

CSE 312

Foundations of Computing II

Lecture 5: Conditional Probability and Bayes Theorem

Review Probability

Definition. A **sample space** Ω is the set of all possible outcomes of an experiment.

Definition. An **event** $E \subseteq \Omega$ is a subset of possible outcomes.

Examples:

- Single coin flip: $\Omega = \{H, T\}$
- Two coin flips: $\Omega = \{HH, HT, TH, TT\}$
- Roll of a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$

Examples:

- Getting at least one head in two coin flips:
 $E = \{HH, HT, TH\}$
- Rolling an even number on a die :
 $E = \{2, 4, 6\}$

Review Probability space

Either finite or infinite countable (e.g., integers)

Definition. A (discrete) **probability space** is a pair (Ω, P) where:

- Ω is a set called the **sample space**.
- P is the **probability measure**, a function $P: \Omega \rightarrow \mathbb{R}$ such that:
 - $P(x) \geq 0$ for all $x \in \Omega$
 - $\sum_{x \in \Omega} P(x) = 1$

Set of possible elementary outcomes

$$A \subseteq \Omega: P(A) = \sum_{x \in A} P(x)$$

Specify Likelihood (or probability) of each elementary outcome

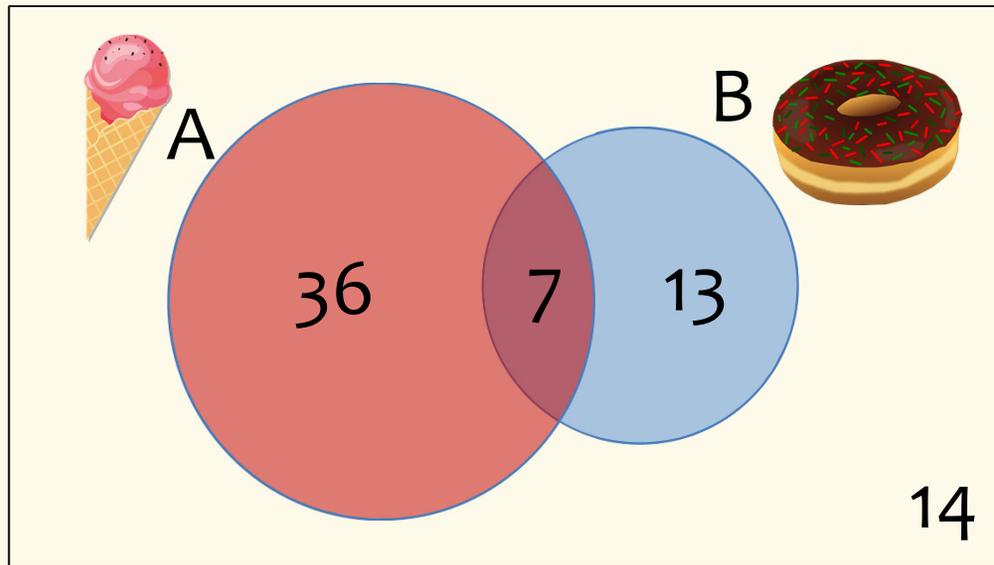
Some outcome must show up

The likelihood (or probability) of each outcome is non-negative.

Agenda

- Conditional Probability ◀
- Bayes Theorem
- Law of Total Probability
- More Examples

Conditional Probability (Idea)



What's the probability that someone likes ice cream **given** they like donuts?

$$\frac{7}{7 + 13} = \frac{7}{20}$$

Conditional Probability

Definition. The **conditional probability** of event A given an event B happened (assuming $P(B) \neq 0$) is

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

An equivalent and useful formula is

$$P(A \cap B) = P(A|B)P(B)$$

Conditional Probability Examples

Suppose that you flip a fair coin twice.

What is the probability that both flips are heads given that you have at least one head?

Let O be the event that at least one flip is heads

Let B be the event that both flips are heads

$$P(O) = 3/4 \quad P(B) = 1/4 \quad P(B \cap O) = 1/4$$

$$P(B|O) = \frac{P(B \cap O)}{P(O)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Ω

HH	HT
TH	TT

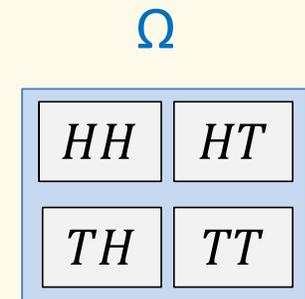
Conditional Probability Examples

Suppose that you flip a fair coin twice.

What is the probability that at least one flip is heads given that at least one flip is tails?

Let H be the event that at least one flip is *heads*

Let T be the event that at least one flip is *tails*



Conditional Probability Examples

Suppose that you flip a fair coin twice.

What is the probability that at least one flip is heads given that at least one flip is tails?

Let H be the event that at least one flip is *heads*

Let T be the event that at least one flip is *tails*

$$P(H) = 3/4 \quad P(T) = 3/4 \quad P(H \cap T) = 1/2$$

$$P(H|T) = \frac{P(H \cap T)}{P(T)} = \frac{1/2}{3/4} = \frac{2}{3}$$

Ω

HH	HT
TH	TT

Reversing Conditional Probability

Question: Does $P(A|B) = P(B|A)$?

No!

- Let A be the event you are wet
- Let B be the event you are swimming

$$P(A|B) = 1$$
$$P(B|A) \neq 1$$

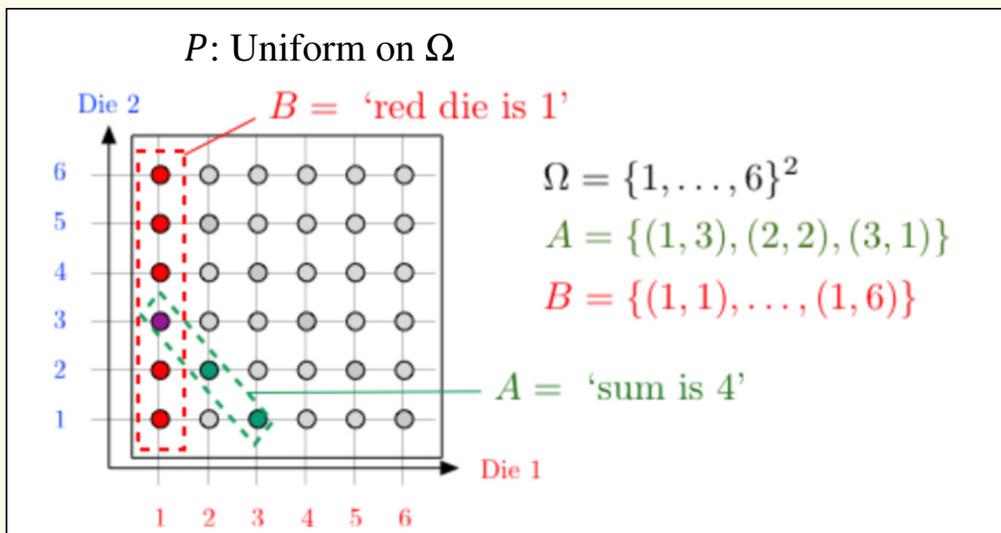
Example with Conditional Probability

pollev.com/paulbeame028

Suppose we toss a red die and a blue die:
both 6 sided and all outcomes equally
likely.

What is $P(B)$? What is $P(B|A)$?

	$P(B)$	$P(B A)$
a)	1/6	1/6
b)	1/6	1/3
c)	1/6	3/36
d)	1/9	1/3



Gambler's fallacy

Assume we toss 51 fair coins.

Assume we have seen 50 coins, and they are all “tails”.

What are the odds the 51st coin is “heads”?

A = first 50 coins are “tails”

B = first 50 coins are “tails”, 51st coin is “heads”

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/2^{51}}{2/2^{51}} = \frac{1}{2}$$

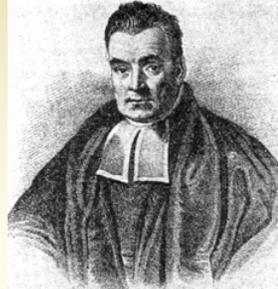
51st coin is independent of
outcomes of first 50 tosses!

Gambler's fallacy = Feels like it's time for “heads”!?

Agenda

- Conditional Probability
- Bayes Theorem ◀
- Law of Total Probability
- More Examples

Bayes Theorem



A formula to let us “reverse” the conditional.

Theorem. (Bayes Rule) For events A and B , where $P(A), P(B) > 0$,

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

$P(A)$ is called the **prior** (our belief without knowing anything)

$P(A|B)$ is called the **posterior** (our belief after learning B)

Bayes Theorem Proof

Claim:

$$P(A), P(B) > 0 \Rightarrow P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes Theorem Proof

Claim:

$$P(A), P(B) > 0 \Rightarrow P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

By definition of conditional probability

$$P(A \cap B) = P(A|B)P(B)$$

Swapping A , B gives

$$P(B \cap A) = P(B|A)P(A)$$

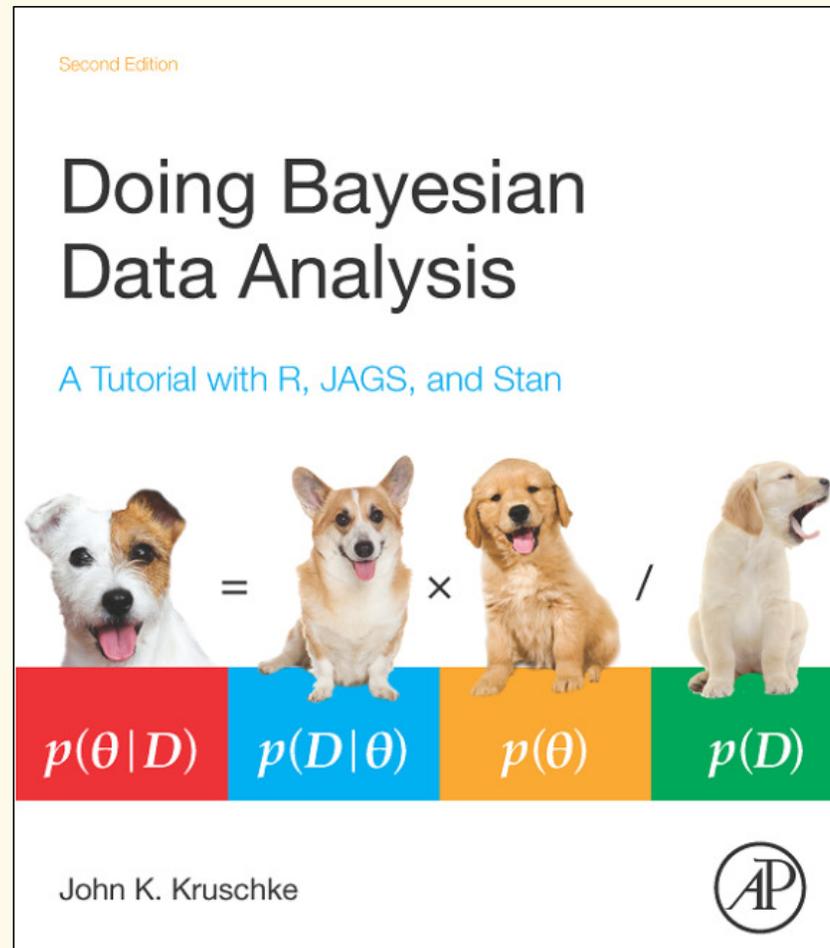
But $P(A \cap B) = P(B \cap A)$, so

$$P(A|B)P(B) = P(B|A)P(A)$$

Dividing both sides by $P(B)$ gives

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

Brain Break



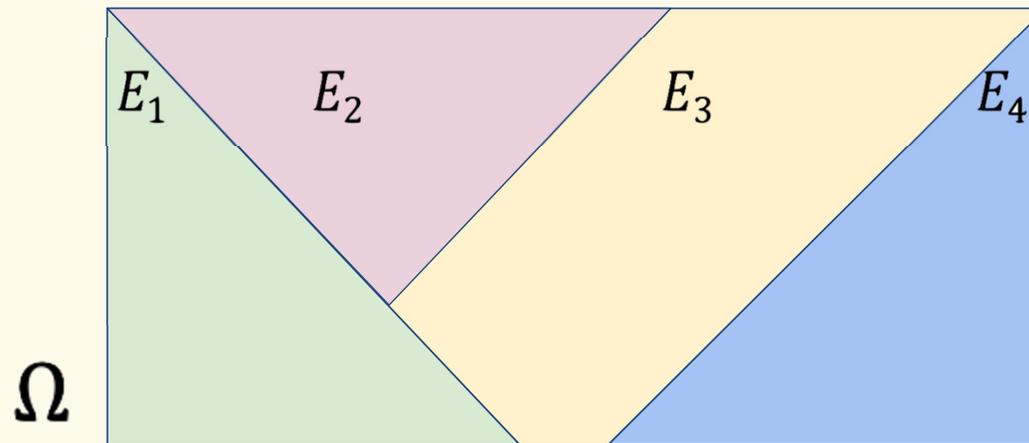
Agenda

- Conditional Probability
- Bayes Theorem
- Law of Total Probability ◀
- More Examples

Partitions (Idea)

These events **partition** the sample space

1. They “cover” the whole space
2. They don’t overlap



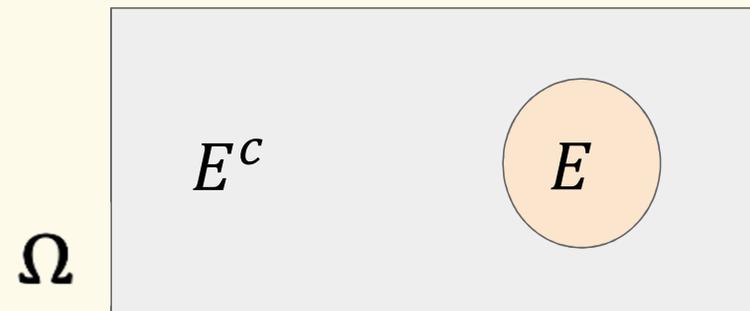
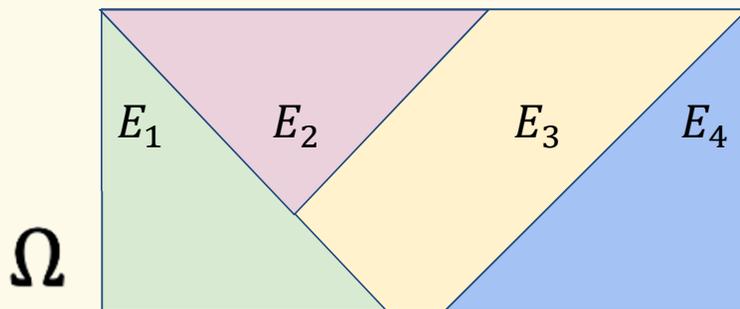
Partition

Definition. Non-empty events E_1, E_2, \dots, E_n **partition** the sample space Ω if
(Exhaustive)

$$E_1 \cup E_2 \cup \dots \cup E_n = \bigcup_{i=1}^n E_i = \Omega$$

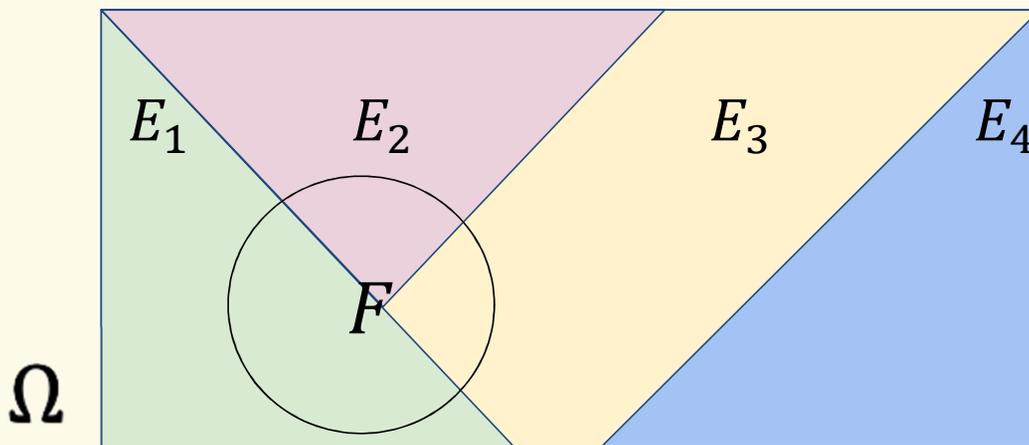
(Pairwise Mutually Exclusive)

$$\forall_i \forall_{i \neq j} E_i \cap E_j = \emptyset$$



Law of Total Probability (Idea)

If we know E_1, E_2, \dots, E_n partition Ω , what can we say about $P(F)$?



Law of Total Probability (LTP)

Definition. If events E_1, E_2, \dots, E_n partition the sample space Ω , then for any event F

$$P(F) = P(F \cap E_1) + \dots + P(F \cap E_n) = \sum_{i=1}^n P(F \cap E_i)$$

Using the definition of conditional probability $P(F \cap E) = P(F|E)P(E)$

We can get the alternate form of this that shows

$$P(F) = \sum_{i=1}^n P(F|E_i)P(E_i)$$

Another Contrived Example

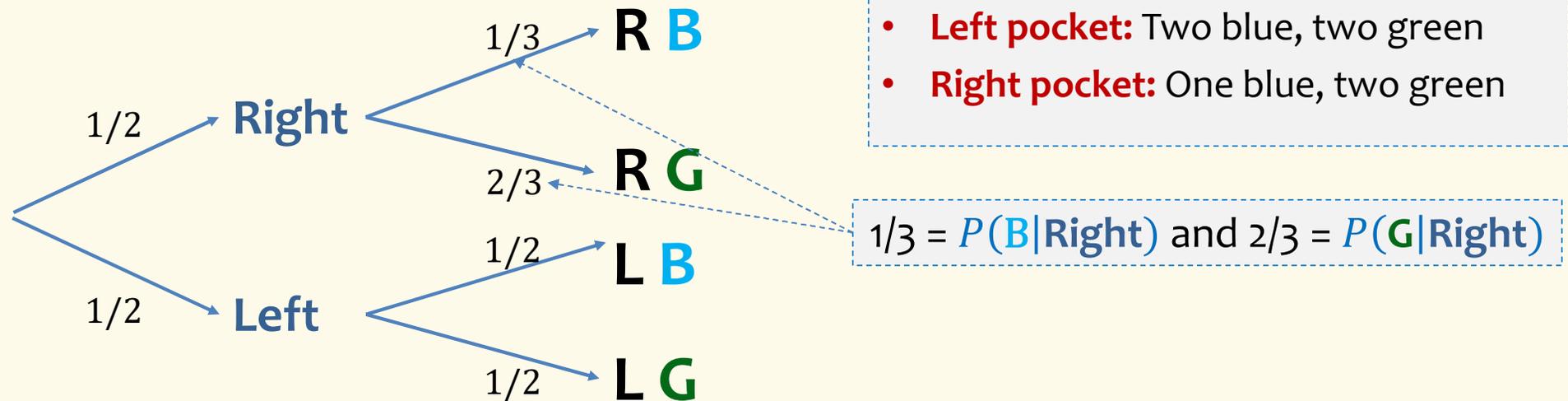
Alice has two pockets:

- **Left pocket:** Two blue balls, two green balls
- **Right pocket:** One blue ball, two green balls.

Alice picks a random ball from a random pocket.

[Both pockets equally likely, each ball equally likely.]

Sequential Process



$$P(\mathbf{B}) = P(\mathbf{B} \cap \text{Left}) + P(\mathbf{B} \cap \text{Right}) \quad (\text{Law of total probability})$$

$$= P(\text{Left}) \times P(\mathbf{B}|\text{Left}) + P(\text{Right}) \times P(\mathbf{B}|\text{Right})$$

$$= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{3} = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$$

Agenda

- Conditional Probability
- Bayes Theorem
- Law of Total Probability
- **More Examples** ◀

Example – Zika Testing

Usually no or mild symptoms (rash); sometimes severe symptoms (paralysis).

During pregnancy: may cause birth defects.

Suppose you took a Zika test, and it returns “positive”, what is the likelihood that you actually have the disease?

- Tests for diseases are rarely 100% accurate.

Zika fever

OVERVIEW SYMPTOMS SPECIALISTS

Fever
Rash
Joint pain
Red eyes



Spread through mosquito bites

Source

A disease caused by Zika virus that's spread through mosquito bites.

The image is a screenshot of a medical information page for Zika fever. It features a light blue header with the title 'Zika fever' and three navigation tabs: 'OVERVIEW', 'SYMPTOMS', and 'SPECIALISTS'. Below the header, there is a list of symptoms: 'Fever', 'Rash', 'Joint pain', and 'Red eyes'. The main visual is a composite image showing a woman with a red rash on her chest and a circular inset showing a mosquito biting her arm. Text at the bottom of the image reads 'Spread through mosquito bites' and 'Source'. Below the image, a caption states: 'A disease caused by Zika virus that's spread through mosquito bites.'

Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika (“true positive”) $P(T|Z)$
- However, the test may yield a “false positive” 1% of the time $P(T|Z^c)$
- 0.5% of the US population has Zika. $P(Z)$

What is the probability you have Zika (event Z) if you test positive (event T)?.

Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika (“true positive”) $P(T|Z)$
- However, the test may yield a “false positive” 1% of the time $P(T|Z^c)$
- 0.5% of the US population has Zika. $P(Z)$

500 have Zika (0.5%)
99,500 do not

What is the probability you have Zika (event Z) if you test positive (event T)?



Suppose we had 100,000 people:

- 490 have Zika and test positive
- 10 have Zika and test negative
- 995 do not have Zika and test positive
- 98,505 do not have Zika and test negative

98% of those with Zika

2% of those with Zika

1% of those without Zika

$$\frac{490}{490 + 995} \approx 0.33$$

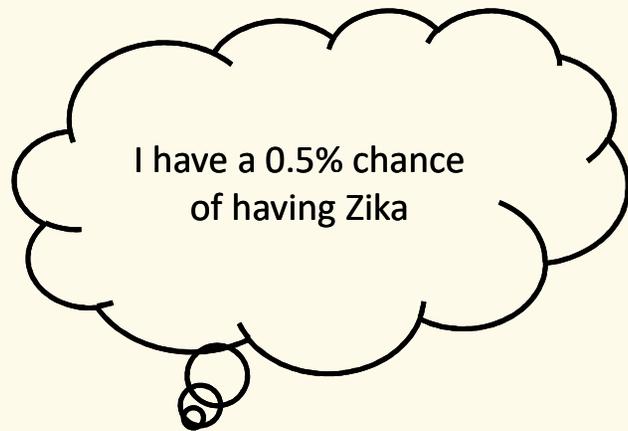
Demo

Philosophy – Updating Beliefs

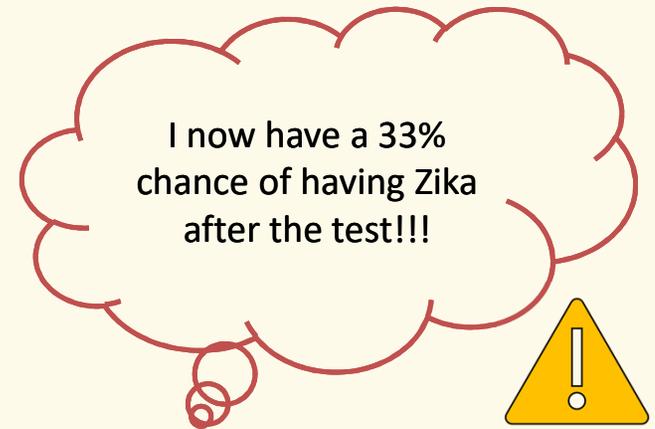
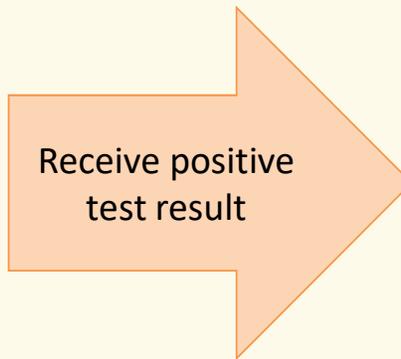
While it's not 98% that you have the disease, your beliefs changed **drastically**

Z = you have Zika

T = you test positive for Zika



Prior: $P(Z)$



Posterior: $P(Z|T)$

Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika (“true positive”) $P(T|Z)$
- However, the test may yield a “false positive” 1% of the time $P(T|Z^c)$
- 0.5% of the US population has Zika. $P(Z)$

What is the probability you test negative (event T^c) if you have Zika (event Z)?

$$P(T^c|Z) = 1 - P(T|Z) = 2\%$$

Conditional Probability Defines a Probability Space

The probability conditioned on \mathcal{A} follows the same properties as (unconditional) probability.

Example. $P(\mathcal{B}^c|\mathcal{A}) = 1 - P(\mathcal{B}|\mathcal{A})$

Formally. (Ω, P) is a probability space and $P(\mathcal{A}) > 0$

 $(\mathcal{A}, P(\cdot | \mathcal{A}))$ is a probability space