Law of Total Probability, Bayes Rule and More on Independence

CSE 312 Spring 26

Lecture 6

Thank you for your feedback!!!

Some people mentioned that I was going too fast in parts and would like more examples in class.

How do we address this?

What I can do and what I will try to do:

- *I will try to be more consistent in my pacing and go over the examples a little bit slower.*
- *I will try to give more examples.*

What you can do:

- *Slow me down! Ask questions!!!*
- *Give me feedback between classes. (I've added feedback as a category on Ed.)*
- *Do the concept checks!!*
- *Please come to office hours.*
- *Ask the 312 Learning Assistant to explain anything that is confusing.*
- *Form study groups and go over all the section problems together.*
- *Do the readings.*

Agenda

- Recap
- Law of Total Probability
- Bayes Theorem
- More on Independence

Review - Probability

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

Examples:

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

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

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 - Axioms of Probability

Let Ω denote the sample space and $E, F \subseteq \Omega$ be events.

Axiom 1 (Non-negativity): $P(E) \geq 0$
Axiom 2 (Normalization): $P(\Omega) = 1$
Axiom 3 (Countable Additivity): If E and F are mutually exclusive events,
then $P(E \cup F) = P(E) + P(F)$

Corollary 1 (Complementation): $P(E^c) = 1 - P(E)$
Corollary 2 (Monotonicity): If $E \subseteq F$, $P(E) \leq P(F)$
Corollary 3 (Inclusion-Exclusion): $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Review: Equally Likely Outcomes

If (Ω, P) is a **uniform** probability space, then for any event $E \subseteq \Omega$, then

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

Examples:

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

Events:

- Getting at least one head in two coin flips: $E = \{HH, HT, TH\}$

$$P(E) = \frac{3}{4}$$

- Rolling an even number on a die : $E = \{2, 4, 6\}$

$$P(E) = \frac{3}{6} = \frac{1}{2}$$

Review - Conditional Probability

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

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

A relevant example

When you ask ChatGPT or Gemini or any of them a question, they are not "looking up" the answer. They are calculating a conditional probability distribution over a vocabulary of ``tokens". It is answering the question:

$$Pr(w_n | w_1, w_2, \ldots, w_{n-1})\ ?$$

``Given the words that have come before, what is the probability of the next word?"

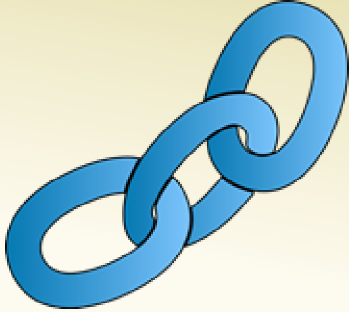
Review - Conditional Probability (uniform case)

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

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

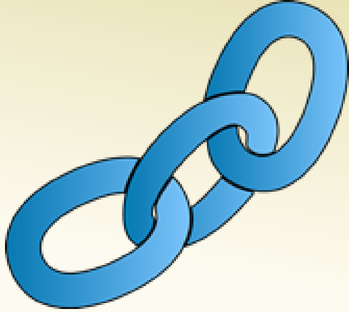
If the probability space is uniform, then $P(E|F) = \frac{|E \cap F|}{|F|}$

Review - Chain Rule



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

Review - Chain Rule (general case)



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

Theorem. (Chain Rule) For events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$,

$$\mathbb{P}(\mathcal{A}_1 \cap \dots \cap \mathcal{A}_n) = \mathbb{P}(\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_2|\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_3|\mathcal{A}_1 \cap \mathcal{A}_2)$$

$$\cdots \mathbb{P}(\mathcal{A}_n|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \dots \cap \mathcal{A}_{n-1})$$

An easy way to remember: We have n tasks and we can do them sequentially, conditioning on the outcome of previous tasks

Review - Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are **independent** if

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}).$$

Alternatively,

- If $\mathbb{P}(\mathcal{A}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(B)$
- If $\mathbb{P}(\mathcal{B}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A})$

Agenda (2)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on independence

Law of Total Probability

For any sets E and F:

$$F = (F \cap E) \cup (F \cap \overline{E})$$

$$\boldsymbol{P}\{F\} = \boldsymbol{P}\{F \cap E\} + \boldsymbol{P}\{F \cap \overline{E}\}$$

$$= \boldsymbol{P}\{F|E\} \cdot \boldsymbol{P}\{E\} + \boldsymbol{P}\{F|\overline{E}\} \cdot \boldsymbol{P}\{\overline{E}\}$$

Law of Total Probability (General case)

For any sets E and F:

$$F = (F \cap E) \cup (F \cap \overline{E})$$

$$\boldsymbol{P}\{F\} = \boldsymbol{P}\{F \cap E\} + \boldsymbol{P}\{F \cap \overline{E}\}$$

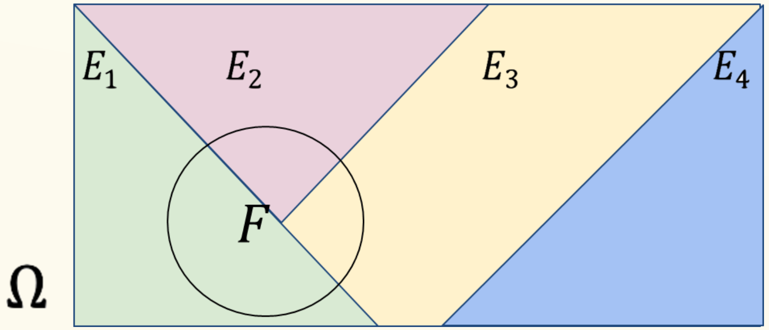
$$= \boldsymbol{P}\{F|E\} \cdot \boldsymbol{P}\{E\} + \boldsymbol{P}\{F|\overline{E}\} \cdot \boldsymbol{P}\{\overline{E}\}$$

Generalizing, we have:

Theorem: **[Law of Total Probability]**

Let $E_1, E_2, \dots, E_n$ partition the state space Ω. Then:

$$\boldsymbol{P}\{F\} = \sum_{i=1}^{n} \boldsymbol{P}\{F \cap E_i\} = \sum_{i=1}^{n} \boldsymbol{P}\{F|E_i\} \cdot \boldsymbol{P}\{E_i\}$$



Dance party?

- Your friend is having a party. The music will either be:
  - Playlist (PL) – probability 0.6
  - DJ - probability 0.3
  - Live band (LB) – probability 0.1

- You are very picky about what music you will dance to. If they go with their playlist, you will dance with probability 0.2, if they go with a DJ, you'll dance with probability 0.6 and if they go with a live band, you'll dance with probability 0.9.

- What's the probability you dance and the music is a DJ?

Dance Party (2)

$Pr(PL) = 0.6$
$Pr(DJ) = 0.3$
$Pr(LB) = 0.1$

D=Dance
$Pr(D \mid PL) = 0.2$
$Pr(D \mid DJ) = 0.6$
$Pr(D \mid LB) = 0.9$

https://pollev.com/ annakarlin185

What is Pr (Dance and DJ)?

A) 0.3
B) 0.3 x 0.6
C) 0.3^2
D) 0.6

What is the probability you dance?

Theorem: **[Law of Total Probability]**

Let $E_1, E_2, \ldots, E_n$ partition the state space Ω. Then:

$$\boldsymbol{P}\{F\} = \sum_{i=1}^{n} \boldsymbol{P}\{F \cap E_i\} = \sum_{i=1}^{n} \boldsymbol{P}\{F | E_i\} \cdot \boldsymbol{P}\{E_i\}$$

Dance party (4)

- Your friend is having a party. The music will either be:
  - Playlist (PL) – probability 0.6
  - DJ - probability 0.3
  - Live band (LB) – probability 0.1

- You are very picky about what music you will dance to. If they go with their playlist, you will dance with probability 0.2, if they go with a DJ, you'll dance with probability 0.6 and if they go with a live band, you'll dance with probability 0.9.

- What's the probability you dance?

$$\Pr(\text{Dance}) = \Pr(D \mid PL)\Pr(PL) + \Pr(D \mid DJ)\Pr(DJ) + \Pr(D \mid LB)\Pr(LB)$$

Agenda (3)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on Independence

Our First Machine Learning Task: Spam Filtering

`Subject: “FREE $$$ CLICK HERE”`

Suppose you know that 80% of emails you receive are spam.

So a priori, our belief is that any email has an 80% chance of being spam.

How do you update that belief when you see that the subject line contains the phrase “FREE $$$”?

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)

Back to Spam Filtering - 1

```
Subject: “FREE $$$ CLICK HERE”
```

What is the probability this email is spam, given the subject contains “FREE”?
Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.

Back to Spam Filtering - 2

```
Subject: “FREE $$$ CLICK HERE”
```

What is the probability this email is spam, given the subject contains “FREE”?
Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.

You receive a random email:

- Let S be event that the email is spam
- Let F be the event that the email contains the word “FREE”.

Back to Spam Filtering - 3

Subject: “FREE $$$ CLICK HERE”

What is the probability this email is spam, given the subject contains “FREE”?

Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.

$$P(F|\overline{S}) = 0.1$$
$$P(F|S) = 0.7$$
$$P(S) = 0.8$$

You receive a random email:

- Let S be event that the email is spam
- Let F be the event that the email contains the word “FREE”.

Back to Spam Filtering - 4

Subject: “FREE $$$ CLICK HERE”

What is the probability this email is spam, given the subject contains “FREE”?

Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.

$$P(F|\overline{S}) = 0.1$$
$$P(F|S) = 0.7$$
$$P(S) = 0.8$$

You receive a random email:

- Let S be event that the email is spam
- Let F be the event that the email contains the word “FREE”.

By Bayes Rule, $P(S|F) = \dfrac{P(F|S)P(S)}{P(F)}$

Back to Spam Filtering - 5

Subject: "FREE $$$ CLICK HERE"

What is the probability this email is spam, given the subject contains "FREE"?

Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word "FREE" in the subject.
- 70% of spam emails contain the word "FREE" in the subject.
- 80% of emails you receive are spam.

$$P(F|\overline{S}) = 0.1$$
$$P(F|S) = 0.7$$
$$P(S) = 0.8$$

You receive a random email:

- Let S be event that the email is spam
- Let F be the event that the email contains the word "FREE".

By Bayes Rule, $P(S|F) = \dfrac{P(F|S)P(S)}{P(F)}$

By the Law of Total Probability $P(F) = P(F|S)P(S) + P(F|\overline{S})P(\overline{S})$

Bayes Theorem with Law of Total Probability

Bayes Theorem with LTP: Let $E_1, E_2, \dots, E_n$ be a partition of the sample space, and F and event. Then,

$$P(E_1|F) = \frac{P(F|E_1)P(E_1)}{P(F)} = \frac{P(F|E_1)P(E_1)}{\sum_{i=1}^{n} P(F|E_i)P(E_i)}$$

Simple Partition: In particular, if E is an event with non-zero probability, then

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

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.

This example and following slides are from Lisa Yan (Stanford).

Example – Zika Testing (1)

Z have Zika
T test positive

Suppose we know the following Zika stats

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

What is the probability a random person in the US has Zika (event Z) given that they test positive (event T)?

Example – Zika Testing (2)

Z have Zika
T test positive

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive") $P(T|Z) = 0.98$
- However, the test may yield a "false positive" 1% of the time $P(T|Z^c) = 0.01$
- 0.5% of the US population has Zika. $P(Z) = 0.005$

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

By Bayes Rule, $P(Z|T) = \dfrac{P(T|Z)P(Z)}{P(T)}$

Example – Zika Testing (3)

Z have Zika
T test positive

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive") $P(T|Z) = 0.98$
- However, the test may yield a "false positive" 1% of the time $P(T|Z^c)=0.01$
- 0.5% of the US population has Zika. $P(Z)=0.005$

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

By Bayes Rule, $P(Z|T) = \dfrac{P(T|Z)P(Z)}{P(T)} = \dfrac{0.98 \cdot 0.005}{P(T)}$

By the Law of Total Probability, $P(T) = P(T|Z)P(Z) + P(T|Z^c)P(Z^c)$

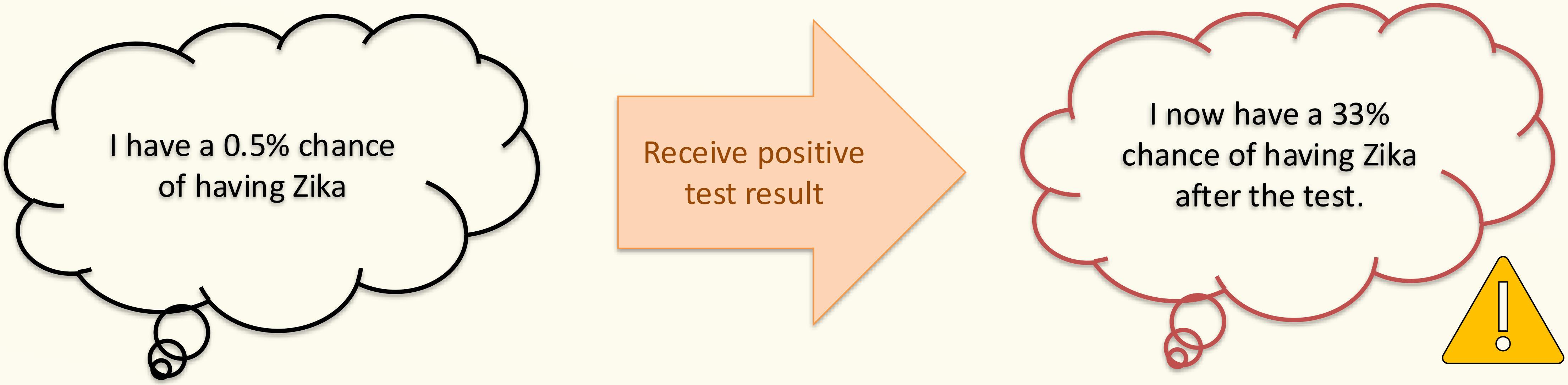
$$= 0.98 \cdot 0.005 + 0.01 \cdot 0.995$$

Philosophy – Updating Beliefs

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

Z = you have Zika

T = you test positive for Zika

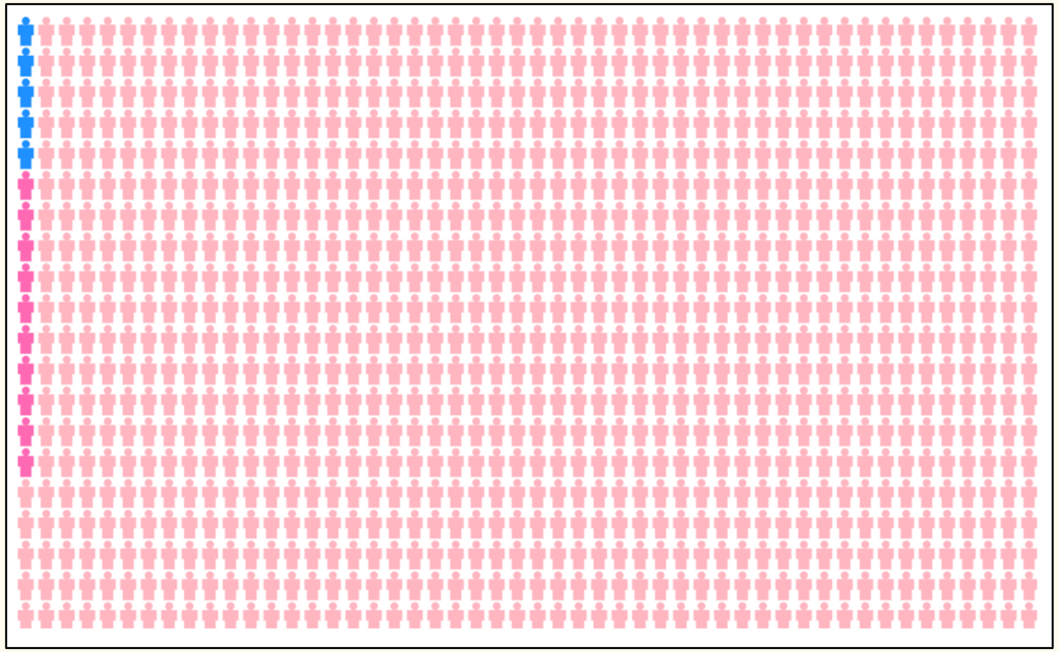


Prior: $P(Z)$

Posterior: $P(Z|T)$

Example – Zika Testing (4)

Have zika blue, don’t pink



Suppose we had 1000 people:

- 5 **have Zika** and all **test positive**
- 985 **do not have Zika** and **test negative**
- 10 **do not have Zika** and **test positive**

Example – Zika Testing (5)

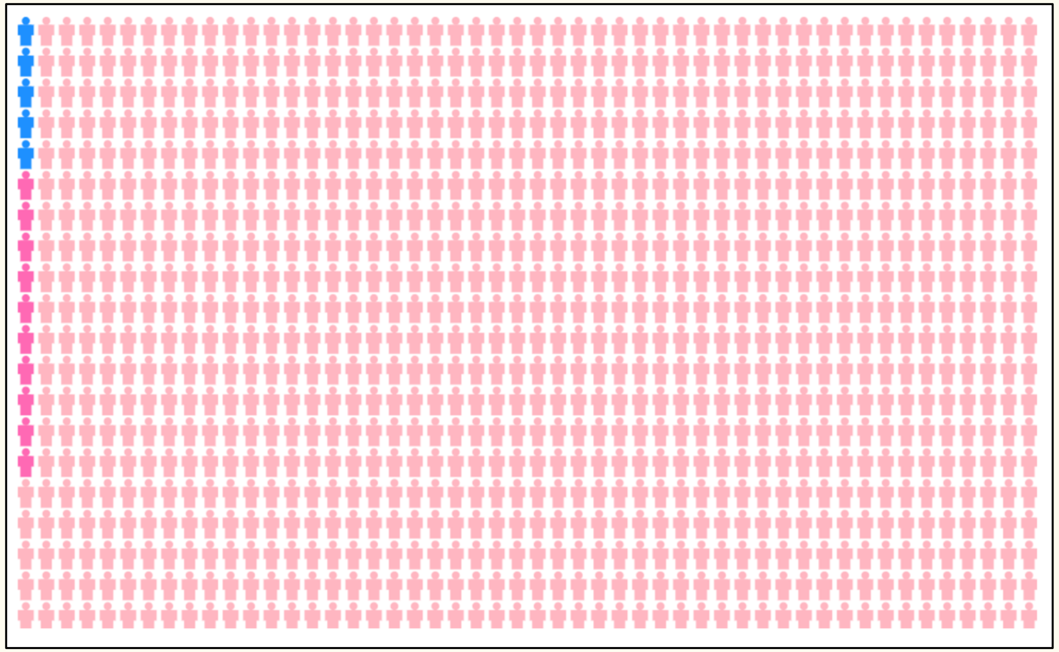
Have zika blue, don't pink

Picture below gives us the following Zika stats

- A test is 100% effective at detecting Zika ("true positive"). $P(T|Z) = 5/5 = 1$
- However, the test may yield a "false positive" 1% of the time $P(T|Z^c) = 10/995$
- 0.5% of the US population has Zika. 5% have it.

$$P(Z) = \frac{995}{1000} = 0.005$$

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



Suppose we had 1000 people:

- 5 **have Zika** and and **test positive**
- 985 **do not have Zika** and **test negative**
- 10 **do not have Zika** and **test positive**

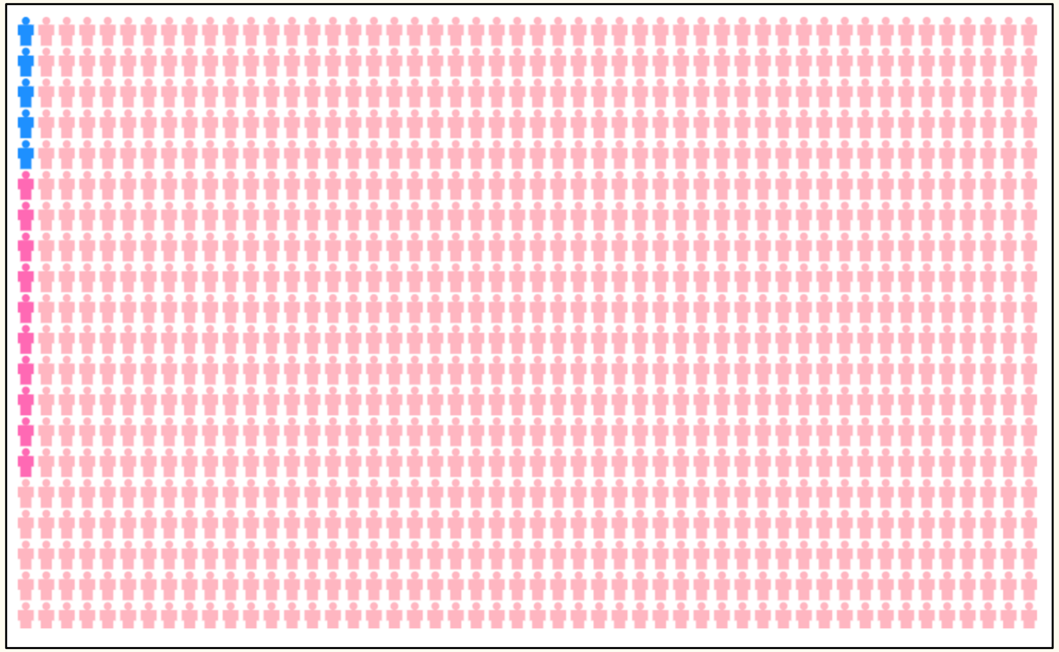
Example – Zika Testing (5)

Have zika blue, don't pink

Picture below gives us the following Zika stats

- A test is 100% effective at detecting Zika ("true positive"). $P(T|Z) = 5/5 = 1$
- However, the test may yield a "false positive" 1% of the time $P(T|Z^c) = 10/995$
- 0.5% of the US population has Zika. 5% have it. $P(Z) = \frac{995}{1000} = 0.005$

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



Suppose we had 1000 people:

- 5 **have Zika** and and **test positive**
- 985 **do not have Zika** and **test negative**
- 10 **do not have Zika** and **test positive**

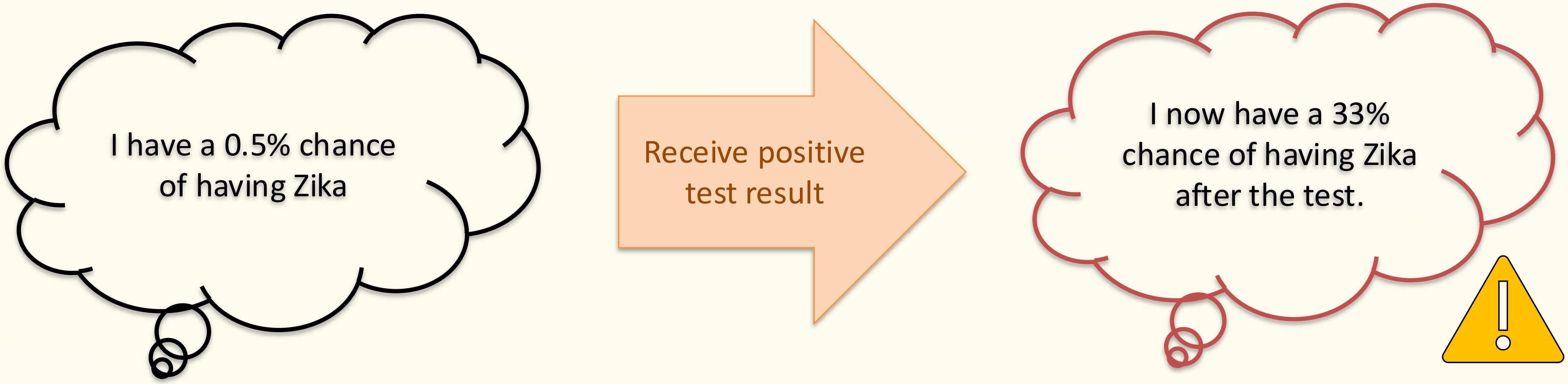
$$\frac{5}{5+10} = \frac{1}{3} \approx 0.33$$

Philosophy – Updating Beliefs - should be clearer

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: $P(Z|T)$

Zika Testing – different question

Z = you have Zika
T = you test positive for Zika

Suppose we know the following Zika stats

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

$P(T \mid Z) = 0.98$
$P(T \mid \overline{Z}) = 0.01$
$P(Z) = 0.005$

What is the probability you test negative (event $\overline{T}$) if you have Zika (event Z)?

https://pollev.com/ annakarlin185

$\Pr(\overline{T} \mid Z)$

A) 0.02
B) 0.99
C) 0.02 x 0.005
D) 0.02 x 0.995

Conditional Probability Define a Probability Space

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

Conditional Probability Define a Probability Space (2)

The probability conditioned on A follows the same properties as (unconditional) probability.

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

Formally. $(\Omega, \mathbb{P})$ is a probability space + $\mathbb{P}(\mathcal{A}) > 0$

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

Axiom 1 (Non-negativity): $P(E) \geq 0.$
Axiom 2 (Normalization): $P(\Omega) = 1$
Axiom 3 (Countable Additivity): If E and F are mutually exclusive,
then $P(E \cup F) = P(E) + P(F)$

Agenda (4)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on independence
  - Example using Law of Total Probability
  - Independence of many events
  - Independence not always obvious
  - Defining a probability space using independence
  - Independence as an assumption not always justified

Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are **independent** if

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}).$$

Alternatively,

- If $\mathbb{P}(\mathcal{A}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{B})$
- If $\mathbb{P}(\mathcal{B}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A})$

"The probability that $\mathcal{B}$ occurs after observing $\mathcal{A}$" -- Posterior

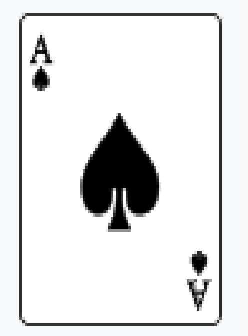
= "The probability that $\mathcal{B}$ occurs" -- Prior

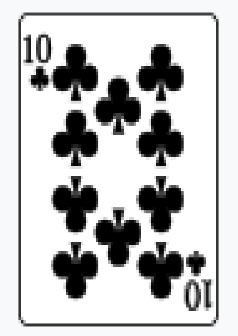
Are A and B independent?

Have a Standard 52-Card Deck. Shuffle It, and draw the top 2 cards **in order:**

Step 1: Draw a uniformly random card.

Step 2: Draw a uniformly random card from remaining cards.





A: Ace of Spades First

B: 10 of Clubs Second

Are the events A and B independent?

$$\mathbb{P}(B|A) = \frac{1}{51}$$

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

https://pollev.com/ annakarlin185

$\Pr\left(2^{nd}\ \mathbf{card} = \mathbf{10\ clubs}\ |\ 1^{st}\ \boldsymbol{card} = \boldsymbol{X}\right)$

A) 1/52
B) 1/51
C) Depends on X
D) (1/52) x (1/51)

Agenda (5)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on independence
  - Example using Law of Total Probability
  - Independence of many events
  - Independence not always obvious
  - Defining a probability space using independence
  - Independence as an assumption not always justified

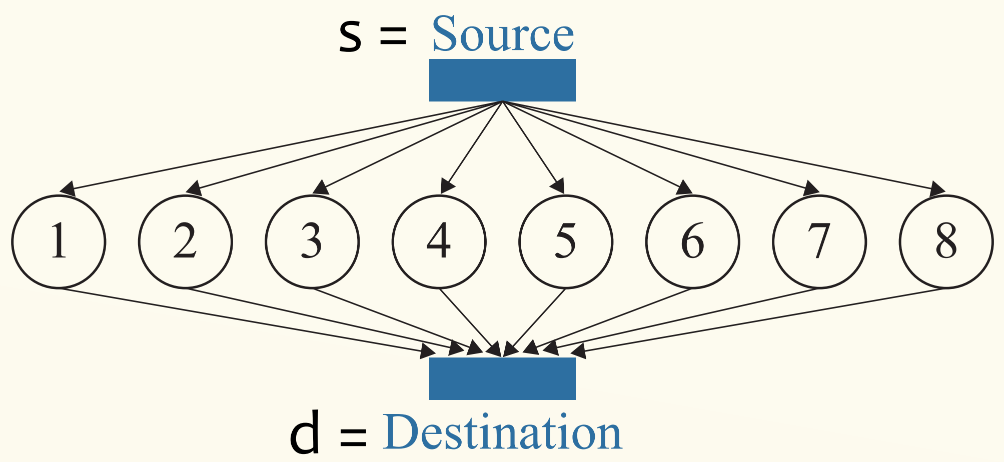
Multiple Events – Mutual Independence

Definition. Events $A_1, \dots, A_n$ are **mutually independent** if for every non-empty subset $I \subseteq \{1, \dots, n\}$, we have

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

An example

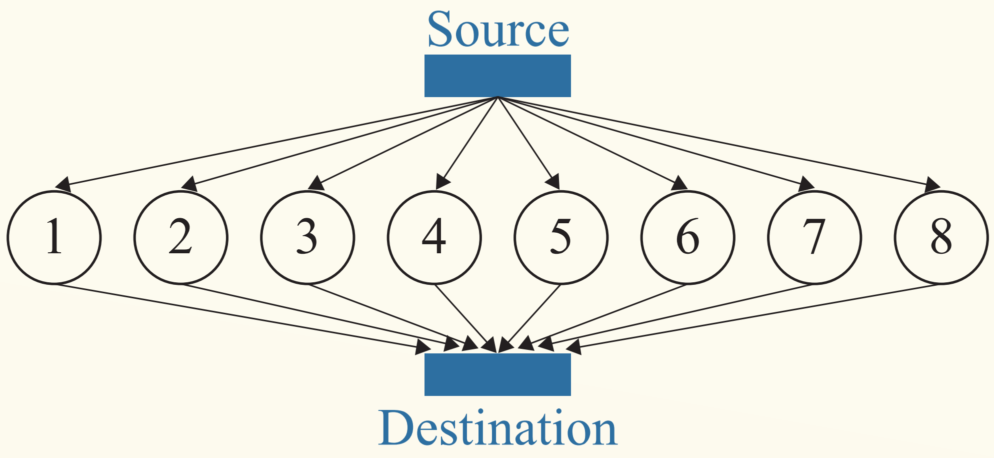
You are routing a packet from the source to the destination.
But each of the 16 edges in the network works with probability p, independently.



Q: Pr(edge (s,3) works, edge (5,d) works, edge (s,7) doesn't work)?

An example (2)

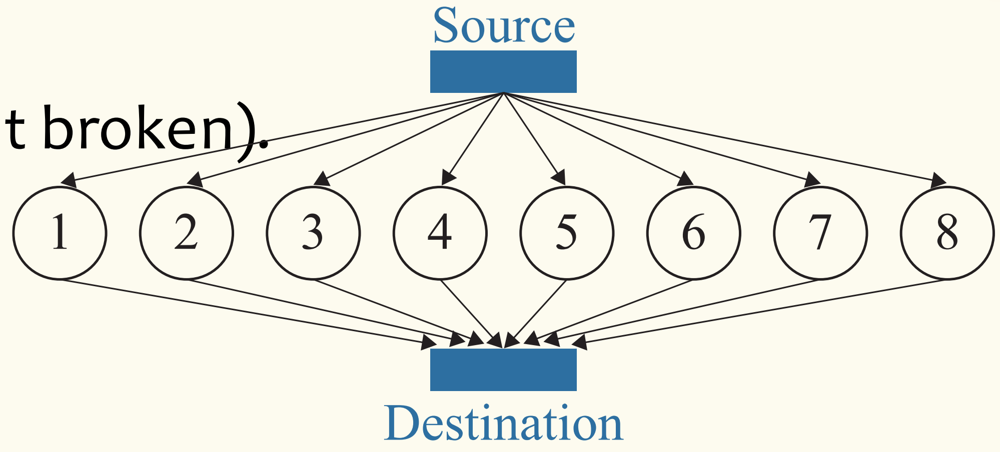
You are routing a packet from the source to the destination.
But each of the 16 edges in the network works with probability p, independently.



Q: What is the probability that you can get the packet from the source to the destination?

Let’s break it down

Each edge works with probability p, independently of other edges. There are 8 paths.
Let E_i denote the event that the i^{th} path is usable (not broken).



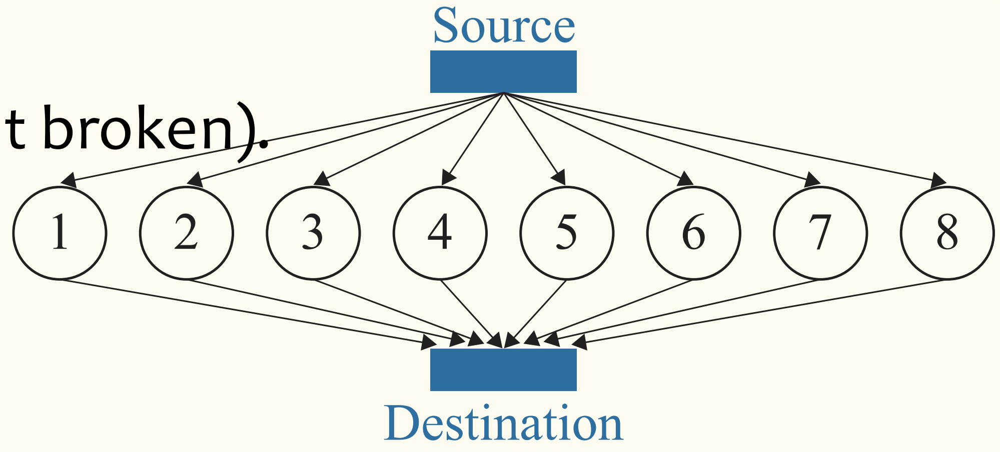
Q1: What is $\boldsymbol{P}\{E_i\}$?

Q2: What is $\boldsymbol{P}\{\overline{E_i}\}$?

More calculations

Each edge works with probability p, independently of other edges. There are 8 paths.

Let E_i denote the event that the i^{th} path is usable (not broken).



$\boldsymbol{P}\{E_i\} = p^2$ $\boldsymbol{P}\{\overline{E_i}\} = 1 - p^2$

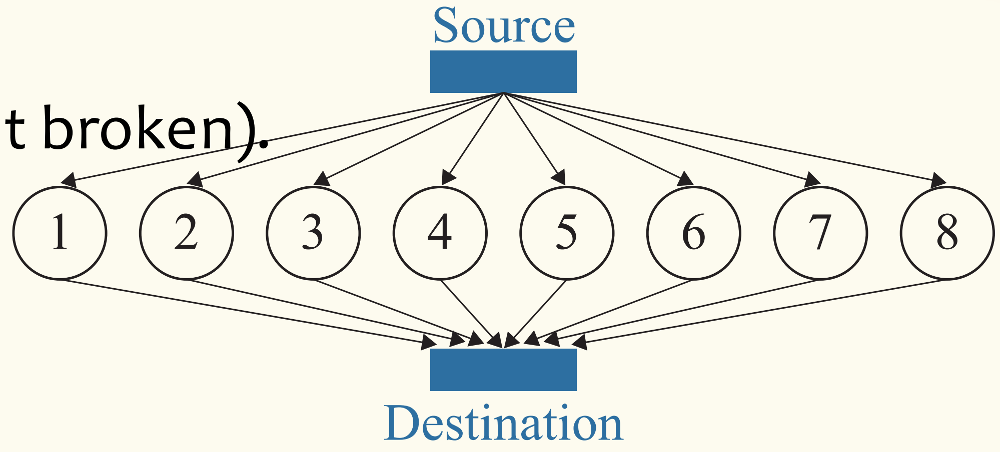
What is $\boldsymbol{P}$\{Can get from source to destination\}?

$\boldsymbol{P}$\{Can get from source to destination\} $=$

Answer

Each edge works with probability p, independently of other edges. There are 8 paths.

Let E_i denote the event that the i^{th} path is usable (not broken).



$\boldsymbol{P}\{E_i\} = p^2$ $\qquad\qquad$ $\boldsymbol{P}\{\overline{E_i}\} = 1 - p^2$

What is $\boldsymbol{P}$\{Can get from source to destination\}?

$$\boldsymbol{P}\{\text{Can get from source to destination}\} = \boldsymbol{P}\{\text{At least one path works}\}$$

$$= \boldsymbol{P}\{E_1 \cup E_2 \cup \cdots \cup E_8\}$$

$$= 1 - \boldsymbol{P}\{\text{All paths are broken}\}$$

$$= 1 - \boldsymbol{P}\{\overline{E_1}\} \cdot \boldsymbol{P}\{\overline{E_2}\} \cdots \boldsymbol{P}\{\overline{E_8}\} = 1 - (1 - p^2)^8$$

Agenda (6)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on independence
  - Example using Law of Total Probability
  - Independence of many events
  - Independence not always obvious
  - Defining a probability space using independence
  - Independence as an assumption not always justified

Independence – Another Look

Definition. Two events A and B **independent** if

$$P(A \cap B) = P(A) \cdot P(B).$$

"Equivalently." $P(A|B) = P(A)$.

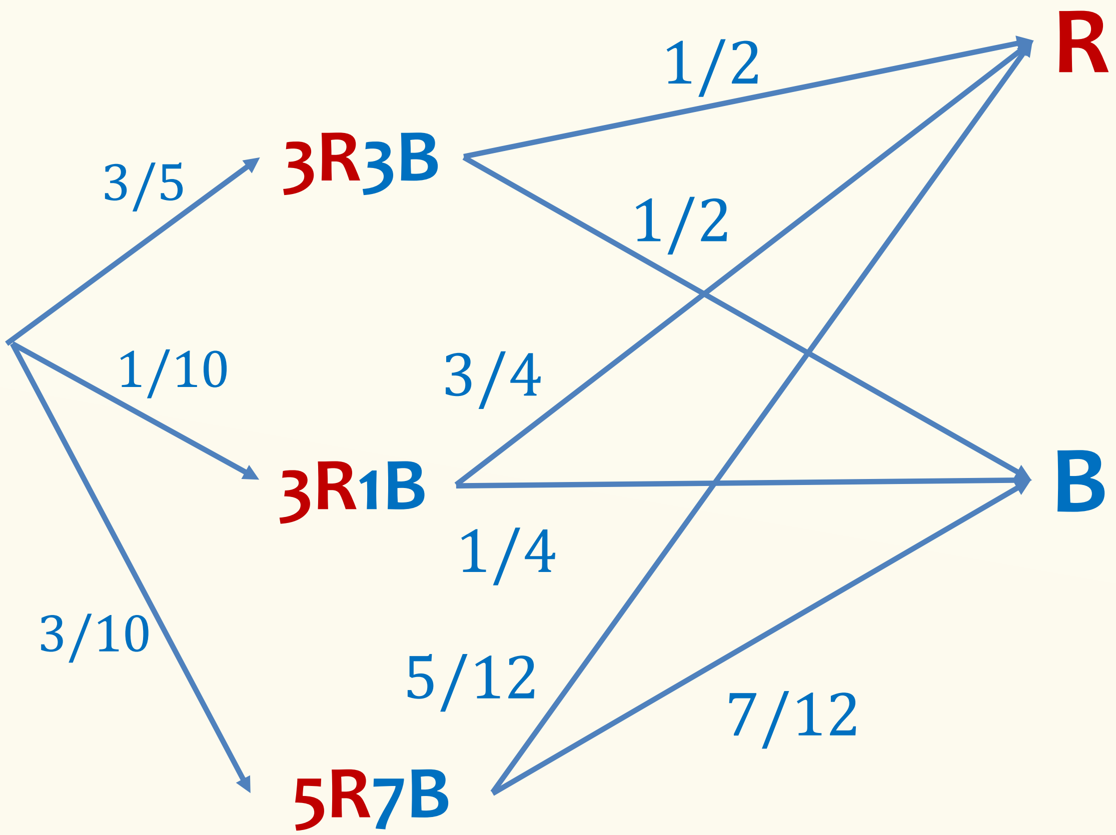
Events generated independently $\rightarrow$ *their probabilities satisfy independence*

But events can be independent without being generated by independent processes.

$\not\Leftarrow$

This can be counterintuitive!

Balls in Urns



Setting: An urn contains:

- 3 **red** and 3 **blue** balls w/ probability 3/5
- 3 **red** and 1 **blue** balls w/ probability 1/10
- 5 **red** and 7 **blue** balls w/ probability 3/10

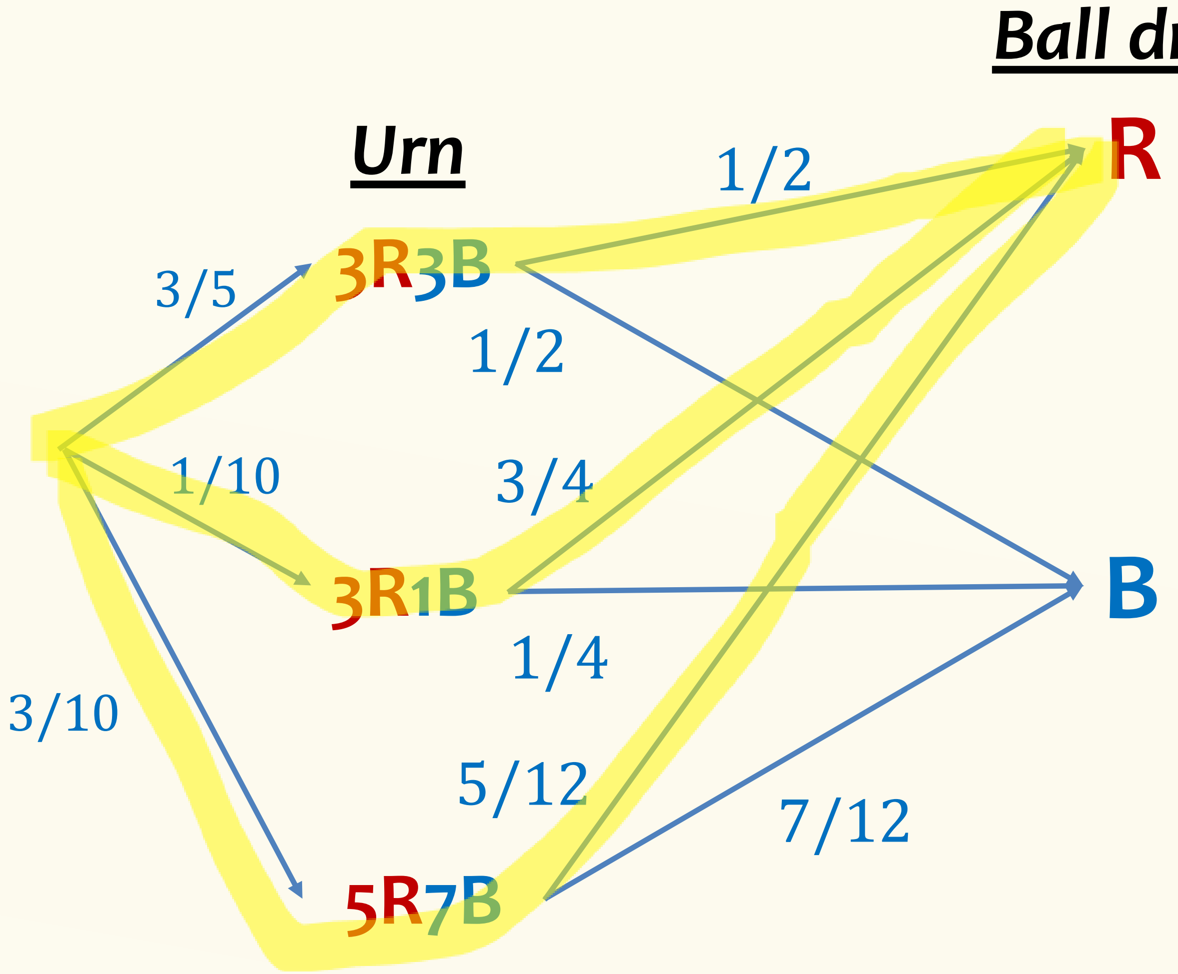
We draw a ball at random from the urn.

Are **R** and **3R3B** independent?

$P(\mathbf{R} \mid \mathbf{3R3B}) =$

$P(\mathbf{R}) =$

Balls in Urns- computation (1)



Setting: An urn contains:

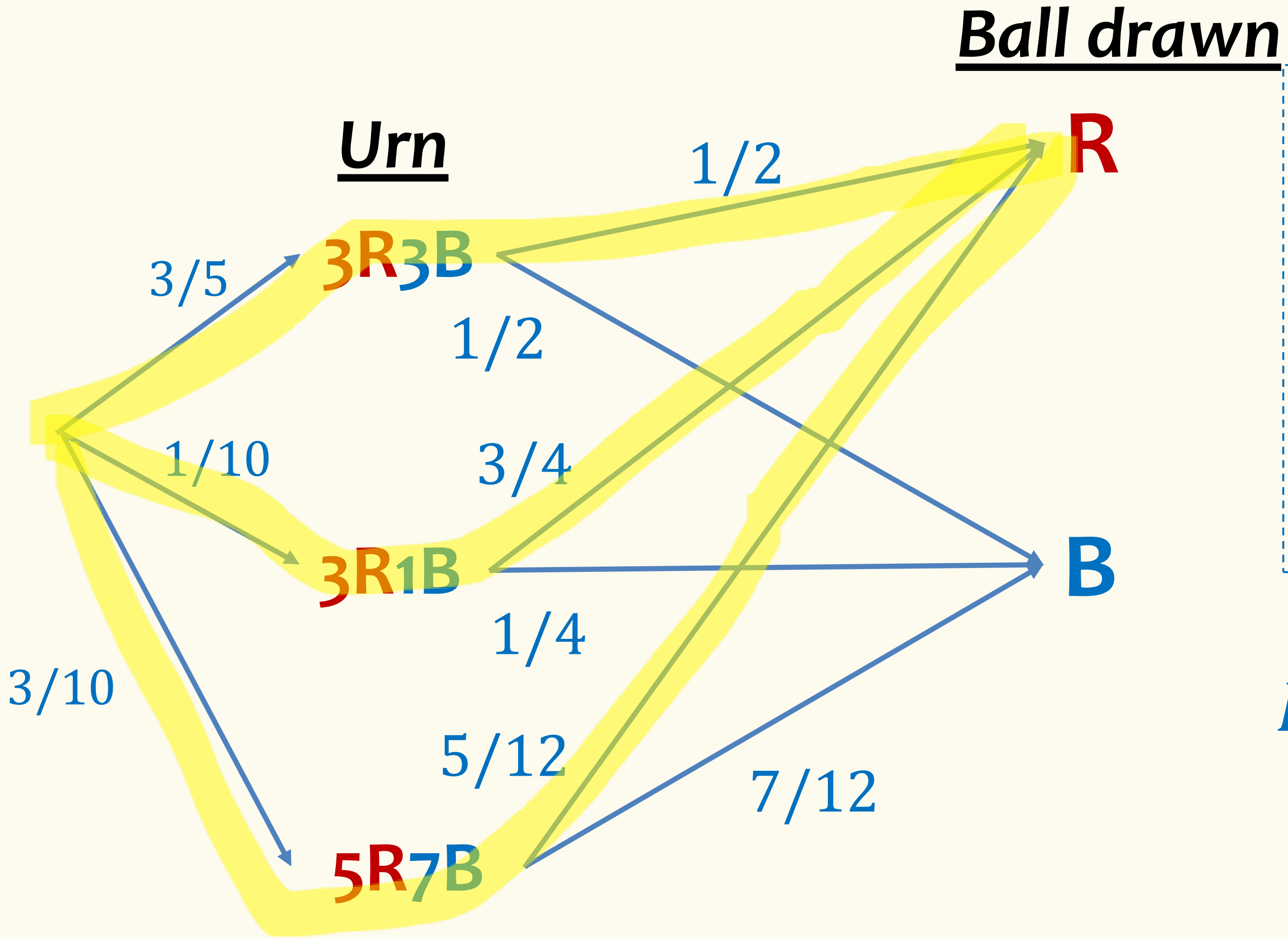
- 3 **red** and 3 **blue** balls w/ probability 3/5
- 3 **red** and 1 **blue** balls w/ probability 1/10
- 5 **red** and 7 **blue** balls w/ probability 3/10

We draw a ball at random from the urn.

$$P(\mathbf{R} \mid \mathbf{3R3B}) = \frac{1}{2}$$

Are **R** and **3R3B** independent?

Balls in Urns - computation (2)



Setting: An urn contains:

- 3 **red** and 3 **blue** balls w/ probability 3/5
- 3 **red** and 1 **blue** balls w/ probability 1/10
- 5 **red** and 7 **blue** balls w/ probability 3/10

We draw a ball at random from the urn.

$$P(\mathbf{R}) = \frac{3}{5} \times \frac{1}{2} + \frac{1}{10} \times \frac{3}{4} + \frac{3}{10} \times \frac{5}{12} = \frac{1}{2}$$

$$P(\mathbf{R} \mid \mathbf{3R3B}) = \frac{1}{2}$$

Are **R** and **3R3B** independent?

Independent! $P(\mathbf{R}) = P(\mathbf{R} \mid \mathbf{3R3B})$

Agenda (7)

- Recap
- Law of Total Probability
- Bayes Theorem
- More on independence
  - Example using Law of Total Probability
  - Independence of many events
  - Independence not always obvious
  - Defining a probability space using independence

Defining a probability space

Often probability space $(\Omega, \mathbb{P})$ is **defined** using independence

Example – Biased coin

We have a biased coin comes up Heads with probability 2/3; Each flip is independent of all other fips. Suppose it is tossed 3 times.

$$\mathbb{P}(HHH) =$$

$$\mathbb{P}(TTT) =$$

$$\mathbb{P}(HTT) =$$

Example – Biased coin – let's do a calculation

We have a biased coin comes up Heads with probability 2/3, independently of other flips. Suppose it is tossed 3 times.

$\mathbb{P}(2 \text{ heads in } 3 \text{ tosses}) =$

https://pollev.com/ annakarlin185

A) $(2/3)^2$ 1/3
B) 2/3
C) 3 $(2/3)^2$ 1/3
D) $(1/3)^2$