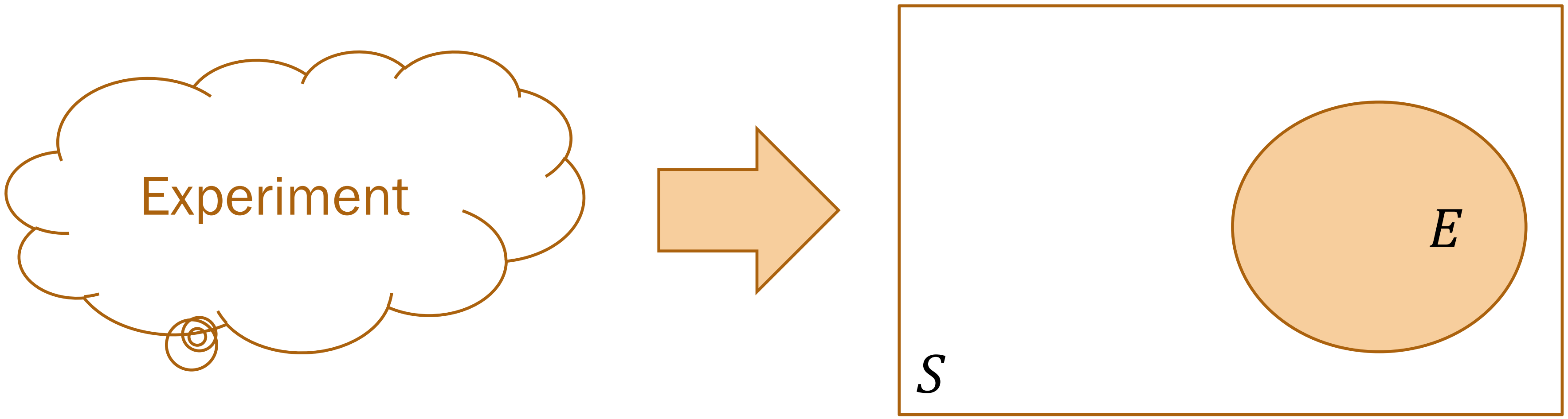
04: Conditional Probability and Bayes

Lisa Yan

September 30, 2019

Key definitions

An experiment in probability:



Sample Space, S: The set of all possible outcomes of an experiment

Event, E: Some subset of S ($E \subseteq S$).

We have the power to redesign our experiment,
provided we can recreate the set of outcomes!

Card Flipping

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

In a 52 card deck, cards are flipped one at a time.
After the first ace (of any suit) appears, consider the next card.

Is P(next card = Ace Spades) < P(next card = 2 Clubs)?

Sample space $S = 52$ in-order cards (shuffle deck)

Event

E_{AS}, next card is Ace Spades

1. Take out Ace of Spades.
2. Shuffle leftover 51 cards.
3. Add Ace Spades after first ace.

$|E_{AS}| = 51! \cdot 1$

E_{2C}, next card is 2 Clubs

1. Take out 2 Clubs.
2. Shuffle leftover 51 cards.
3. Add 2 Clubs after first ace.

$|E_{2C}| = 51! \cdot 1$

$$P(E_{AS}) = P(E_{2C})$$

Today's plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes' Theorem

Dice, our misunderstood friends

Roll two 6-sided dice, yielding values D_1 and D_2.

Let E be event: $D_1 + D_2 = 4$.

What is $P(E)$?

$|S| = 36$

$E = \{(1,3), (2,2), (3,1)\}$

$P(E) = 3/36 = 1/12$

Let F be event: $D_1 = 2$.

What is $P(E,$ given F already observed$)$?

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

$E = \{(2,2)\}$

$P(E,$ given F already observed$) = 1/6$

Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

Written as: $P(E|F)$

Means: "$P(E, \text{given } F \text{ already observed})$"

Sample space → all possible outcomes consistent with F (i.e. $S \cap F$)

Event space → all outcomes in E consistent with F (i.e. $E \cap F$)

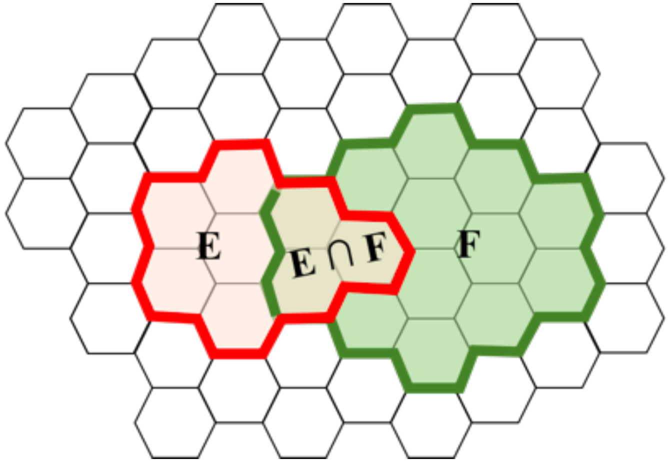
Conditional Probability, equally likely outcomes

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

With equally likely outcomes:

$$P(E|F) = \frac{\text{\# of outcomes in E consistent with F}}{\text{\# of outcomes in S consistent with F}} = \frac{|E \cap F|}{|S \cap F|}$$

$$= \frac{|EF|}{|F|}$$



$$P(E) = \frac{8}{50} \approx 0.16$$

$$P(E|F) = \frac{3}{14} \approx 0.21$$

Slicing up the spam

$P(E|F) = \frac{|EF|}{|F|}$ Equally likely outcomes

24 emails are sent, 6 each to 4 users.

- 10 of the 24 emails are spam.
- All possible outcomes are equally likely.

Let E = user 1 receives 3 spam emails.

What is $P(E)$?

$$P(E) = \frac{\binom{10}{3}\binom{14}{3}}{\binom{24}{6}}$$

$$\approx 0.3245$$

Let F = user 2 receives 6 spam emails.

What is $P(E|F)$?

$$P(E|F) = \frac{\binom{4}{3}\binom{14}{3}}{\binom{18}{6}}$$

$$\approx 0.0784$$

Let G = user 3 receives 5 spam emails.

What is $P(G|F)$?

⚠️

$$P(G|F) = \frac{\binom{4}{5}\binom{14}{1}}{\binom{18}{6}}$$

$$= 0$$

No way to choose 5 spam from 4 remaining spam emails!

Quick check

You have a flowering plant.

Let E = Flowers bloom
F = It gets watered
G = It gets sun

In English, how do you interpret $P(E|FG)$?

The probability that…

A. …flowers bloom given the probability that it gets water and it gets sun
B. …flowers bloom given it gets watered given it gets sun
C. …flowers bloom given (it gets watered and it gets sun)
D. All/none/other

Quick check

You have a flowering plant.

Let E = Flowers bloom
F = It gets watered
G = It gets sun

In English, how do you interpret $P(E|FG)$?

The probability that…

A. …flowers bloom given the probability that it gets water and it gets sun

B. …flowers bloom given it gets watered given it gets sun

C. …flowers bloom given (it gets watered and it gets sun)

D. All/none/other

Conditional probability in general

General definition of conditional probability:

$$P(E|F) = \frac{P(EF)}{P(F)}$$

The Chain Rule (aka Product rule):

$$P(EF) = P(F)P(E|F)$$

👉 These properties hold even when outcomes are not equally likely.

NETFLIX

and Learn

Netflix and Learn

$P(E|F) = \frac{P(EF)}{P(F)}$ Definition of Cond. Probability

Let E = a user watches Life is Beautiful.

What is $P(E)$?

❌ Equally likely outcomes?

$S = \{\text{watch, not watch}\}$

$E = \{\text{watch}\}$

$P(E) = 1/2$?

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n} \approx \frac{\text{\# people who have watched movie}}{\text{\# people on Netflix}}$$

$$= 10{,}234{,}231 \;/\; 50{,}923{,}123 \approx 0.20$$

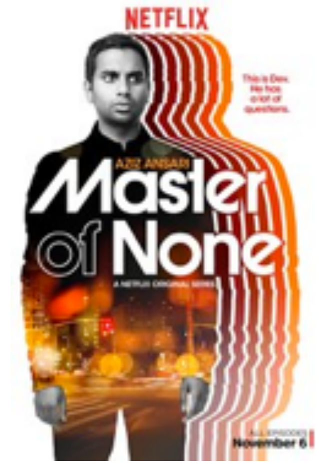
Netflix and Learn

$$P(E|F) = \frac{P(EF)}{P(F)}$$ Definition of Cond. Probability

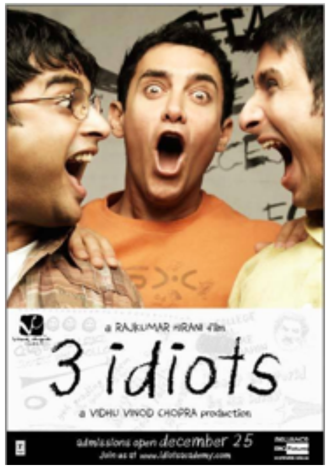
Let E be the event that a user watches the given movie.

$P(E) = 0.19$

$P(E) = 0.32$



$P(E) = 0.20$



$P(E) = 0.09$

$P(E) = 0.20$

Netflix and Learn

$P(E|F) = \frac{P(EF)}{P(F)}$ Definition of Cond. Probability

Let E = a user watches Life is Beautiful.

Let F = a user watches Amelie.

What is the probability that a user watches Life is Beautiful, given they watched Amelie?

$$P(E|F)$$

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{\frac{\text{\# people who have watched both}}{\text{\# people on Netflix}}}{\frac{\text{\# people who have watched Amelie}}{\text{\# people on Netflix}}}$$

$$= \frac{\text{\# people who have watched both}}{\text{\# people who have watched Amelie}}$$

$$\approx 0.42$$

Netflix and Learn

$P(E|F) = \frac{P(EF)}{P(F)}$ Definition of Cond. Probability

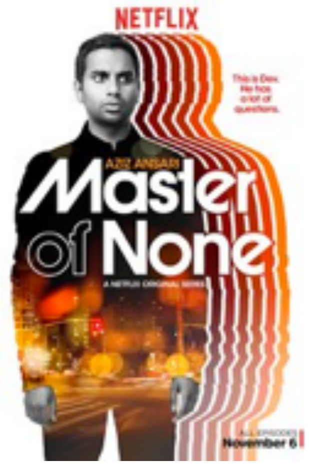
Let E be the event that a user watches the given movie.
Let F be the event that the same user watches Amelie.

$P(E) = 0.19$

$P(E|F) = 0.14$

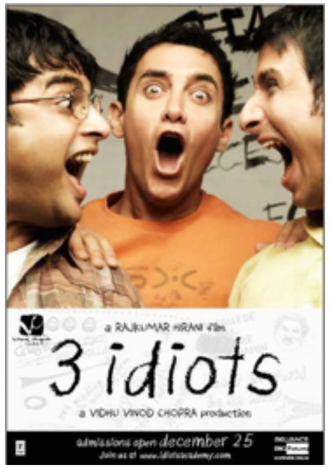
$P(E) = 0.32$

$P(E|F) = 0.35$



$P(E) = 0.20$

$P(E|F) = 0.20$



$P(E) = 0.09$

$P(E|F) = 0.72$

$P(E) = 0.20$

$P(E|F) = 0.42$

Today’s plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes’ Theorem

So far

(just now)

$$P(EF)$$

Chain rule
(Product rule)

Definition of
conditional probability

$$P(E|F)$$

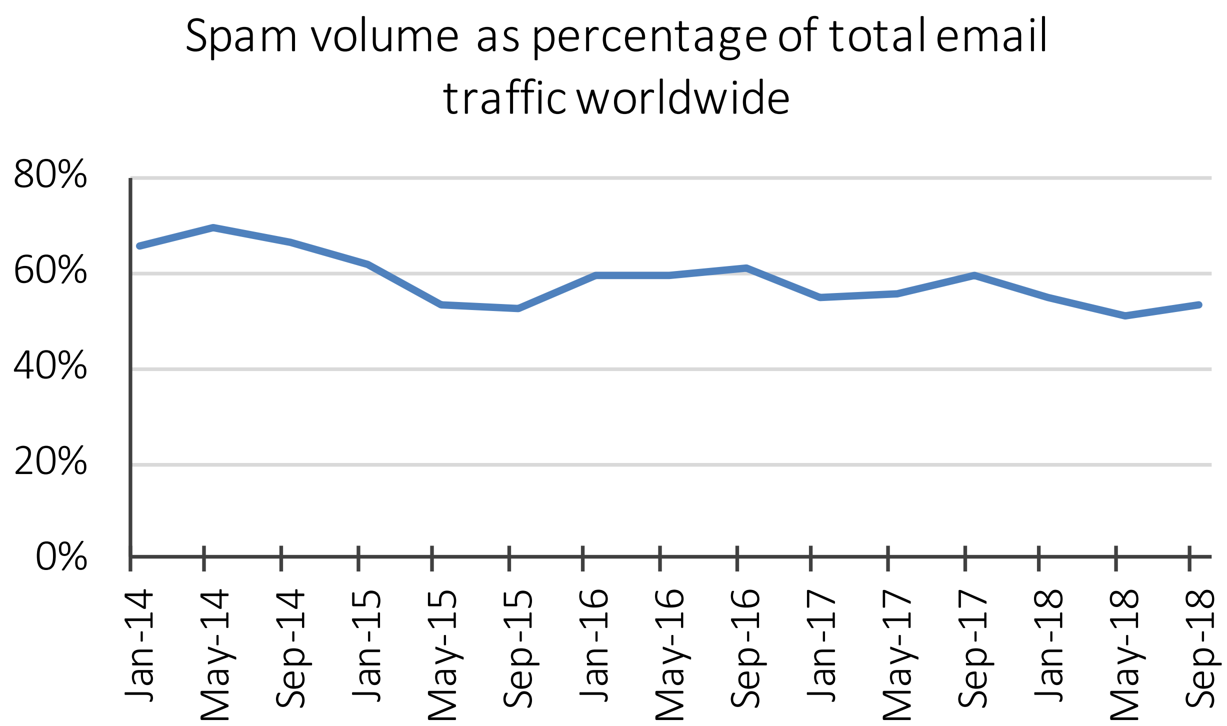
Law of Total
Probability

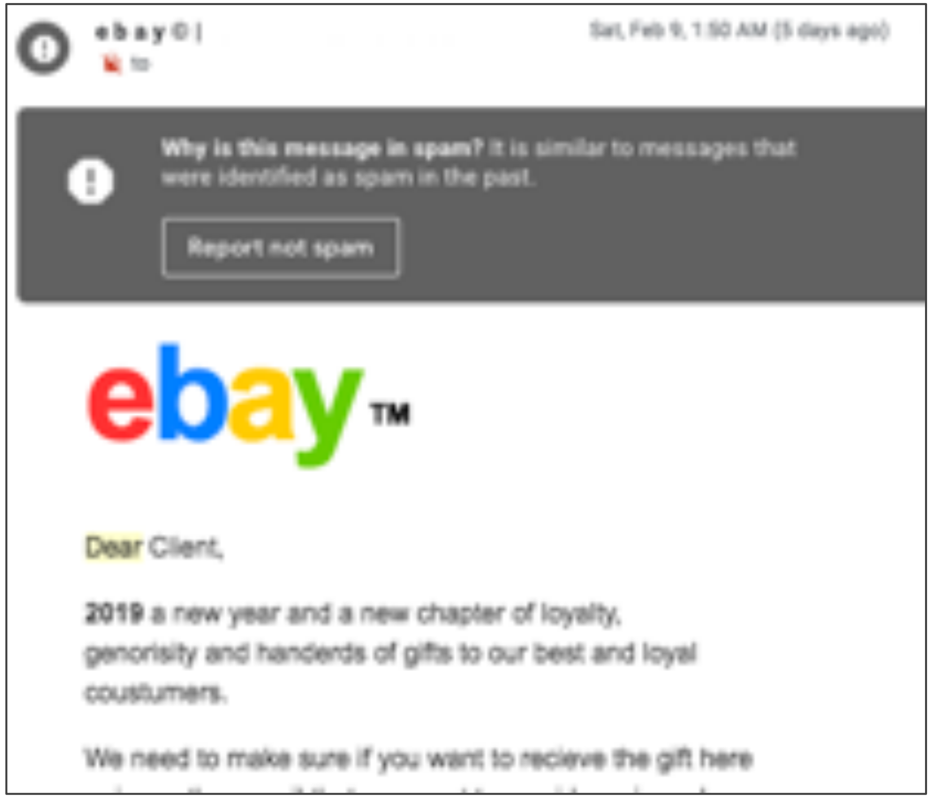
Bayes'
Theorem

$$P(E)$$

$$P(F|E)$$

Detecting spam email





We can easily calculate how many spam emails contain "Dear":

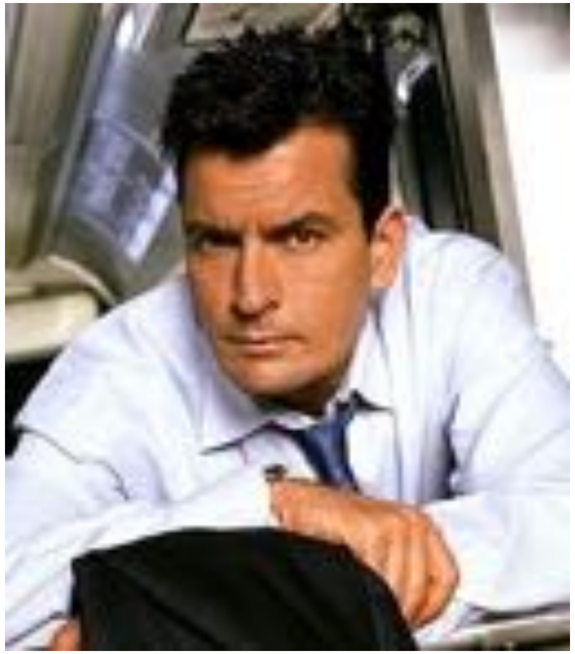
$$P(E|F) = P\left(\text{“Dear”} \,\middle|\, \begin{matrix}\text{Spam}\\ \text{email}\end{matrix}\right)$$

But what is the probability that an email containing "Dear" is spam?

$$P(F|E) = P\left(\begin{matrix}\text{Spam}\\ \text{email}\end{matrix} \,\middle|\, \text{“Dear”}\right)$$

Thomas Bayes

Rev. Thomas Bayes (~1701-1761):
British mathematician and Presbyterian minister



He looked remarkably similar to Charlie Sheen
(but that's not important right now)

Today's tasks

$P(EF)$

Chain rule
(Product rule)

Definition of
conditional probability

$P(E|F)$

Law of Total
Probability

Bayes'
Theorem

$P(E)$

$P(F|E)$

Law of Total Probability

$P(E|F) \Rightarrow P(E)$

<u>Thm</u> Let F be an event where $P(F) > 0$. For any event E,

$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$

<u>Proof</u>

1. F and F^C are disjoint s.t. $F \cup F^C = S$ — Def. of complement
2. $E = (EF) \cup (EF^C)$ — (see board)
3. $P(E) = P(EF) + P(EF^C)$ — Additivity axiom
4. $P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$ — Chain rule (product rule)

General Law of Total Probability

<u>Thm</u> For disjoint events $F_1, F_2, \ldots, F_n$ s.t. $F_1 \cup F_2 \cup \cdots \cup F_n = S$,

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

Finding $P(E)$ from $P(E|F)$

$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$ Law of Total Probability

- Flip a fair coin.
- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

You win if you roll a 6. What is P(winning)?

1. Define events & state goal
2. Identify known probabilities
3. Solve

Let: E: win, F: flip heads
Want: $P(\text{win})$
$= P(E)$

$P(\text{win}|\text{H}) = P(E|F) = 1/6$
$P(\text{H}) = P(F) = 1/2$
$P(\text{win}|\text{T}) = P(E|F^C) = 0$
$P(\text{T}) = P(F^C) = 1 - 1/2$

$$P(E) = (1/6)(1/2) + (0)(1/2) = \frac{1}{12} \approx 0.083$$

Break for jokes/ announcements

Announcements

Section sign-ups

Results: later today
Late signups/changes: later today

Concept checks

Was due: Today 1:00pm
Now due: Tomorrow 1:00pm
(due Tuesdays from now on)

Problem set 1

Gradescope portal: soon
LaTeX tutorial: online

Today's plan

Conditional Probability and Chain Rule

Law of Total Probability

Bayes' Theorem

Today's tasks

$P(EF)$

Chain rule
(Product rule)

Definition of
conditional probability

$P(E|F)$

Law of Total
Probability

Bayes'
Theorem

$P(E)$

$P(F|E)$

(silent drumroll)

Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

<u>Thm</u> For any events E and F where $P(E) > 0$ and $P(F) > 0$,

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

<u>Proof</u>

2 steps! See board

<u>Expanded form</u>:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

<u>Proof</u>

1 more step! See board

Detecting spam email

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- 60% of all email in 2016 is spam. $P(F)$
- 20% of spam has the word "Dear" $P(E|F)$
- 1% of non-spam (aka ham) has the word "Dear" $P(E|F^C)$

You get an email with the word "Dear" in it.

What is the probability that the email is spam? Want: $P(F|E)$

1. Define events & state goal
2. Identify <u>known</u> probabilities
3. Solve

Let: E: "Dear", F: spam

Want: $P(\text{spam}\,|\,\text{“Dear”})$

$= P(F|E)$

$P(\text{spam}) = 0.60 = P(F)$

$P(\text{“Dear”}\,|\,\text{spam}) = 0.20 = P(E|F)$

$P(\text{“Dear”}\,|\,\text{spam}^C) = 0.01 = P(E|F^C)$

Detecting spam email

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- 60% of all email in 2016 is spam. $P(F)$
- 20% of spam has the word "Dear" $P(E|F)$
- 1% of non-spam (aka ham) has the word "Dear" $P(E|F^C)$

You get an email with the word "Dear" in it.

What is the probability that the email is spam? Want: $P(F|E)$

1. Define events & state goal
2. Identify known probabilities
3. Solve

Let: E: "Dear", F: spam

Want: P(spam | "Dear") $= P(F|E)$

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

$$= \frac{\boxed{(0.20)(0.6)}}{\boxed{(0.20)(0.6)} + \boxed{(0.01)(0.4)}} \approx 0.967$$

Spam ∩ "Dear" Spam^C ∩ "Dear"

Bayes' Theorem terminology

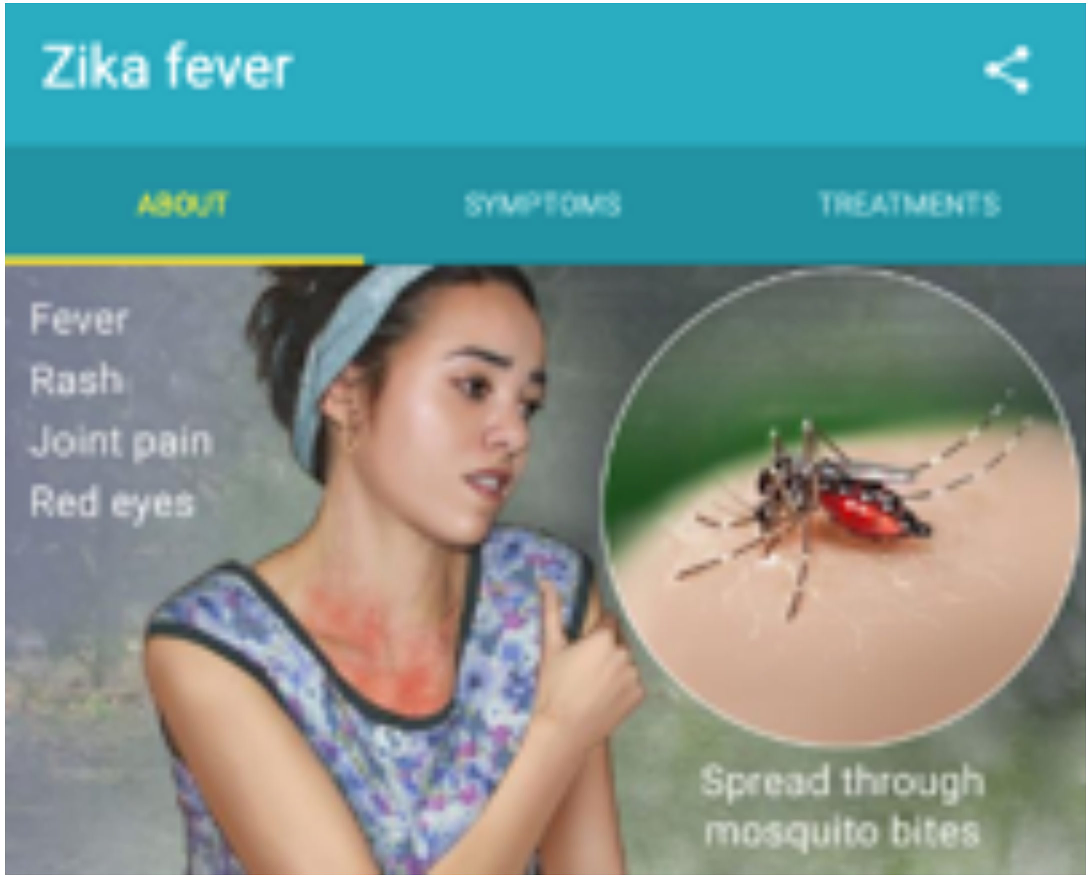
- 60% of all email in 2016 is spam. $P(F)$ prior
- 20% of spam has the word "Dear" $P(E|F)$ likelihood
- 1% of non-spam (aka ham) has the word "Dear" $P(E|F^C)$

You get an email with the word "Dear" in it.

What is the probability that the email is spam? Want: $P(F|E)$ posterior

$$\underbrace{P(F|E)}_{\text{posterior}} = \frac{\overbrace{P(E|F)}^{\text{likelihood}}\overbrace{P(F)}^{\text{prior}}}{\underbrace{P(E)}_{\text{normalization constant}}}$$

Zika, an autoimmune disease



A disease spread through mosquito bites.
Usually no symptoms; worst case paralysis. During pregnancy: may cause birth defects

Ziika Forest, Uganda

Rhesus monkeys

If a test returns positive, what is the likelihood you have the disease?

Taking tests: Confusion matrix

Take test

Fact, F Has disease
or F^C No disease

Evidence, E Test positive
or E^C Test negative

		Fact	
		F, disease +	F^C, disease –
Evidence	E, Test +	True positive $P(E \mid F)$	False positive $P(E \mid F^C)$
	E^C, Test –	False negative $P(E^C \mid F)$	True negative $P(E^C \mid F^C)$

If a test returns positive, what is the likelihood you have the disease?

Taking tests: Confusion matrix

Take test

Fact, F	Has disease
or F^C	No disease

Evidence, E	Test positive
or E^C	Test negative

		Fact	
		F, disease +	F^C, disease –
Evidence	E, Test +	True positive $P(E\|F)$	False positive $P(E\|F^C)$
	E^C, Test –	False negative $P(E^C\|F)$	True negative $P(E^C\|F^C)$

If a test returns positive, what is the likelihood you have the disease?

Zika Testing

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive").
- However, the test has a "false positive" rate of 1%.
- 0.5% of the US population has Zika.

What is the likelihood you have Zika if you test positive? Want: $P(F|E)$

1. Define events & state goal
2. Identify <u>known</u> probabilities
3. Solve

Let: E = you test positive
F = you actually have the disease

Want:
P(disease | test+)
$= P(F|E)$

A. $P(E|F)$, $P(E^C|F)$, $P(F)$

B. $P(F|E)$, $P(F^C|E)$, $P(F)$

C. $P(E|F)$, $P(E|F^C)$, $P(F)$

Zika Testing

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive").
- However, the test has a "false positive" rate of 1%.
- 0.5% of the US population has Zika.

What is the likelihood you have Zika if you test positive? Want: $P(F|E)$

1. Define events & state goal

Let: E = you test positive
F = you actually have the disease

Want:
P(disease | test+)
$= P(F|E)$

2. Identify known probabilities

A. $P(E|F)$, $P(E^C|F)$, $P(F)$

B. $P(F|E)$, $P(F^C|E)$, $P(F)$

C. $P(E|F)$, $P(E|F^C)$, $P(F)$

3. Solve

Zika Testing

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

What is the likelihood you have Zika if you test positive? Want: $P(F|E)$

1. Define events & state goal
2. Identify known probabilities
3. Solve

Let: E = you test positive
F = you actually have the disease

Want:
P(disease | test+)
$= P(F|E)$

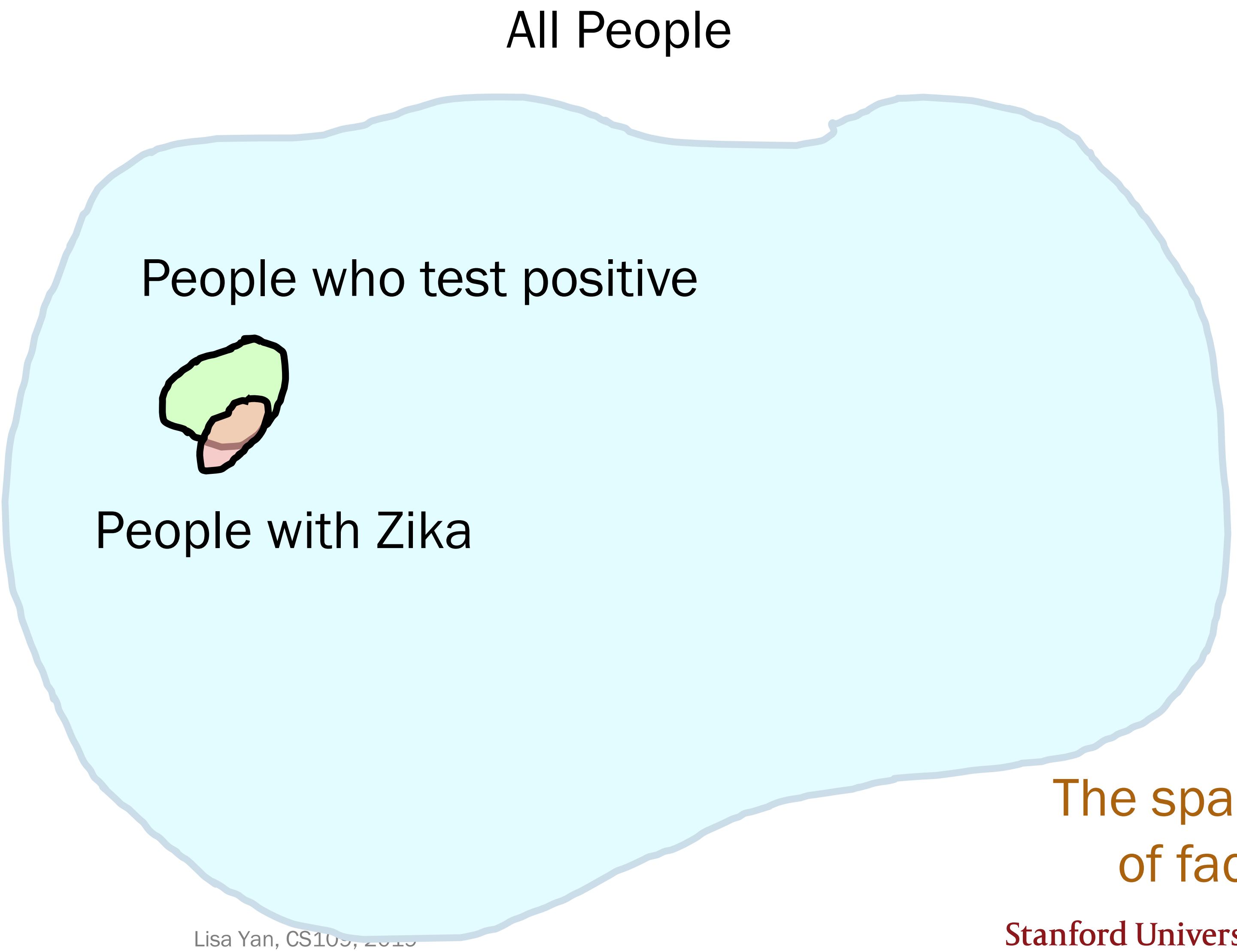
$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

$$= \frac{(0.98)(0.005)}{(0.98)(0.005) + (0.01)(1 - 0.005)} \approx 0.330$$

Bayes' Theorem intuition

<u>Original question</u>:

What is the likelihood you have Zika if you test positive for the disease?



The space of facts

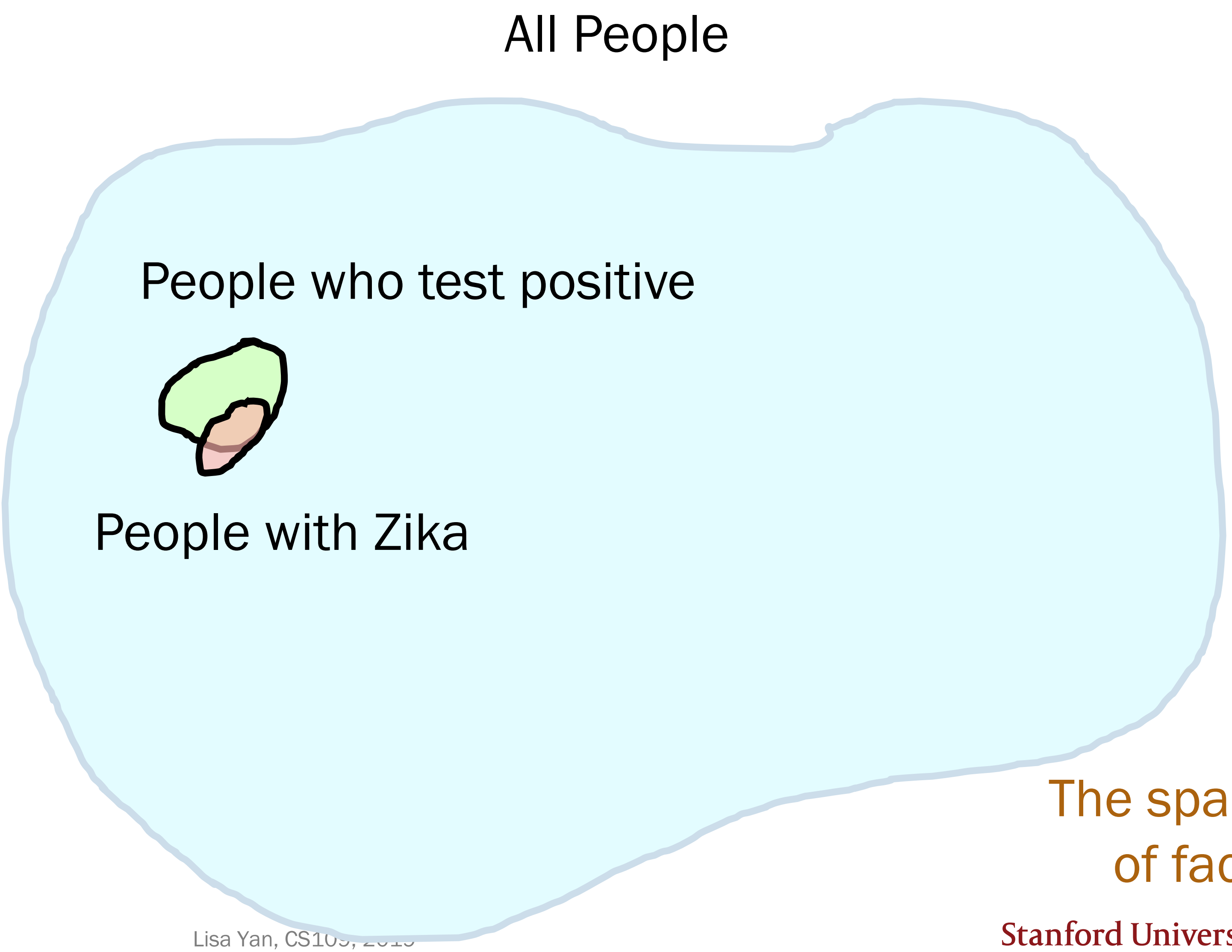
Bayes' Theorem intuition

Original question:

What is the likelihood you have Zika if you test positive for the disease?

Interpret

Interpretation:
Of the people who test positive, how many actually have Zika?



The space of facts

Bayes' Theorem intuition

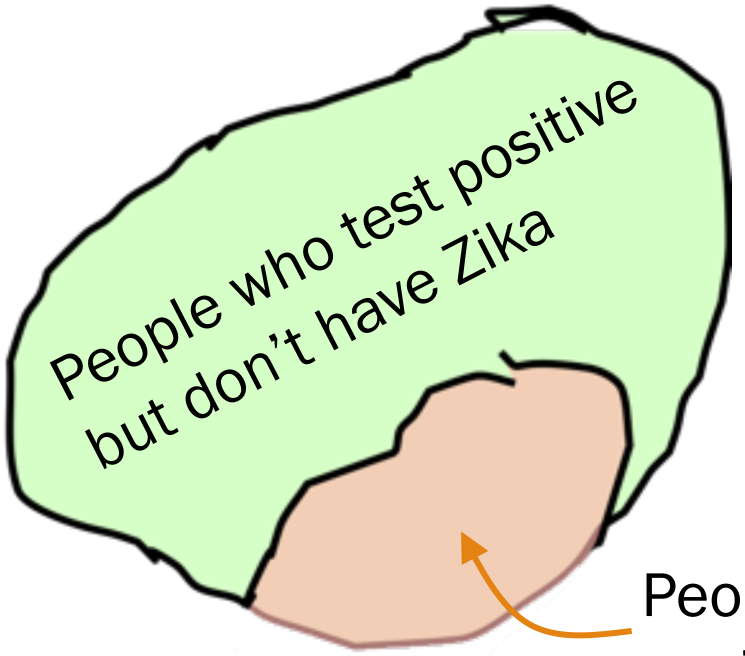
<u>Original question</u>:

What is the likelihood you have Zika if you test positive for the disease?

Interpret

<u>Interpretation</u>:
Of the people who test positive, how many actually have Zika?

People who test positive



People who test positive and have Zika

The space of facts,
<u>**conditioned**</u> on a positive test result

Zika Testing

- A test is 98% effective at detecting Zika ("true positive").
- However, the test has a "false positive" rate of 1%.
- 0.5% of the US population has Zika.

What is the likelihood you have Zika if you test positive?

Say we have 1000 people:

5 have Zika
and test positive
985 do not have Zika
and test negative.
10 do not have Zika
and test positive.

≈ 0.333

Demo (class website)

Why is Bayes' so important?

👉 It links belief to evidence in probability!

Update your beliefs with Bayes' Theorem

E = you test positive for Zika
F = you actually have the disease

I have a 0.5% chance of having Zika.

Take test, results positive

With these test results, I now have a 33% chance of having Zika!!!

$P(F)$

$P(F|E)$

Why it's still good to get tested

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

Let: E = you test positive
F = you actually have the disease

Let: E^C = you test negative for Zika with this test.

What is $P(F|E^C)$?

A. Complement: $P(F|E^C) = 1 - P(F|E)$
B. Find $P(E^C|F)$, $P(E^C|F^C)$, then do Bayes'
C. Calculate $P(E^C F)/P(E^C)$ directly
D. Give up/stay healthy

Why it's still good to get tested

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

Let: E = you test positive
F = you actually have the disease

Let: E^C = you test negative for Zika with this test.

What is $P(F|E^C)$?

A. Complement: $P(F|E^C) = 1 - P(F|E)$ don't do this
B. Find $P(E^C|F)$, $P(E^C|F^C)$, then do Bayes'
C. Calculate $P(E^CF)/P(E^C)$ directly ⚠ No $P(E^CF)$
D. Give up/stay healthy

Why it's still good to get tested

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

Let: E = you test positive
F = you actually have the disease

Let: E^C = you test negative for Zika with this test.

	F, disease +	F^C, disease –
E, Test +	True positive $P(E\|F) = 0.98$	False positive $P(E\|F^C) = 0.01$

What is $P(F|E^C)$?

B. Find $P(E^C|F)$, $P(E^C|F^C)$, then do Bayes'

$$P(F|E^C) = \frac{P(E^C|F)P(F)}{P(E^C|F)P(F) + P(E^C|F^C)P(F^C)}$$

Why it's still good to get tested

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

Let: E = you test positive
F = you actually have the disease

Let: E^C = you test negative for Zika with this test.

What is $P(F|E^C)$?

	F, disease +	F^C, disease –
E, Test +	True positive $P(E|F) = 0.98$	False positive $P(E|F^C) = 0.01$
E^C, Test –	False negative $P(E^C|F) = 0.02$	True negative $P(E^C|F^C) = 0.99$

B. Find $P(E^C|F)$, $P(E^C|F^C)$, then do Bayes'

$$P(F|E^C) = \frac{P(E^C|F)P(F)}{P(E^C|F)P(F) + P(E^C|F^C)P(F^C)}$$

Why it's still good to get tested

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$ Bayes' Theorem

- A test is 98% effective at detecting Zika ("true positive"). $P(E|F)$
- However, the test has a "false positive" rate of 1%. $P(E|F^C)$
- 0.5% of the US population has Zika. $P(F)$

Let: E = you test positive
F = you actually have the disease

Let: E^C = you test negative for Zika with this test.

What is $P(F|E^C)$?

	F, disease +	F^C, disease –
E, Test +	True positive $P(E\|F) = 0.98$	False positive $P(E\|F^C) = 0.01$
E^C, Test –	False negative $P(E^C\|F) = 0.02$	True negative $P(E^C\|F^C) = 0.99$

$$P(F|E^C) = \frac{P(E^C|F)P(F)}{P(E^C|F)P(F) + P(E^C|F^C)P(F^C)} = \frac{(0.02)(0.005)}{(0.02)(0.005) + (0.99)(1 - 0.005)} \approx 0.0001$$

Why it's still good to get tested

E = you test positive for Zika
F = you actually have the disease
E^C = you test negative for Zika

I have a **0.5%** chance of having Zika disease.

$P(F)$

Take test, results positive

With these test results, I now have a **33%** chance of having Zika!!!

$P(F|E)$

Take test, results negative

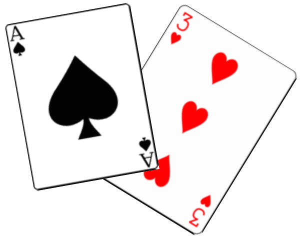
With these test results, I now have a **0.01%** chance of having Zika disease!!!

$P(F|E^C)$

This class going forward

<u>Last week</u>

Equally likely
events



$P(E \cap F)$ $P(E \cup F)$

(counting, combinatorics)

<u>Today and for most of this course</u>

Not equally likely events

$P(E = \text{Evidence} \mid F = \text{Fact})$

(collected from data)

Bayes'

$P(F = \text{Fact} \mid E = \text{Evidence})$

(categorize
a new datapoint)

Another conditional probability example

Monty Hall Problem

and Wayne Brady

Monty Hall Problem aka Let's Make a Deal

Behind one door is a prize (equally likely to be any door).

Behind the other two doors is nothing

1. We choose a door
2. Host opens 1 of other 2 doors, revealing nothing
3. We are given an option to change to the other door.

Doors A,B,C

Should we switch?

Note: If we don't switch, P(win) = 1/3 (random)

We are comparing P(win) and P(win|switch).

If we switch

Without loss of generality, say we pick A (out of Doors A,B,C).

1/3 1/3 1/3

A = prize	B = prize	C = prize
• Host opens B or C • We switch • We always lose	• Host must open C • We switch to B • We always win	• Host must open B • We switch to C • We always win
P(win | A prize, picked A, switched) = 0	P(win | B prize, picked A, switched) = 1	P(win | C prize, picked A, switched) = 1

P(win | picked A, switched) = 1/3 * 0 + 1/3 * 1 + 1/3 * 1 = 2/3

You should switch.

Monty Hall, 1000 envelope version

Start with 1000 envelopes
(of which 1 is the prize).

1. You choose 1 envelope.

$$\frac{1}{1000} = \text{P(envelope is prize)}$$

$$\frac{999}{1000} = \text{P(other 999 envelopes have prize)}$$

2. I open 998 of remaining 999
(showing they are empty).

$$\frac{999}{1000} = \text{P(998 empty envelopes had prize)} + \text{P(last other envelope has prize)}$$

$$= \text{P(last other envelope has prize)}$$

3. Should you switch?

$$\text{P(you win without switching)} = \frac{1}{\text{original \# envelopes}}$$

$$\text{P(you win with switching)} = \frac{\text{original \# envelopes - 1}}{\text{original \# envelopes}}$$