more formally)

E_1, E_2, E_3, E_4 partition the sample space Ω

$\mathbb{P}(A) =$

$\mathbb{P}(A \cap \Omega) =$ Because $A \subseteq \Omega$

$\mathbb{P}(A \cap (\mathrm{E}_1 \cup \mathrm{E}_2 \cup E_3 \cup E_4)) =$ Definition of partitions

$\mathbb{P}((A \cap E_1) \cup (A \cap E_2) \cup (A \cap E_3) \cup (A \cap E_4)) =$ Distributive property of intersection over union

$\mathbb{P}(A \cap E_1) \quad + \mathbb{P}(A \cap E_2) \quad + \mathbb{P}(A \cap E_3) \quad + \mathbb{P}(A \cap E_4) =$ By def. of partition E_i's are all mutually exclusive, so each $A \cap E_i$ is also mutually exclusive. Then, applying additive property of probability for mutually exclusive events

$\mathbb{P}(A|E_1)\mathbb{P}(E_1) + \mathbb{P}(A|E_2)\mathbb{P}(E_2) + \mathbb{P}(A|E_3)\mathbb{P}(E_3) + \mathbb{P}(A|E_4)\mathbb{P}(E_4)$

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

Law of Total Probability

To generalize…

Law of Total Probability

Let $E_1, E_2, \dots, E_k$ be events that partition the sample space Ω.

For any event A,

$$\mathbb{P}(A) = \sum_{\text{all } i} \mathbb{P}(A \cap E_i) = \sum_{\text{all } i} \mathbb{P}(E_i | A_i)\mathbb{P}(A_i)$$

Back to the Wonka Bars! 🍫

Now that we’ve totally forgotten why we needed this 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?

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

What probability are we looking for?

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

What probabilities are each of these from the problem?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.	$\mathbb{P}(B) = 0.001$
If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.	$\mathbb{P}(A\|B) = 0.999$
If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.	$\mathbb{P}(A\|\bar{B}) = 0.01$

Bayes Rule: relates $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$

Bayes' Rule

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

$\mathbb{P}(B) = 0.1\%$
$\mathbb{P}(A|B) = 99.9\%$
$\mathbb{P}(A|\bar{B}) = 1\%$
$\mathbb{P}(B|A) = ???$

What do we know about Wonka Bars?

$$0.999 = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{.001}$$

We're looking for $\mathbb{P}(B|A)$, so to solve for that, all we need left is $\mathbb{P}(A)$

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

We know the probability of A *conditioned on* whether the bar has the ticket (B and $\bar{B}$)

How can we use those conditional probabilities to find the probability of A when we're not conditioning on anything?

We'll use a trick called "the **law of total probability (LoTP)**":

The events B and $\bar{B}$ partition the sample space. Then, by **LoTP**:

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

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

$$= .010989$$

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

Now, we plug in and solve for $\mathbb{P}(B|A)$….

$$0.999 = \frac{\mathbb{P}(B|A) \cdot .010989}{.001}$$

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

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

Wait a minute…

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

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

Let's say we tested all 1000 bars.

Approximately…..

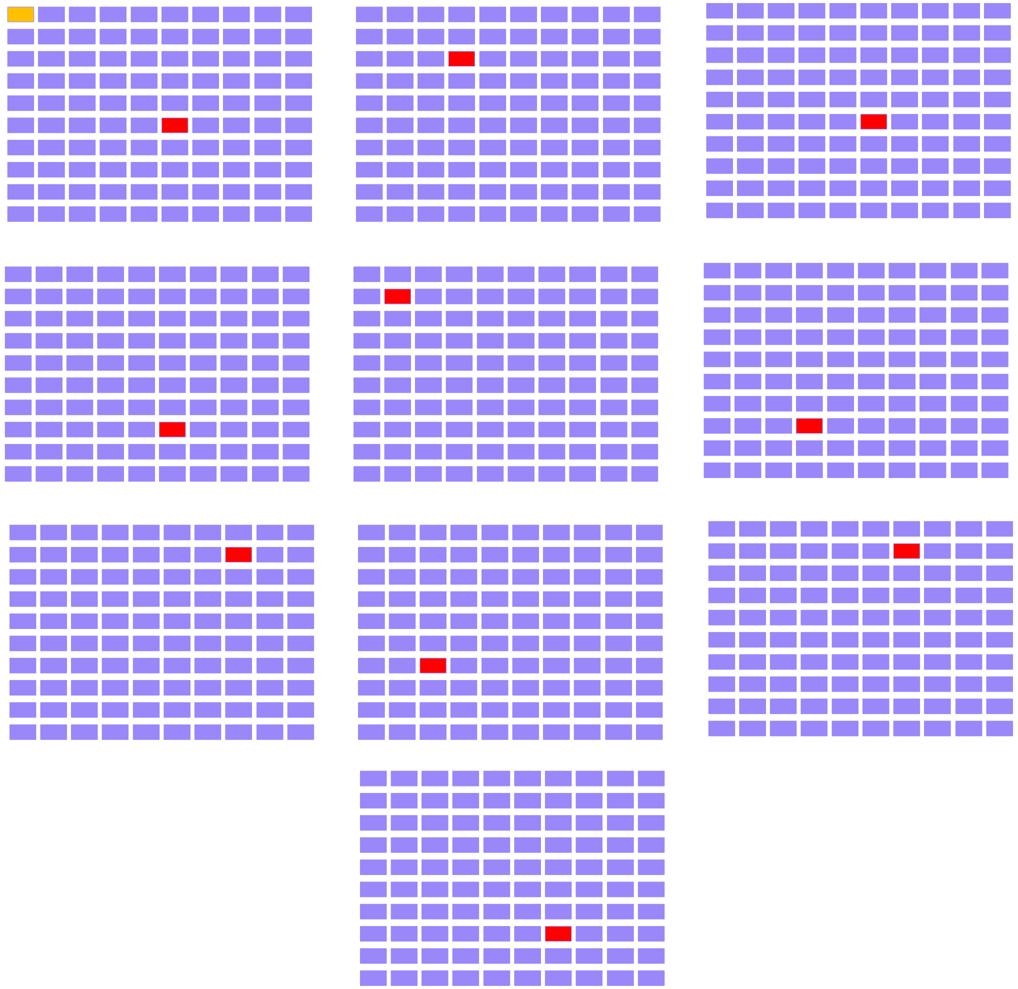
1 has a golden ticket, 999 do not have a golden ticket

Let's say the scale correctly alerts on the golden ticket

About 1% of the 999 without a golden ticket would be a positive. Let's say the scale alerted on 10 of the bars without a golden ticket

But, in only 1 out of the $10 + 1$ positives, we had the golden ticket

Visually



Gold bar is the one (true) golden ticket bar.
True positive

Purple bars don't have a ticket, tested negative.
True negative

Red bars don't have a ticket, tested positive.
False positive

The test is, in a sense, doing really well.
It's almost always right (only red is incorrect)

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

Updating Your Intuition

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

🔥 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

> 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.

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

🔥 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.

Updating Your Intuition

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

🔥 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.

The Technical Stuff

Proof of Bayes' Rule

Bayes' Rule

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

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

We also know $\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \rightarrow \mathbb{P}(B \cap A) = \mathbb{P}(B|A) \cdot \mathbb{P}(A)$

Intersections are commutative, so $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$

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

A Technical Note

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

$(\Omega, \mathbb{P})$ is a valid probability space and, $A \subseteq \Omega$

$(A, \mathbb{P}(\cdot\,|A))$ is valid probability space

Sum of probabilities of outcomes in a sample space is 1:
$\sum_{\omega \in A} \mathbb{P}(\omega|A) = 1$

$\mathbb{P}(\bar{E}) = 1 - \mathbb{P}(E)$

Sum of probabilities in this conditioned probability space is 1:
$\sum_{\omega \in A} \mathbb{P}(\omega|A) = 1$

$\mathbb{P}(\bar{E}|A) = 1 - \mathbb{P}(E|A)$

⚠️ **Careful!** $\mathbb{P}(E|\bar{A}) \neq 1 - \mathbb{P}(E|A)$ because this changes the sample space!

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

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

Thinking something like “probability of A given B given we also know C

This is not a thing.

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

“probability of A given both B and 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.

Summary + Tips

- Today, we talked about **conditional probability**
  - Definition of conditional probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
  - Bayes' Theorem: $P(A|B) = \frac{P(A|B)P(B)}{P(B)}$
  - Law of Total Probability: $P(A) = P(A|E_1)P(E_1) + \cdots + P(A|E_n)P(E_n)$ if $E_1, E_2, \ldots E\_n$ partition the sample space
- Now that we're getting into more complex probabilities, it's very helpful to clearly define events and then write
  - Writing down all the information given in the problem/what we're asked for in terms of those events is helpful to figure out what rules we can use to relate them together

Extra Practice

Where There’s Smoke There’s…

There is a dangerous (you-need-to-call-the-fire-department-dangerous) fire in your area 1% of the time.

If there is a dangerous fire, you’ll smell smoke 95% of the time;

If there is not a dangerous fire, you’ll smell smoke 10% of the time (barbecues are popular in your area)

If you smell smoke, should you call the fire department?

S be the event you smell smoke

F be the event there is a dangerous fire

$$\mathbb{P}(F|S) = \frac{\mathbb{P}(S|F) \cdot \mathbb{P}(F)}{\mathbb{P}(S)} = \frac{\mathbb{P}(S|F) \cdot \mathbb{P}(F)}{\mathbb{P}(S|F)\mathbb{P}(F) + \mathbb{P}\left(S\middle|\bar{F}\right)\mathbb{P}(\bar{F})}$$

$$= \frac{.95 \cdot .01}{.95 \cdot .01 + .1 \cdot .99} \approx .088$$

Probably not time yet to call the fire department.

A contrived example

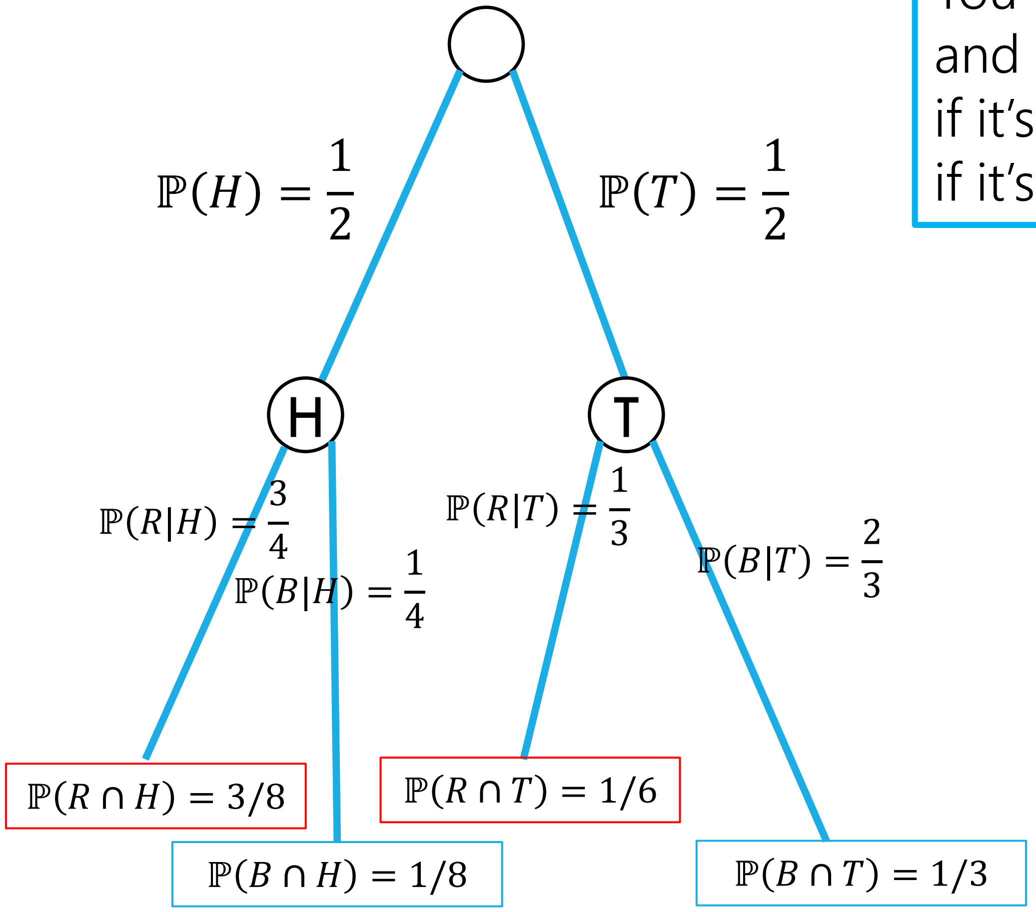
You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.

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

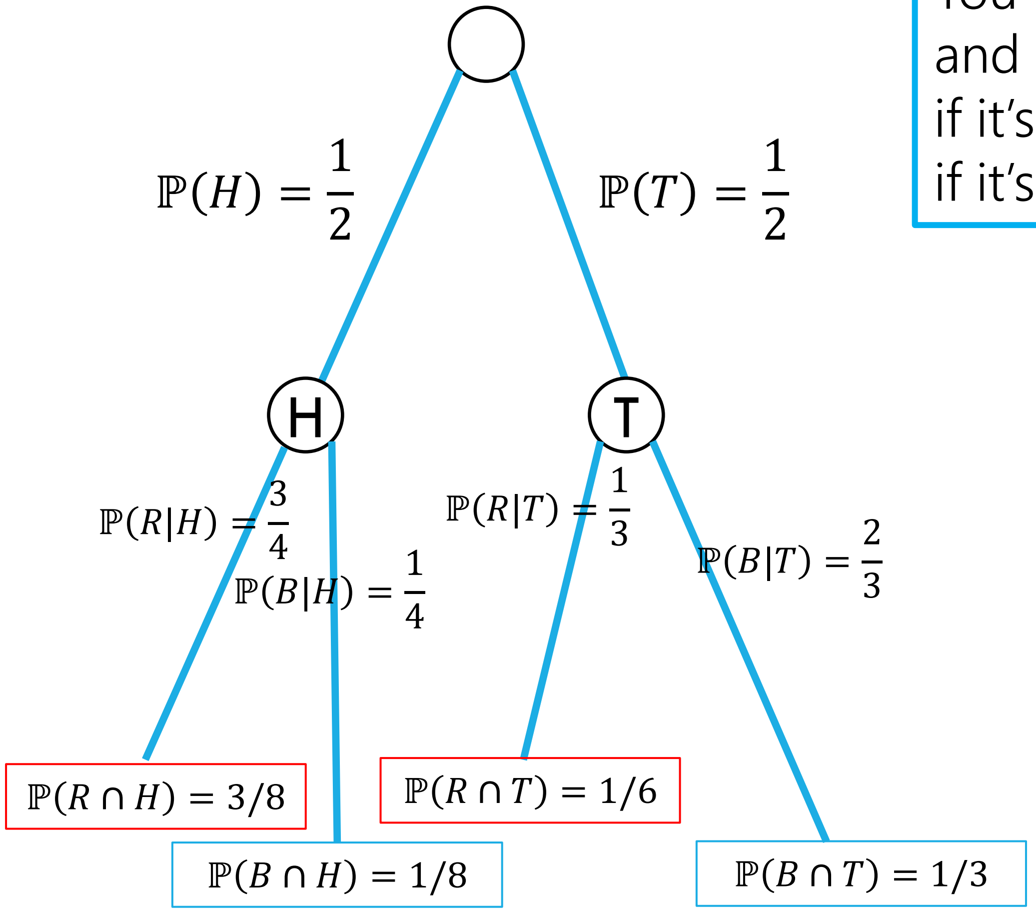


You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.
if it's heads, you'll draw a marble (uniformly) from your left pocket,
if it's tails, you'll draw a marble (uniformly) from your right pocket.

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

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

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:

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

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