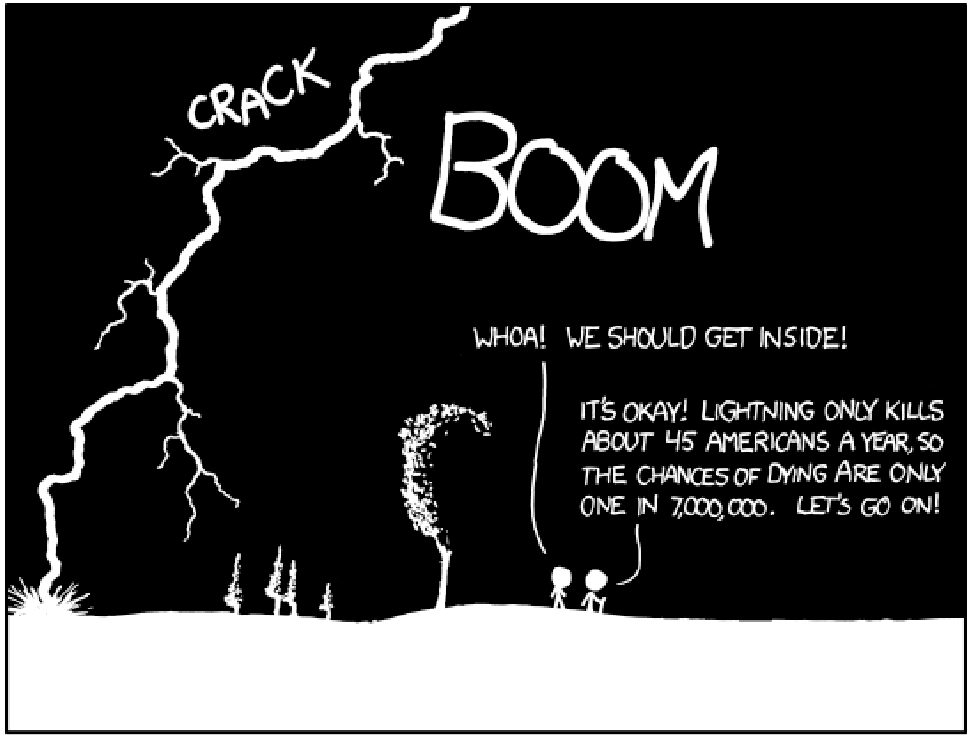


THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

Title text: 'Dude, wait -- I'm not American! So my risk is basically zero!'

Conditional Probability

CSE 312 23Su

Lecture 5

https://xkcd.com/795/

Announcements

- HW1 is due tonight (11:59pm)
- HW2 will be released later today

Where Are We?

Last time:

- Probability
- Sample Space and Events
- Probability Axioms
- Uniform Probability Space
- More Examples
- Birthday Paradox

Today:

- Conditional Probability
- Bayes Theorem
- Law of Total Probability

Recap

- Probability allows us to assign a value between 0 and 1 to outcomes
- *Sample space* Ω is the set of all possible outcomes of an experiment
- *Event* $E \subseteq \Omega$ is a subset of the sample space (some set of outcomes)
- *(Discrete) probability space* is the pair $(\Omega, \mathbb{P})$ where:
  - Ω is the sample space
  - $\mathbb{P}$ is the probability measure, a *function* $\mathbb{P} : \Omega \to [0, 1]$ such that:
    - $\mathbb{P}(\omega) \geq 0$ for every outcome $\omega \in \Omega$ in the sample space
    - $\sum_{\{\omega \in \Omega\}} \mathbb{P}(\omega) = 1$
- *Uniform probability space* is a type of probability space where every outcome is equally likely.

$P(E) = \frac{|E|}{|\Omega|}$

  - 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

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.**

(R, B)

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

$A \cap B$

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

P(A given B)
A cond on B

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

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…

blue

red

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"

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

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

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

	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"

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

red=5 blue=3

$= \frac{P(A \cap C)}{P(C)}$

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

	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)

C

A

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

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

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

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

Work on this with the people around you!
PollEv.com/jolie312

	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

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

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

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

$= 1/6$

	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

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

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

$= {}^{1}/_{36} \,/\, {}^{4}/_{36}$

$= 1/4$

	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

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

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

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

$= 1/6$

	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

$P(A|B)$

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

$P(B|A)$

$= \frac{1}{6}$

	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.

A   B

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

$P(B|A)$

$\mathbb{P}$("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!

$A \rightarrow B$

$B \rightarrow A$

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

$P(A|B) = P(B|A) = \frac{1}{6}$ independence

★ next lecture

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?

scale unhelpful

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

yay scale

PollEv.com/jolie312

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**?"*

$P(B|A)$

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

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. $P(B) = \frac{1}{1000}$

If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time. $P(A|B)\ 0.999$

If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time. $P(A|\bar{B}) = 0.01$

Conditioning

Let A be the event the scale ALERTS you

Let B be the event your bar has a ticket.

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?

$$\boxed{\begin{aligned} \mathbb{P}(B) &= 0.1\% \\ \mathbb{P}(A|B) &= 99.9\% \\ \mathbb{P}(A|\bar{B}) &= 1\% \\ \mathbb{P}(B|A) &= ??? \end{aligned}}$$

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 $\mathbb{P}(A)$?

$A|B$

$A|\bar{B}$

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**"

$P(A)$

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

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**":

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

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

$$= .010989$$

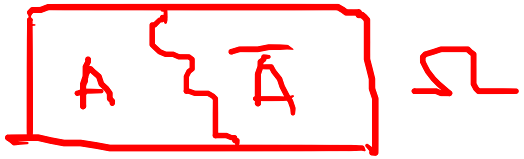
(detour) Law of Total Probability

Partitions *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

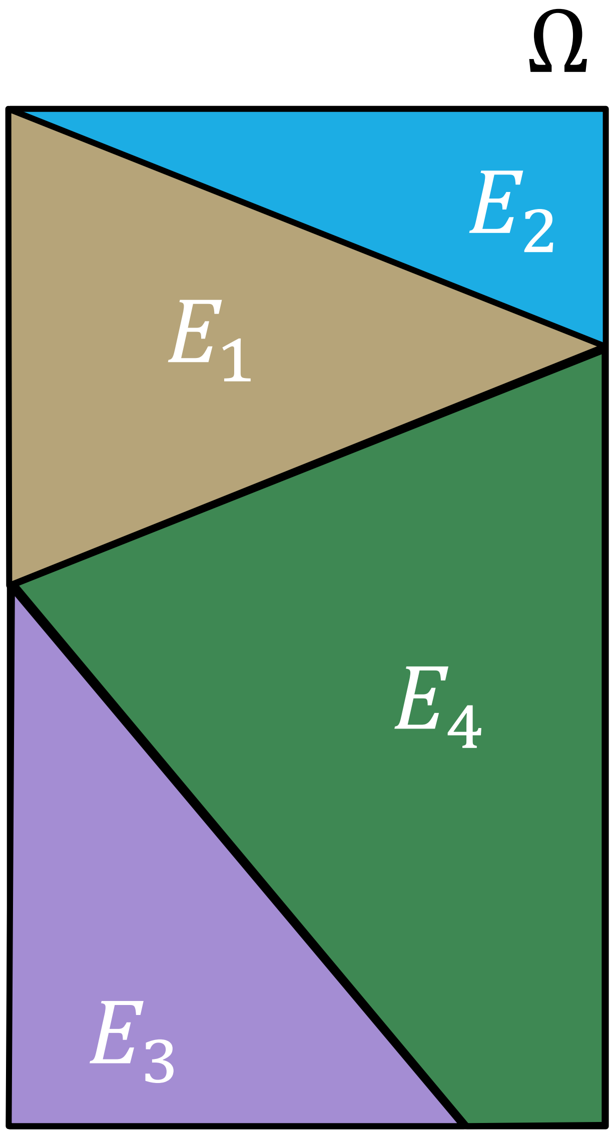
2. They cover/**exhaust the entire sample space Ω**
   $E_1 \cup E_2 \cup \cdots \cup E_n = \Omega$



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

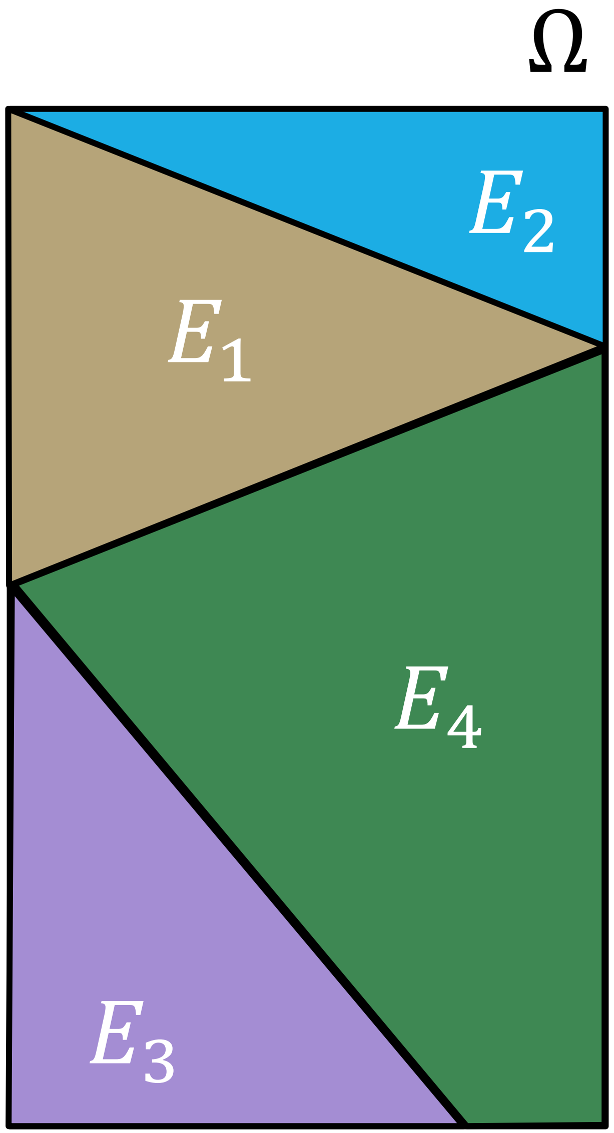
Because (1) they are mutually exclusive

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



Law of Total Probability

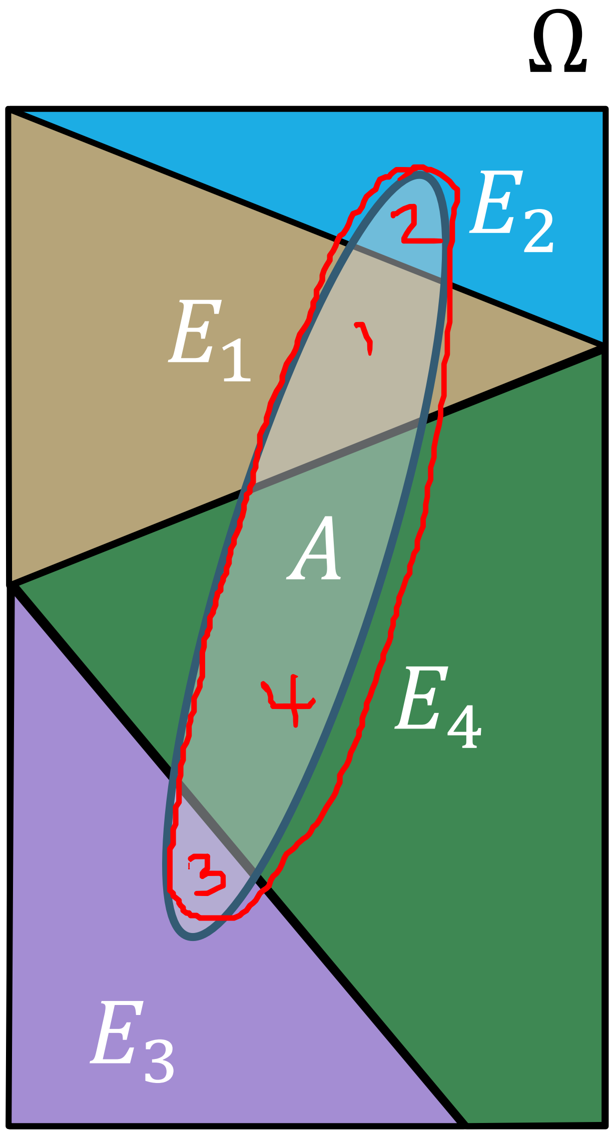
E_1, E_2, E_3, E_4 partition the sample space Ω



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

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

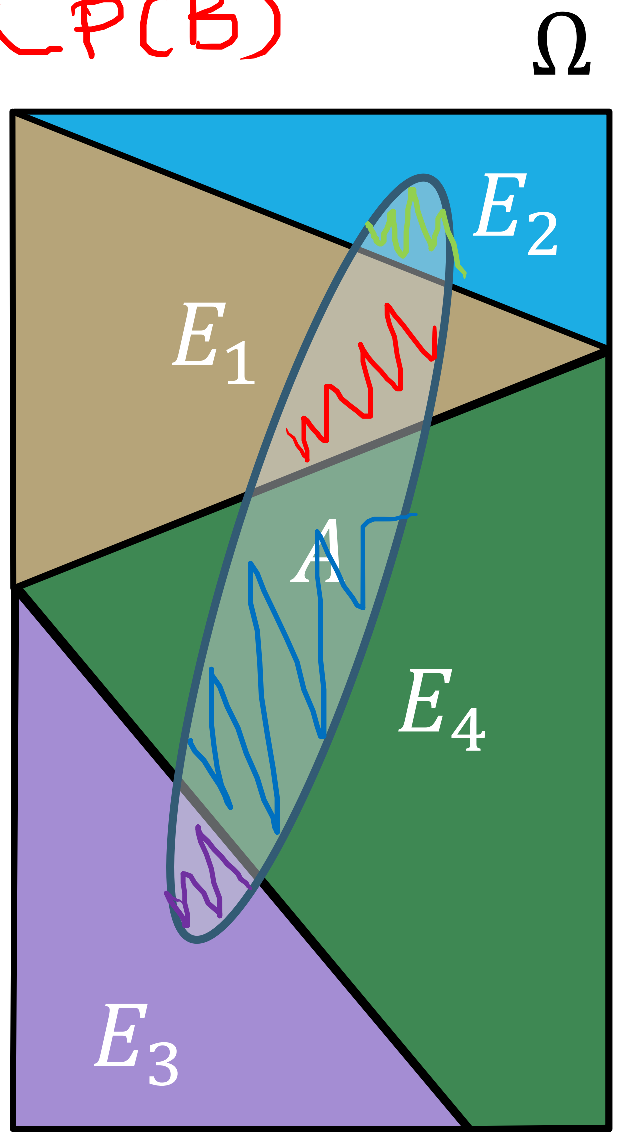
E_1, E_2, E_3, E_4 partition the sample space Ω

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

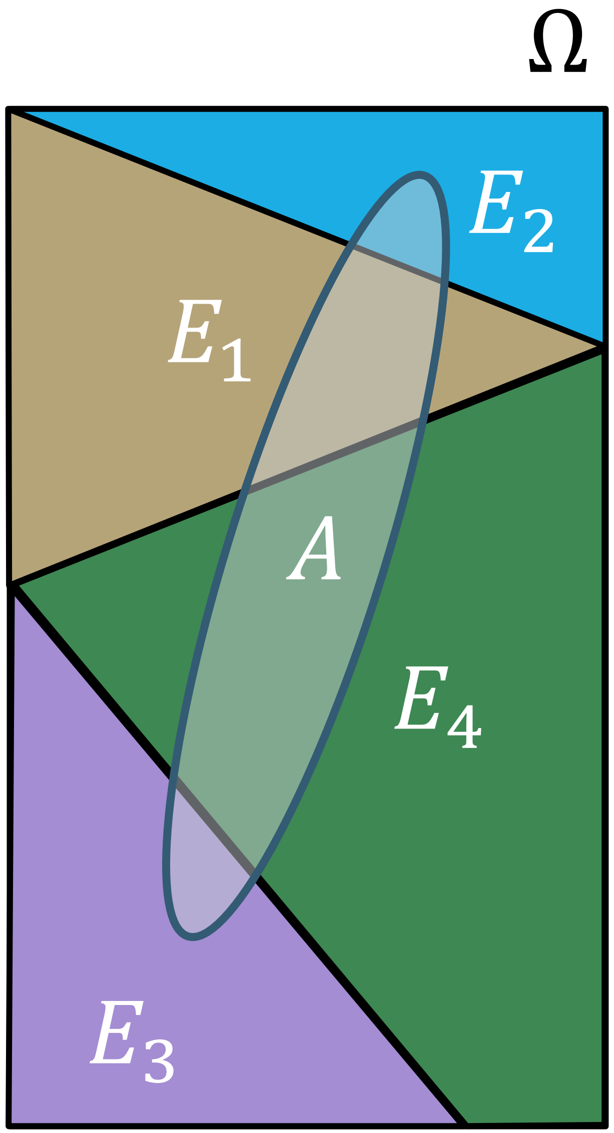
E_1, E_2, E_3, E_4 partition the sample space Ω

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 - $\mathbb{P}(A)$

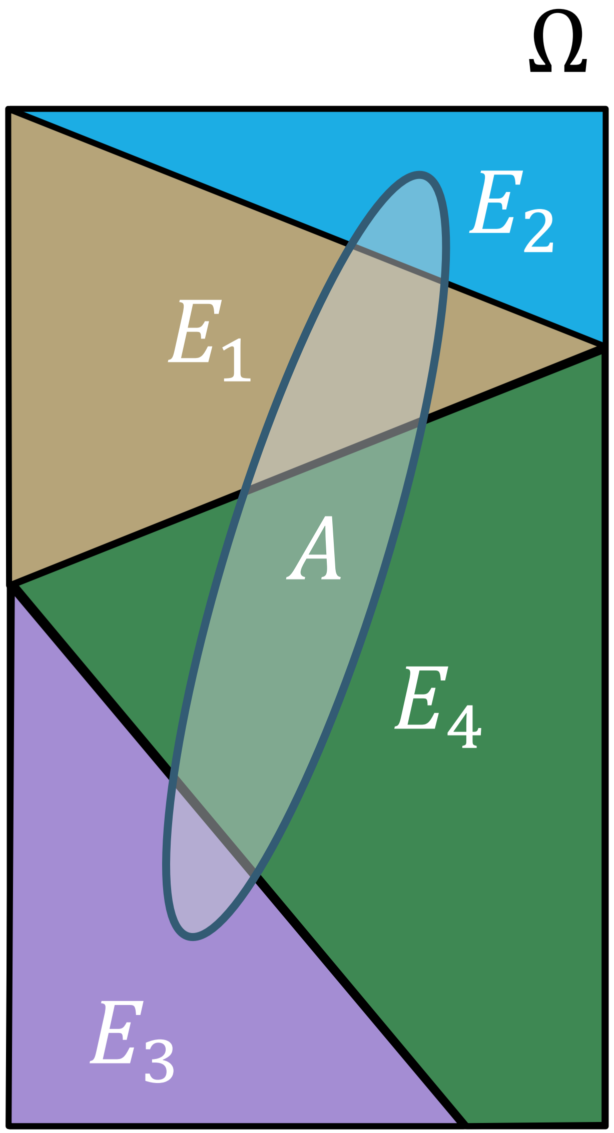
Because it might overlap with the partitions:

$\mathbb{P}(A) =$

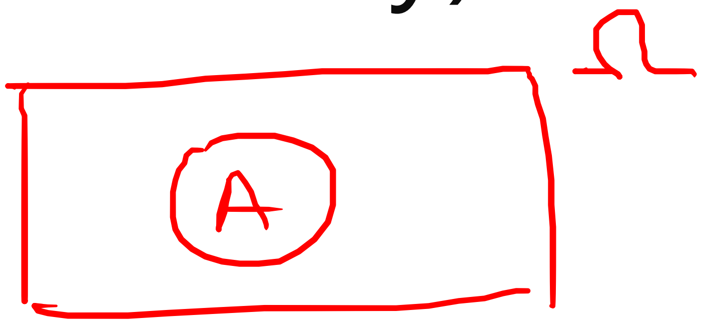
$\mathbb{P}(A \cap E_1) \quad + \mathbb{P}(A \cap E_2) \quad + \mathbb{P}(A \cap E_3) \quad + \mathbb{P}(A \cap E_4) =$

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

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



Law of Total Probability (more formally)



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

A, B mut. ex.

$P(A \cup B) = P(A) + P(B)$

$\mathbb{P}(A) =$

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

$\mathbb{P}(A \cap (E_1 \cup E_2 \cup E_3 \cup E_4)) =$ — Definition of partitions

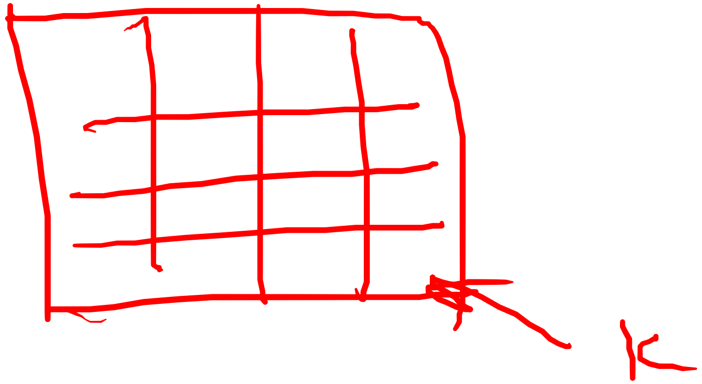
$\mathbb{P}((A \cap E_1) \cup (A \cap E_2) \cup (A \cap E_3) \cup (A \cap E_4)) =$ — Distributive property of intersection over union

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

all k events

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

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...

Alert 99.9%

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

1 correct

10 false alarm

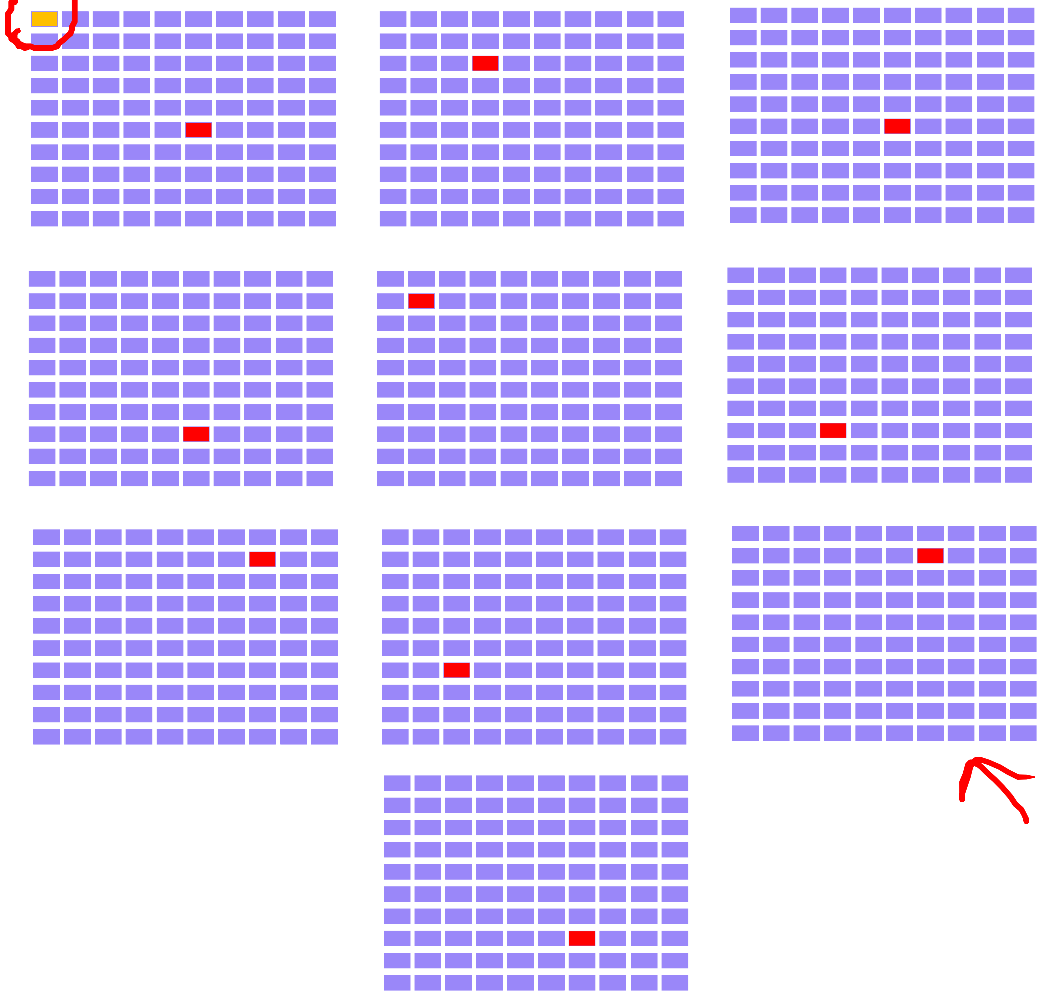
Let's say the scale correctly alerts on the golden ticket

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

1000



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.

1/1000 golden → 1/11

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

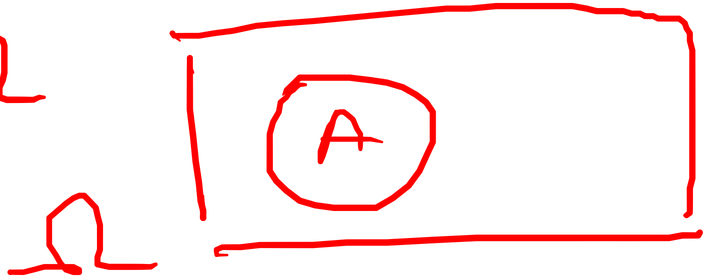
Intersections are commutative, so $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$

def commutative

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

A Technical Note

restricting Ω



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

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

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

placeholder

1

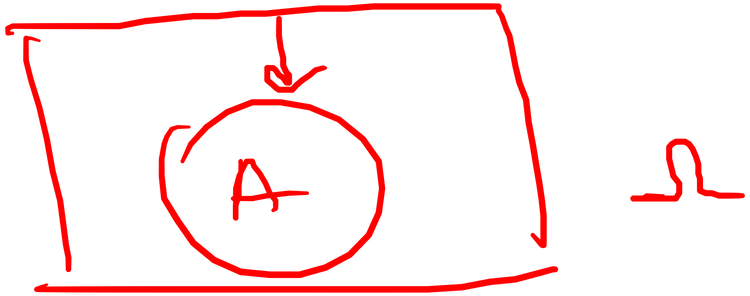
Sum of probabilities of outcomes in a sample space is 1:
$\sum_{\omega\in\Omega} \mathbb{P}(\omega) = 1$

normalization

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

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

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:
$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$

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

normalization

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

$P(A|B \cup S)$ Let $D = B \cup C$

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

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

LHS:

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

RHS:

$$\frac{\dfrac{P(B \cap A \cap S)}{P(A \cap S)} \cdot \dfrac{P(A \cap S)}{P(S)}}{\dfrac{P(B \cap S)}{P(S)}}$$

$$= \frac{P(B \cap A \cap S) \cdot \cancel{P(A \cap S)} \cdot \cancel{P(S)}}{\cancel{P(A \cap S)}\ \cancel{P(S)}\ P(B \cap S)}$$

A Quick Technical Remark

I often see students write things like

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

A given B

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

This is not a thing.

$\cap$

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

Quiz resource

- Today, we talked about **conditional probability**
  - Definition of conditional probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

  Bayes Rule
  - Bayes' Theorem: $P(A|B) = \frac{P(A|B)P(B)}{P(B)}$ ← $\frac{P(B|A)P(A)}{P(B)}$
  - 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

Let A - alert

Let B — golden