Conditional Probability

CSE 312 Spring 21
Lecture 6

More Examples!

Suppose you roll two dice. Each die is fair and they don't affect each other. What is the probability of both dice being even?

What is your sample space?

What is your probability measure $\mathbb{P}$?

What is your event?

What is the probability?

More Examples!

Suppose you roll two dice. Each die is fair and they don't affect each other. What is the probability of both dice being even?

What is your sample space? $\{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$

What is your probability measure $\mathbb{P}$? $\mathbb{P}(\omega) = 1/36$ for all $\omega \in \Omega$

What is your event? $\{2,4,6\} \times \{2,4,6\}$

What is the probability? $3^2/6^2$

More Examples!

Suppose you roll two dice. Each die is fair and they don't affect each other. What is the probability of both dice being even?

What if we can't tell the dice apart and always put the dice in increasing order by value.

What is your sample space?

$\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,2), (2,3), (2,4), (2,5), (2,6)$

$(3,3), (3,4), (3,5), (3,6), (4,4), (4,5), (4,6), (5,5), (5,6), (6,6)\}$

What is your probability measure $\mathbb{P}$?

$\mathbb{P}((x,y)) = 2/36$ if $x \neq y$, $\mathbb{P}(x,x) = 1/36$

What is your event? $\{(2,2), (4,4), (6,6), (2,4), (2,6), (4,6)\}$

What is the probability? $3 \cdot \dfrac{1}{36} + 3 \cdot \dfrac{2}{36} = \dfrac{9}{36}$

Takeaways

There is often more than one sample space possible! But one is probably easier than the others.

Finding a sample space that will make the uniform measure correct will probably make finding the probabilities easier to calculate.

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

Probability Measure

Event

Probability

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: $\{(x, y): x \text{ and } y \text{ are different cards }\}$

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

Event: all pairs with equal values

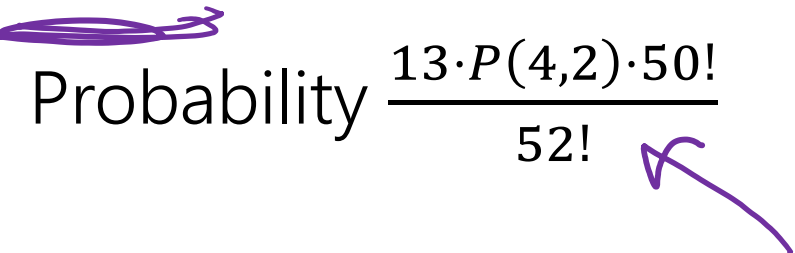Probability $\dfrac{13 \cdot P(4,2)}{52 \cdot 51}$

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 $\mathbb{P}(\omega) = \dfrac{1}{52!}$

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

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

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.

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).

"complimenting"

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

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?

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

$\mathbb{P}(\text{Red die is } 5 \cap \text{Sum is } 4)$

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

Conditional Probability

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

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

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 happened) – $\mathbb{P}(A|B)$ is <u>undefined</u> 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"

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

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

$$\mathbb{P}(B) = 3/36$$

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

	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"

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

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

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

	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

Red die 6 conditioned on sum 9

Sum 7 conditioned on red die 6

Fill out the poll everywhere so Robbie knows how long to explain
Go to pollev.com/cse312

	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

$\frac{1/36}{4/36} = \frac{1}{4}$

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

$\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$ are different quantities.

$\mathbb{P}$("I'm using an umbrella" | "it's raining") is pretty small [I usually forget it at home.

$\mathbb{P}$("it's raining" | "I'm using an umbrella") is 1 (or close to it); I don't use an umbrella for anything else.

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

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 of these is closest to the right answer?

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

Conditioning

Let A be the event you get ALERTED

Let B be the event your bar has a 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.

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

Conditioning

Let A be the event you get ALERTED

Let B be the event your bar has a ticket.

What conditional probabilities are each of these?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars. $\mathbb{P}(B)$

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

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

If you pick up a bar and it alerts, what is the probability you have a golden ticket? $\mathbb{P}(B|A)$

Reversing the Conditioning

All of our information conditions on whether B 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

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

Bayes Rule

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

What do we know about Wonka Bars?

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

Filling In

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

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

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

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

$$= .010989$$

Bayes Rule

What do we know about Wonka Bars?

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

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…

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.

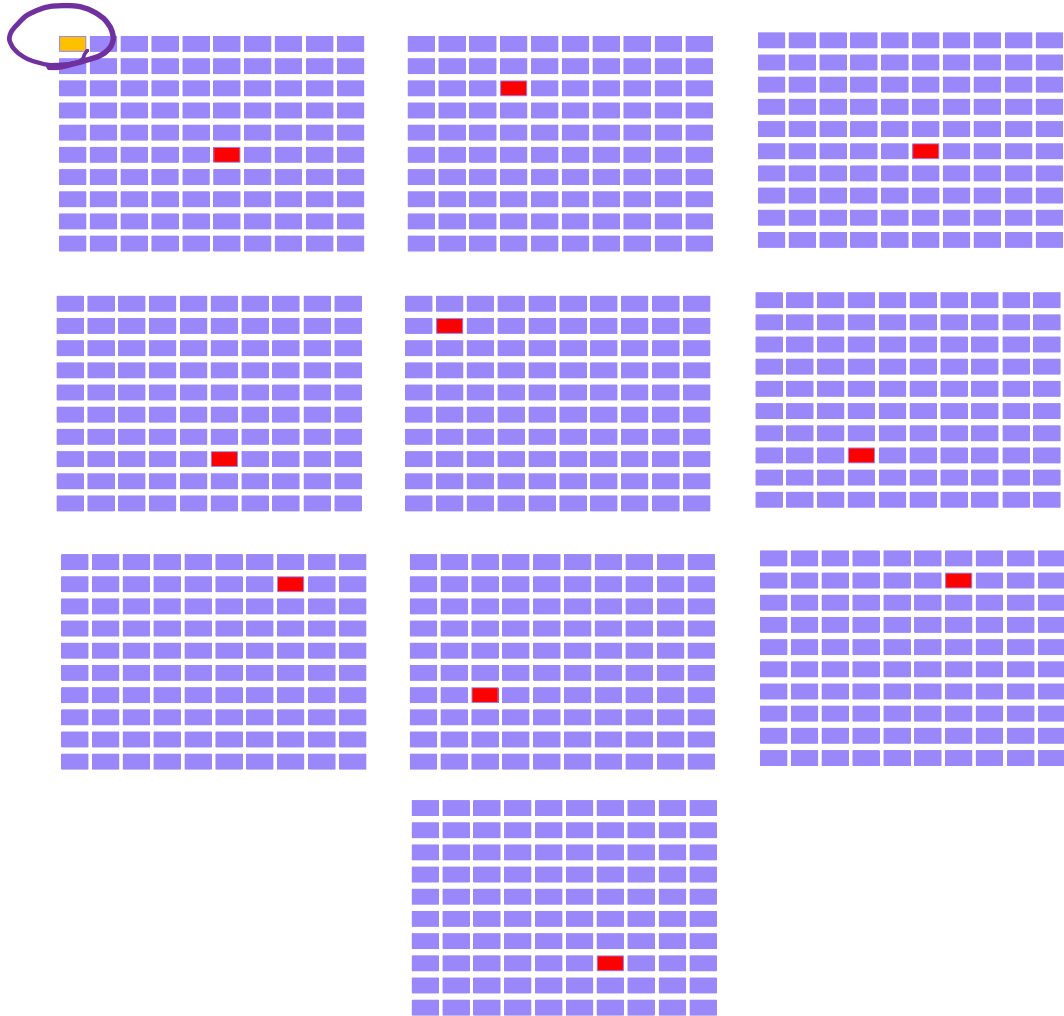That doesn't fit with many of our guesses. What's going on?

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. Lets say those are exactly right.

Lets just say that one golden is truly found

(about) 1% of the 999 without would be a positive. Lets 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

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.