

# Copyright Notice

These slides are distributed under the Creative Commons License.

[DeepLearning.ai](#) makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite [DeepLearning.AI](#) as the source of the slides.

For the rest of the details of the license, see <https://creativecommons.org/licenses/by-sa/2.0/legalcode>

# W1 Lesson 1

# Introduction to Probability



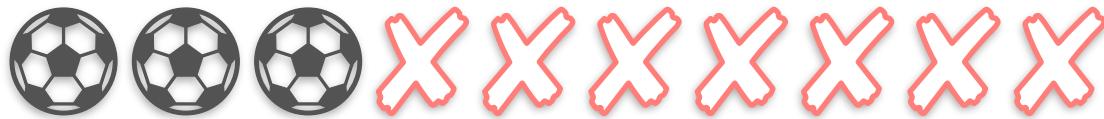
DeepLearning.AI

# Introduction to probability

---

## What is Probability?

# Introduction to Probability



Find the probability that a child picked at random plays soccer.

# Introduction to Probability



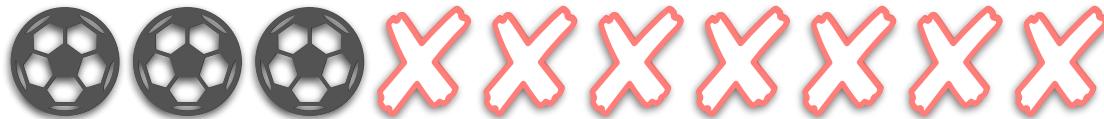
Find the probability that a child picked at random plays soccer.

**The probability that a child picked at random plays soccer.**

$$P(\text{soccer})$$

A teal curved arrow points from the text above down to the mathematical expression below.

# Introduction to Probability

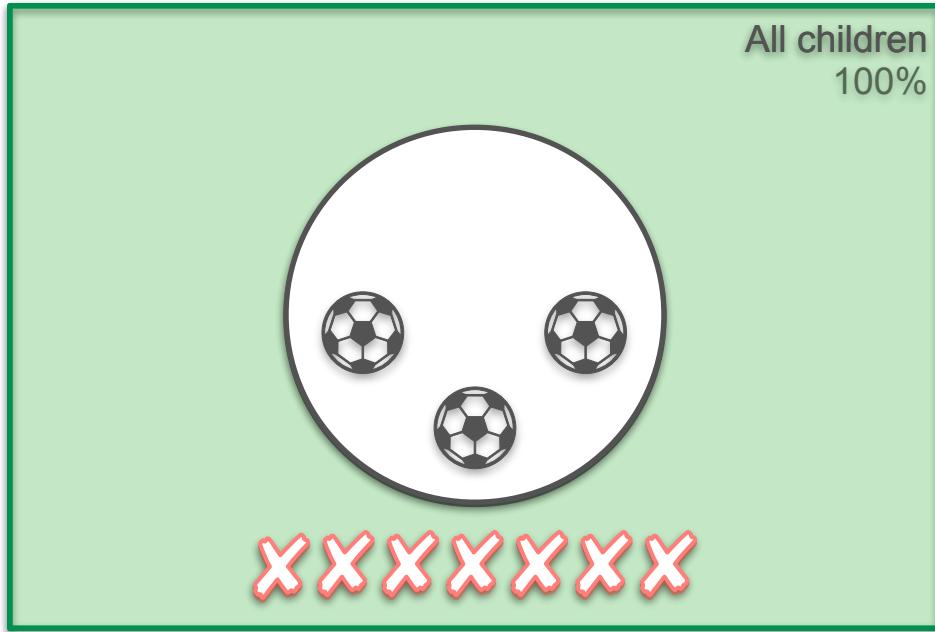


Find the probability that a child picked at random plays soccer.

$$P(\text{soccer}) = \frac{\text{soccer}}{\text{total}} = \frac{\text{Event}}{\text{Sample space}} = \frac{3}{10} = 0.3$$

The diagram illustrates the calculation of probability. A horizontal bar represents the 'Sample space' containing three soccer balls and seven 'X' marks. Above this bar, a smaller horizontal bar represents the 'Event' containing the three soccer balls. Arrows point from the labels 'Event' and 'Sample space' to their respective bars.

# Introduction to Probability: Venn Diagram



# Introduction to Probability: Coin Example 1

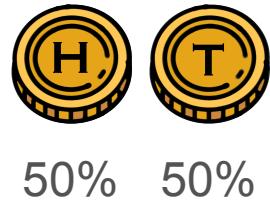


## Experiment

Probability of landing on heads

$$P(\text{heads})$$

# Introduction to Probability: Coin Example 1



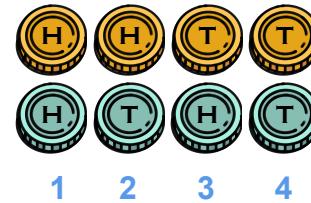
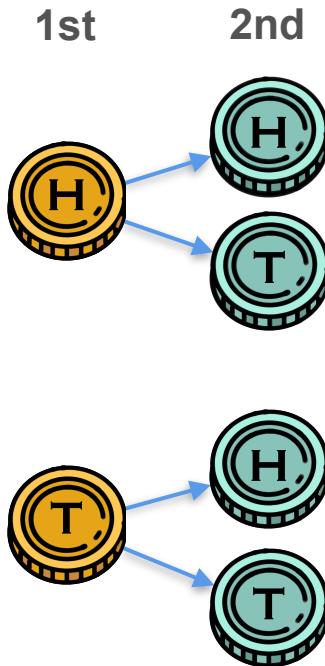
$$P(\text{heads}) = \frac{\text{Number of heads}}{\text{Total number of outcomes}} = \frac{1}{2} = 0.5$$
The equation shows the probability of getting heads as the ratio of the number of heads (1) to the total number of outcomes (2). Above the fraction, there is a single gold coin with 'H' (heads) facing up. Below the fraction, there are two gold coins: one with 'H' (heads) facing up and one with 'T' (tails) facing up.

# Introduction to Probability: Coin Example 2



50% 50%

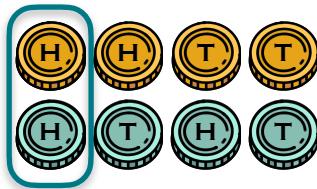
What is the probability of landing on heads twice?



# Introduction to Probability: Coin Example 2



50%    50%

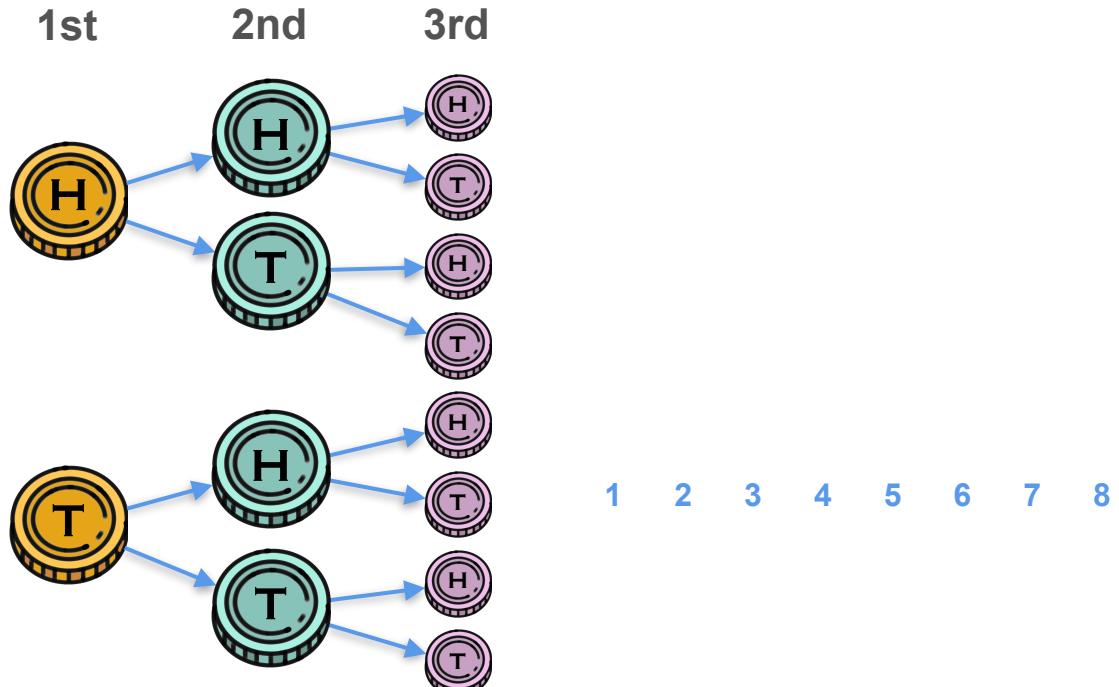


$$P(HH) = \frac{1}{4} = 0.25$$

# Introduction to Probability: Coin Example 3



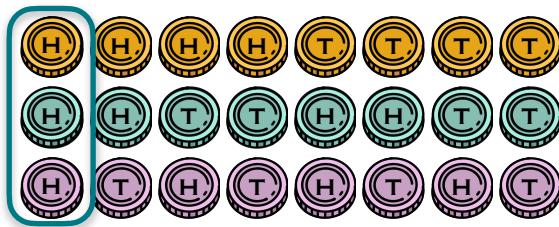
What is the probability of landing on heads 3 times?



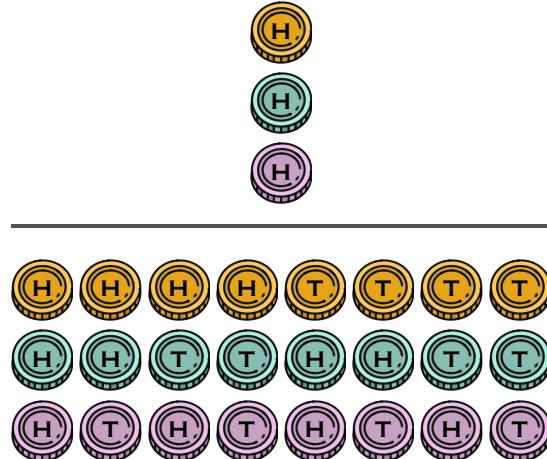
# Introduction to Probability: Coin Example 3



50% 50%



$$P(HHH) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$



$$= \frac{1}{8} = 0.125$$



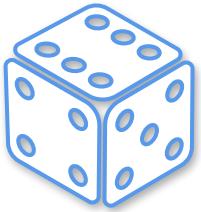
DeepLearning.AI

# Introduction to Probability

---

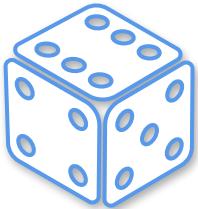
**What is Probability? - Dice Example**

# Introduction to Probability: Dice Example 1

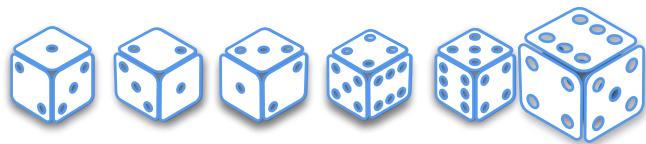


What is the probability of obtaining 6?

# Introduction to Probability: Dice Example 1

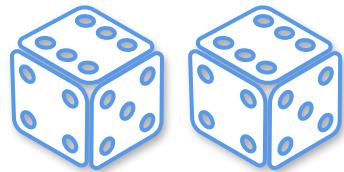


What is the probability of obtaining 6?



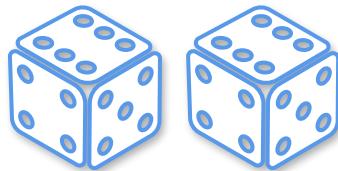
$$P(6) = \text{_____} = \frac{1}{6}$$

# Introduction to Probability: Dice Example 2



What is the probability of obtaining 6,6?

# Introduction to Probability: Dice Example 2



What is the probability of obtaining 6,6?

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$$P(6,6) = \frac{1}{36} = \frac{1}{36}$$

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6



DeepLearning.AI

## Introduction to probability

---

### Complement of Probability

# Complement of Probability



30%

What is the probability of a child NOT playing soccer?

# Complement of Probability

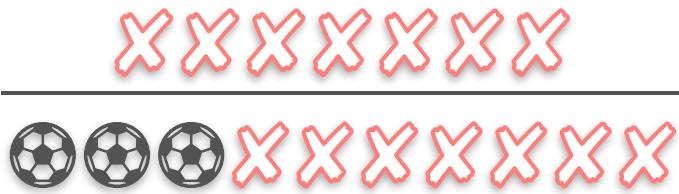


30%

What is the probability of a child NOT playing soccer?

$$P(\text{not soccer}) = \frac{\text{not soccer}}{\text{total}} = \frac{\text{XXXXXXX}}{\text{Soccer Balls XXXXXXXX}} = \frac{7}{10} = 0.7$$

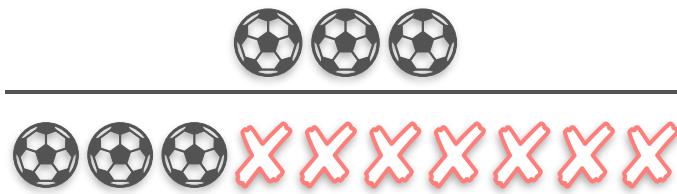
# Complement of Probability



$P(\text{not soccer})$

0.7

+

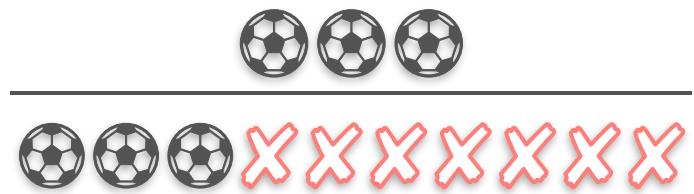
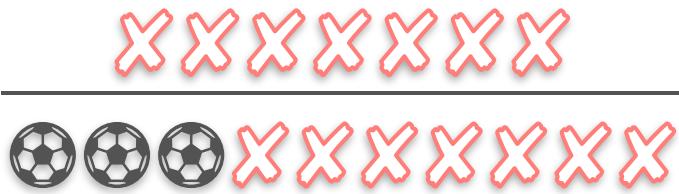


$P(\text{soccer})$

0.3

= 1

# Complement of Probability



P(not soccer)

P(soccer)

0.7

= 1 —

0.3

Complement Rule

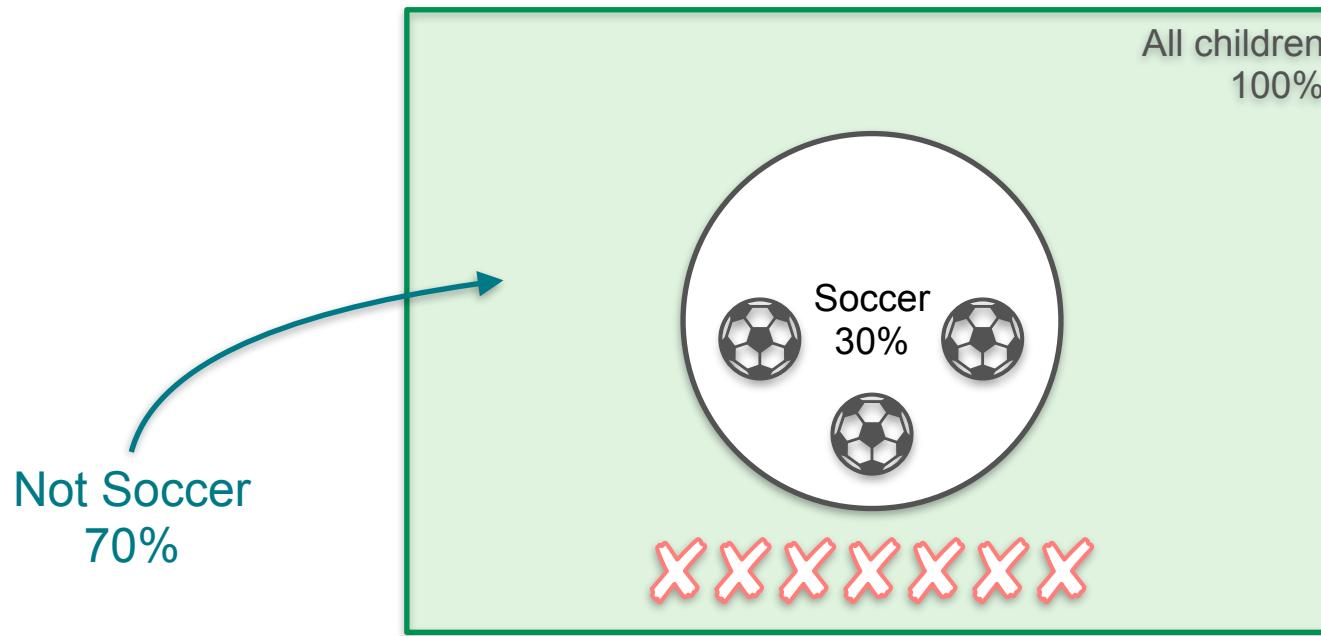
# Complement of Probability

$$P(\text{not soccer}) = 1 - P(\text{soccer})$$

$$P(A') = 1 - P(A)$$

Complement Rule

# Complement of Probability: Venn Diagram



# Complement of Probability: Coin Example 1



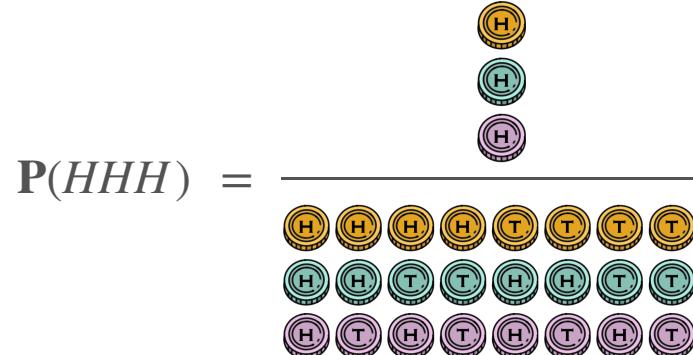
What is the probability of not landing on heads 3 times?

# Complement of Probability: Coin Example 1



What is the probability of not landing on heads 3 times?

$$P(\text{not } HHH) = 1 - P(HHH)$$



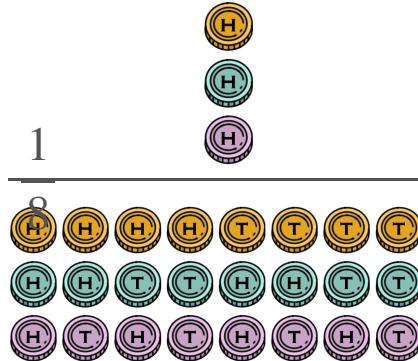
# Complement of Probability: Coin Example 1



What is the probability of not landing on heads 3 times?

$$P(\text{not } HHH) = 1 - P(HHH)$$

$$P(\text{not } HHH) = 1 - \frac{1}{8}$$



# Complement of Probability: Coin Example 1

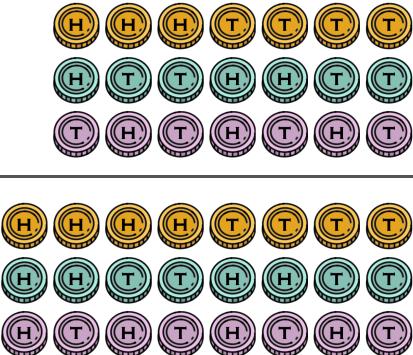


What is the probability of not landing on heads 3 times?

$$P(\text{not } HHH) = 1 - P(HHH)$$

$$P(\text{not } HHH) = 1 - \frac{1}{8}$$

$$= \frac{7}{8}$$



# Complement of Probability: Dice Example 1



What is the probability of obtaining anything other than 6?

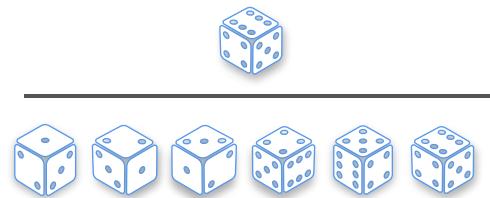
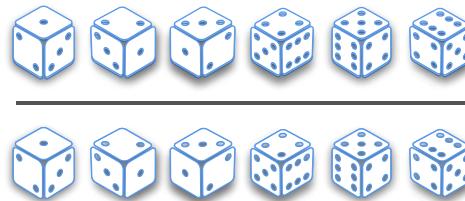
# Complement of Probability: Dice Example 1



What is the probability of obtaining anything other than 6?



$$P(\text{not } 6) =$$



$$= \underline{\hspace{2cm}}$$

$$= \frac{5}{6}$$



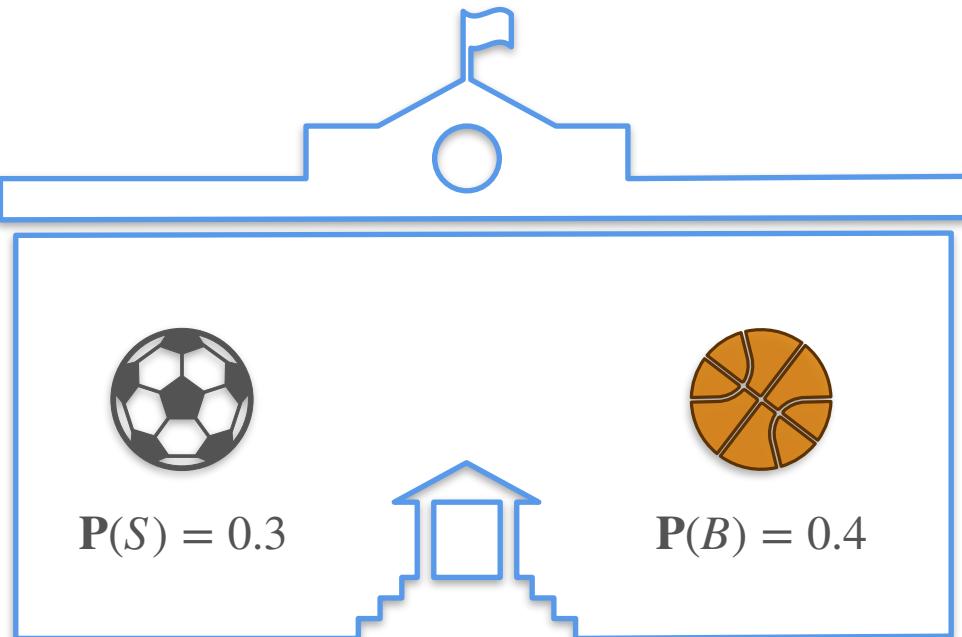
DeepLearning.AI

# Introduction to probability

---

**Sum of Probabilities  
(Disjoint Events)**

# Sum of Probabilities: Quiz 1

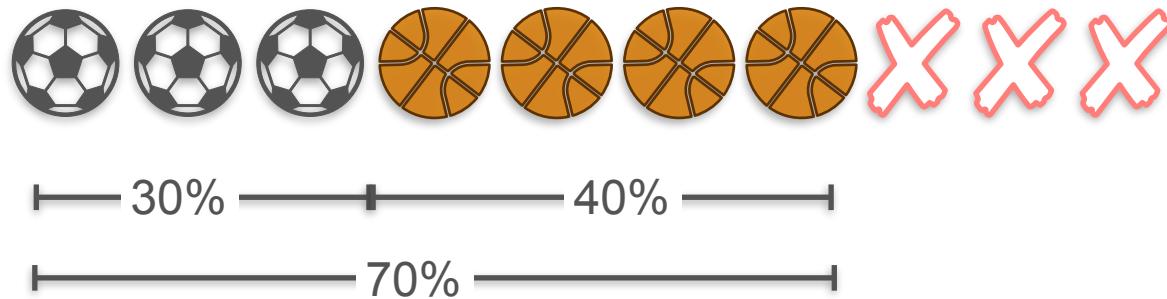


At a school, kids can only play one sport.

What is the probability that a kid plays soccer or basketball?

Hint: What if there were only 10 kids?

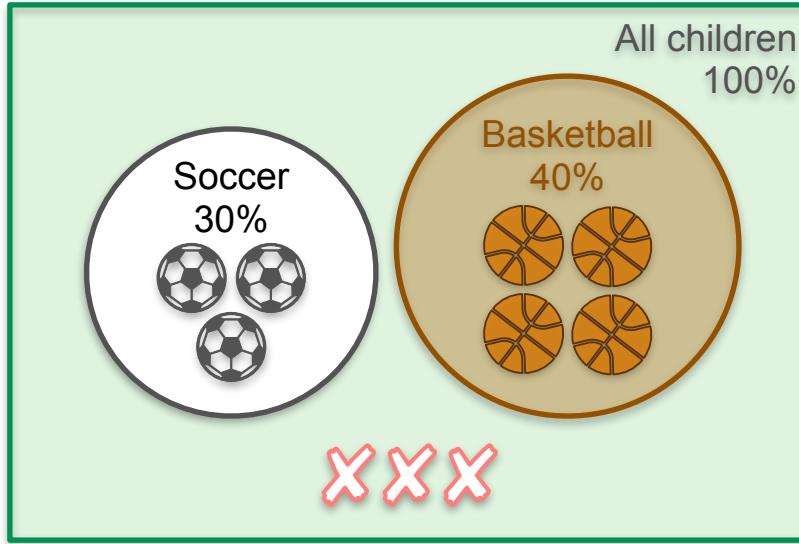
# Sum of Probabilities: Quiz 1 Solution



$$P(\text{soccer or basketball}) = \frac{\text{soccer or basketball}}{\text{total}} = \frac{3 + 4}{10} = 0.7$$

$$P(\text{soccer or basketball}) = P(\text{soccer}) + P(\text{basketball})$$

# Sum of Probabilities: Quiz 1 Solution



$$P(\text{soccer} \cup \text{basketball}) = P(\text{soccer}) + P(\text{basketball})$$

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

# Sum of Probabilities: Dice Example 1



What is the probability of obtaining  
an even number or a 5?

# Sum of Probabilities: Dice Example 1



What is the probability of obtaining an even number or a 5?

A



\_\_\_\_\_ + \_\_\_\_\_

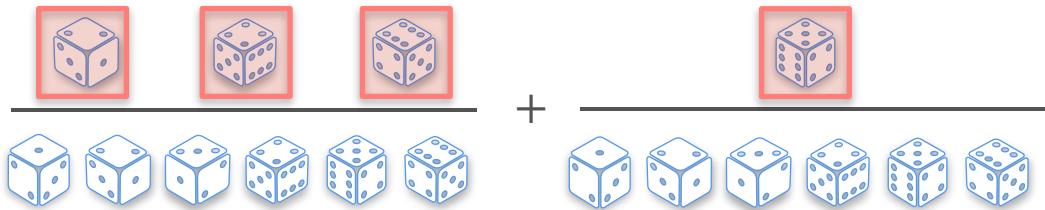
B



# Sum of Probabilities: Dice Example 1

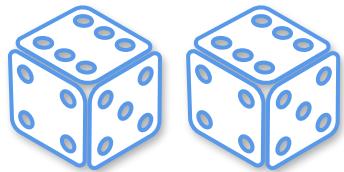


What is the probability of obtaining an even number or a 5?



$$P(\text{even number or } 5) = P(\text{even number}) + P(5) = \frac{3}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

# Sum of Probabilities: Dice Example 2



What is the probability of obtaining a sum of 7 or a sum of 10?

# Sum of Probabilities: Dice Example 2

A

sum of 7

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

B

sum of 10

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

# Sum of Probabilities: Dice Example 2

*A* or *B*

sum of 7 or sum of 10

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

# Sum of Probabilities: Dice Example 2

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

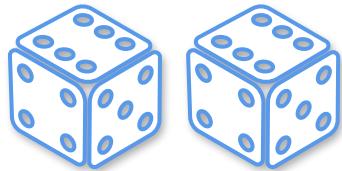
P(sum of 7 or sum of 10)

$$= \boxed{P(\text{sum of 7})} + \boxed{P(\text{sum of 10})}$$

$$= \frac{6}{36} + \frac{3}{36}$$

$$= \frac{9}{36} = \frac{1}{4}$$

# Sum of Probabilities: Dice Example 3



What is the probability of obtaining  
a difference of 2 or a difference of 1?

# Sum of Probabilities: Dice Example 3

A

diff = 2

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,1	6,2	6,3	6,4	6,5	6,6

B

diff = 1

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,1	6,2	6,3	6,4	6,5	6,6

# Sum of Probabilities: Dice Example 3

*A* or *B*

diff = 2 or diff = 1

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

Dice icons are placed along the top row and left column of the grid.

# Sum of Probabilities: Dice Example 3



$P(\text{diff} = 2 \text{ or } \text{diff} = 1)$

$$= P(\text{diff} = 2) + P(\text{diff} = 1)$$

$$= \frac{8}{36} + \frac{10}{36}$$

$$= \frac{18}{36}$$



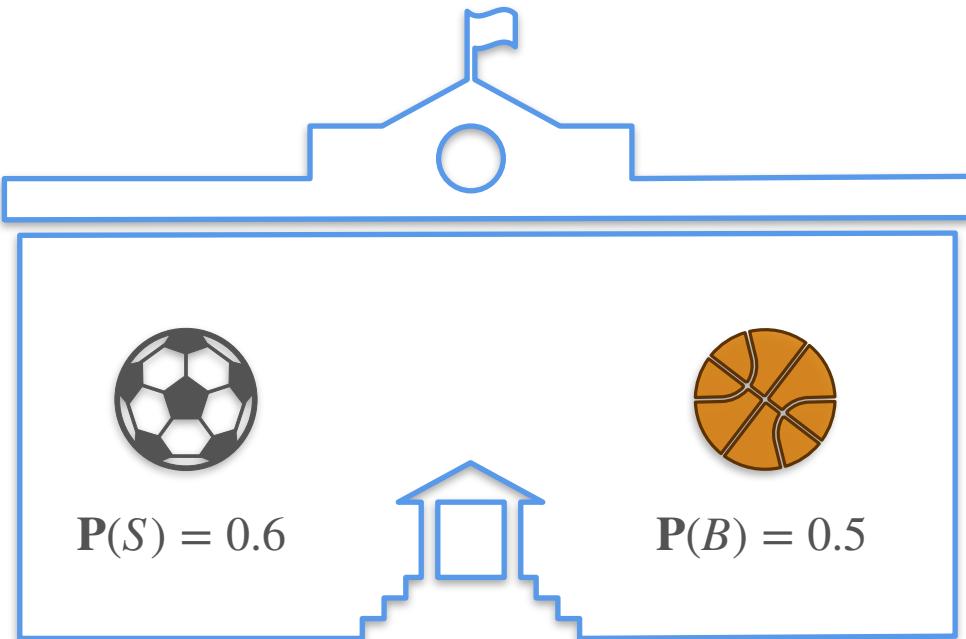
DeepLearning.AI

# Introduction to probability

---

**Sum of Probabilities  
(Joint Events)**

# Sum of Probabilities (Joint Events): Quiz 1

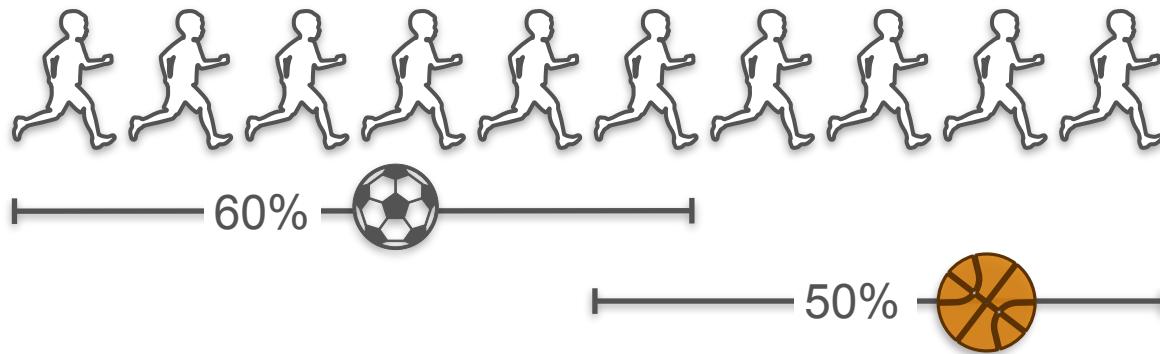


At a school, kids can play as many sports as they want.

What is the probability that a kid plays soccer or basketball?

Hint: What if there were only 10 kids?

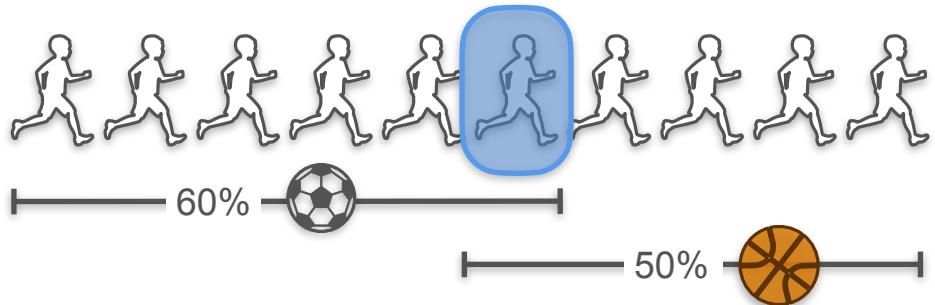
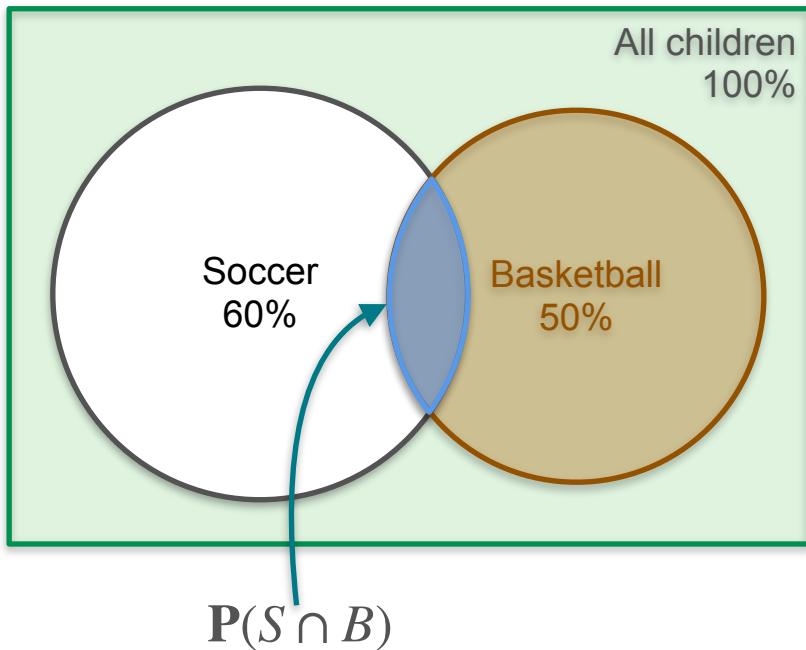
# Sum of Probabilities (Joint Events): Quiz 1 Solution



$$P(\text{soccer} \cup \text{basketball}) = ?$$

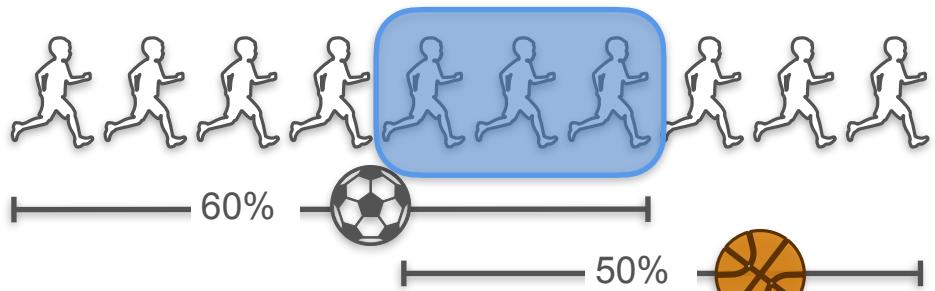
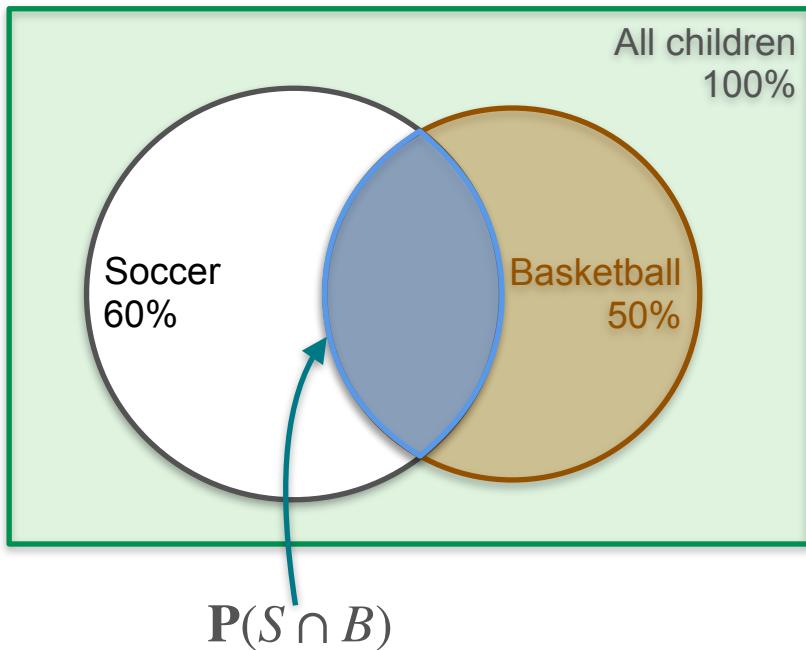
We don't know how many children play multiple sports

# Sum of Probabilities (Joint Events): Quiz 1 Solution



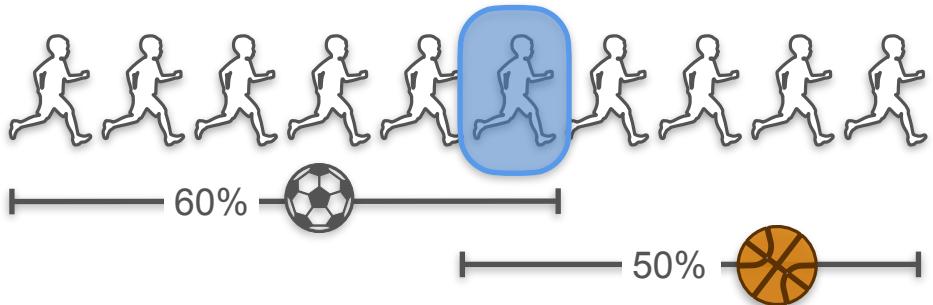
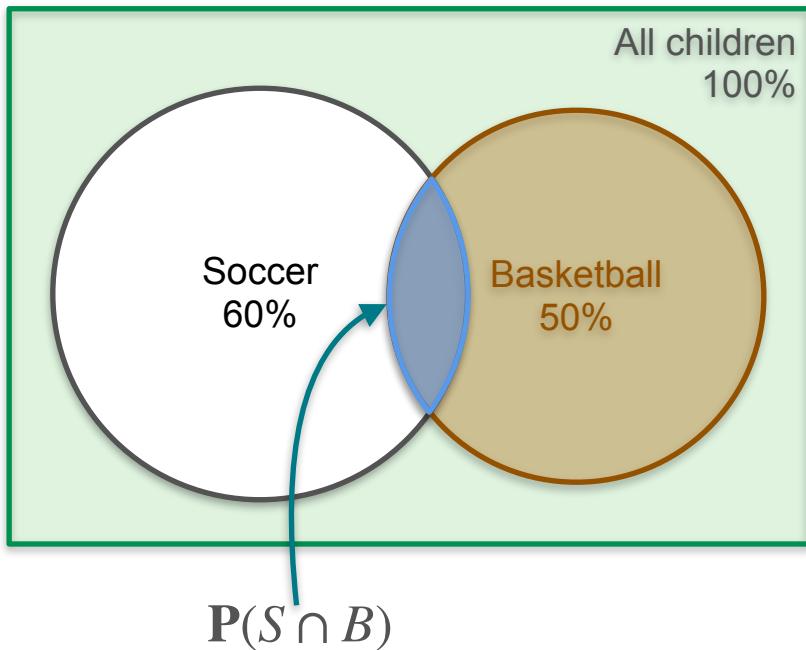
$$P(S \cup B) = P(S) + P(B)$$

# Sum of Probabilities (Joint Events): Quiz 1 Solution



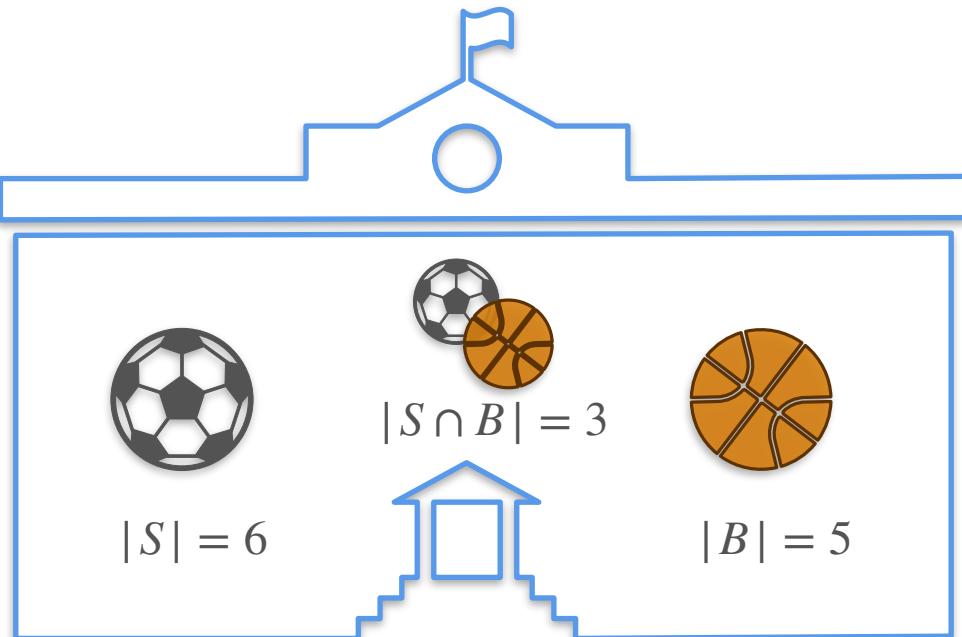
$$P(S \cup B) = P(S) + P(B)$$

# Sum of Probabilities (Joint Events): Quiz 1 Solution



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

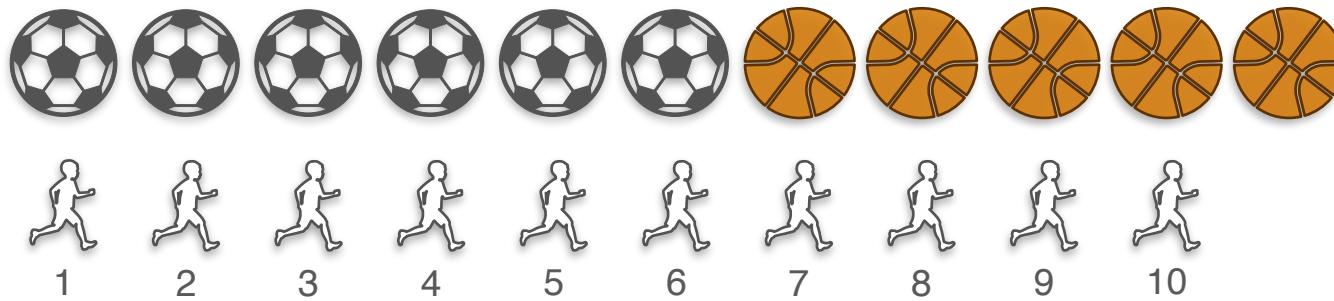
# Sum of Probabilities (Joint Events): Quiz 2



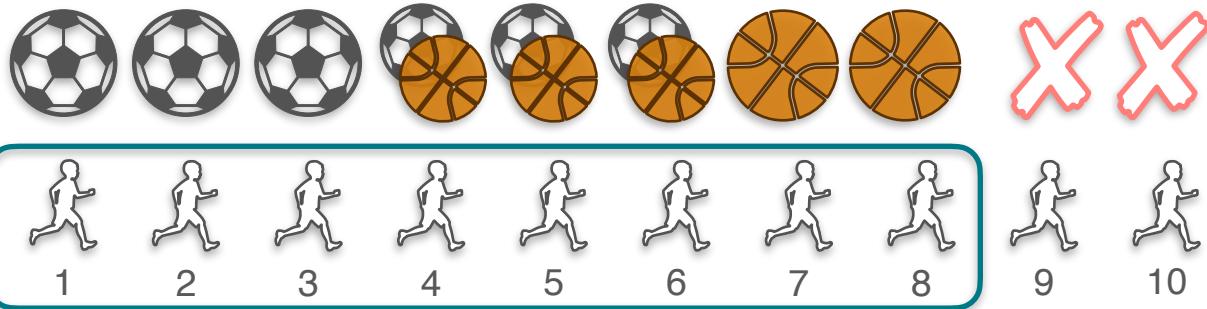
How many kids play soccer or basketball?

Hint: What if there were only 10 kids?

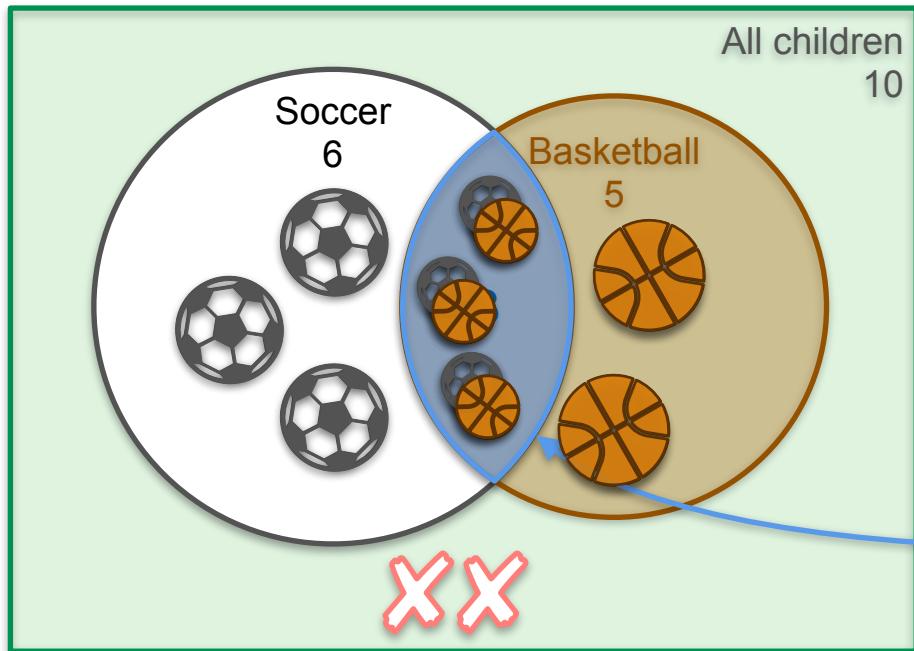
# Sum of Probabilities (Joint Events): Quiz 2 Solution



# Sum of Probabilities (Joint Events): Quiz 2 Solution

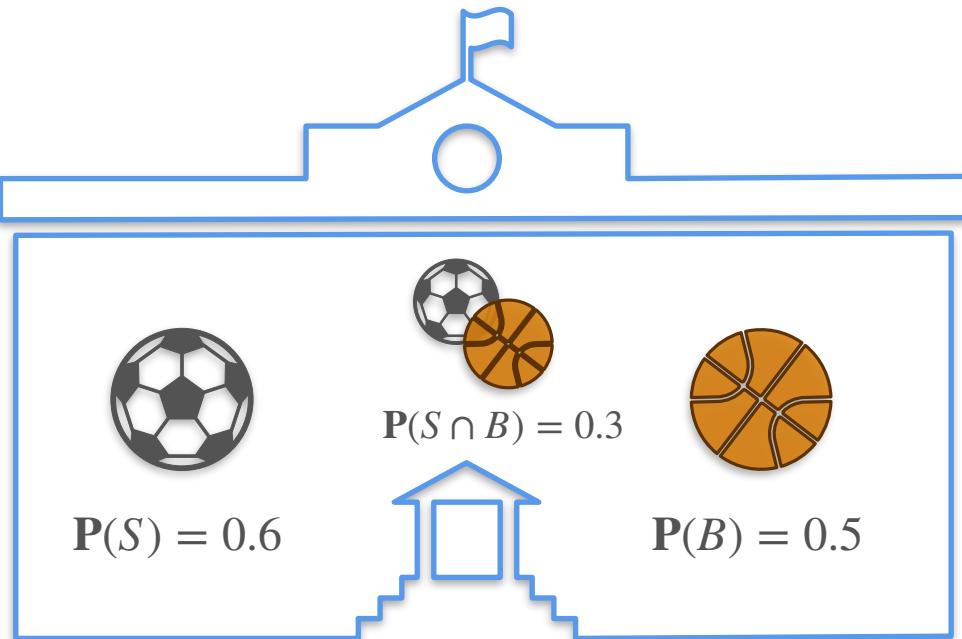


# Sum of Probabilities (Joint Events): Venn Diagram



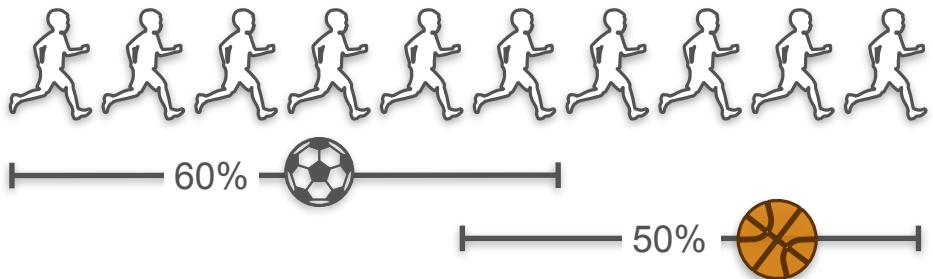
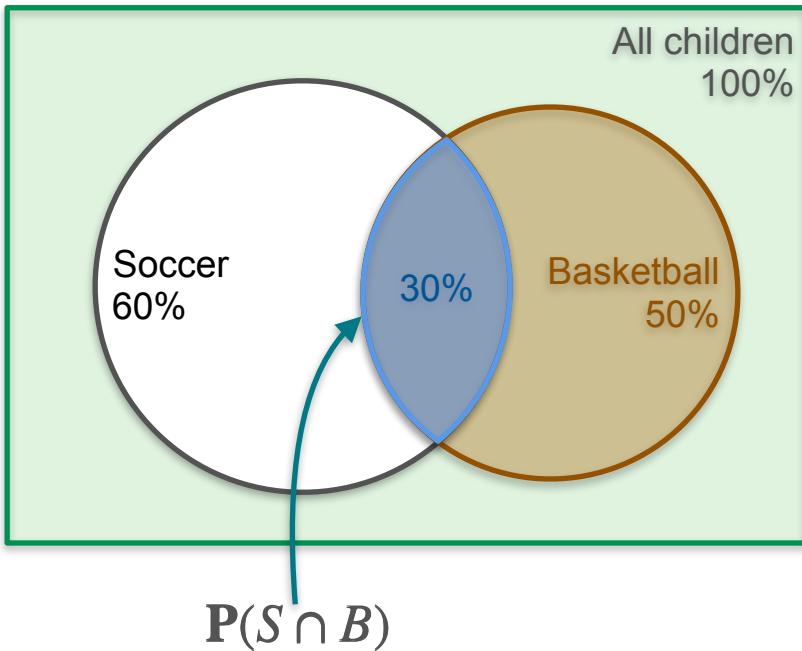
$$\begin{aligned}|S \cup B| &= |S| + |B| - |S \cap B| \\&= 6 + 5 - 3 \\&= 8\end{aligned}$$

# Sum of Probabilities (Joint Events): Quiz 3



What is the probability that a child plays soccer or basketball?

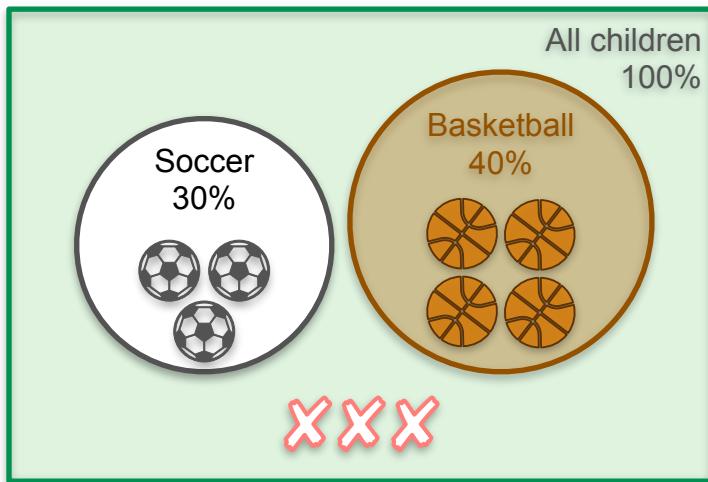
# Sum of Probabilities (Joint Events): Quiz 3 Solution



$$\begin{aligned}\mathbf{P}(S \cup B) &= \mathbf{P}(S) + \mathbf{P}(B) - \mathbf{P}(S \cap B) \\ &= 0.6 + 0.5 - 0.3 \\ &= 0.8\end{aligned}$$

# Disjoint Events vs Joint Events

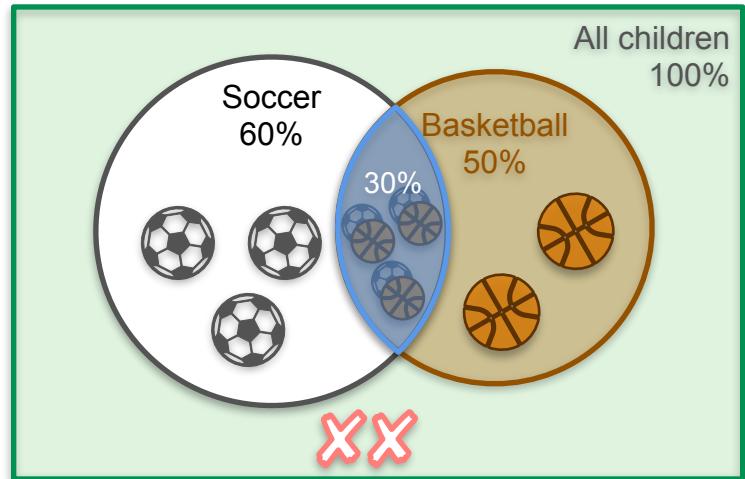
Disjoint



Mutually exclusive

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

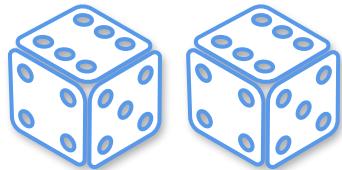
Joint



Non-mutually exclusive

$$P(S \cup B) = P(S) + P(B) - P(S \cap B)$$

# Sum of Probabilities (Joint Events): Dice Example 1



What is the probability of obtaining a sum of 7 or a difference of 1?

# Sum of Probabilities (Joint Events): Dice Example 1

A

sum = 7

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

B

diff = 1

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

# Sum of Probabilities (Joint Events): Dice Example 1

A or B

sum = 7 or diff = 1



sum = 7 and diff = 1

# Sum of Probabilities (Joint Events): Dice Example 1



1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$P(\text{sum} = 7 \text{ or } \text{diff} = 1)$

$$= P(\text{sum} = 7) + P(\text{diff} = 1) - P(\text{sum} = 7 \cap \text{diff} = 1)$$

$$= \frac{6}{36} + \frac{10}{36} - \frac{2}{36}$$

$$= \frac{14}{36}$$



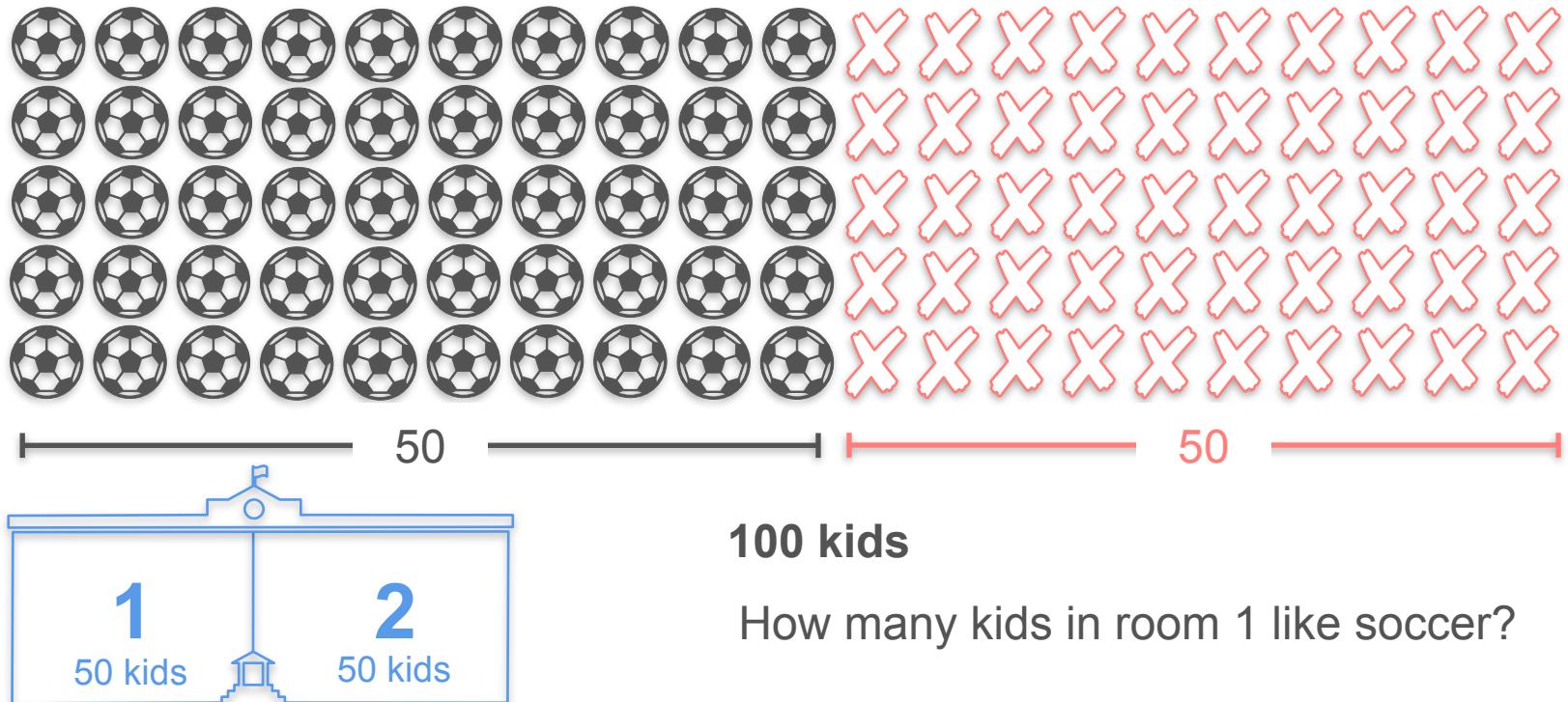
DeepLearning.AI

# Introduction to probability

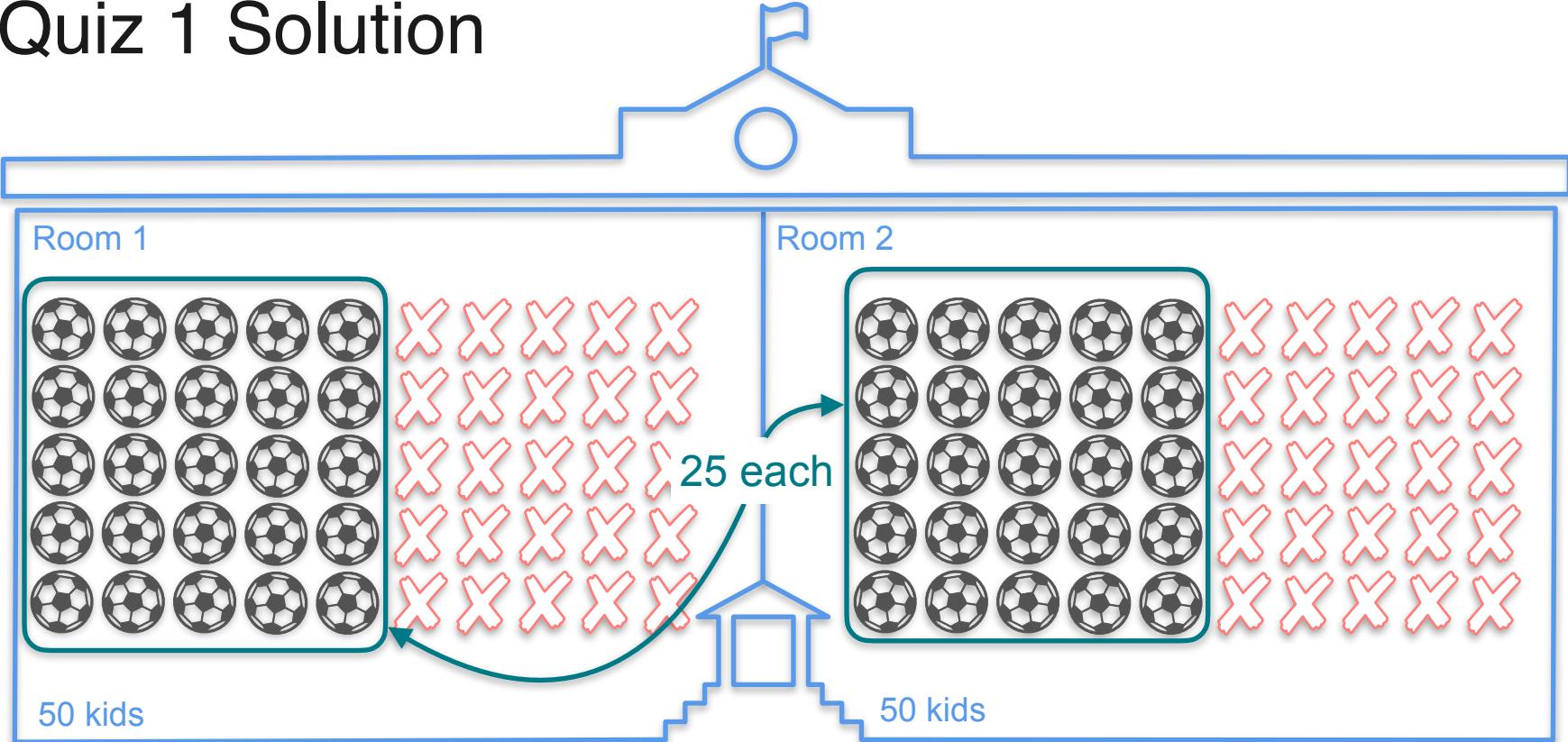
---

## Independence

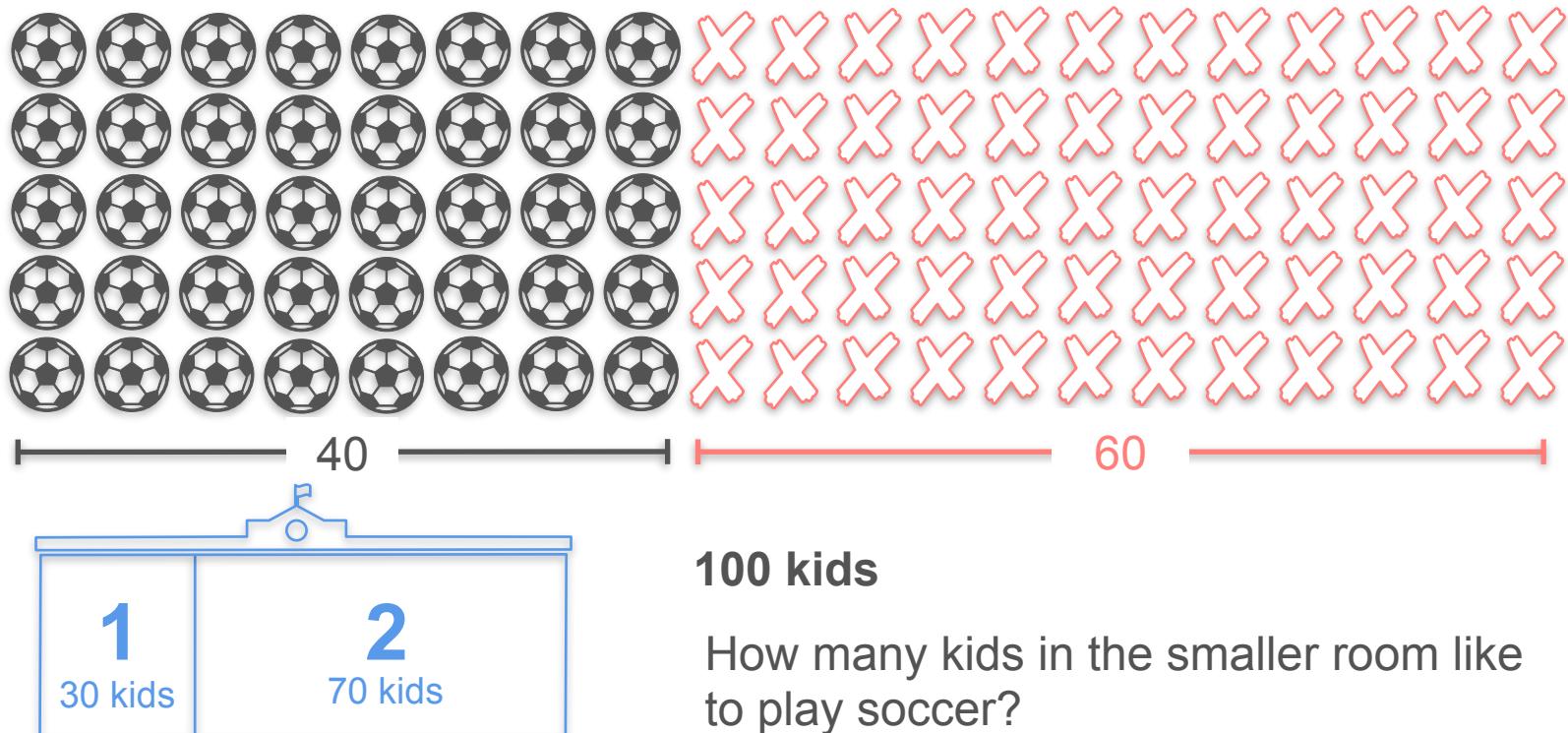
# Independence: Quiz 1



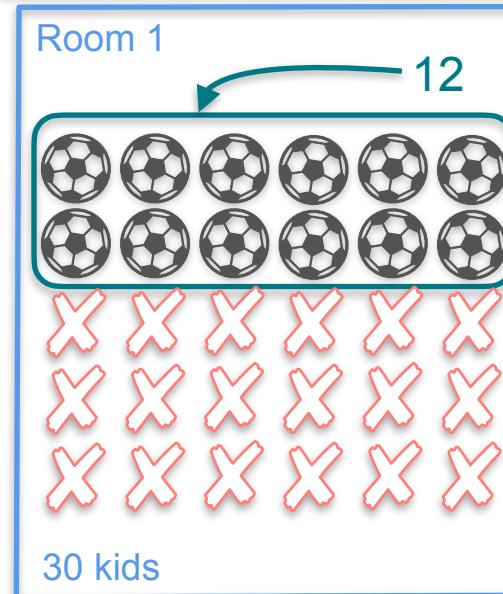
# Quiz 1 Solution



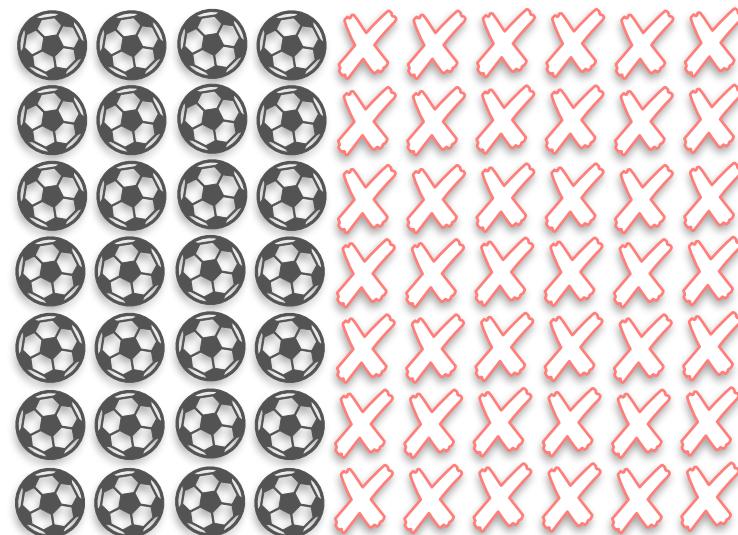
# Independence: Quiz 2



# Quiz 2 Solution



Room 2



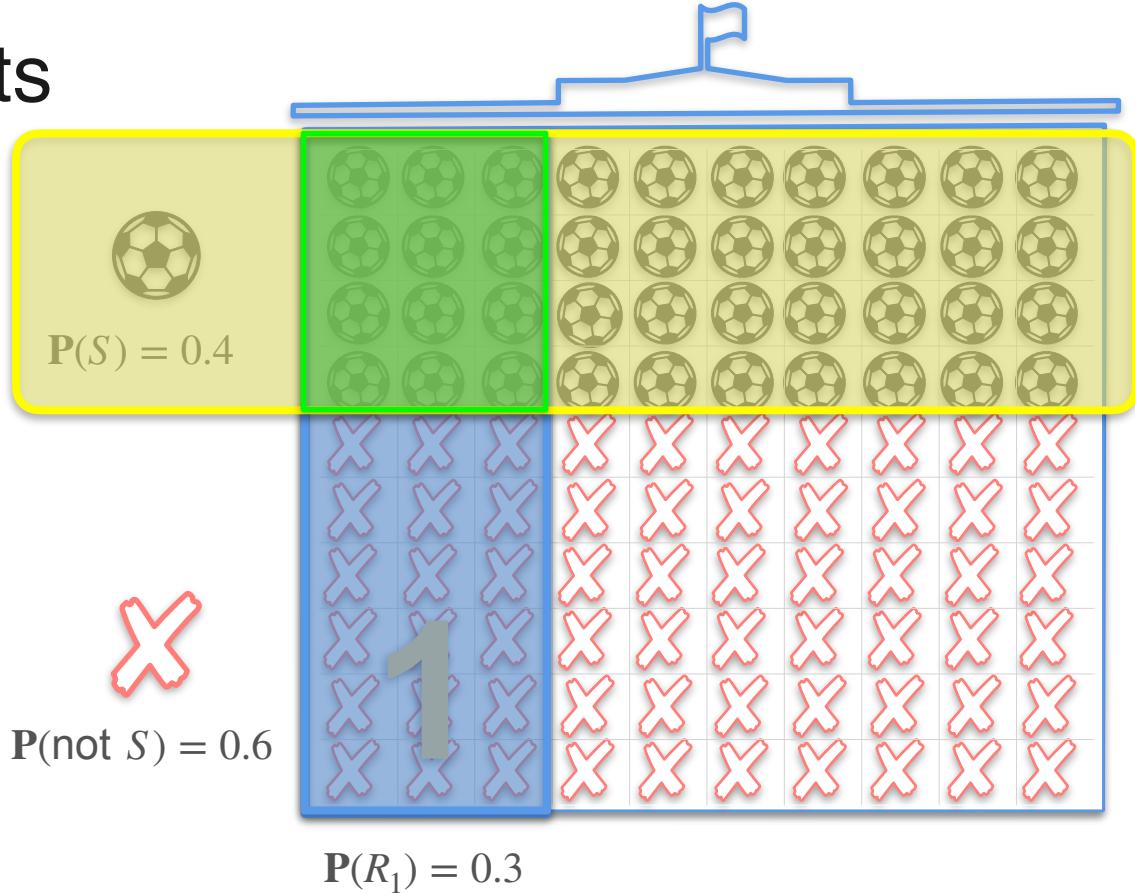
# Independent Events

$P(\text{Soccer and Room 1})$

$$P(S \cap R_1) = P(S) \bullet P(R_1)$$

$$= 0.4 \bullet 0.3$$

$$= 0.12$$



# Product Rule (for Independent Events)

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

# Independent Events: Coin Example 1

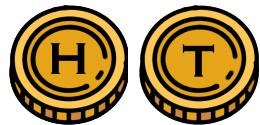


50% 50%

What is the probability of landing on heads five times?



# Independent Events: Coin Example 1



50% 50%

What is the probability of landing on heads five times?

$$P(5 \text{ heads}) =$$

$$\frac{\text{Diagram}}{\text{Probability}} \cdot \frac{\text{Diagram}}{\text{Probability}} \cdot \frac{\text{Diagram}}{\text{Probability}} \cdot \frac{\text{Diagram}}{\text{Probability}} \cdot \frac{\text{Diagram}}{\text{Probability}}$$

The diagram illustrates the calculation of the probability of getting 5 heads in a row. It shows five separate coin tosses, each with two possible outcomes: heads (H) or tails (T). The top row shows the outcome of each toss. The bottom row shows the probability of each outcome, which is  $\frac{1}{2}$ . The five probabilities are multiplied together to find the total probability of all five events happening simultaneously.

$\frac{1}{2} \quad \bullet \quad \frac{1}{2} \quad \bullet \quad \frac{1}{2} \quad \bullet \quad \frac{1}{2} \quad \bullet \quad \frac{1}{2}$

# Independent Events: Coin Example 2



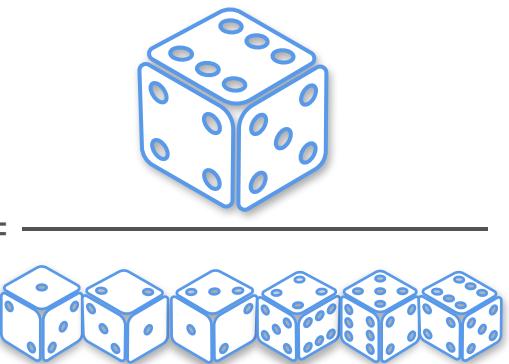
What is the probability of landing 5n heads five times?

$$P(5 \text{ heads}) = \left( \frac{\text{Diagram of 5 heads}}{\text{Diagram of 5 tails}} \right)$$

The numerator shows three coins stacked vertically, all showing heads (H). The denominator shows three coins stacked vertically, with the top coin showing heads (H) and the bottom two showing tails (T).

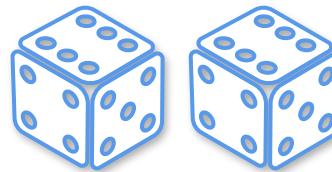
$$\left( \frac{1}{2} \right)^5 = \frac{1}{32}$$

# Independent Events: Dice Example 1

$$P(6) = \frac{1}{6}$$


# Independent Events: Dice Example 1

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6



2 dice

$$P(6,6) = \text{_____} = \frac{1}{36}$$

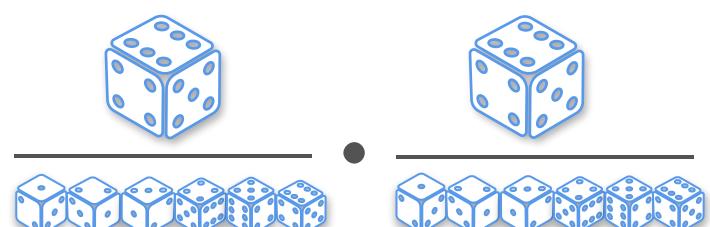
# Independent Events: Dice Example 1

						$\frac{1}{6}$
	1,1	1,2	1,3	1,4	1,5	1,6
	2,1	2,2	2,3	2,4	2,5	2,6
	3,1	3,2	3,3	3,4	3,5	3,6
	4,1	4,2	4,3	4,4	4,5	4,6
	5,1	5,2	5,3	5,4	5,5	5,6
$\frac{1}{6}$	6,1	6,2	6,3	6,4	6,5	6,6

2 dice

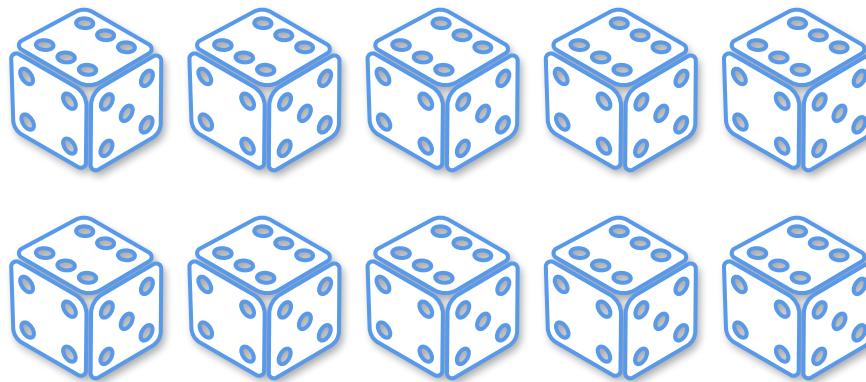
$$P(6,6) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \frac{1}{6} \cdot \frac{1}{6} = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$$

$P(6,6) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \frac{1}{6} \cdot \frac{1}{6} = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$



# Independent Events: Dice Example 2

What is the probability of getting 10 sixes?



# Independent Events: Dice Example 2

What is the probability of getting 10 sixes?

$$\begin{aligned} P(10 \text{ sixes}) &= \left( \frac{\text{one die showing 6}}{\text{one die showing 6}} \right)^{10} \\ &= \left( \frac{1}{6} \right)^{10} \end{aligned}$$



DeepLearning.AI

## Introduction to probability

---

### Birthday problem

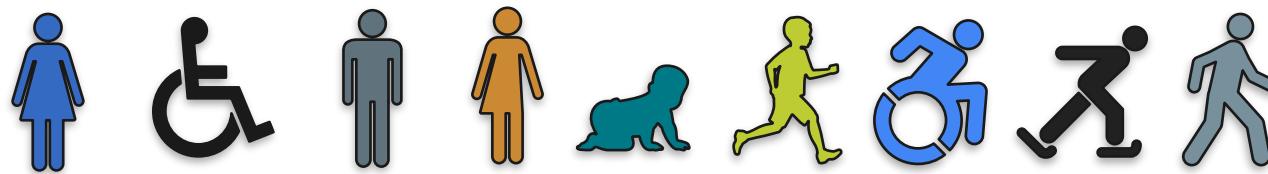
# Quiz

- You have 30 friends at a party. What do you think is more likely:
  - That there exist two people with the same birthday
  - That no two of them have the same birthday
- (Assume the year has 365 days, nobody has a birthday on Feb 29).

# Quiz

- Answer: It's more likely that 2 people have the same birthday.
- In fact, the probability of no two people having the same birthday is around 0.3.

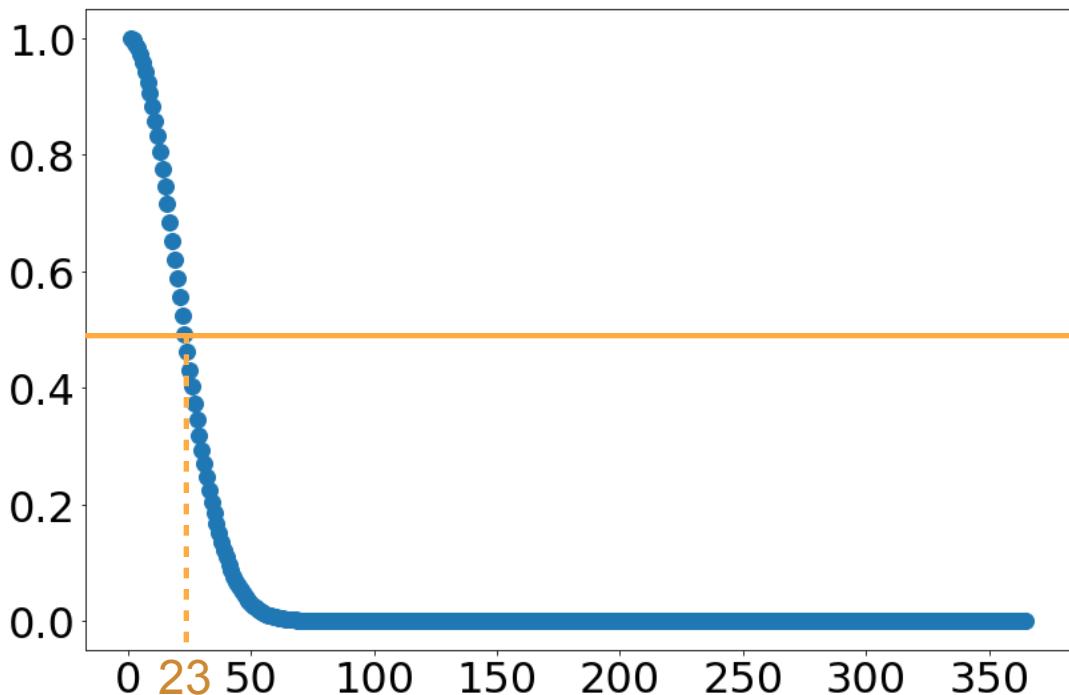
# Probability That Everyone Has a Different Birthday



$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \frac{361}{365} \cdot \frac{360}{365} \cdot \frac{359}{365} \cdot \frac{358}{365} \cdot \frac{357}{365} = 0.905$$

# Probability That no Two People Have the Same Birthday

1 person: 1  
2 people: 0.997  
3 people: 0.992  
4 people: 0.984  
5 people: 0.973  
10 people: 0.883  
20 people: 0.589  
**23 people: 0.493**  
30 people: 0.294  
50 people: 0.030  
100 people: 0.0000003  
365 people: 0





DeepLearning.AI

# Introduction to Probability

---

## Conditional probability

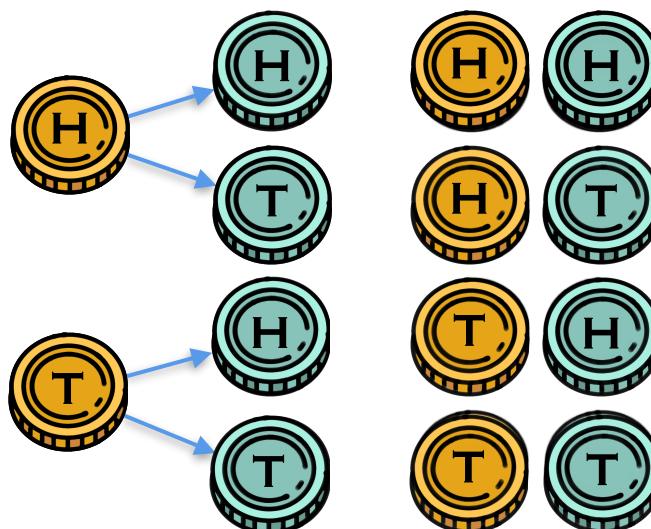
# Conditional Probability: Coin Example 1



50% 50%

What is the probability of landing on heads twice?

1st      2nd



$$P(HH) = \frac{1}{4} = \frac{1}{4}$$

The equation shows the probability of getting heads on both the 1st and 2nd coin tosses given that the first one is heads. The numerator is 1, representing the favorable outcome (HH), and the denominator is 4, representing the total number of outcomes in the sample space (HH, HT, TH, TT).

# Conditional Probability: Coin Example 1

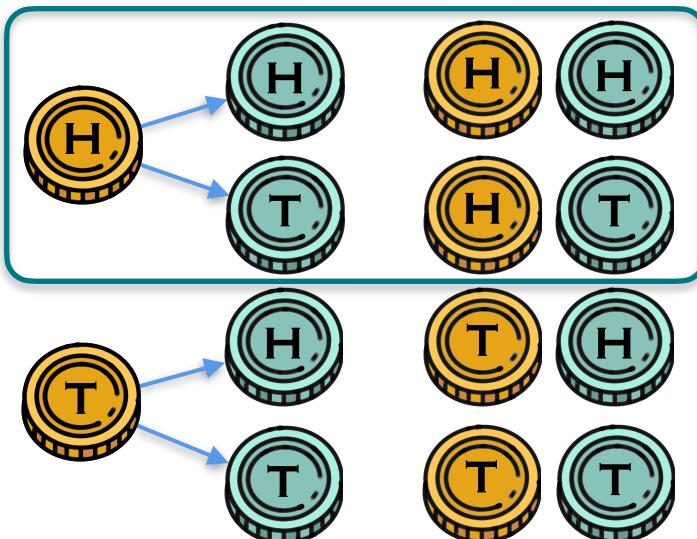


50% 50%

What is the probability of landing on heads twice?

1st      2nd

**GIVEN** that the first one is heads



$$P(HH) = \frac{\text{Number of HH outcomes}}{\text{Total number of outcomes}} = \frac{1}{4}$$
A diagram showing the calculation of the probability P(HH). It shows two coins, both yellow with 'H' (heads), followed by a fraction line. Below the fraction line, there are four possible outcomes arranged in a 2x2 grid: H, H; H, T; T, H; T, T. The top-left outcome (H, H) is shaded in light blue, while the others are light yellow.

# Conditional Probability: Coin Example 1

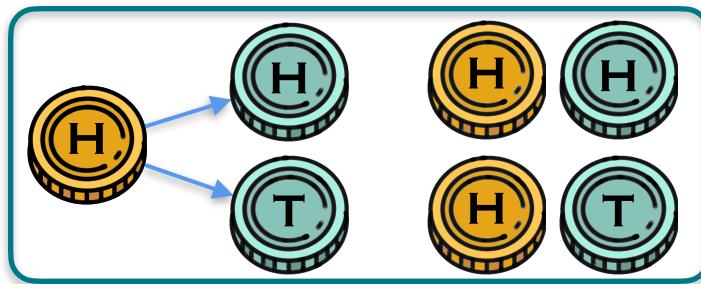


50% 50%

What is the probability of landing on heads twice?

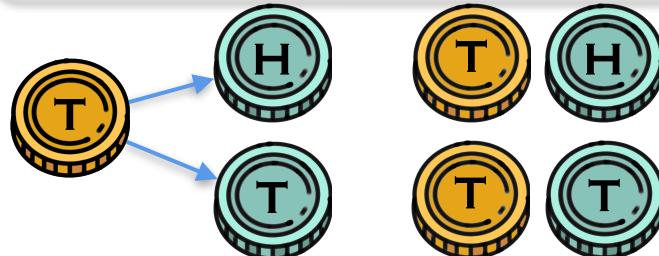
1st      2nd

**GIVEN** that the first one is heads



$$P(HH | \text{1st is } H) =$$

$$= \frac{1}{2}$$



# Conditional Probability: Coin Example 1

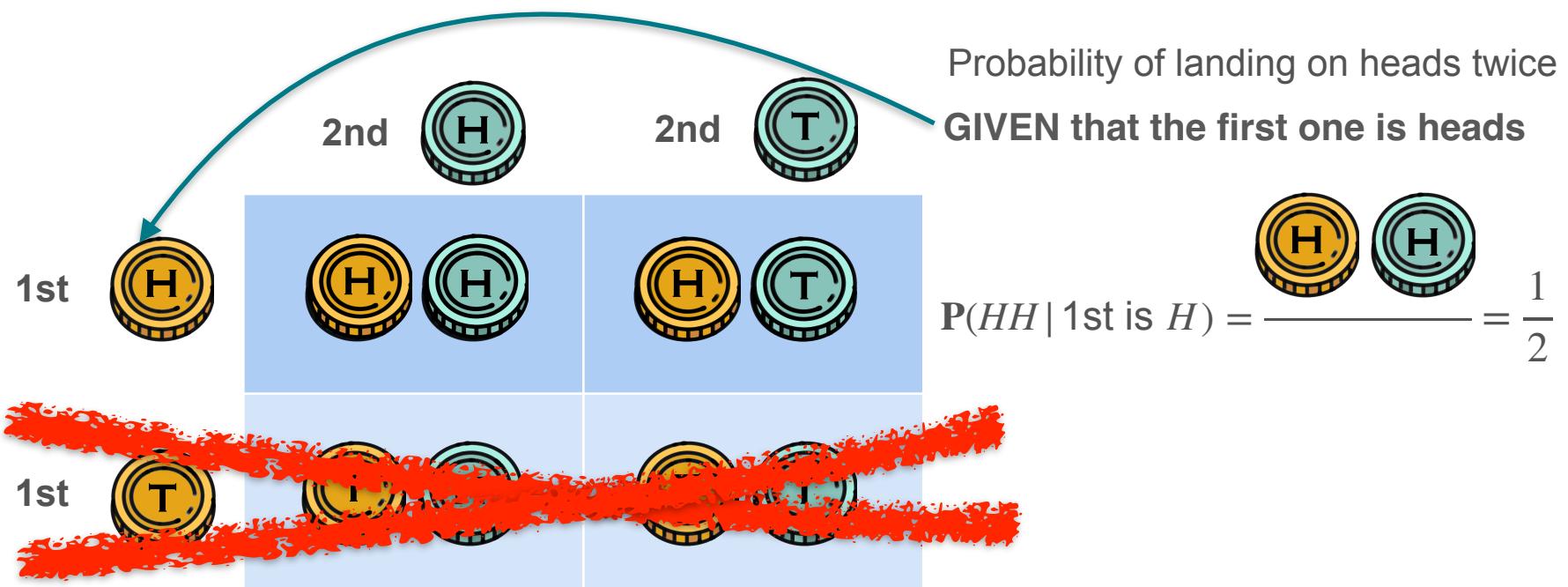
What is the probability of landing on heads twice?

**GIVEN** that the first one is heads

$$P(HH | \text{1st is } H)$$



# Conditional Probability: Coin Example 1



# Conditional Probability: Coin Example 2

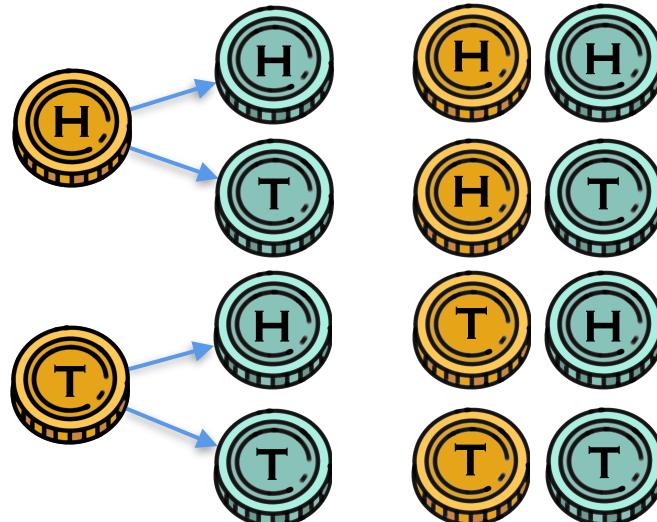


50% 50%

What is the probability of landing on heads twice?

1st      2nd

**GIVEN** that the first one is tails



$$P(HH) = \frac{1}{4}$$

The probability of getting heads twice given that the first flip is tails is  $\frac{1}{4}$ . There are 4 possible outcomes when flipping two coins: HH, HT, TH, TT. Only one of these outcomes (HH) satisfies the condition of the first flip being tails.

# Conditional Probability: Coin Example 2

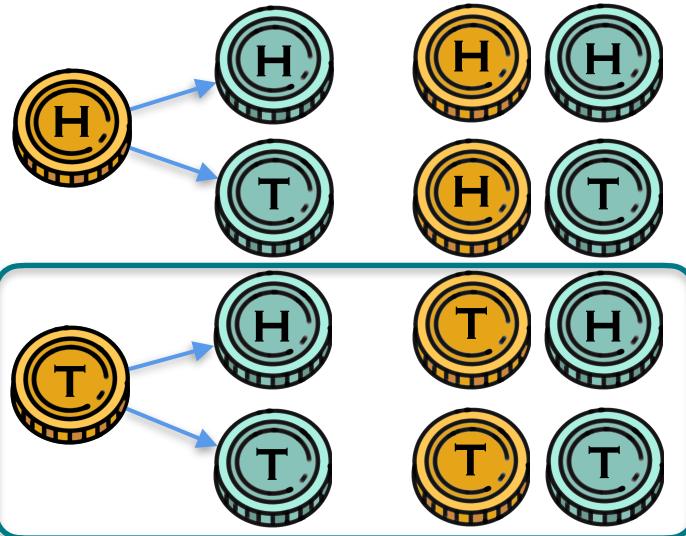


50% 50%

What is the probability of landing on heads twice?

1st      2nd

**GIVEN** that the first one is tails



$$P(HH) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \frac{1}{4}$$
Two coins are shown side-by-side. The left coin is yellow with 'H' on top and 'T' on the bottom. The right coin is blue with 'H' on top and 'T' on the bottom.



# Conditional Probability: Coin Example 2

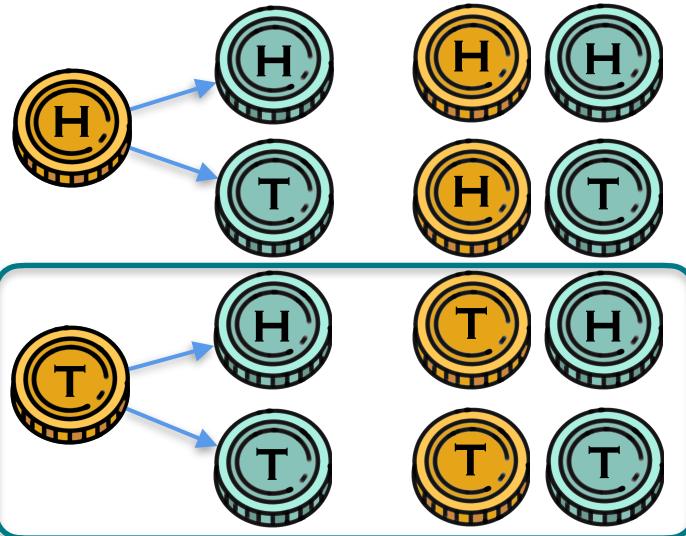


50% 50%

What is the probability of landing on heads twice?

1st      2nd

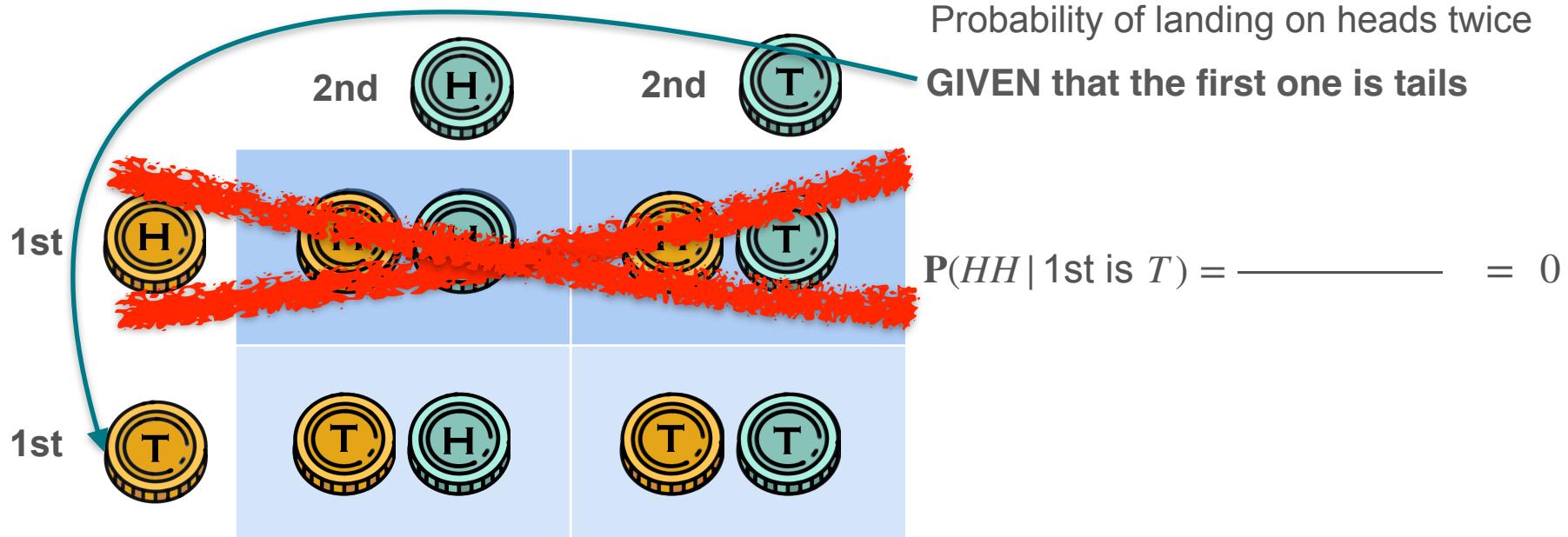
**GIVEN** that the first one is tails



$$P(HH \mid \text{1st is } T) = \frac{0}{8} = 0$$



# Conditional Probability: Coin Example 2

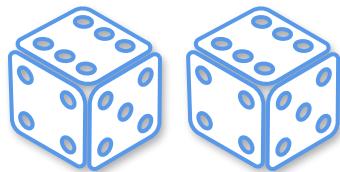


# Product Rule (for Independent Events)

When A and B independent

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

# Conditional Probability: Dice Example 3



What is the probability that  
the first is 6 **AND** the sum = 10?

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{1st is 6} \cap \text{sum} = 10) =$$

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

# Conditional Probability: Dice Example 3

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

$$P(1\text{st is } 6 \cap \text{sum} = 10) =$$

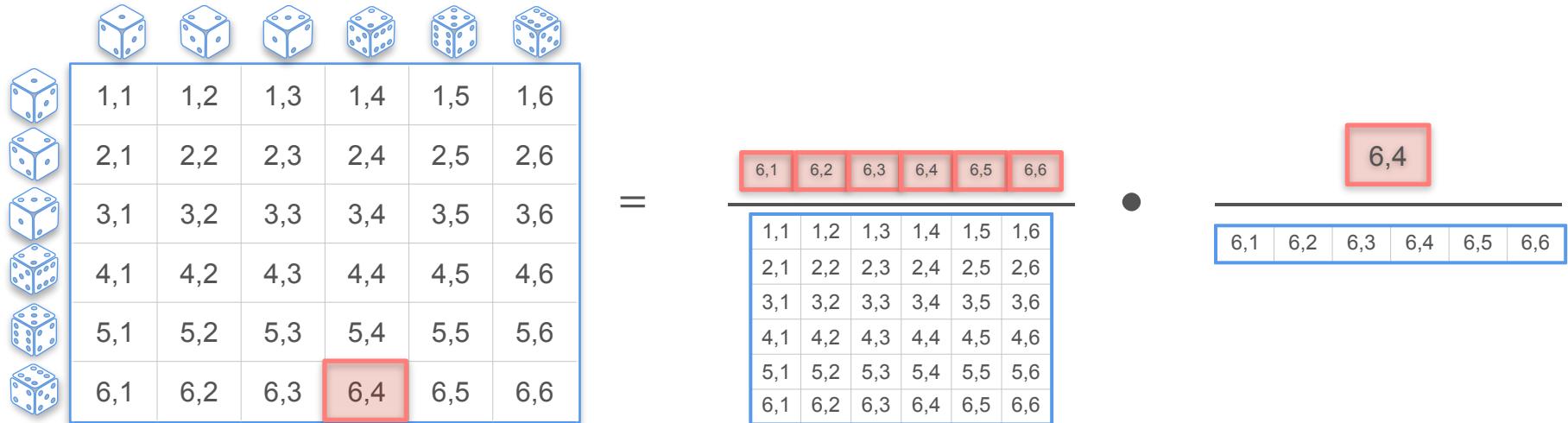
$$P(1\text{st is } 6)$$



$$P(\text{sum} = 10 | 1\text{st } 6)$$

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

# Conditional Probability: Dice Example 3



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

# The General Product Rule

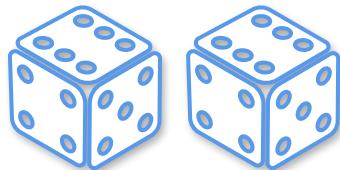
$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B | A)$$



When independent

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

# Conditional Probability: Dice Example 1

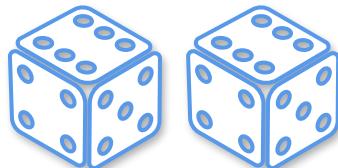


What is the probability that the sum is 10?

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10) = \frac{3}{36} = \frac{1}{12}$$

# Conditional Probability: Dice Example 1



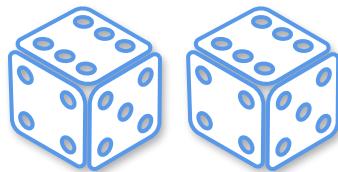
What is the probability that the sum is 10?

**GIVEN** that the first one is 6

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
2,1	3,1	3,2	3,3	3,4	3,5	3,6
3,1	4,1	4,2	4,3	4,4	4,5	4,6
4,1	5,1	5,2	5,3	5,4	5,5	5,6
5,1	6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10 \mid \text{1st is } 6) = \underline{\hspace{10cm}}$$

# Conditional Probability: Dice Example 1



What is the probability that the sum is 10?

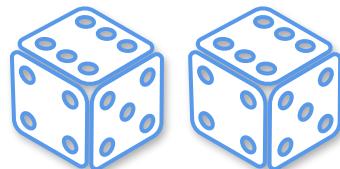
**GIVEN** that the first one is 6

	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6	
3,1	3,2	3,3	3,4	3,5	3,6	
4,1	4,2	4,3	4,4	4,5	4,6	
5,1	5,2	5,3	5,4	5,5	5,6	
6,1	6,2	6,3	6,4	6,5	6,6	

$$P(\text{sum} = 10 \mid \text{1st is } 6) = \underline{\hspace{10cm}}$$

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

# Conditional Probability: Dice Example 2



What is the probability that the sum is 10?

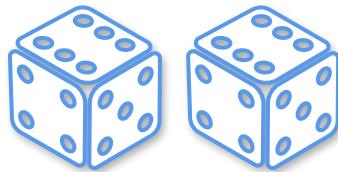
1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \frac{3}{36}$$

A staircase diagram showing the 36 possible outcomes of two dice rolls. The outcomes are represented as pairs of dice faces. The outcomes where the sum is 10 are highlighted with red boxes: (4,6), (5,5), and (6,4).

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

# Conditional Probability: Dice Example 2



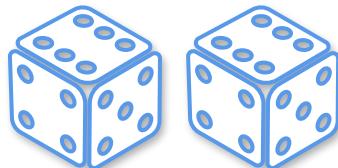
What is the probability that the sum is 10?

**GIVEN** that the first one is 1

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10 \mid \text{1st is } 1) = \underline{\hspace{10cm}}$$

# Conditional Probability: Dice Example 2



What is the probability that the sum is 10?

**GIVEN** that the first one is 1

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	2,1	2,2	2,3	2,4	2,5	2,6
1,2	3,1	3,2	3,3	3,4	3,5	3,6
1,3	4,1	4,2	4,3	4,4	4,5	4,6
1,4	5,1	5,2	5,3	5,4	5,5	5,6
1,5	6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10 \mid \text{1st is } 1) = \frac{0}{\dots} = 0$$



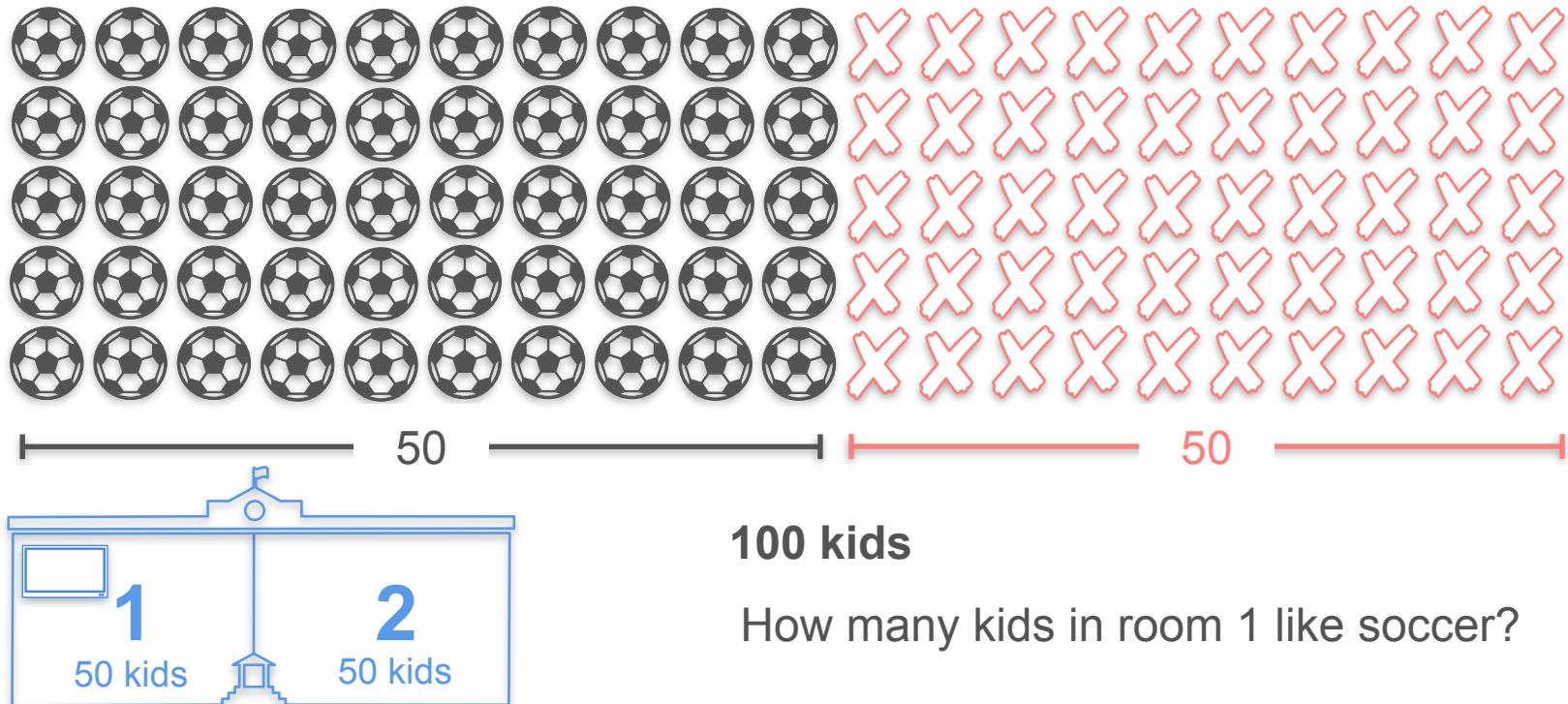
DeepLearning.AI

# Introduction to Probability

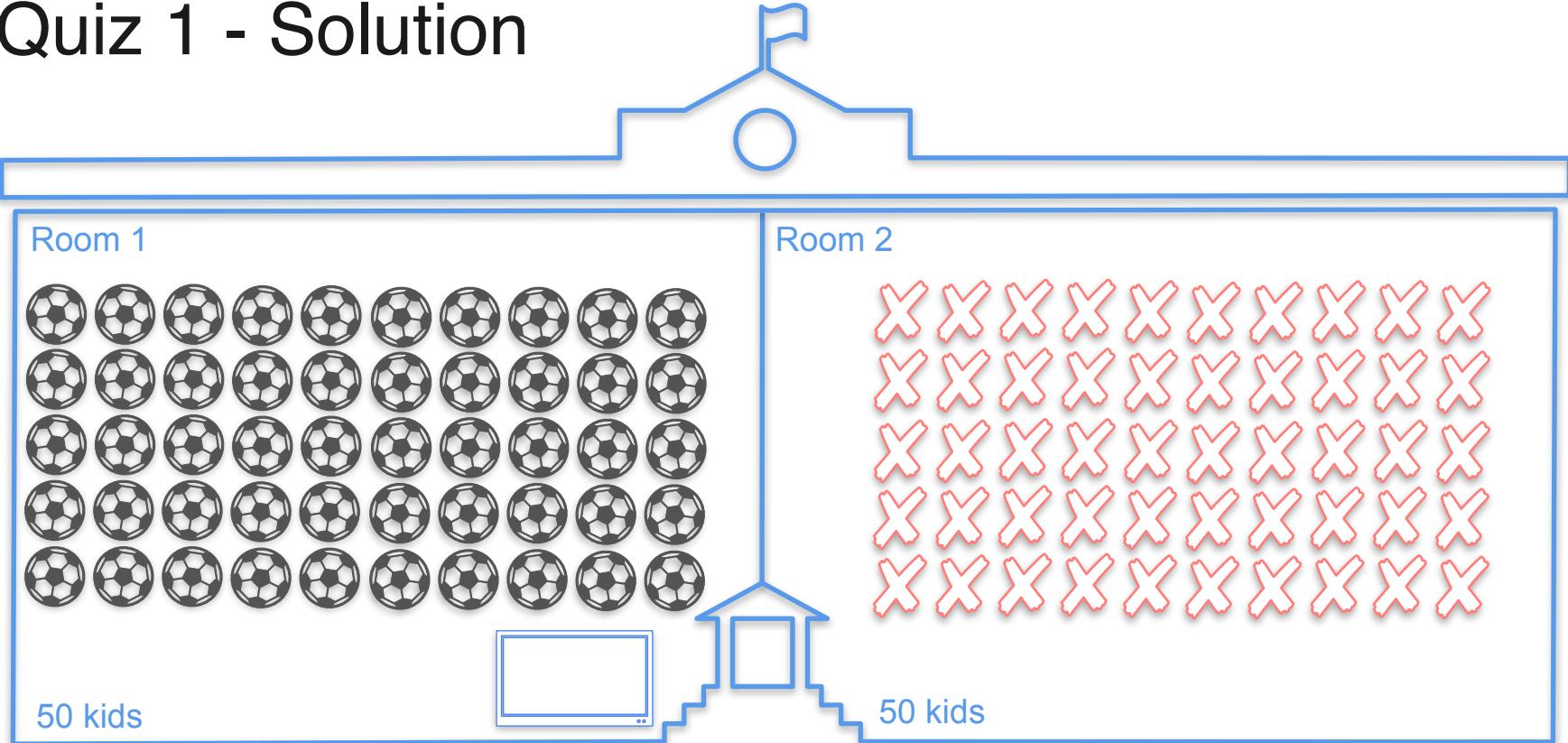
---

## Conditional probability - Part 2

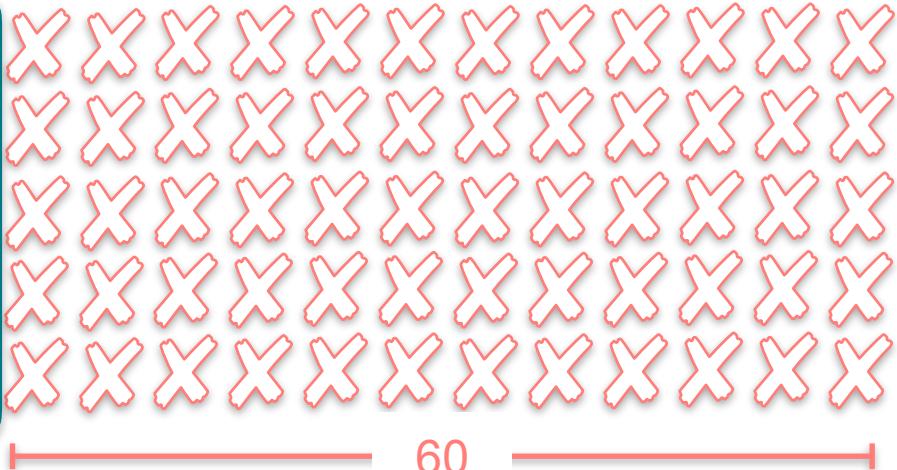
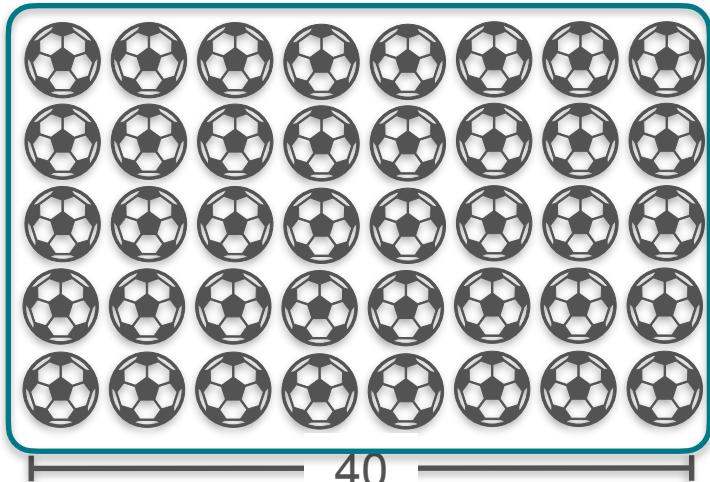
# Quiz 1



# Quiz 1 - Solution



## Quiz 2

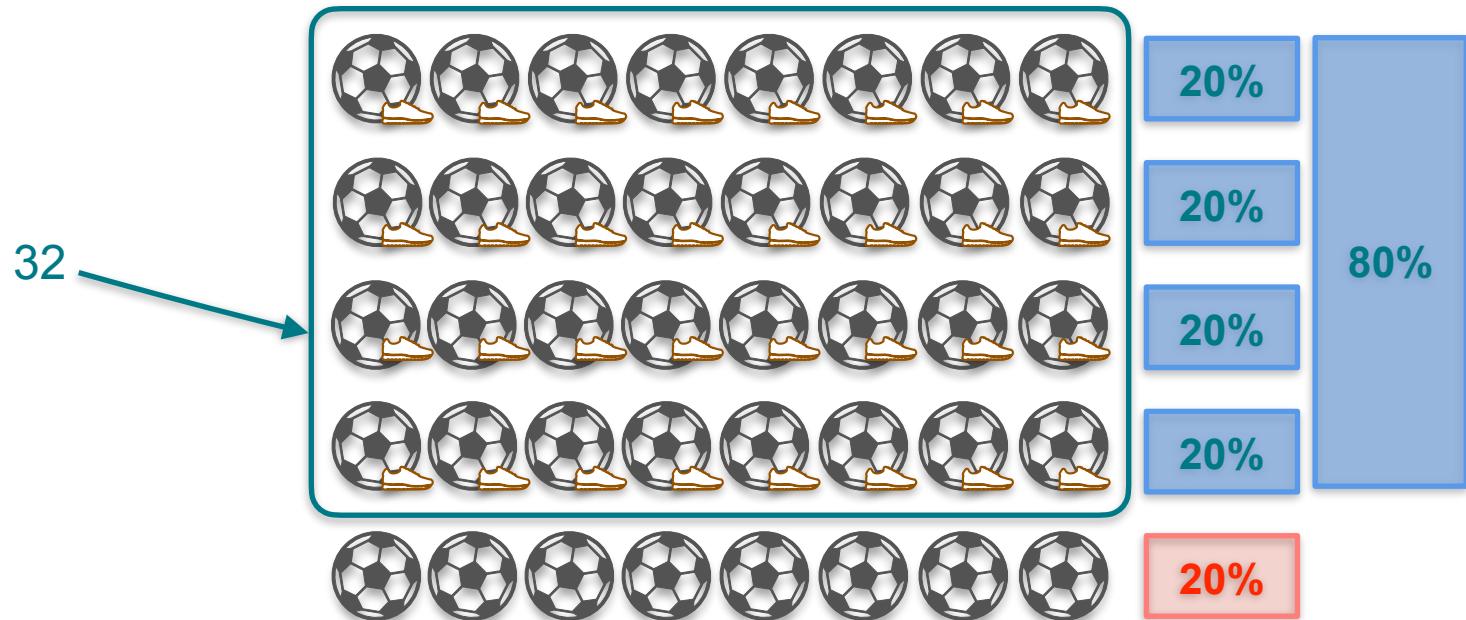


80%

**100 kids**

How many kids play soccer and wear running shoes

# Quiz 2 - Solution



# Conditional Probability

$P(\text{Soccer and Running shoes})$

$$P(S \cap R) = P(S) \bullet P(R|S)$$

$$= 0.4 \bullet 0.8$$

$$= 0.32$$



$$P(\text{not } S) = 0.6$$

$$P(R | S) = 0.8$$



# Conditional Probability

$P(\text{Soccer and Running shoes})$



$$P(S \cap R) = 0.32$$

$$P(S) = 0.4$$

$P(\text{not Soccer and Running shoes})$

$$P(\text{not } S \cap R) = P(\text{not } S) \bullet P(R | \text{not } S)$$

$$= 0.6 \bullet 0.5$$

$$= 0.3$$

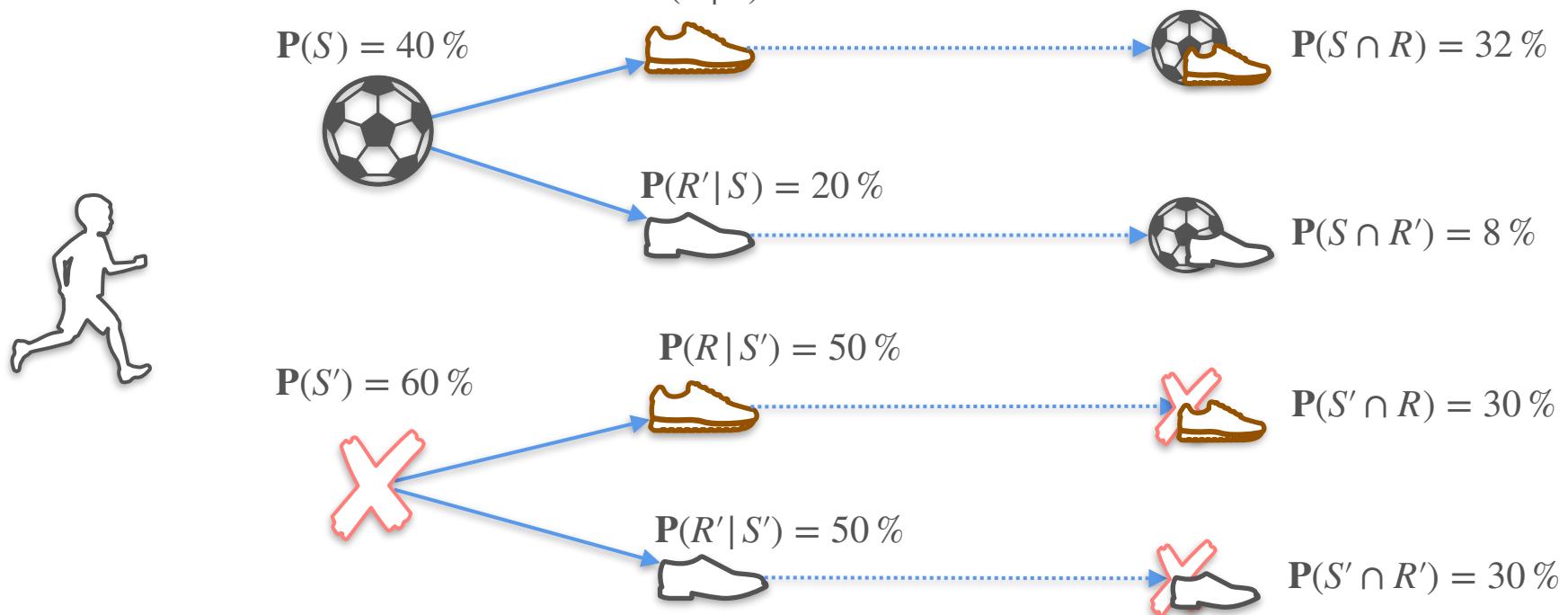


$$P(\text{not } S) = 0.6$$

$$P(R | S) = 0.8$$



# Conditional Probability

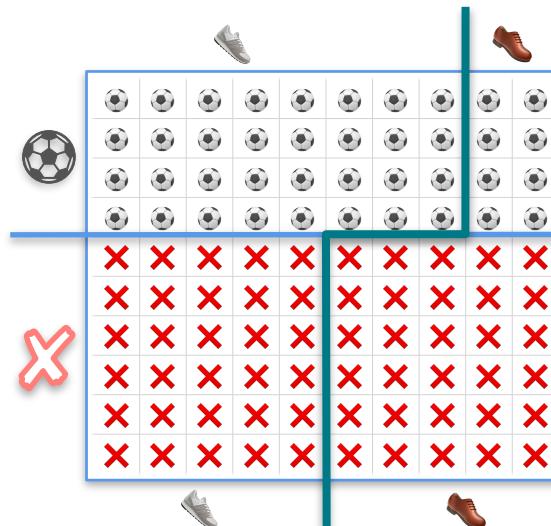


# Independent vs Dependent Events

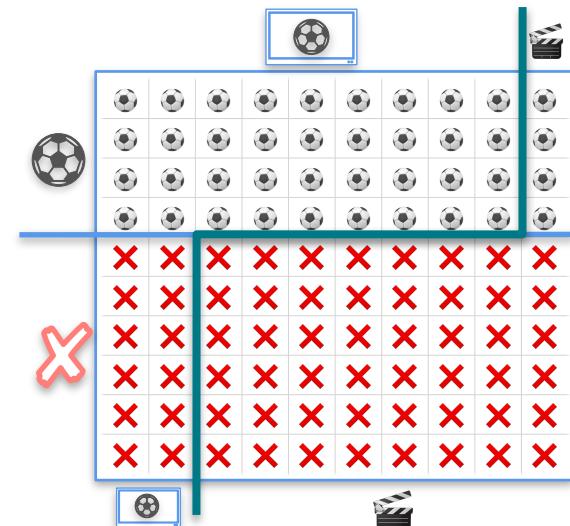
Independent



Dependent



Dependent





DeepLearning.AI

# Introduction to probability

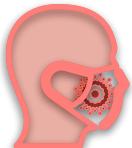
---

## Bayes Theorem - Intuition

# Bayes Theorem: Intuition



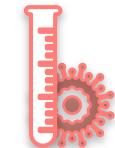
1,000,000 people



1 / 10,000 people



100 people



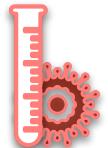
99



1



99% Effective



Tested Sick



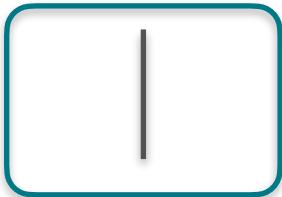
100 people



1

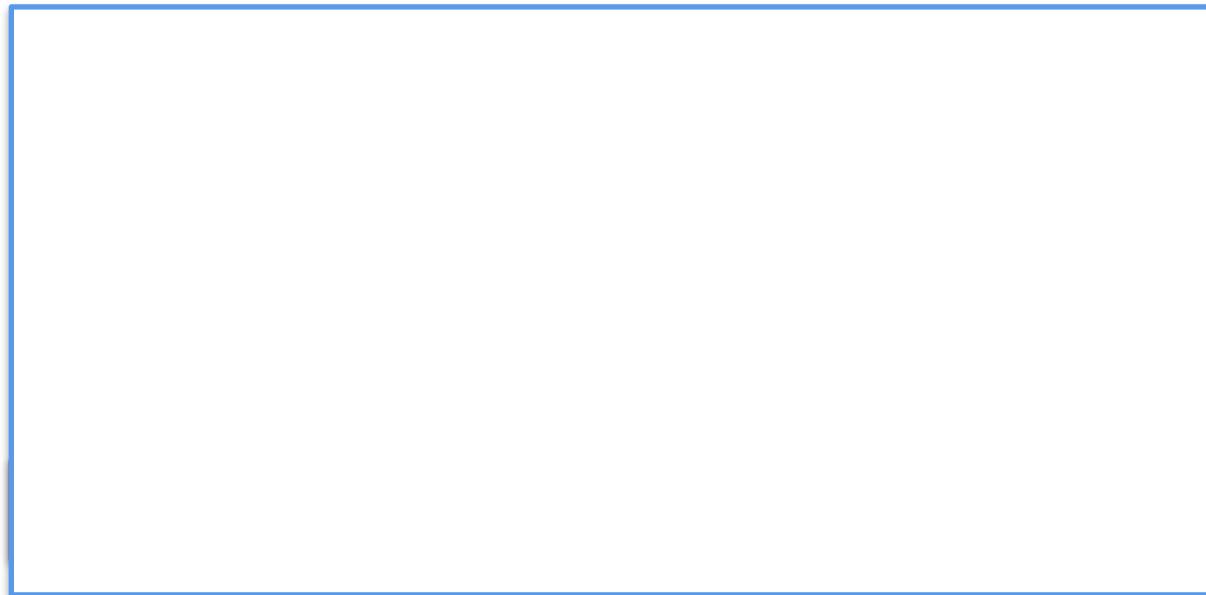
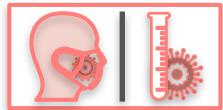


99



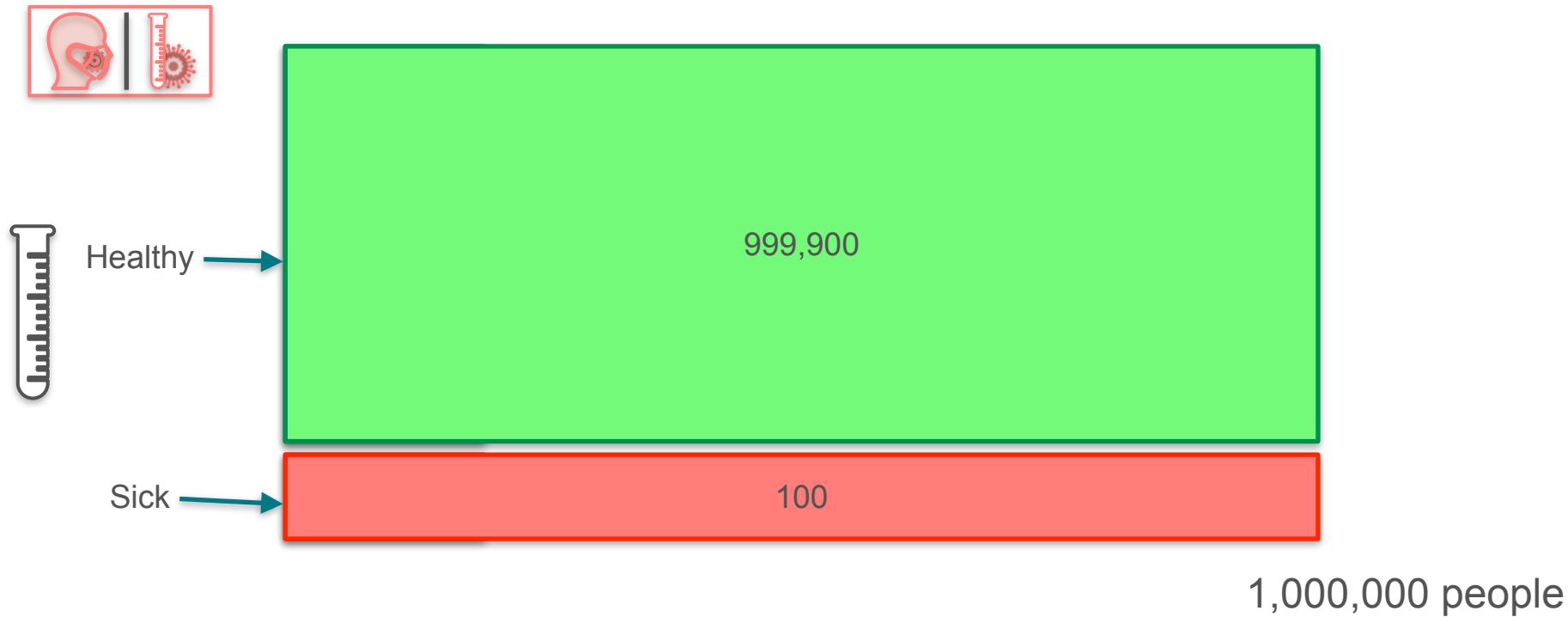
What's the probability that **you are sick**  
**GIVEN that you tested sick?**

# Bayes Theorem: Intuition

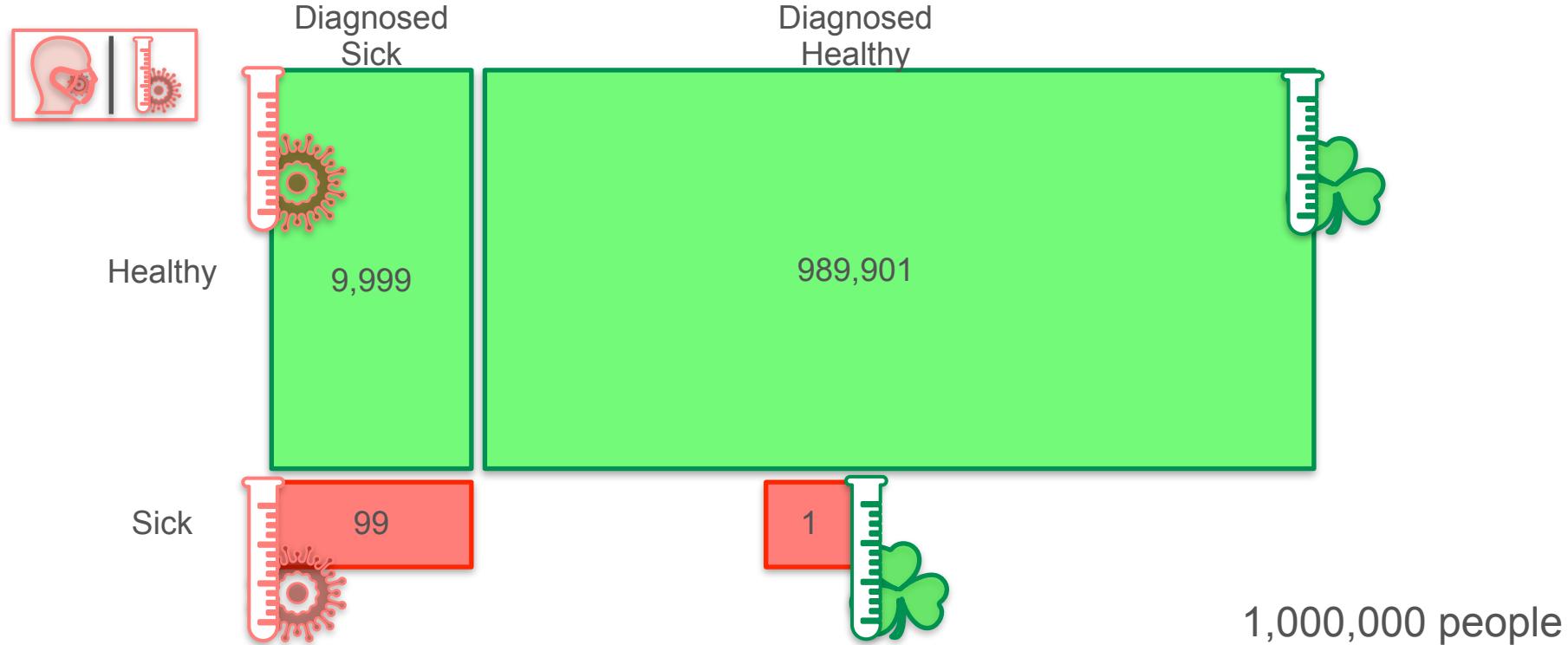


1,000,000 people

# Bayes Theorem: Intuition



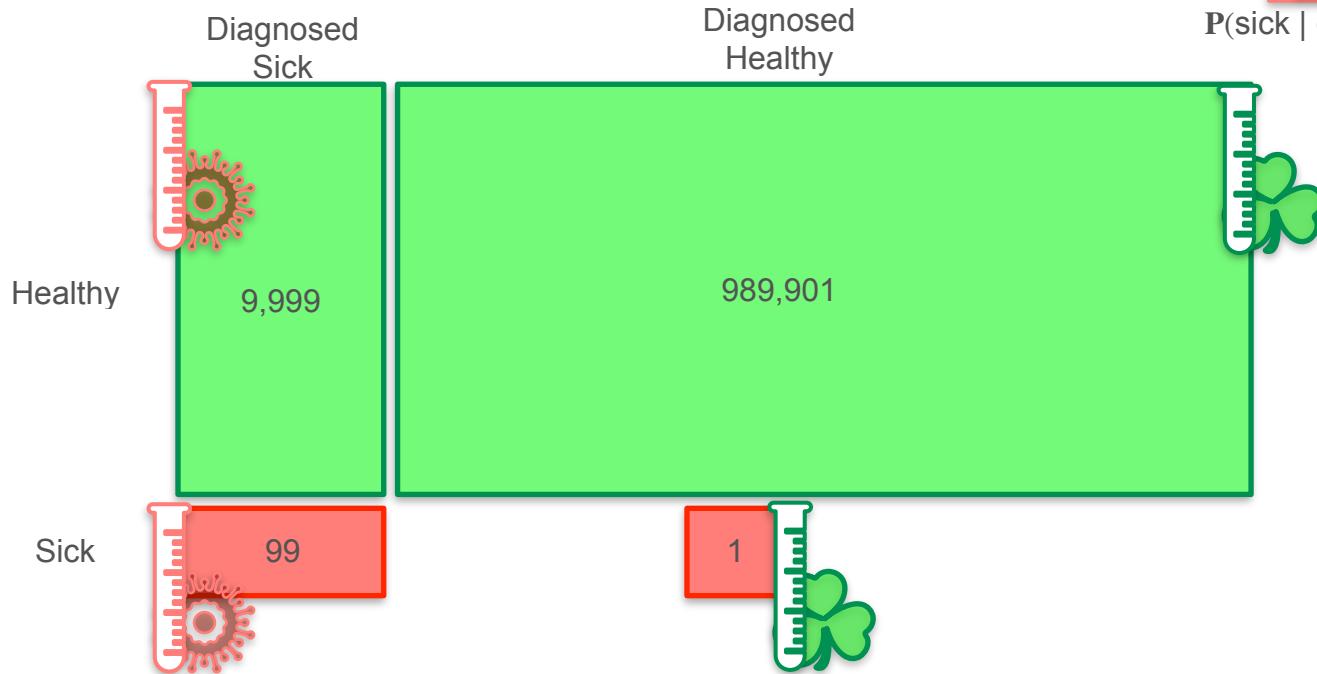
# Bayes Theorem: Intuition



# Bayes Theorem: Intuition



$P(\text{sick} | \text{diagnosed sick}) = \underline{\hspace{2cm}}$

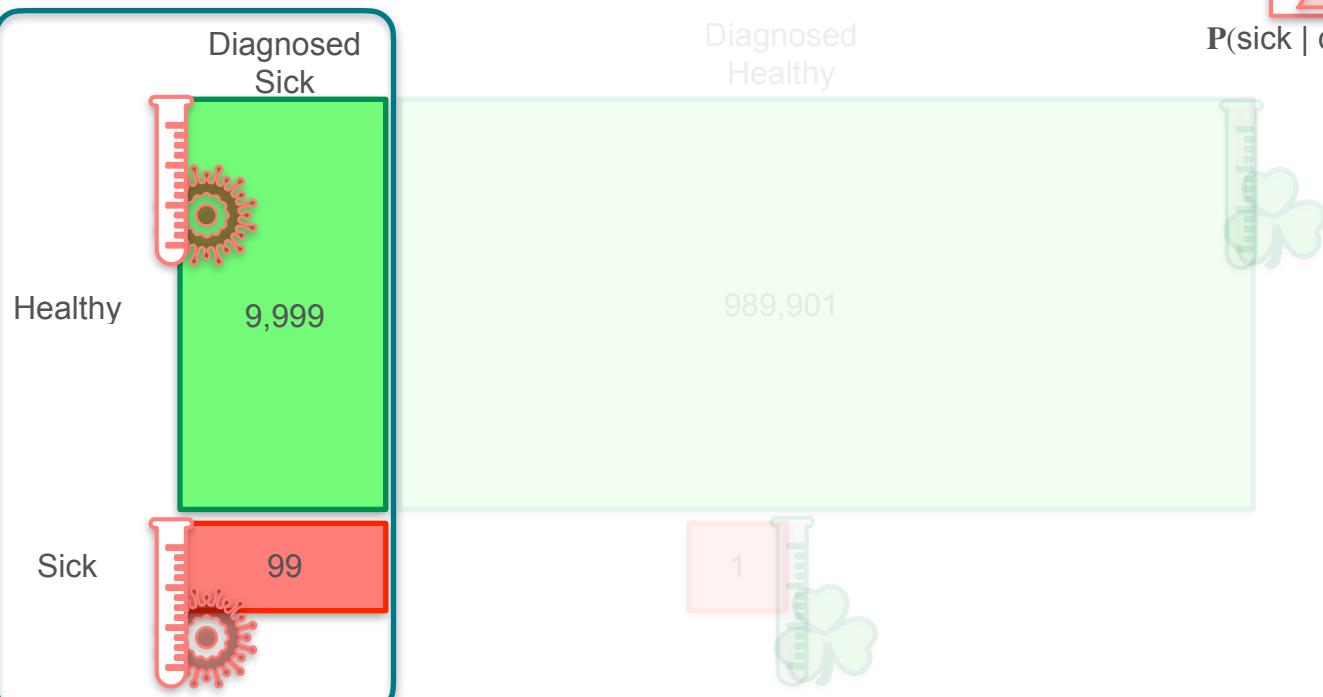


1,000,000 people

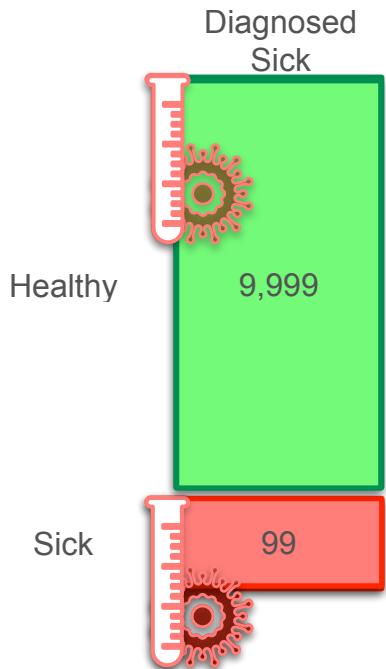
# Bayes Theorem: Intuition



$P(\text{sick} | \text{diagnosed sick}) = \underline{\hspace{2cm}}$



# Bayes Theorem: Intuition

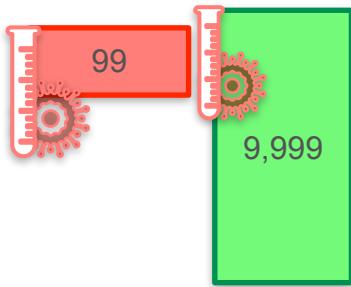


$$\begin{aligned} P(\text{sick} | \text{diagnosed sick}) &= \frac{\text{sick and diagnosed sick}}{\text{sick and diagnosed sick} + \text{healthy and diagnosed sick}} \\ &= \frac{99}{99 + 9999} \\ &= \frac{99}{10098} = 0.0098 \end{aligned}$$

# Bayes Theorem: Intuition



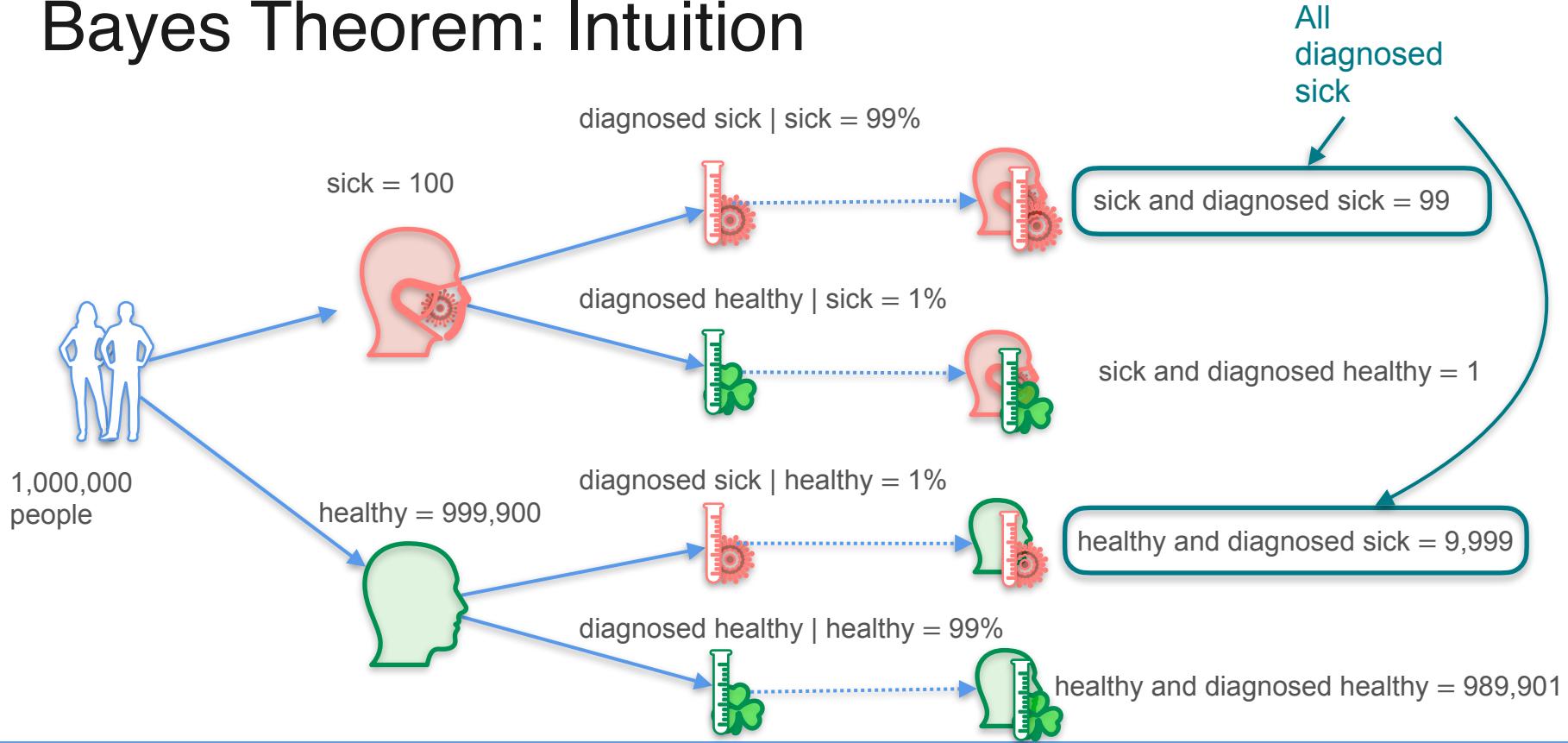
$$P(\text{sick} \mid \text{diagnosed sick}) = \frac{\text{ }}{\text{ }}$$



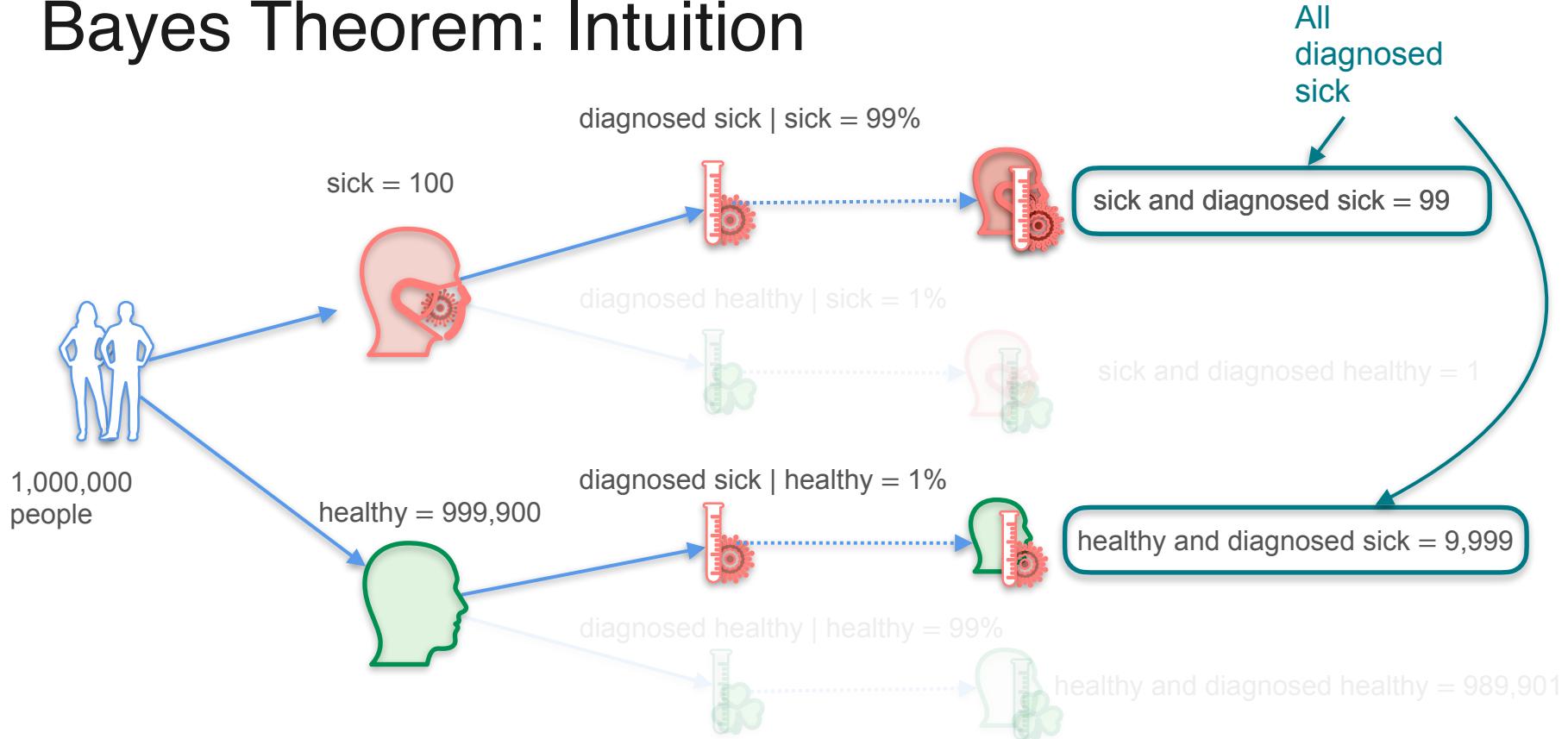
$$= \frac{\text{sick and diagnosed sick}}{\text{sick and diagnosed sick + healthy and diagnosed sick}}$$

$$= \frac{\text{sick and diagnosed sick}}{\text{everyone diagnosed sick}}$$

# Bayes Theorem: Intuition

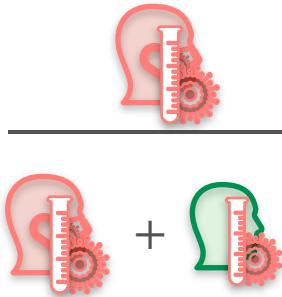


# Bayes Theorem: Intuition



# Bayes Theorem: Intuition

$$P(\text{sick} \mid \text{diagnosed sick}) = \frac{\text{ }}{\text{ }}$$



$$P(\text{sick} \mid \text{diagnosed sick}) = \frac{\text{sick and diagnosed sick} = 99}{\text{healthy and diagnosed sick} = 9,999 + \text{sick and diagnosed sick} = 99}$$

$$P(\text{sick} \mid \text{diagnosed sick}) = \frac{99}{10098} = 0.0098$$



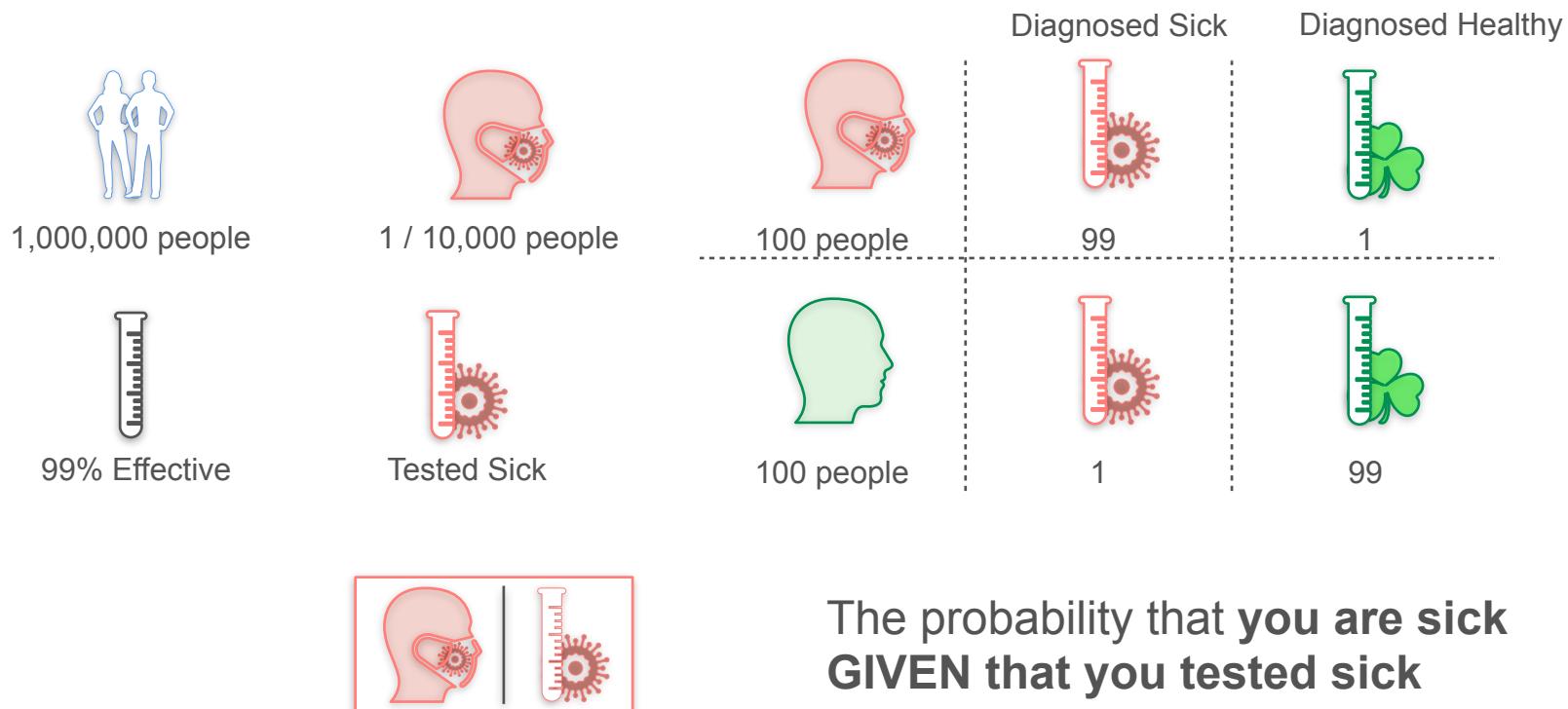
DeepLearning.AI

# Introduction to probability

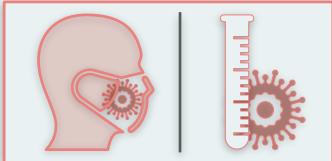
---

**Bayes Theorem -  
Mathematical Formula**

# Bayes Theorem: Formula



# Bayes Theorem: Formula



$$P(\text{sick} \mid \text{diagnosed sick}) = ?$$

$$P(\text{sick}) = 0.01\%$$

$$P(\text{not sick}) = 99.99\%$$

$$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$$

$$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$$

The probability that **you are sick**  
**GIVEN** that you tested sick



1,000,000

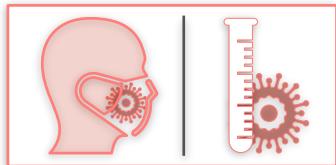


1 / 10,000



99% Effective

# Bayes Theorem: Formula



$P(\text{sick} \mid \text{diagnosed sick}) = ?$

$P(\text{sick}) = 0.01\%$

$P(\text{not sick}) = 99.99\%$

$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$

$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$

$A: \text{sick}$

$B: \text{diagnosed sick}$

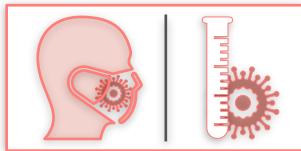
$P(A \mid B) = ?$

**From Conditional Probability**

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

$$P(A \mid B) = \frac{P(\text{sick and diagnosed sick})}{P(\text{diagnosed sick})}$$

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

$$P(\text{sick} \mid \text{diagnosed sick}) = ?$$

$$P(\text{sick}) = 0.01\%$$

$$P(\text{not sick}) = 99.99\%$$

$$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$$

$$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$$

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

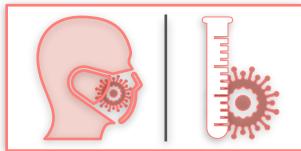
$$P(A \mid B) = \frac{P(\text{sick and diagnosed sick})}{P(\text{diagnosed sick})}$$

$$P(\text{sick and diagnosed sick}) = ?$$

$$P(\text{diagnosed sick}) = ?$$

**BAYES THEOREM FORMULA CAN HELP**

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

$P(\text{sick} \mid \text{diagnosed sick}) = ?$

$P(\text{sick}) = 0.01\%$

$P(\text{not sick}) = 99.99\%$

$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$

$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$

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

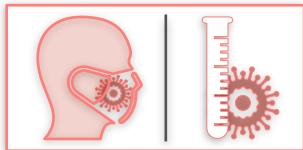
$$P(\text{sick and diagnosed sick}) = ?$$
$$P(A \mid B) = \frac{P(\text{sick and diagnosed sick})}{P(\text{diagnosed sick})} = ?$$

**From Conditional Probability**

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

$$P(\text{sick and diagnosed sick}) = P(\text{sick}) \cdot P(\text{diagnosed sick} \mid \text{sick})$$

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

$$P(\text{sick} \mid \text{diagnosed sick}) = ?$$

$$P(\text{sick}) = 0.01\%$$

$$P(\text{not sick}) = 99.99\%$$

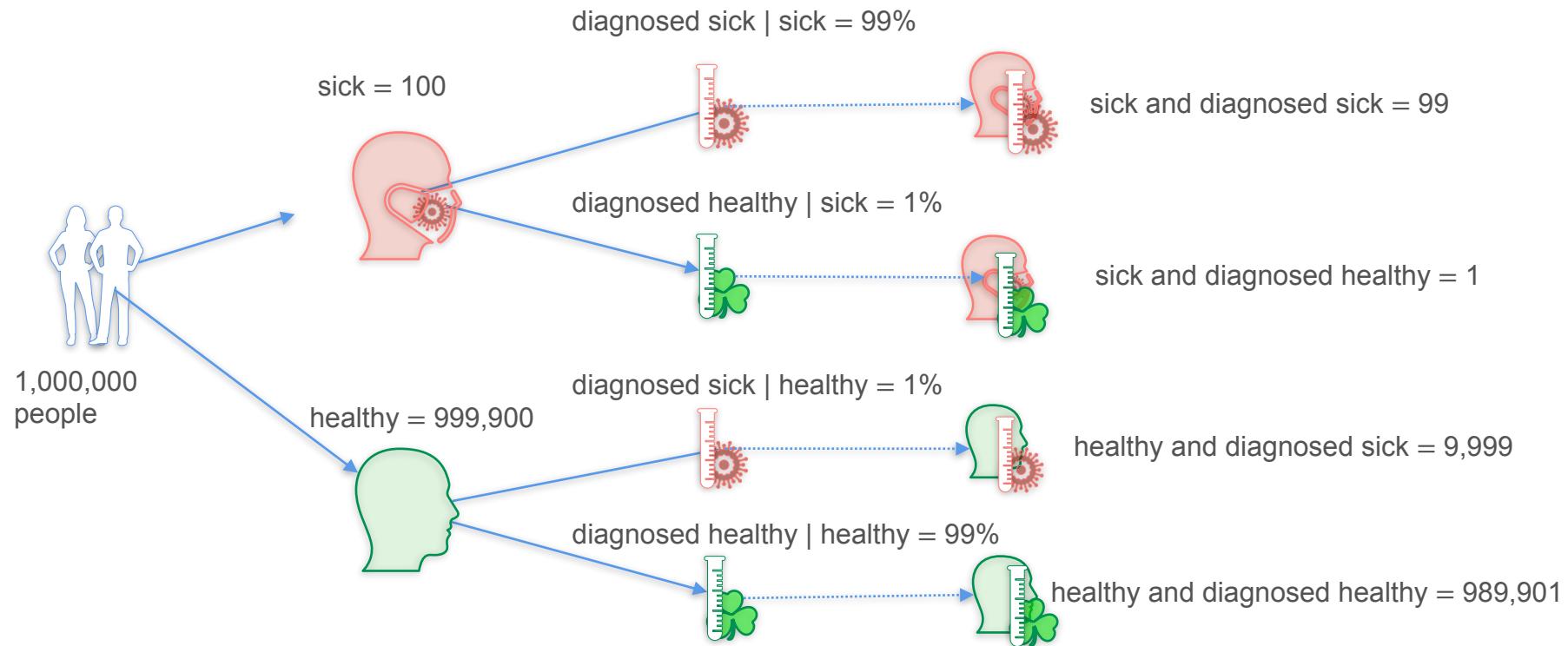
$$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$$

$$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$$

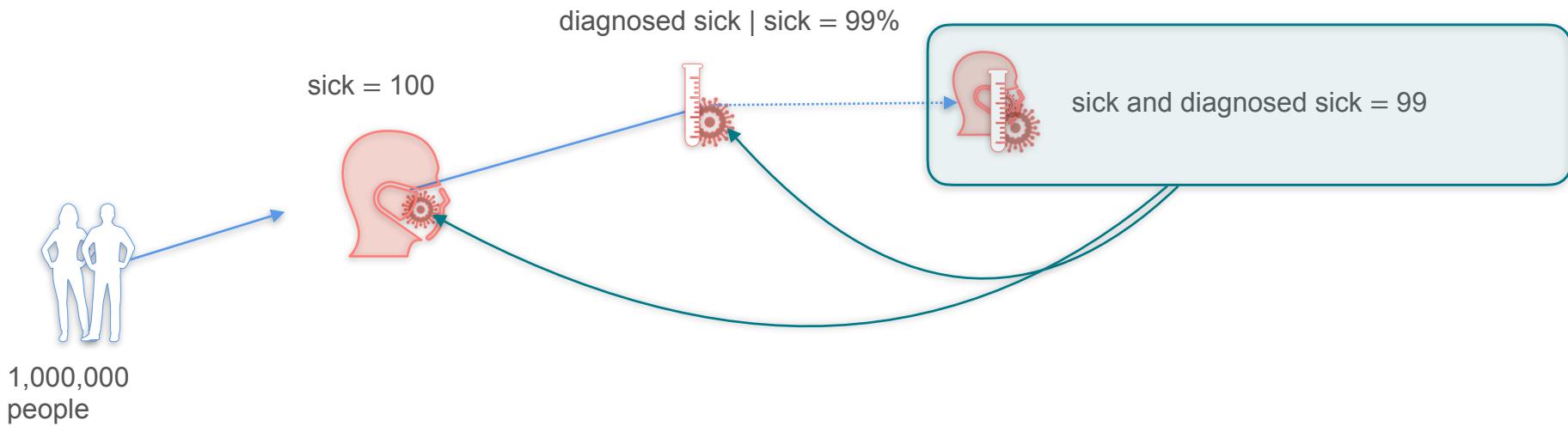
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \mid B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} \mid \text{sick})}{P(\text{diagnosed sick})} = ?$$

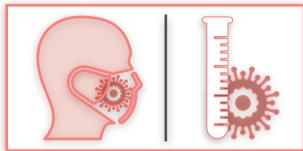
# Bayes Theorem: Formula



# Bayes Theorem: Formula



# Bayes Theorem: Formula



A: sick

B: diagnosed sick

$$P(\text{sick} \mid \text{diagnosed sick}) = ?$$

$$P(\text{sick}) = 0.01\%$$

$$P(\text{not sick}) = 99.99\%$$

$$P(\text{diagnosed sick} \mid \text{sick}) = 99\%$$

$$P(\text{diagnosed sick} \mid \text{not sick}) = 1\%$$

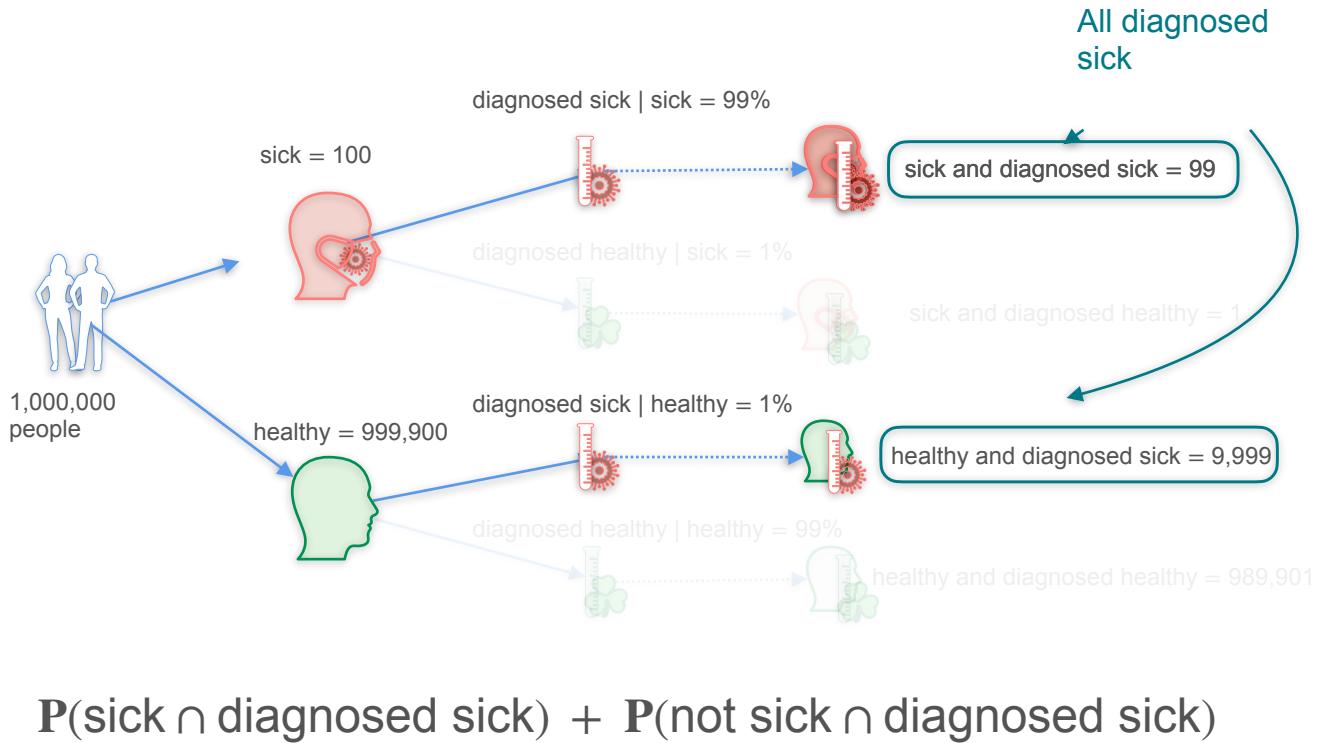
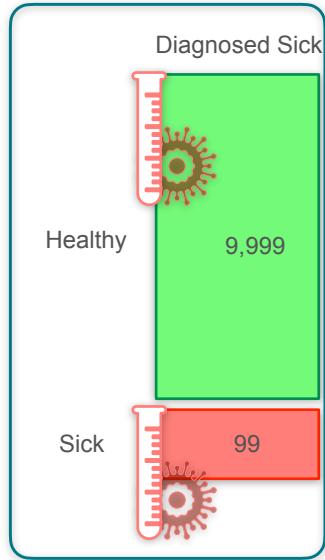
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \mid B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} \mid \text{sick})}{P(\text{diagnosed sick})} = ?$$

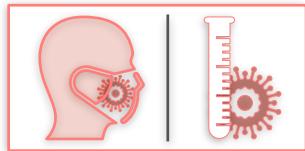
$$P(\text{diagnosed sick}) =$$

# Bayes Theorem: Formula

$P(\text{diagnosed sick}) =$



# Bayes Theorem: Formula



$A$ : sick

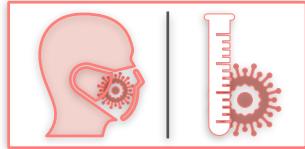
$B$ : diagnosed sick

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

$$P(A | B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick})}{P(\text{diagnosed sick}) = ?}$$

$$P(\text{diagnosed sick}) = P(\text{sick} \cap \text{diagnosed sick}) + P(\text{not sick} \cap \text{diagnosed sick})$$

# Bayes Theorem: Formula



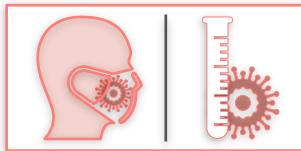
$A$ : sick

$B$ : diagnosed sick

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

$$P(A | B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick})}{P(\text{sick} \cap \text{diagnosed sick}) + P(\text{not sick} \cap \text{diagnosed sick})}$$

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

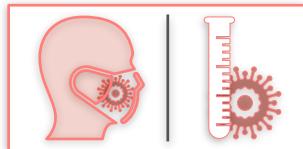
$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A | B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick})}{P(\text{sick} \cap \text{diagnosed sick}) + P(\text{not sick} \cap \text{diagnosed sick})}$$

$$\begin{aligned} P(\text{sick} \cap \text{diagnosed sick}) &= P(A \cap B) \\ &= P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick}) \end{aligned}$$

$$\begin{aligned} P(\text{not sick} \cap \text{diagnosed sick}) &= P(A' \cap B) \\ &= P(\text{not sick}) \cdot P(\text{diagnosed sick} | \text{not sick}) \end{aligned}$$

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

?

$$P(A | B) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick})}{P(\text{sick}) \cdot P(\text{diagnosed sick} | \text{sick}) + P(\text{not sick}) \cdot P(\text{diagnosed sick} | \text{not sick})}$$

$$P(\text{sick}) = 0.01\%$$

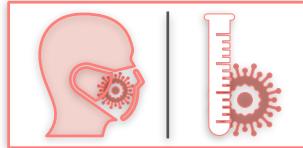
$$P(\text{not sick}) = 99.99\%$$

$$P(\text{diagnosed sick} | \text{sick}) = 99\%$$

$$P(\text{diagnosed sick} | \text{not sick}) = 1\%$$

**BAYES THEOREM  
FORMULA**

# Bayes Theorem: Formula



$A$ : sick

$B$ : diagnosed sick

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

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

$$P(A) = 0.01\%$$

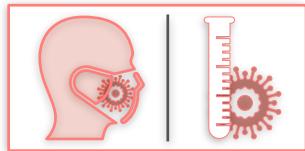
$$P(A') = 99.99\%$$

$$P(B | A) = 99\%$$

$$P(B | A') = 1\%$$

**BAYES THEOREM  
FORMULA**

# Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

$$P(A) = 0.01\%$$

$$P(A') = 99.99\%$$

$$P(B | A) = 99\%$$

$$P(B | A') = 1\%$$

$$P(A | B) = \frac{0.0001 \times 0.99}{(0.0001 \times 0.99) + (0.9999 \times 0.01)}$$

$$P(A | B) = 0.0098$$

**BAYES THEOREM  
FORMULA**



DeepLearning.AI

# Introduction to probability

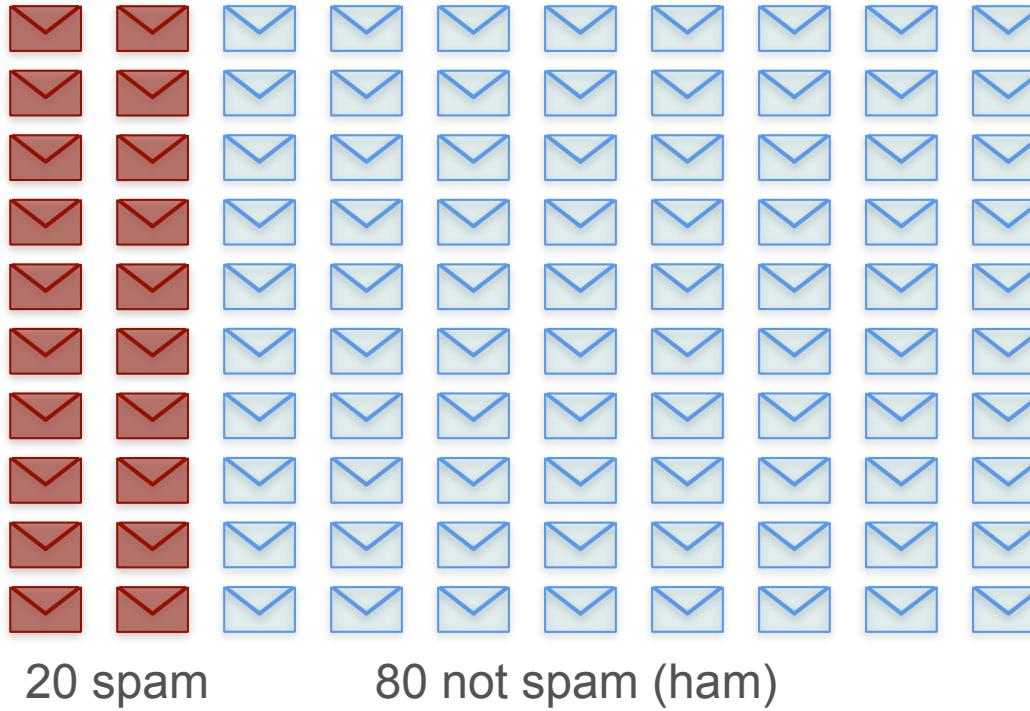
---

## Bayes Theorem - Spam Example

# Bayes Theorem: Spam Example



# Bayes Theorem: Spam Example



“lottery”

# Bayes Theorem: Spam Example

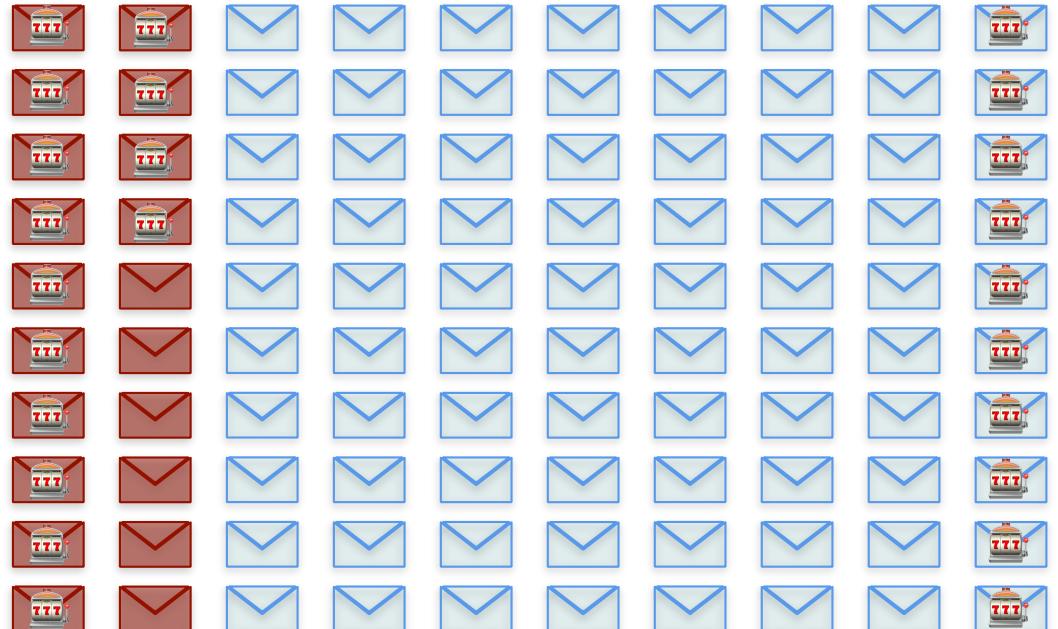
10



What is the probability that an email containing lottery is a spam?

$P(\text{spam} \mid \text{lottery})$

14



20 spam

80 not spam (ham)

# Bayes Theorem: Spam Example (Intuition Solution)



$P(\text{spam} \mid \text{lottery})$

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



14

$$P(\text{spam} \mid \text{lottery}) = \frac{\text{spam and lottery}}{\text{all lottery}}$$

$$= \frac{14}{24}$$

$$= \frac{7}{12} = 0.583$$

10



# Bayes Theorem: Spam Example (Formula Solution)

$P(\text{spam} \mid \text{lottery})$

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(A) \cdot P(B \mid A) + P(A') \cdot P(B \mid A')}$$

$A$ : Email is spam       $B$ : Email contains lottery

$$P(\text{spam} \mid \text{lottery}) = \frac{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam})}{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam}) + P(\text{not spam}) \cdot P(\text{lottery} \mid \text{not spam})}$$

# Bayes Theorem: Spam Example (Formula Solution)

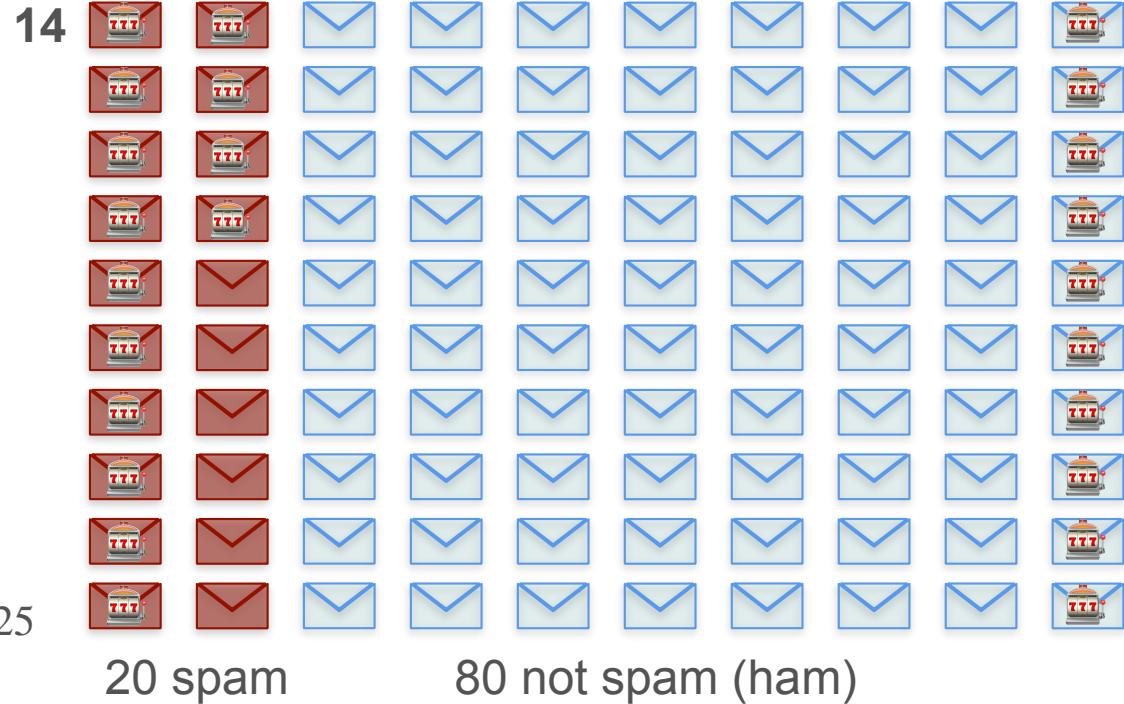
10

$$P(\text{spam}) = \frac{20}{100} = 0.2$$

$$P(\text{not spam}) = \frac{80}{100} = 0.8$$

$$P(\text{lottery} \mid \text{spam}) = \frac{14}{20} = 0.7$$

$$P(\text{lottery} \mid \text{not spam}) = \frac{10}{80} = 0.125$$



# Bayes Theorem: Spam Example (Formula Solution)

$$P(\text{spam}) = 0.2$$

$$P(\text{lottery} \mid \text{spam}) = 0.7$$

$$P(\text{not spam}) = 0.8$$

$$P(\text{lottery} \mid \text{not spam}) = 0.125$$

$$P(\text{spam} \mid \text{lottery}) = \frac{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam})}{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam}) + P(\text{not spam}) \cdot P(\text{lottery} \mid \text{not spam})}$$

$$P(\text{spam} \mid \text{lottery}) = \frac{0.2 \times 0.7}{(0.2 \times 0.7) + (0.8 \times 0.125)} = 0.583$$



DeepLearning.AI

# Introduction to probability

---

**Bayes Theorem -  
Prior and Posterior**

# Bayes Theorem

PRIOR

$$\mathbf{P}(A)$$

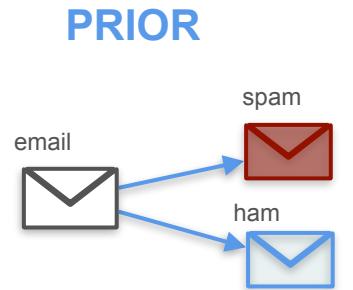
EVENT

$$E$$

POSTERIOR

$$\mathbf{P}(A | E)$$

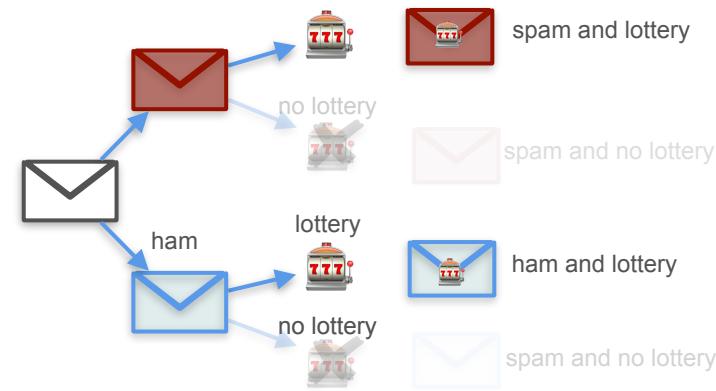
# Prior and Posterior



**EVENT**



Email contains lottery



$$P(\text{spam}) = \frac{\text{spam}}{\text{spam} + \text{ham}}$$

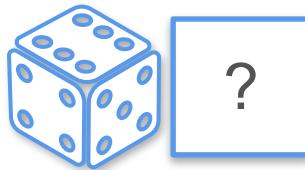
$$P(\text{spam} | \text{lottery}) = \frac{\text{spam and lottery}}{\text{spam and lottery} + \text{ham and lottery}}$$

# Prior and Posterior

PRIOR

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,1	6,2	6,3	6,4	6,5	6,6

EVENT



1st dice is 6

POSTERIOR

	1,1	1,2	1,3	1,4	1,5	1,6
1,1	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,1	6,2	6,3	6,4	6,5	6,6

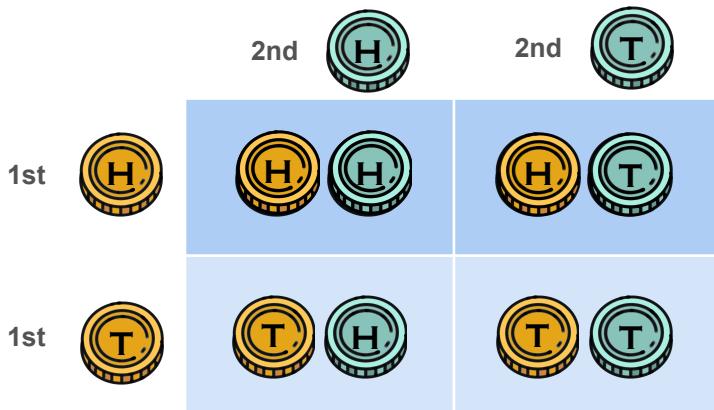
	1,1	1,2	1,3	1,4	1,5	1,6
1,1	1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,1	6,2	6,3	6,4	6,5	6,6

$$P(\text{sum} = 10) = \frac{3}{36}$$

$$P(\text{sum} = 10 | \text{1st is } 6) = \frac{1}{6}$$

# Prior and Posterior

PRIOR



$$P(HH) = \frac{1}{4}$$

EVENT



1st coin is  $H$

POSTERIOR



$$P(HH | \text{1st is } H) = \frac{1}{2}$$



DeepLearning.AI

## Introduction to probability

---

# **Bayes Theorem - The Naive Bayes Model**

# What About 2 Events?

PRIOR

$$P(\text{spam}) = \frac{\text{Red Envelope}}{\text{Red Envelope} + \text{Blue Envelope}}$$

EVENT



Email contains 'lottery'



Email contains 'winning'

POSTERIOR

$$P(\text{spam} | \text{lottery}) = \frac{\text{Red Envelope with lottery icon}}{\text{Red Envelope with lottery icon} + \text{Blue Envelope with lottery icon}}$$

$$P(\text{spam} | \text{winning}) = \frac{\text{Red Envelope with winning icon}}{\text{Red Envelope with winning icon} + \text{Blue Envelope with winning icon}}$$

?



Email contains 'lottery' and 'winning'

# What About 2 Events?

EVENT



Email contains 'lottery' and 'winning'

POSTERIOR



$$P(\text{spam} | \text{lottery} \& \text{winning}) = \frac{\text{# Spam emails with 'lottery' and 'winning'}}{\text{# Spam emails}}$$

$$P(\text{spam} | \text{lottery} \& \text{winning}) = \frac{P(\text{spam}) \cdot P(\text{lottery} \& \text{winning} | \text{spam})}{P(\text{spam}) \cdot P(\text{lottery} \& \text{winning} | \text{spam}) + P(\text{ham}) \cdot P(\text{lottery} \& \text{winning} | \text{ham})}$$

# What About More Than 2 Events?

EVENT

POSTERIOR

Email contains  $w_1, w_2, \dots, w_{100}$

$$\frac{\# \text{ Spam emails with } w_1, \dots, w_{100}}{\# \text{Spam emails}} = \frac{P(\text{spam}) \cdot P(w_1, \dots, w_{100} | \text{spam})}{P(\text{spam}) \cdot P(w_1, \dots, w_{100} | \text{spam}) + P(\text{ham}) \cdot P(w_1, \dots, w_{100} | \text{ham})}$$

The fraction above is highlighted with a red box. A red arrow points from the numerator to the denominator, with a question mark at the end of the arrow.

# Spam emails with  $w_1, \dots, w_{100}$       0  
#Spam emails

P(spam |  $w_1, \dots, w_{100}$ ) =  $\frac{P(\text{spam}) \cdot P(w_1, \dots, w_{100} | \text{spam})}{P(\text{spam}) \cdot P(w_1, \dots, w_{100} | \text{spam}) + P(\text{ham}) \cdot P(w_1, \dots, w_{100} | \text{ham})}$

# Is There a Quicker Way To Estimate the Probability?

## Naive assumption



The appearances of 'lottery' and 'winning' are independent

$$P(\text{spam} \mid \text{lottery \& winning}) = \frac{P(\text{spam}) \cdot P(\text{lottery \& winning} \mid \text{spam})}{P(\text{spam}) \cdot P(\text{lottery \& winning} \mid \text{spam}) + P(\text{ham}) \cdot P(\text{lottery \& winning} \mid \text{ham})}$$

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

↓

The terms  $P(\text{lottery \& winning} \mid \text{spam})$  and  $P(\text{lottery \& winning} \mid \text{ham})$  are circled in red.

# Is There a Quicker Way To Estimate the Probability?

## Naive assumption



The appearances of ‘lottery’ and ‘winning’ are independent

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

$$\mathbf{P}(\text{spam} \mid \text{lottery} \& \text{winning}) = \frac{\mathbf{P}(\text{spam}) \cdot \mathbf{P}(\text{lottery} \mid \text{spam}) \cdot \mathbf{P}(\text{winning} \mid \text{spam})}{\mathbf{P}(\text{spam}) \cdot \mathbf{P}(\text{lottery} \mid \text{spam}) \cdot \mathbf{P}(\text{winning} \mid \text{spam}) + \mathbf{P}(\text{ham}) \cdot \mathbf{P}(\text{lottery} \mid \text{ham}) \cdot \mathbf{P}(\text{winning} \mid \text{ham})}$$

# Is There a Quicker Way To Estimate the Probability?

## Naive assumption

The appearances of the words  $w_1, w_2, \dots, w_n$  are independent

$$P(\text{spam} | w_1, \dots, w_n) = \frac{P(\text{spam}) \cdot P(w_1 | \text{spam}) \cdots P(w_n | \text{spam})}{P(\text{spam}) \cdot P(w_1 | \text{spam}) \cdots P(w_n | \text{spam}) + P(\text{ham}) \cdot P(w_1 | \text{ham}) \cdots P(w_n | \text{ham})}$$

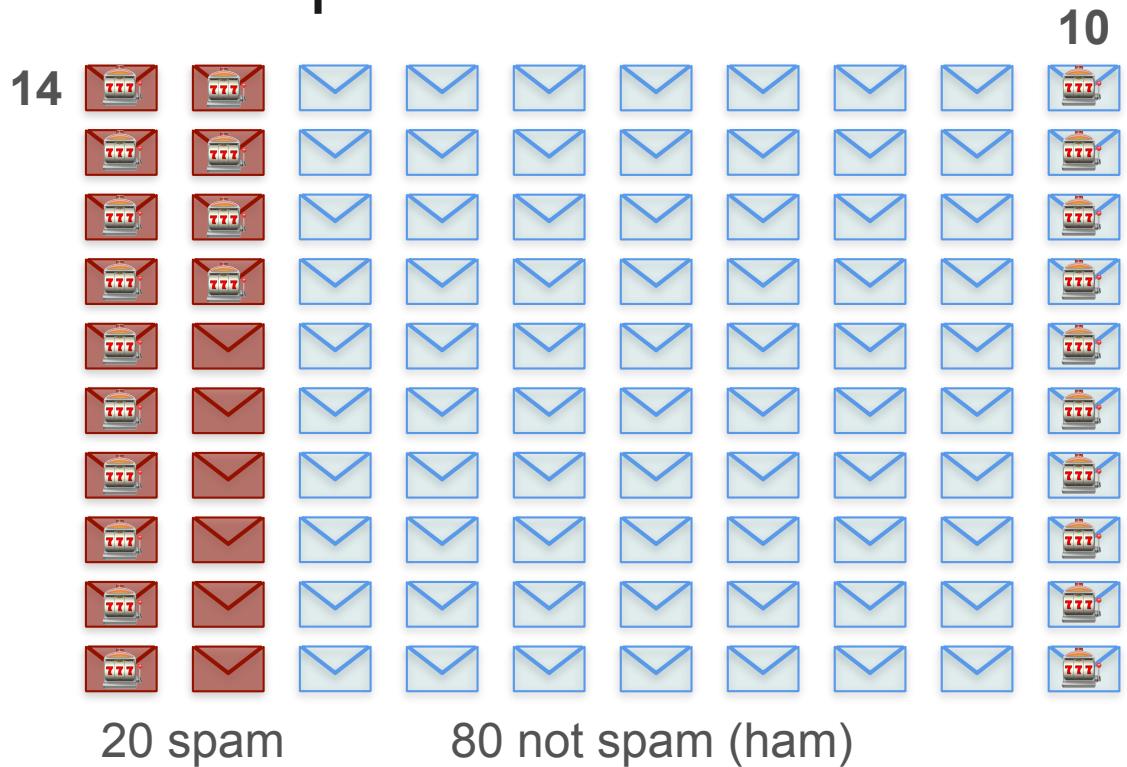
# Naive Bayes: Spam Example

$$P(\text{spam}) = \frac{20}{100} = 0.2$$

$$P(\text{ham}) = \frac{80}{100} = 0.8$$

$$P(\text{lottery} | \text{spam}) = \frac{14}{20} = 0.7$$

$$P(\text{lottery} | \text{ham}) = \frac{10}{80} = 0.125$$



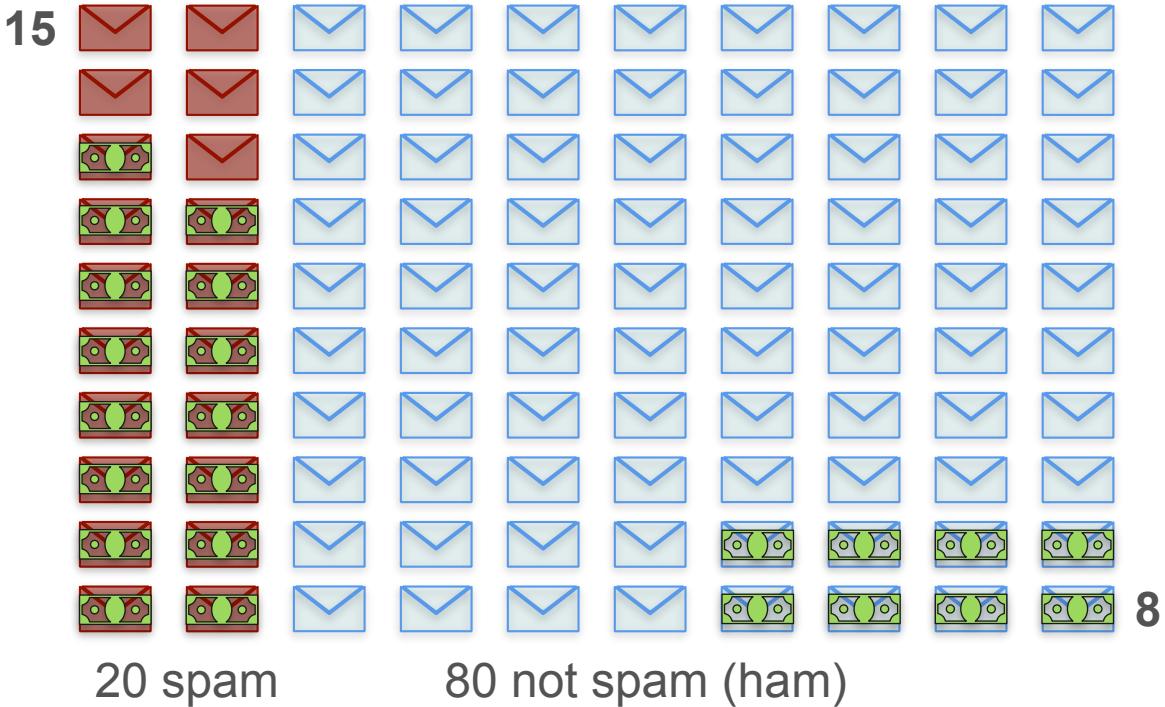
# Naive Bayes: Spam Example

$$P(\text{spam}) = \frac{20}{100} = 0.2$$

$$P(\text{ham}) = \frac{80}{100} = 0.8$$

$$P(\text{winning} \mid \text{spam}) = \frac{15}{20} = 0.75$$

$$P(\text{winning} \mid \text{ham}) = \frac{8}{80} = 0.1$$



# Naive Bayes: Spam Example

$$P(\text{spam}) = 0.2$$

$$P(\text{lottery} \mid \text{spam}) = 0.7$$

$$P(\text{winning} \mid \text{spam}) = 0.75$$

$$P(\text{not spam}) = 0.8$$

$$P(\text{lottery} \mid \text{ham}) = 0.125$$

$$P(\text{winning} \mid \text{ham}) = 0.1$$

$$P(\text{spam} \mid \text{lottery}) = \frac{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam}) \cdot P(\text{winning} \mid \text{spam})}{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam}) \cdot P(\text{winning} \mid \text{spam}) + P(\text{ham}) \cdot P(\text{lottery} \mid \text{ham}) \cdot P(\text{winning} \mid \text{ham})}$$

$$P(\text{spam} \mid \text{lottery} \& \text{winning}) = \frac{0.2 \times 0.7 \times 0.75}{(0.2 \times 0.7 \times 0.75) + (0.8 \times 0.125 \times 0.1)} = 0.913$$



DeepLearning.AI

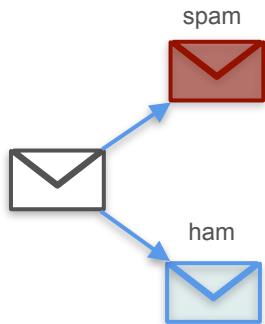
## Introduction to probability

---

# Probability in Machine Learning

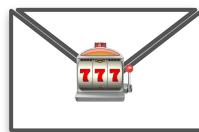
# Bayes Theorem

PRIOR

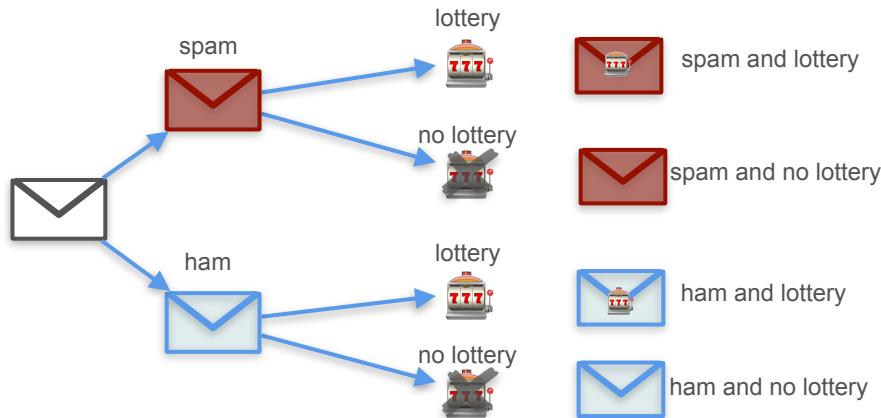


$$P(\text{spam}) = \frac{\text{red envelope}}{\text{red envelope} + \text{blue envelope}}$$

EVENT



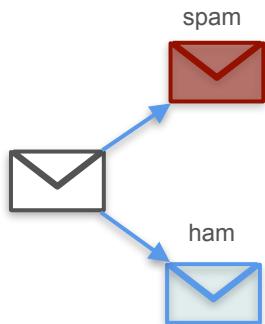
POSTERIOR



$$P(\text{spam} | \text{lottery}) =$$

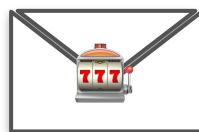
# Bayes Theorem

PRIOR

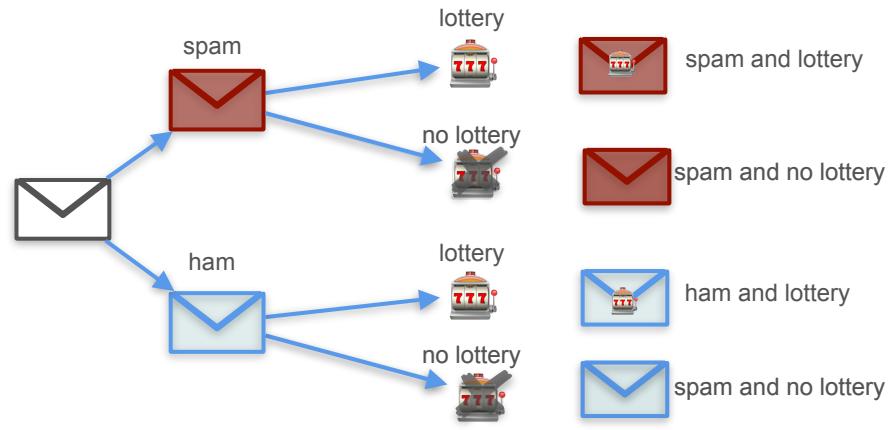


$$P(\text{spam}) = \frac{\text{red envelope}}{\text{red envelope} + \text{blue envelope}}$$

EVENT



POSTERIOR



$$P(\text{spam} | \text{lottery}) = \frac{\text{red envelope and lottery}}{\text{red envelope and lottery} + \text{blue envelope and lottery}}$$

# Example Problem

Image recognition

- What is the probability that there is a cat in the image
- $P(\text{cat} \mid \text{image}) = P(\text{cat} \mid \text{pixel}_1, \text{pixel}_2, \dots, \text{pixel}_n)$



# Example Problem

## Classification

Patient 1		
A	Age	29
G	Gender	Female
H	Height	169 cm
W	Weight	62 kg
S	Smoker	No
...	...	...
B	Heart rate	63
B	Blood pressure	120 90

- Is this patient healthy?
- Calculate  $P(\text{healthy} \mid \text{symptoms and history})$

# Example Problem

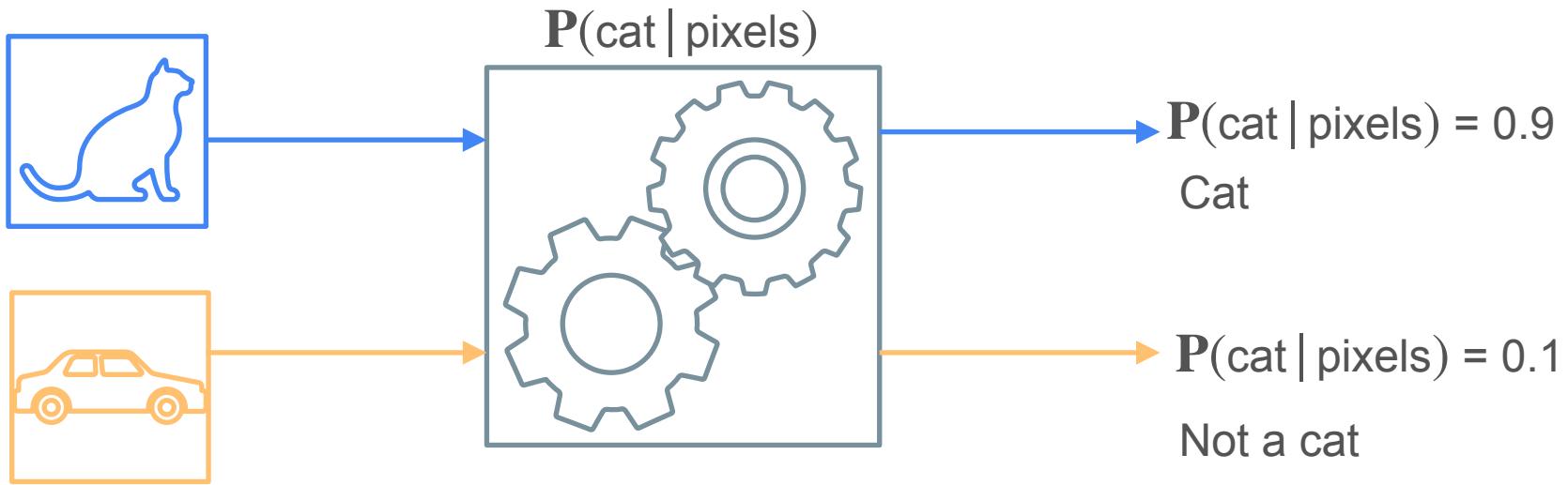
Sentiment analysis

the first cold shower  
even the monkey seems to want  
a little coat of straw

Matsuo Bashō

- Is this a happy sentence?
- Calculate  $P(\text{happy} \mid \text{words in the sentence})$

# Example Solution



# Example Problem: Generative Models

Face generation

- Generate a group of pixels such that the resulting image looks like a human face.
- Goal: generate images such that  $P(\text{face} \mid \text{pixels})$  is high.



Image generated by a StyleGAN

# W1 Lesson 2

# Probability Distributions



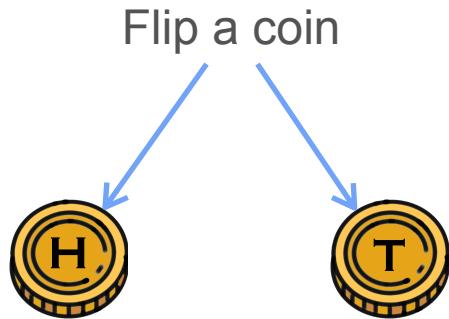
DeepLearning.AI

# Probability Distributions

---

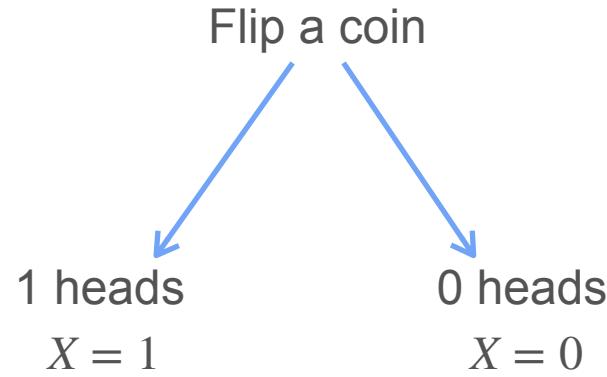
## Random Variables

# From Events to Random Variables



$$\mathbf{P}(H) = 0.5 \quad \mathbf{P}(T) = 0.5$$

$X$  = Number of heads



$$\mathbf{P}(X = 1) = 0.5 \quad \mathbf{P}(X = 0) = 0.5$$

$X$  is a random variable

# From Events to Random Variables

$X$  = Number of heads in 10 coin tosses

$$\mathbf{P}(H) = 0.5$$



$$X = 10$$

$$0.5^{10}$$

$$\mathbf{P}(X = 0)?$$



$$X = 9$$

$$0.5^9 0.5$$

$$\mathbf{P}(X = 1)?$$

...



$$X = 9$$

$$0.5^9 0.5$$

$$\mathbf{P}(X = 10)?$$

# From Events to Random Variables

$X$  = Number of heads in 10 coin tosses



Possible outcomes:

0	4	8
1	5	9
2	6	10
3	7	

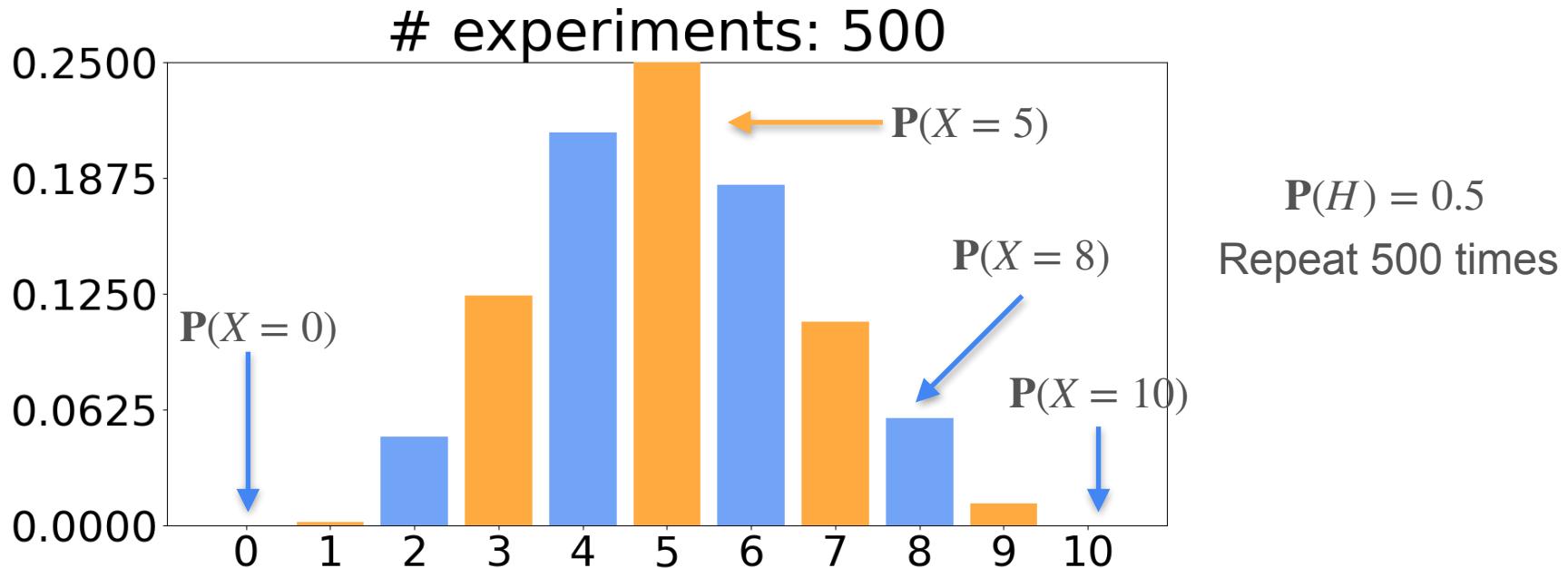


⋮

$$P(H) = 0.5$$

Repeat 500 times

# Flipping a Fair Coin 500 Times



# Why Random Variables?

- Random variables allow you to model the whole experiment at once



$X$  = Number of heads



$X$  = Number of sick patients



$X$  = Number of 1's

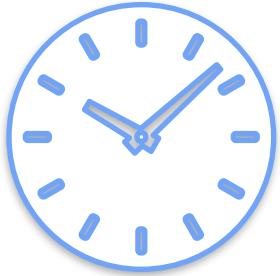


$$\mathbf{P}(X = 1) = 0.5$$

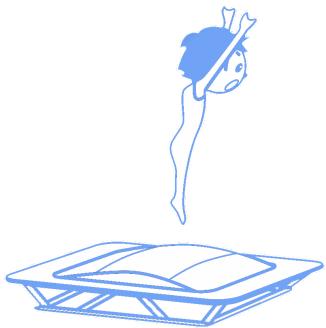
$$\mathbf{P}(X = -7) = 0.2$$

$$\mathbf{P}(X = 3.14159) = 0.3$$

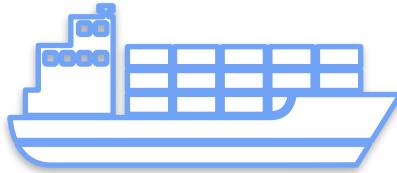
# Other Random Variables



Wait time until the  
next bus arrives



Height of an  
gymnast's jump



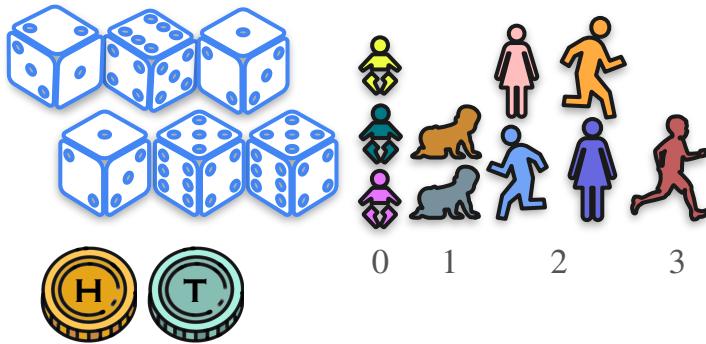
Number of  
defective products  
in a shipment



mm. of rain in  
November

# Discrete and Continuous Random Variables

## Discrete random variables



~~Finite number of values~~

(Could be infinite too)

Can take only a **countable** number of values

## Continuous random variables



Infinite number of values

Takes values on an interval

# Random Variable Vs. Deterministic Variable

## Deterministic

$$x = 2, f(x) = x^2$$

Fixed outcome

## Random

$X$  = number of defective items in  
a shipment

Uncertain outcome



DeepLearning.AI

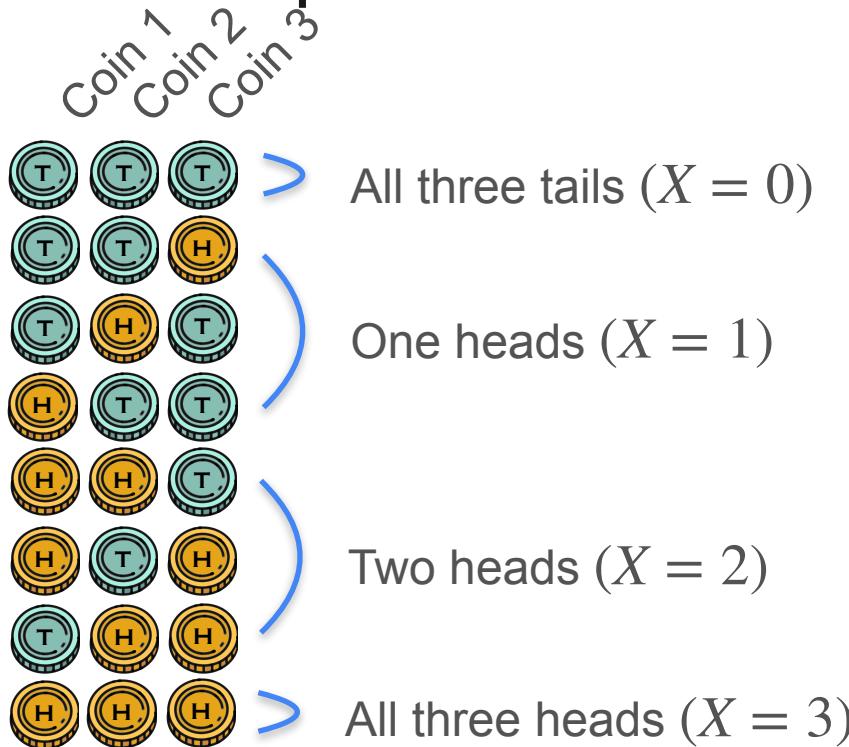
# Probability Distributions

---

## Probability Distributions (Discrete)

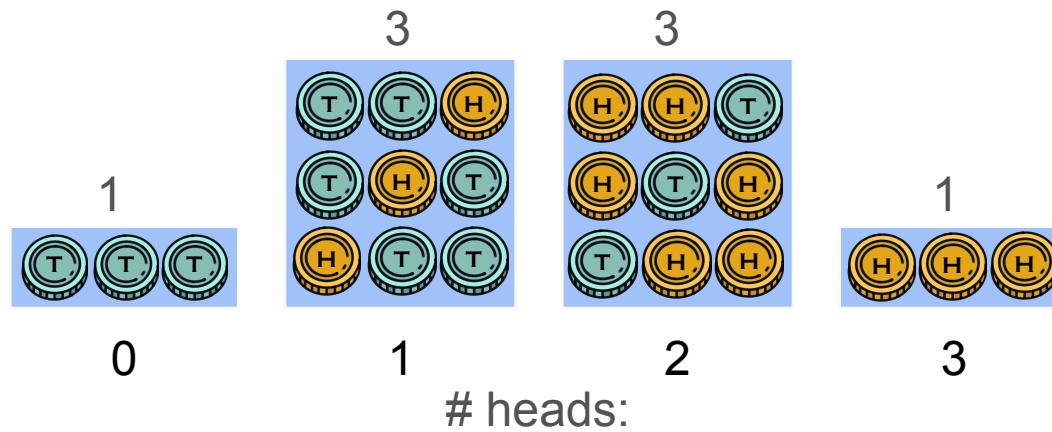
# Discrete Distributions: Flip Three Coins

$X_1$ : number of heads in 3 coin tosses



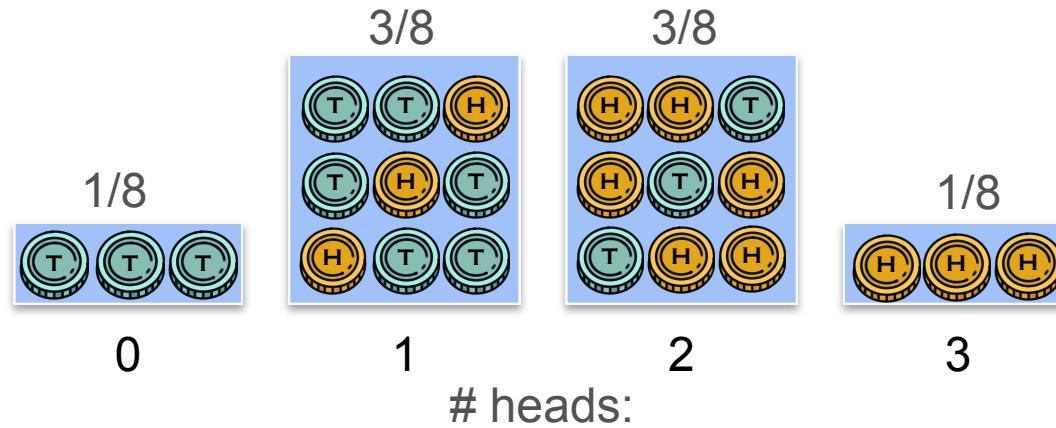
# Discrete Distributions: Flip Three Coins

$X_1$ : number of heads in 3 coin tosses



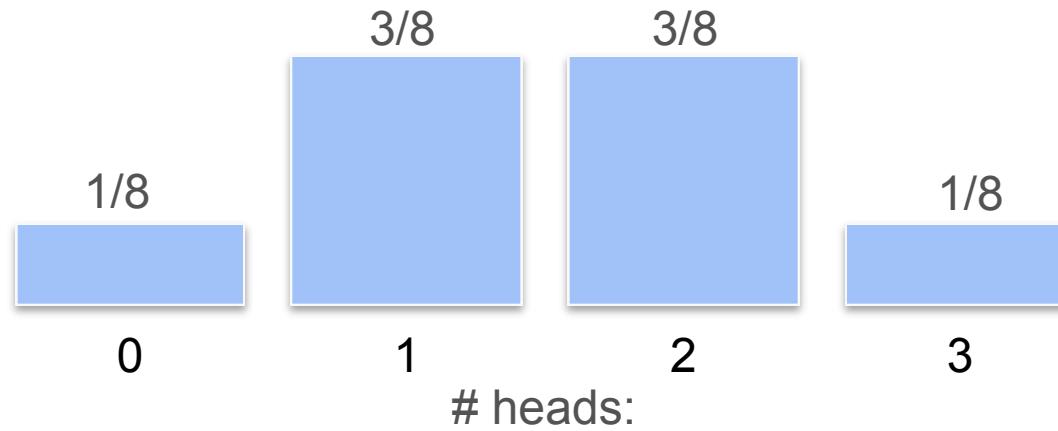
# Discrete Distributions: Flip Three Coins

$X_1$ : number of heads in 3 coin tosses



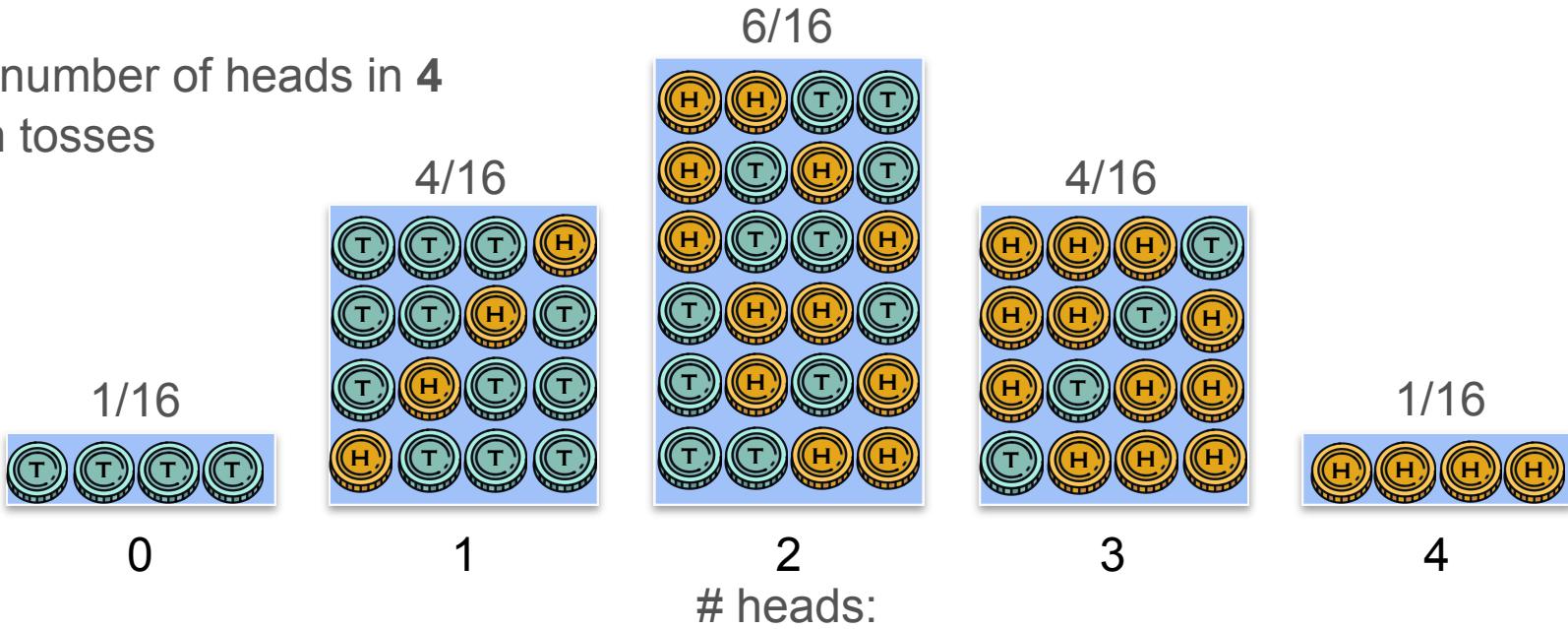
# Discrete Distributions: Flip Three Coins

$X_1$ : number of heads in 3 coin tosses



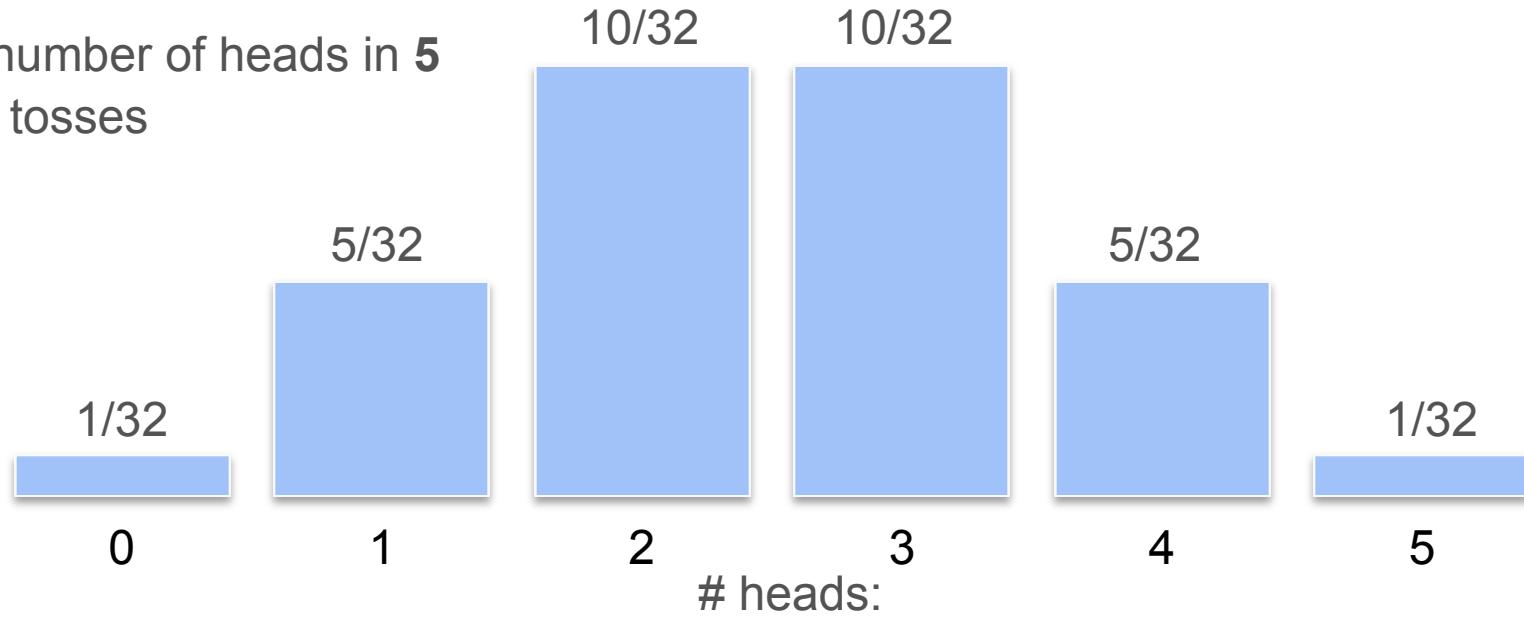
# Discrete Distributions: Flip Four Coins

$X_2$ : number of heads in 4 coin tosses



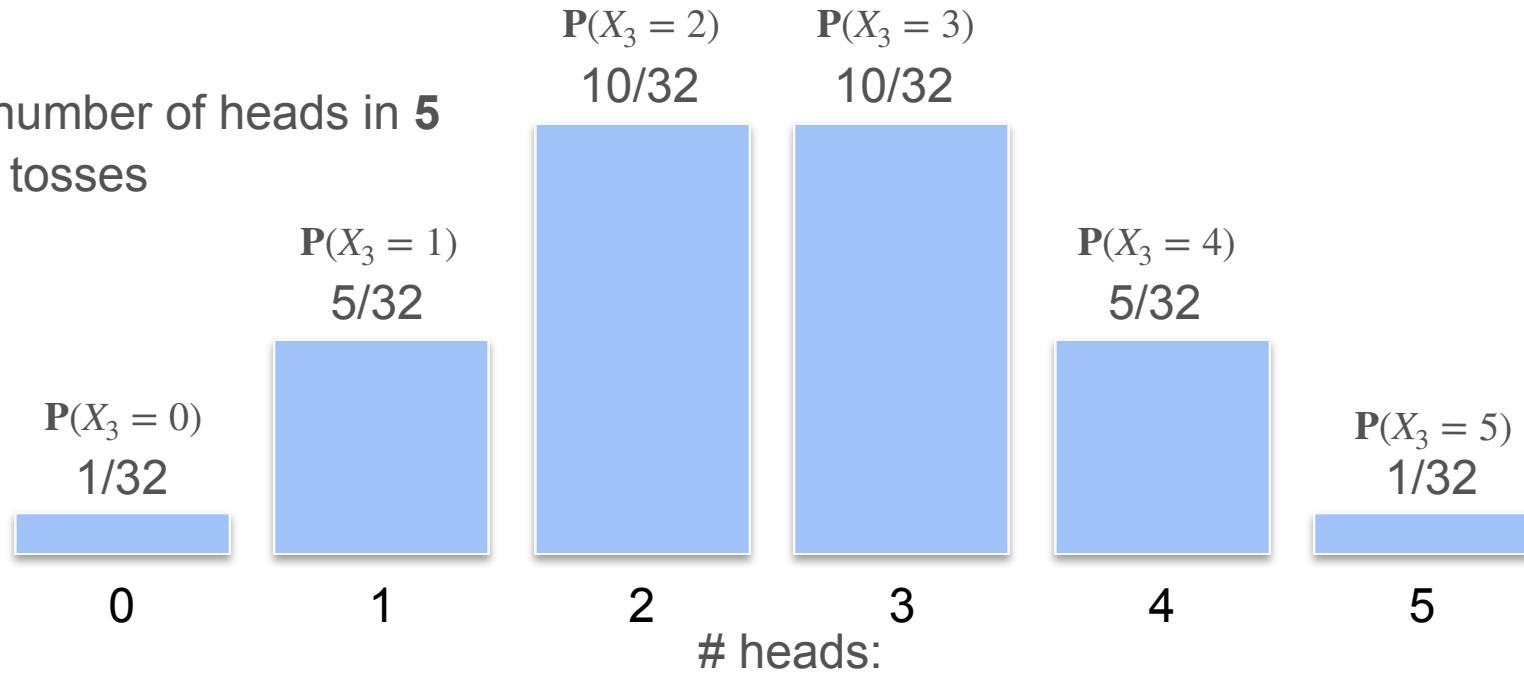
# Discrete Distributions: Flip Five Coins

$X_3$ : number of heads in 5 coin tosses



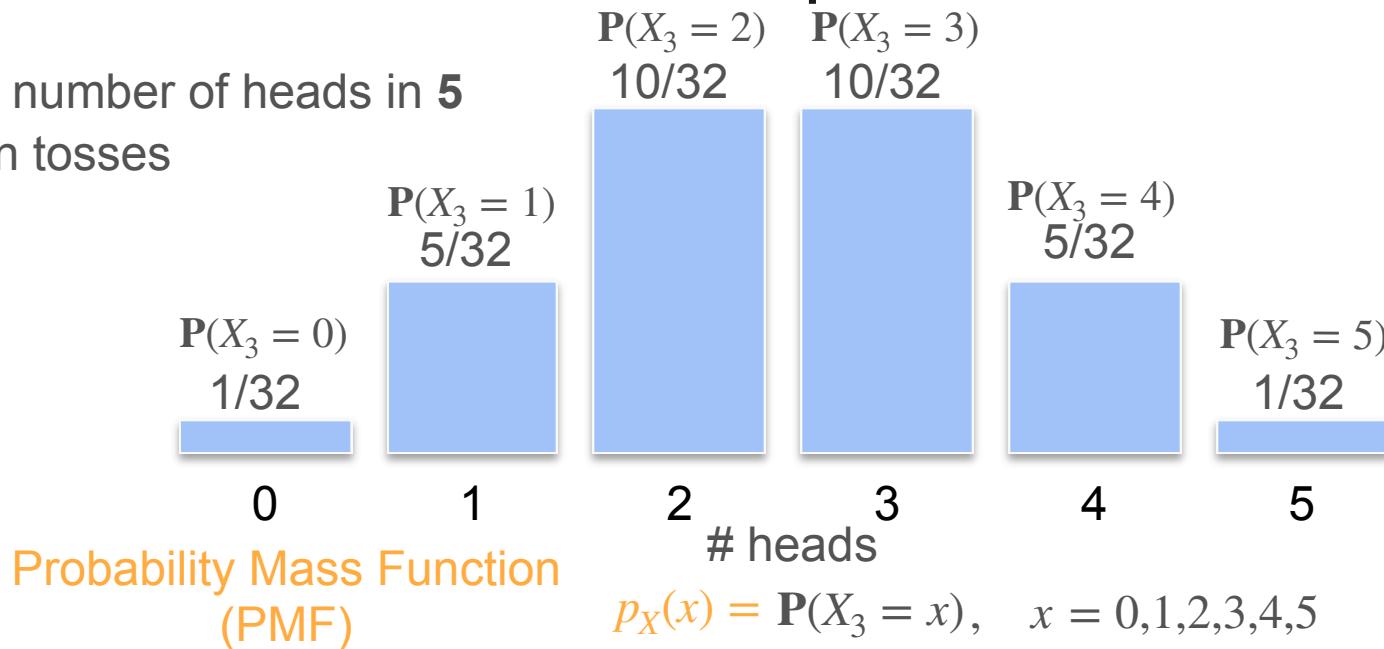
# Discrete Distributions: Flip Five Coins

$X_3$ : number of heads in 5 coin tosses



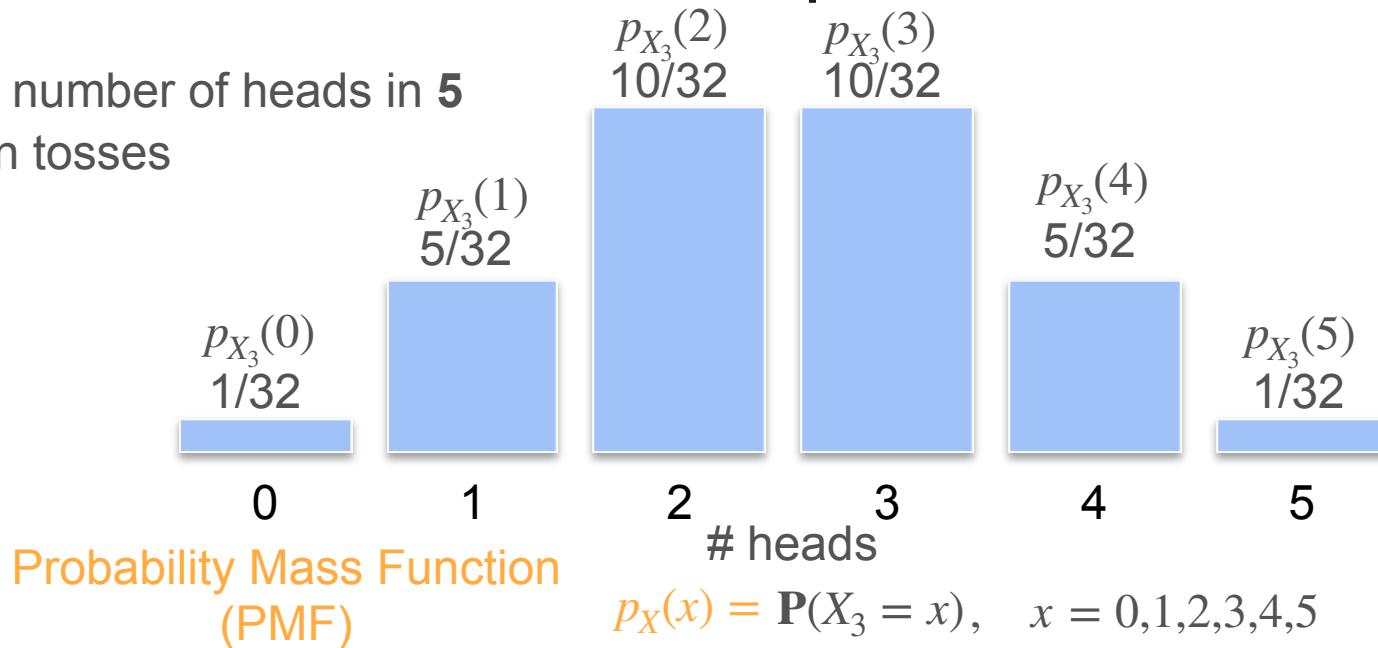
# Discrete Distributions: Flip Five Coins

$X_3$ : number of heads in 5 coin tosses



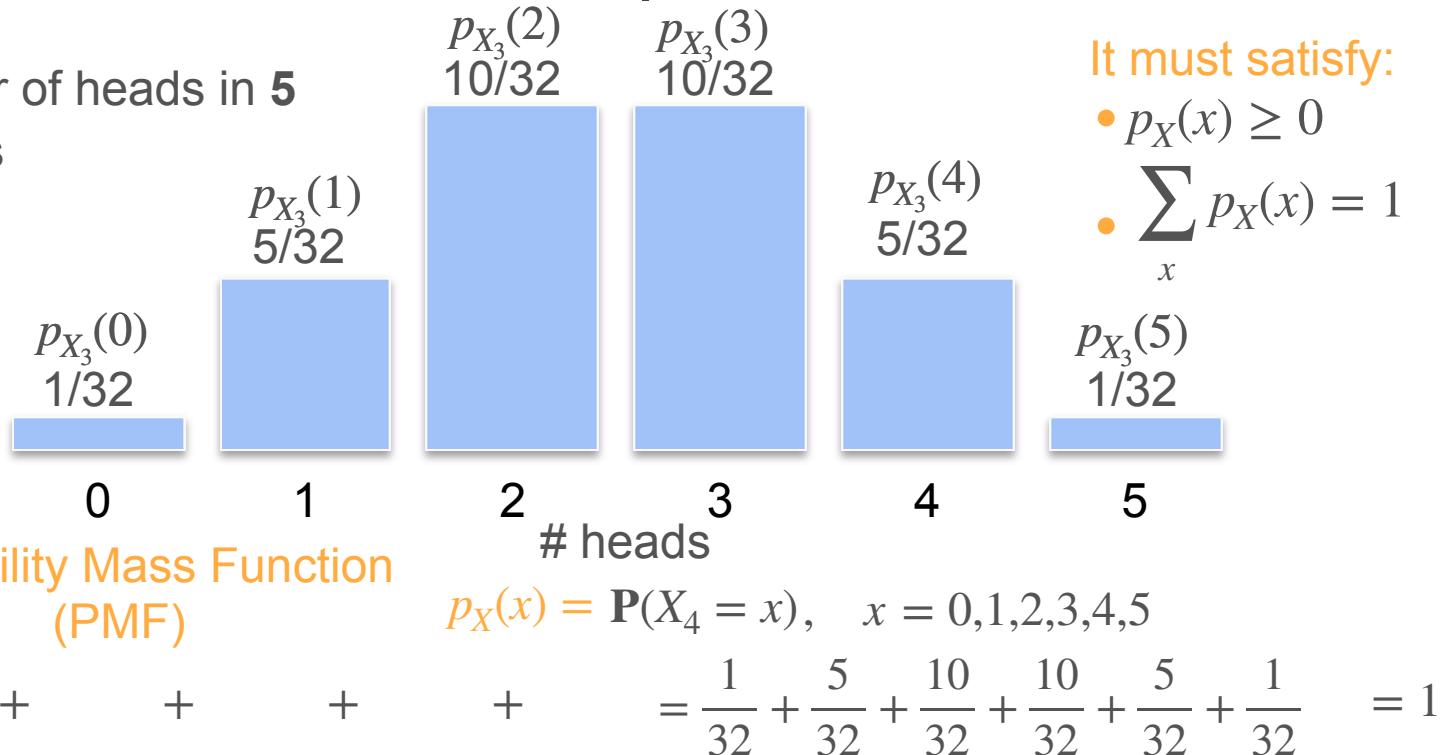
# Discrete Distributions: Flip Five Coins

$X_3$ : number of heads in 5 coin tosses



# Discrete Distributions: Flip Five Coins

$X_3$ : number of heads in 5 coin tosses



# Can You See a Pattern?

$X_1, X_2, X_3, X_4$  are very similar

They all represent **number of heads in  $n$  experiments**

The way the probability distributes along the possible outcomes seems to have a similar pattern

Could there be a **single model** to represent all this variables?



Binomial distribution



DeepLearning.AI

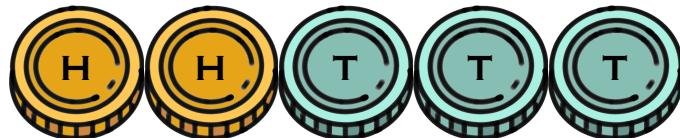
# Probability Distributions

---

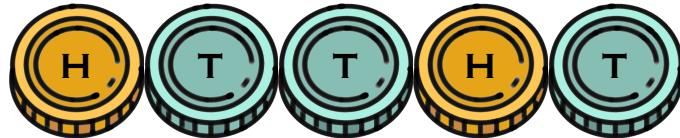
## Binomial Distribution

# Binomial Distribution: Example

What is the probability that if I flip 5 coins, 2 of them land in heads?



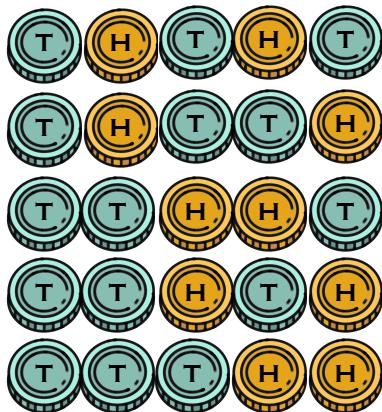
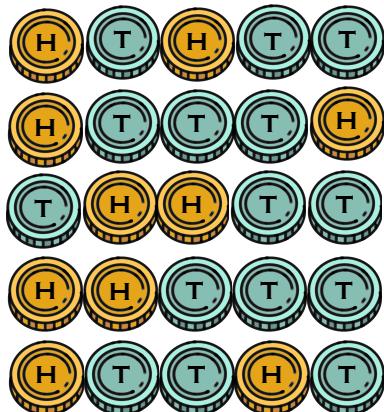
$$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 = \frac{1}{32}$$



$$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 = \frac{1}{32}$$

# Binomial Distribution: Example

What is the probability that if I flip 5 coins, 2 of them land in heads?



$$10 = \frac{5!}{2!}$$

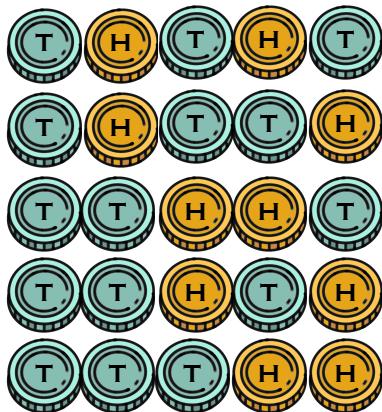
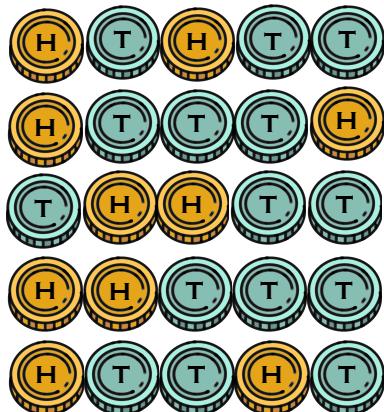
Number of ways you can order 5 coins

Number of H



# Binomial Distribution: Example

What is the probability that if I flip 5 coins, 2 of them land in heads?



$$10 = \frac{5!}{2!(5-2)!}$$

Number of ways you can order 5 coins

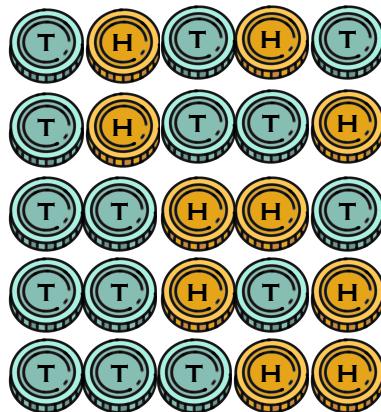
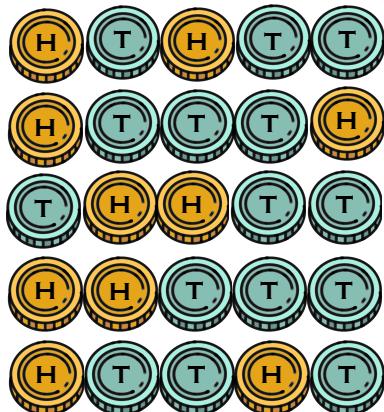
Number of H      Number of T

The equation  $10 = \frac{5!}{2!(5-2)!}$  represents the number of ways to order 5 coins, where 5! is the total number of permutations of 5 items, and 2!(5-2)! is the number of permutations of 2 heads and 3 tails. Blue arrows point from the labels "Number of ways you can order 5 coins", "Number of H", and "Number of T" to the corresponding parts of the equation.



# Binomial Distribution: Example

What is the probability that if I flip 5 coins, 2 of them land in heads?



$$10 = \frac{5!}{2!(5-2)!} = \binom{5}{2}$$

Binomial coefficient

Number of ways you can get 2 heads in 5 coin tosses

# Binomial Distribution: Binomial Coefficient

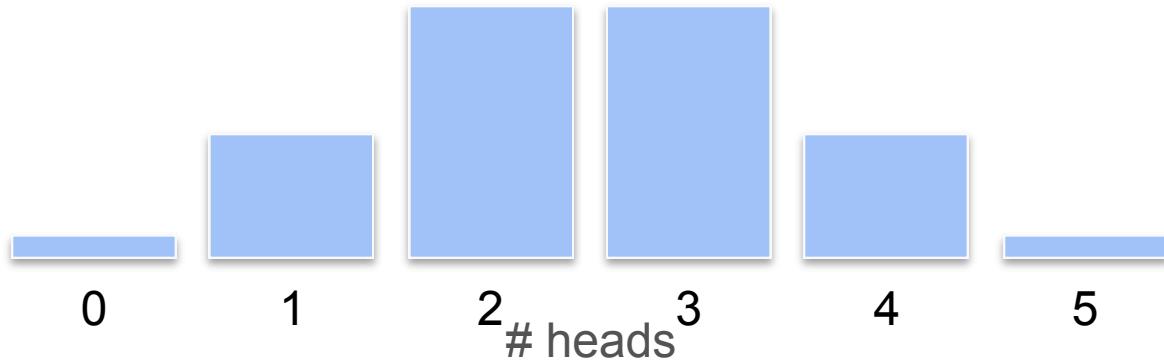
In general:

$\binom{n}{k}$  counts all the combinations for landing  $k$  heads in  $n$  coin tosses

Property:

The PMF with a fair coin is symmetrical

$$\binom{n}{k} = \binom{n}{n-k}$$



# Binomial Distribution: Binomial Coefficient

General PMF for  $X$  : number of heads in 5 coin tosses?

Your coin has  $\mathbf{P}(H) = p$

Event:  $X = x$ :  $x$  heads in 5 tosses

$$\binom{5}{x} p^x (1 - p)^{5-x}$$

All the possible orders

Probability of seeing  $x$  heads

Probability of seeing  $5 - x$  tails

# Binomial Distribution

General PMF for  $X$  : number of heads in 5 coin tosses?

Your coin has  $\mathbf{P}(H) = p$

Event:  $X = x$ :  $x$  heads in 5 tosses

$$p_X(x) = \binom{5}{x} p^x (1-p)^{5-x}, \quad x = 0,1,2,3,4,5$$

**$X$  follows a binomial distribution**

$X \sim \text{Binomial}(5, p)$

Number of flips       $\mathbf{P}(H)$

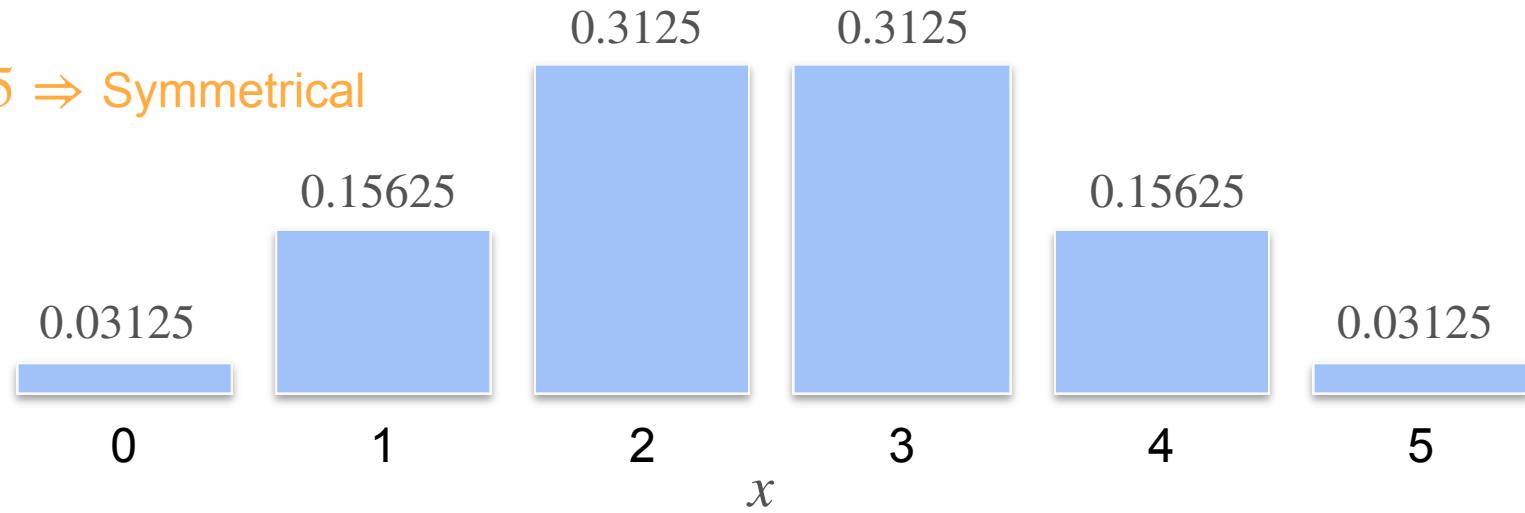
The diagram illustrates the connection between the mathematical definition of a binomial distribution and its statistical representation. On the left, the probability mass function  $p_X(x) = \binom{5}{x} p^x (1-p)^{5-x}$  is shown, with a green bracket grouping the terms. A blue curved arrow points from this formula to the distribution on the right. Below the formula, the text "X follows a binomial distribution" is enclosed in a green box. To the right of the distribution, two blue boxes labeled "Number of flips" and "P(H)" have arrows pointing to the parameters 5 and  $p$  respectively, indicating their roles in defining the binomial distribution.

# Binomial Distribution

$$\begin{array}{|l|} \hline n = 5 \\ p = 0.5 \\ \hline \end{array}$$

$$p_X(x) = \mathbf{P}(X = x) = \binom{5}{k} 0.5^k 0.5^{5-k}$$

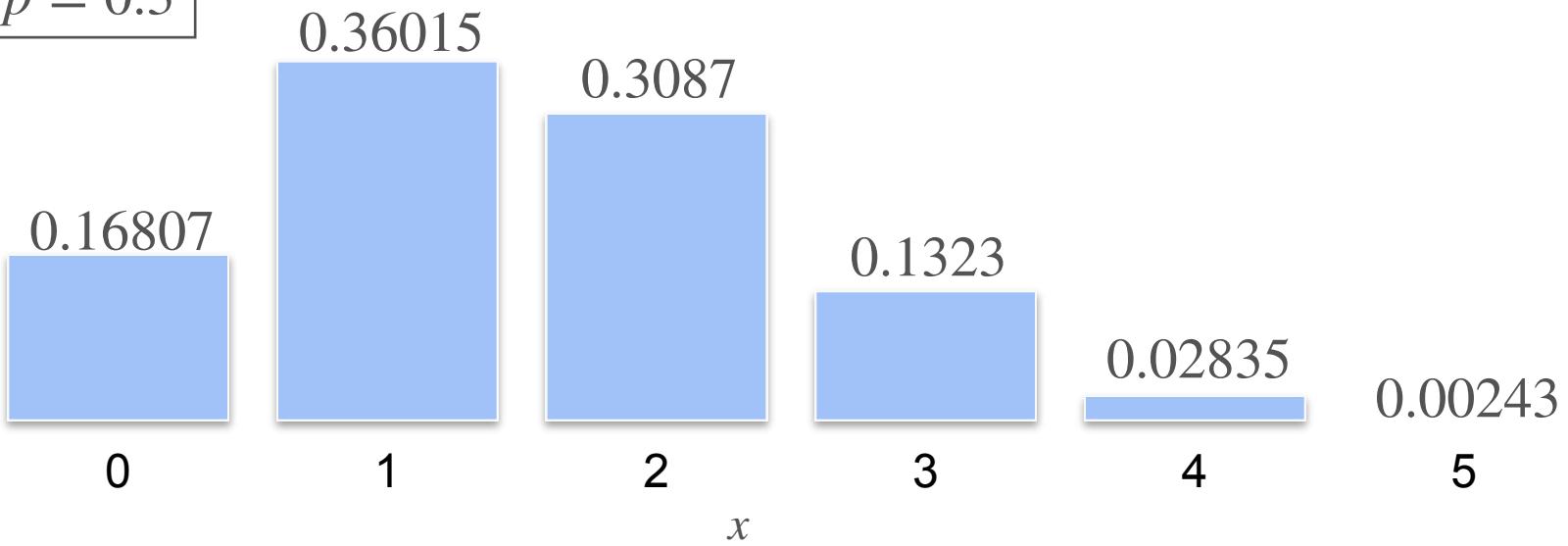
$p = 0.5 \Rightarrow$  Symmetrical



# Binomial Distribution

$$\begin{aligned} n &= 5 \\ p &= 0.3 \end{aligned}$$

$$p_X(x) = \mathbf{P}(X = x) = \binom{5}{k} 0.3^k 0.7^{5-k}$$



# Binomial Distribution

General PMF for  $X$  : number of heads in 5 coin tosses?

Your coin has  $\mathbf{P}(H) = p$

Event:  $X = x$ :  $x$  heads in 5 tosses

$$p_X(x) = \binom{5}{x} p^x (1-p)^{5-x}, \quad x = 0,1,2,3,4,5$$

$$X \sim \text{Binomial}(5,p)$$

# Binomial Distribution

General PMF for  $X$  : number of heads in  $n$  coin tosses?

Your coin has  $\mathbf{P}(H) = p$

Event:  $X = x$ :  $x$  heads in  $n$  tosses

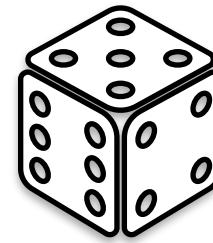
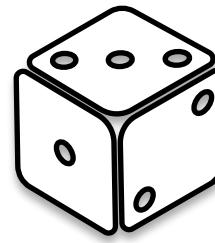
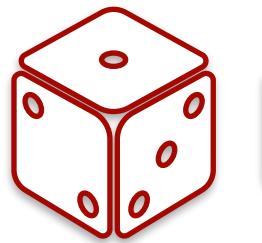
$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

$X \sim \text{Binomial}(n, p)$

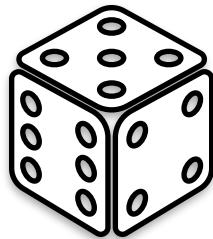
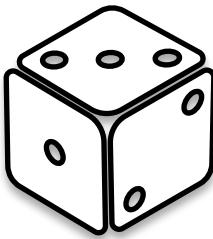
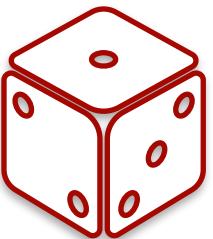
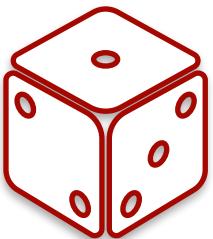
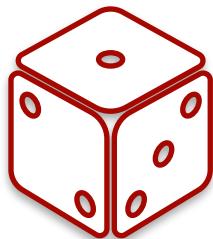
$n$  and  $p$  are called the **parameters** of the binomial distribution

# Binomial Distribution: Quiz

What is the probability of getting three ones when rolling a dice five times (no matter on which dice)?



# Binomial Distribution: Dice Is a Biased Coin!



one  
 $p = \frac{1}{6}$



heads  
 $p = \frac{1}{6}$

one  
 $p = \frac{1}{6}$



heads  
 $p = \frac{1}{6}$

one  
 $p = \frac{1}{6}$



heads  
 $p = \frac{1}{6}$

not one  
 $p = \frac{5}{6}$



not heads  
 $p = \frac{5}{6}$

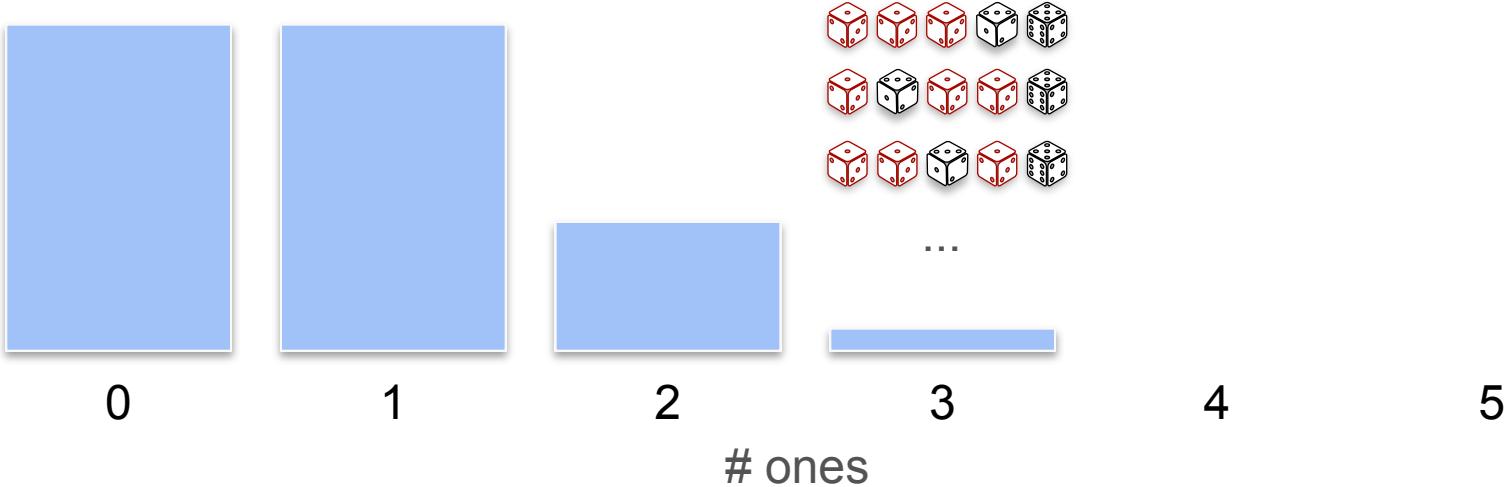
not one  
 $p = \frac{5}{6}$



not heads  
 $p = \frac{5}{6}$

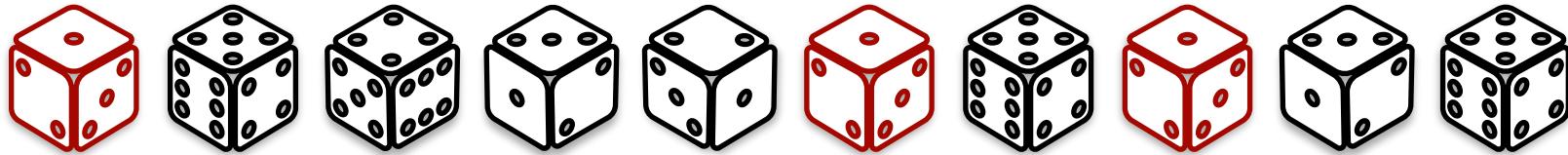
# Binomial Distribution: Dice Is a Biased Coin!

$n = 5$   
 $p = 0.1666$



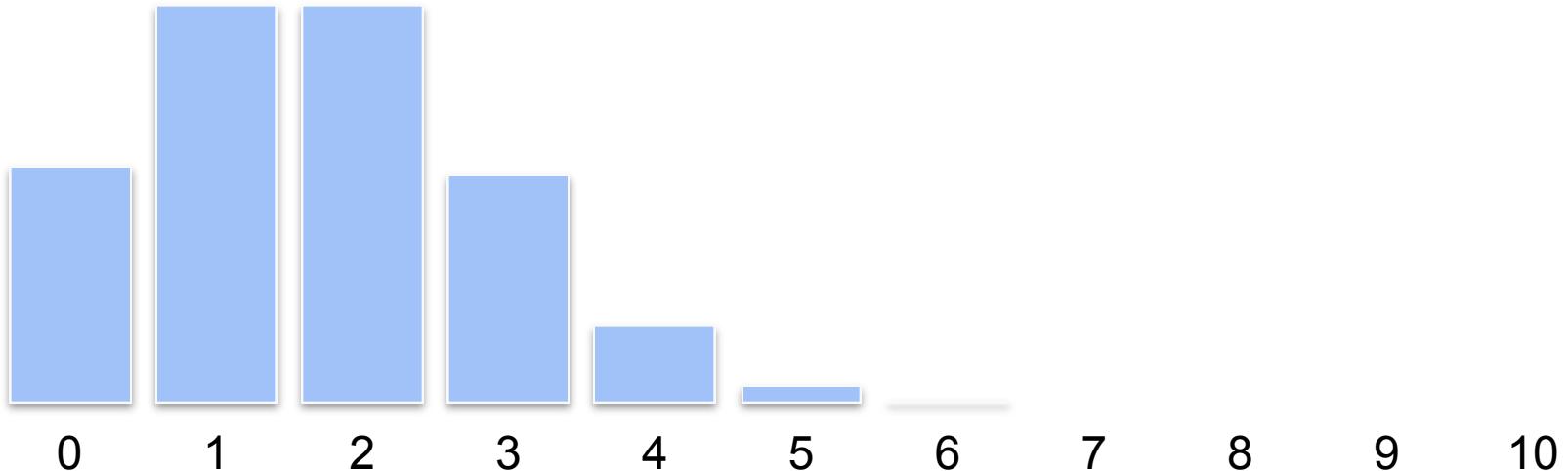
# Binomial Distribution: Quiz

- Quiz: What are the parameters for the following binomial distribution:
  - I roll 10 dice
  - I want to record the number of times I obtain the number 1



# Binomial Distribution: Quiz

$$\begin{aligned}n &= 10 \\p &= 0.1666\end{aligned}$$





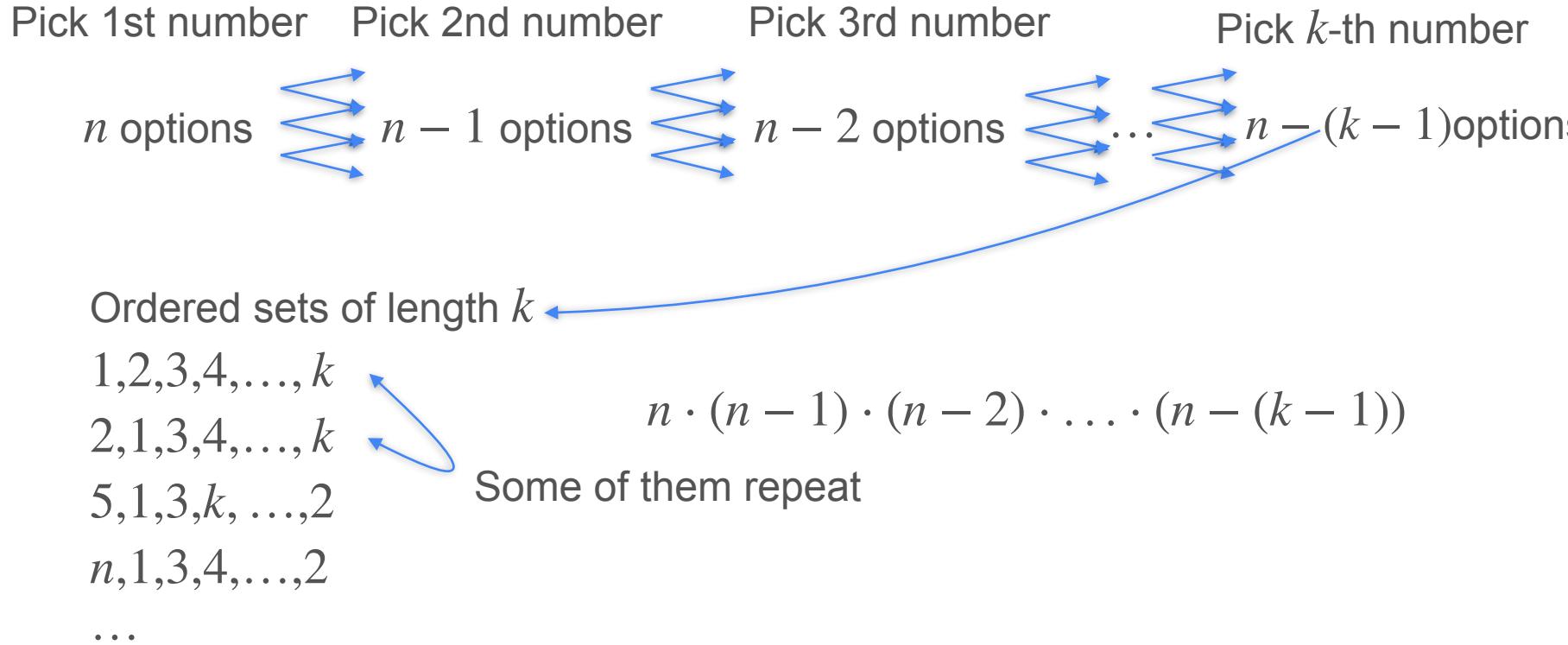
DeepLearning.AI

# Probability Distributions

---

**(Optional)**  
**Binomial Coefficient**

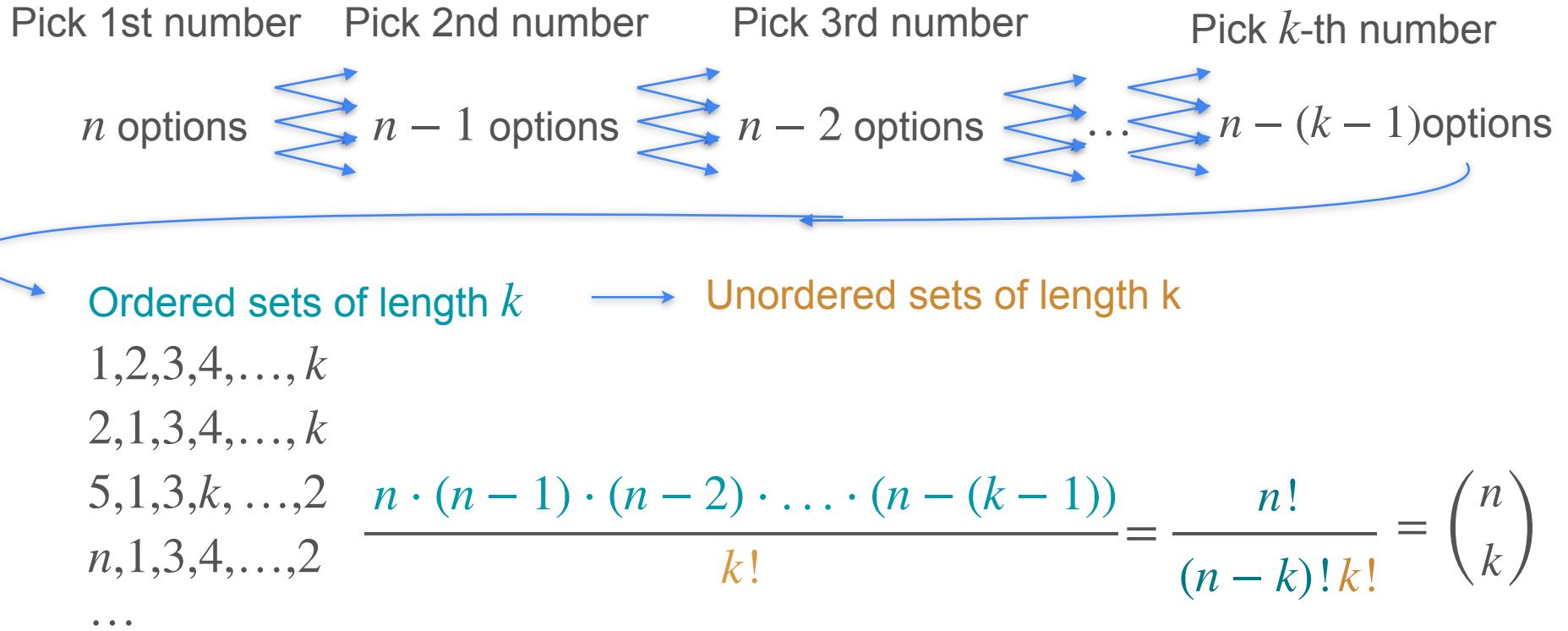
# Binomial Coefficient



# Binomial Coefficient

	Pick 1st	Pick 2nd	Pick 3rd	Pick 4th
1,2,3,4	4 options	3 options	2 options	1 option
1,2,4,3				
1,3,2,4		$4 \cdot 3 \cdot 2 \cdot 1 = 4!$		
1,3,4,2		For five numbers:		
		$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$		
...				
4,3,2,1		General solution:	$k!$	

# Binomial Coefficient

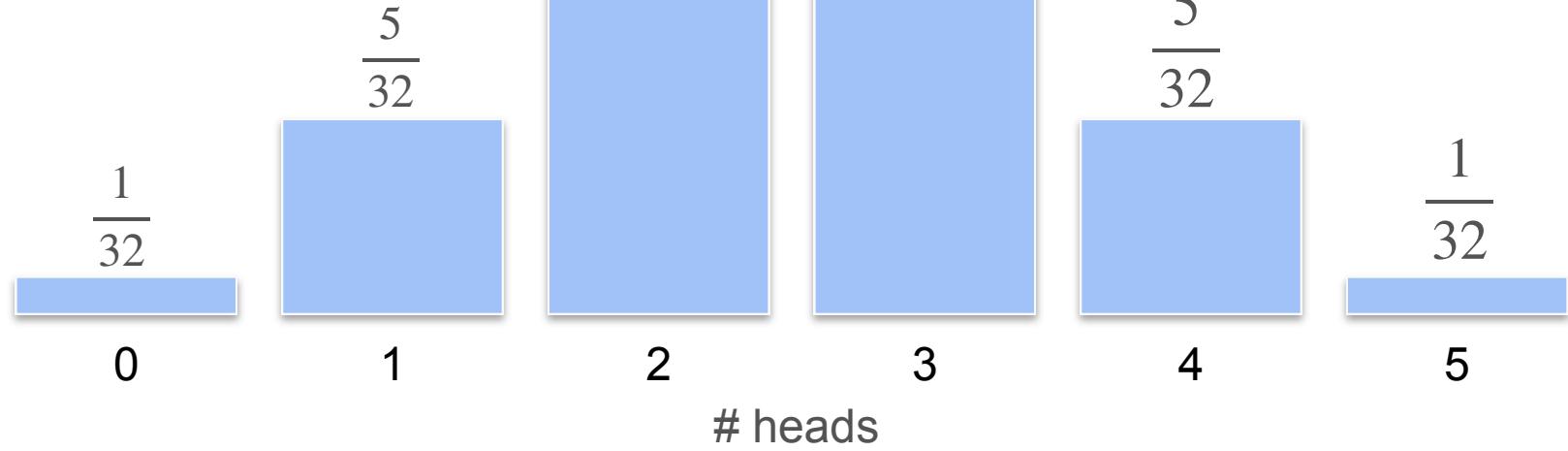


# Binomial Distribution: Fair Coins

$$\binom{5}{0} = \frac{5!}{0!5!} = 1$$

$$\frac{10}{32}$$

$$\frac{10}{32}$$



# Binomial Distribution: Fair Coins



50 %

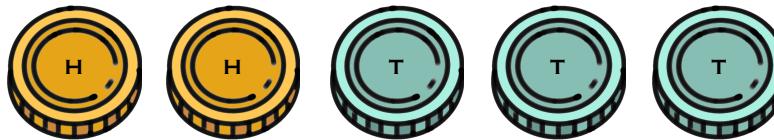


50 %



$$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5$$

$$= \frac{1}{32}$$



$$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5$$

$$= \frac{1}{32}$$

# Binomial Distribution: Biased Coins



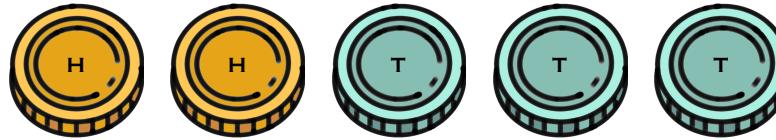
30 %



70 %



$$0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 = 0.00243$$



$$0.3 \cdot 0.3 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.01323$$

# Binomial Distribution: Biased Coins


$$0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 = 0.3^5 \cdot 0.7^0$$


$$0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.7 = 0.3^4 \cdot 0.7^1$$


$$0.3 \cdot 0.3 \cdot 0.3 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.3^3 \cdot 0.7^2$$


$$0.3 \cdot 0.3 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.3^2 \cdot 0.7^3$$


$$0.3 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.3^1 \cdot 0.7^4$$


$$0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.3^0 \cdot 0.7^5$$

$$= 0.3^k \cdot 0.7^{n-k}$$

# Binomial Distribution: Biased Coins

	$0.3^5 \cdot 0.7^0$
	$0.3^4 \cdot 0.7^1$
	$0.3^3 \cdot 0.7^2$
	$0.3^2 \cdot 0.7^3$
	$0.3^1 \cdot 0.7^4$
	$0.3^0 \cdot 0.7^5$

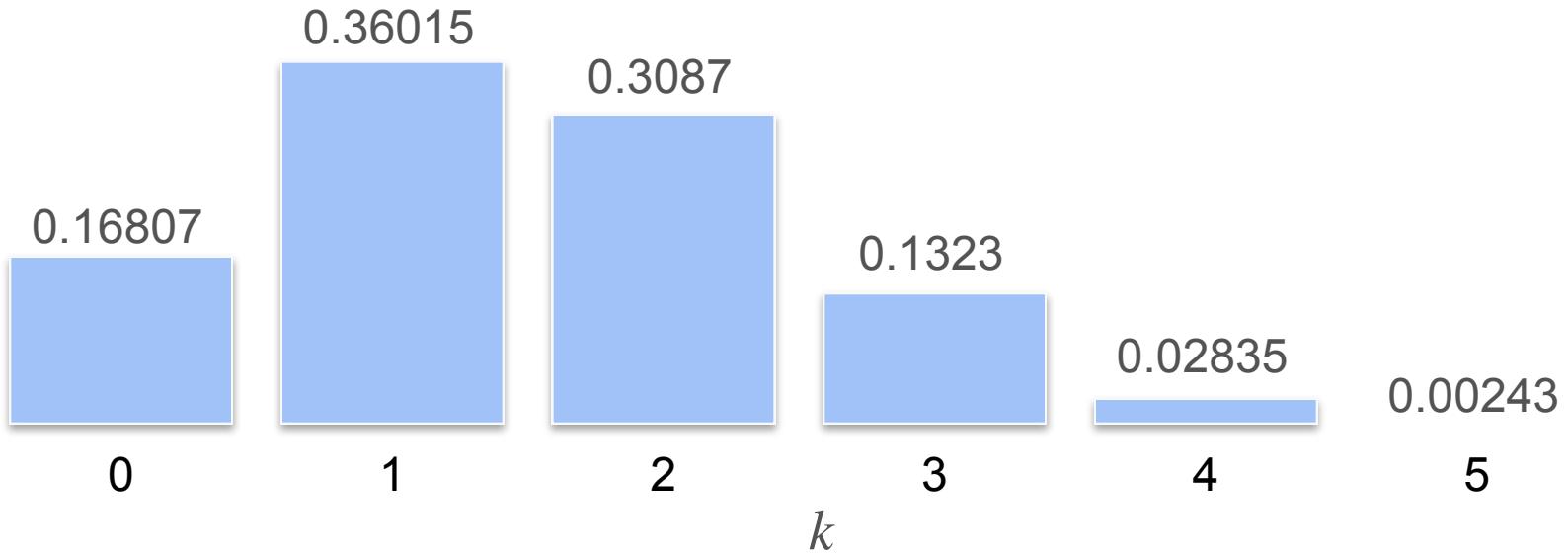
$$= 0.3^k \cdot 0.7^{n-k} \rightarrow \binom{n}{k} 0.3^k \cdot 0.7^{n-k}$$

Account for all possible orders of heads and tails

# Binomial Distribution: Biased Coins

$$= \binom{n}{k} 0.3^k \cdot 0.7^{n-k}$$

n = 5  
p = 0.3





DeepLearning.AI

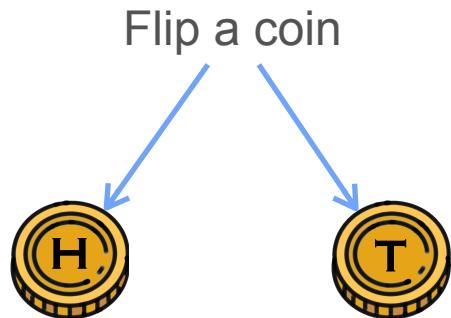
# Probability Distributions

---

## Bernoulli Distribution

# Bernoulli Distribution

$X$  = Number of heads



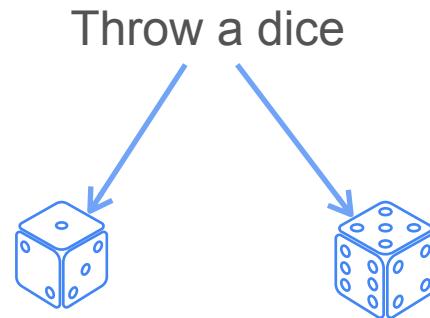
$$P(X = 1) = 0.5$$

Success

$$P(X = 0) = 0.5$$

Failure

$X$  = Number of 1's



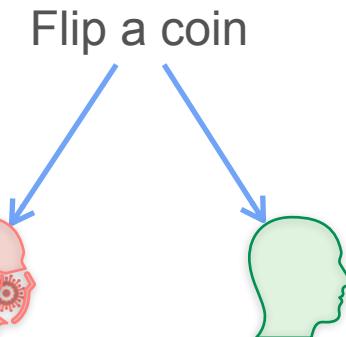
$$P(X = 1) = \frac{1}{6}$$

Success

$$P(X = 0) = \frac{5}{6}$$

Failure

$X$  = Number of sick patients



$$P(X = 1) = p$$

Success

$$P(X = 0) = 1 - p$$

Failure

$$X \sim \text{Bernoulli}(p)$$

$p$  is the parameter of the Bernoulli distribution



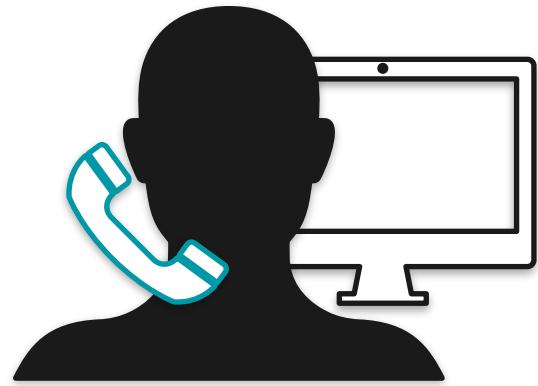
DeepLearning.AI

# Probability Distributions

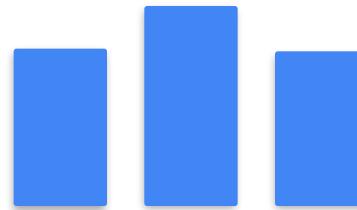
---

## Probability Distributions (Continuous)

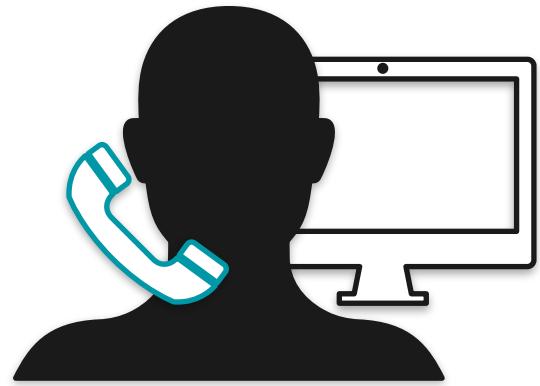
# From Discrete to Continuous Distributions



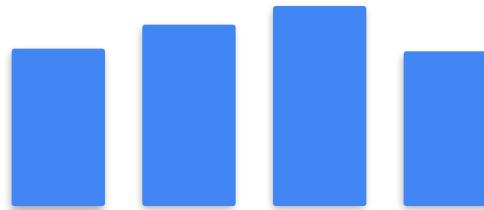
Waiting time: 1 2 3 (min)



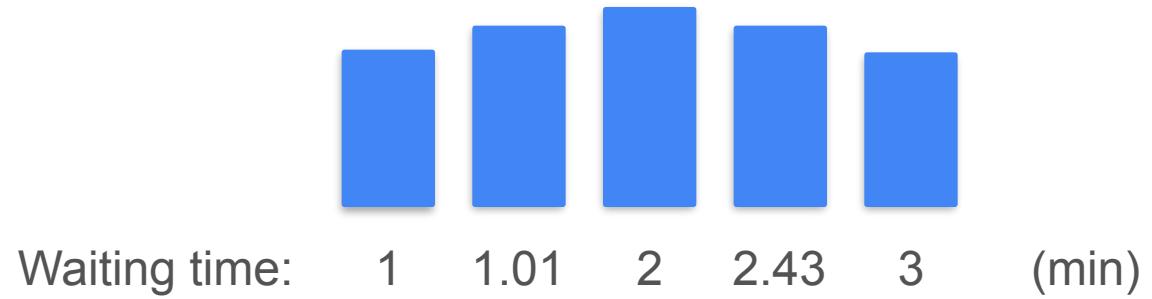
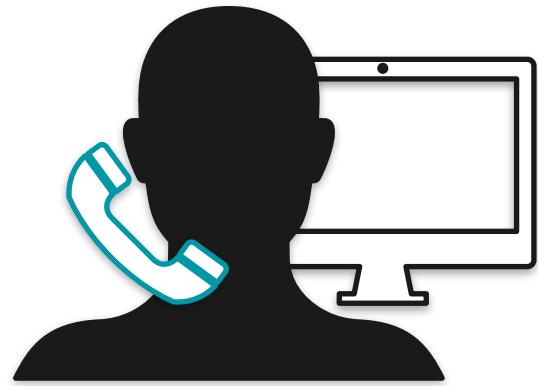
# From Discrete to Continuous Distributions



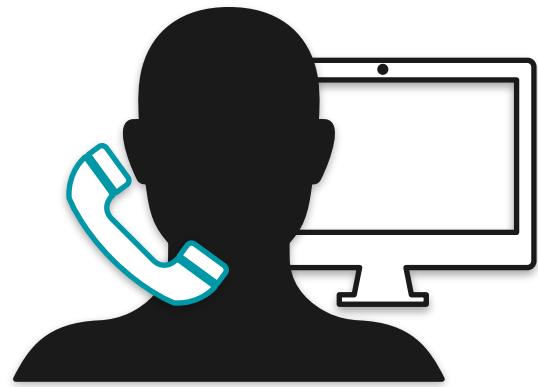
Waiting time: 1 1.01 2 3 (min)



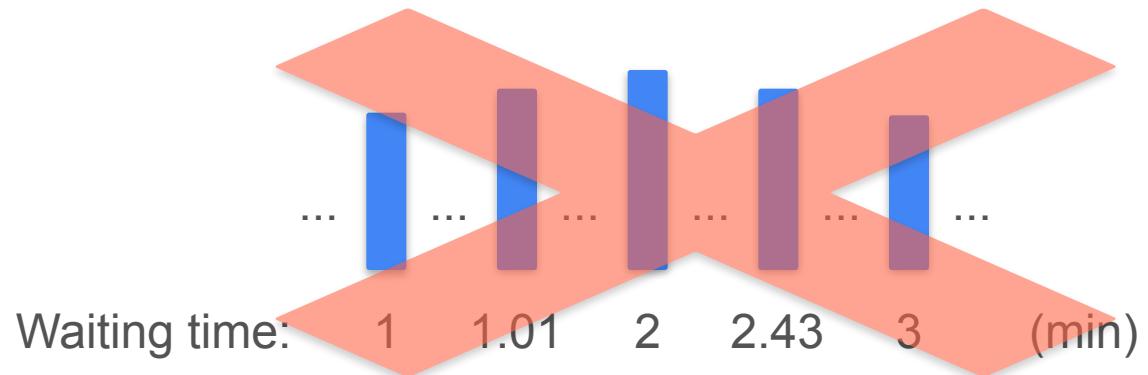
# From Discrete to Continuous Distributions



# From Discrete to Continuous Distributions



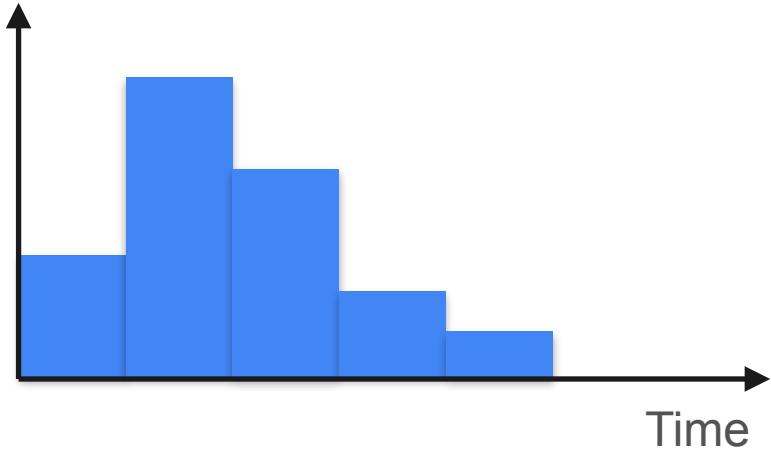
This is a continuous distribution!



What is the probability that you will wait EXACTLY one minute for the call?

Answer: ZERO

# From Discrete to Continuous Distributions



$P(\text{between 0 and 1 mins})$

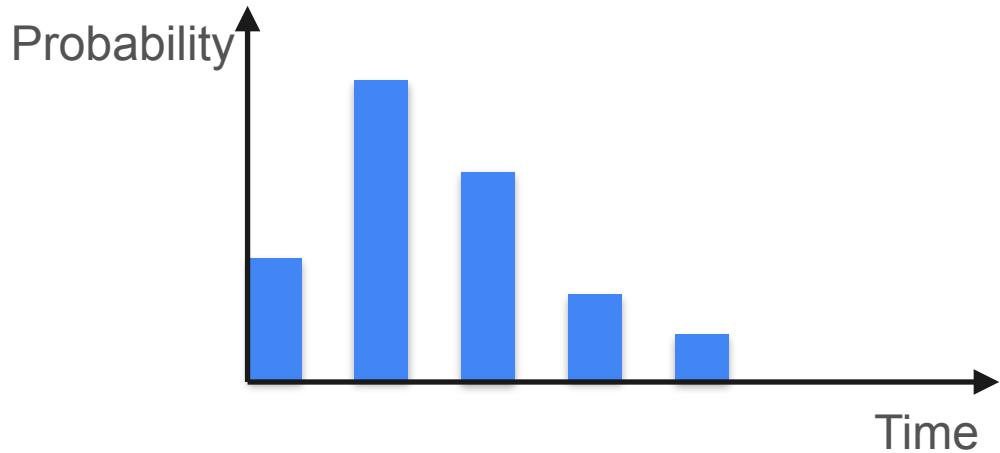
$P(\text{between 1 and 2 mins})$

$P(\text{between 2 and 3 mins})$

$P(\text{between 3 and 4 mins})$

$P(\text{between 4 and 5 mins})$

# From Discrete to Continuous Distributions



$P(\text{between 0 and 1 mins})$

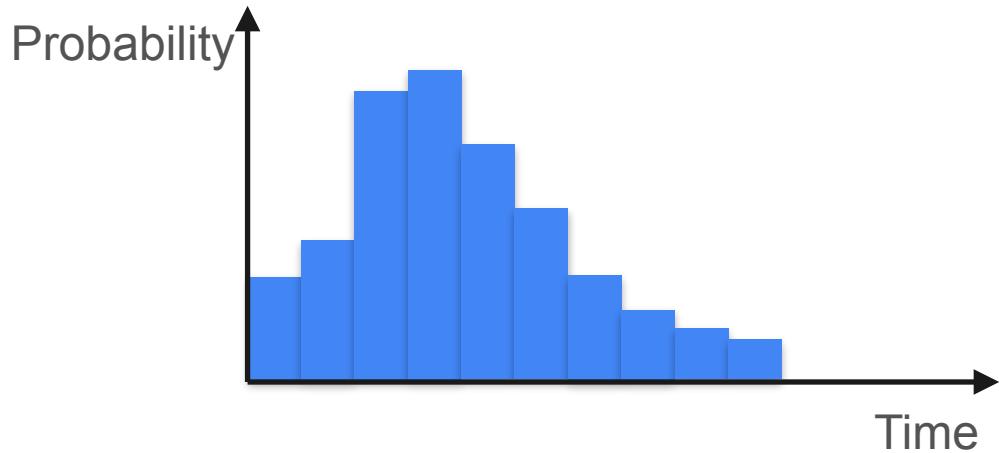
$P(\text{between 1 and 2 mins})$

$P(\text{between 2 and 3 mins})$

$P(\text{between 3 and 4 mins})$

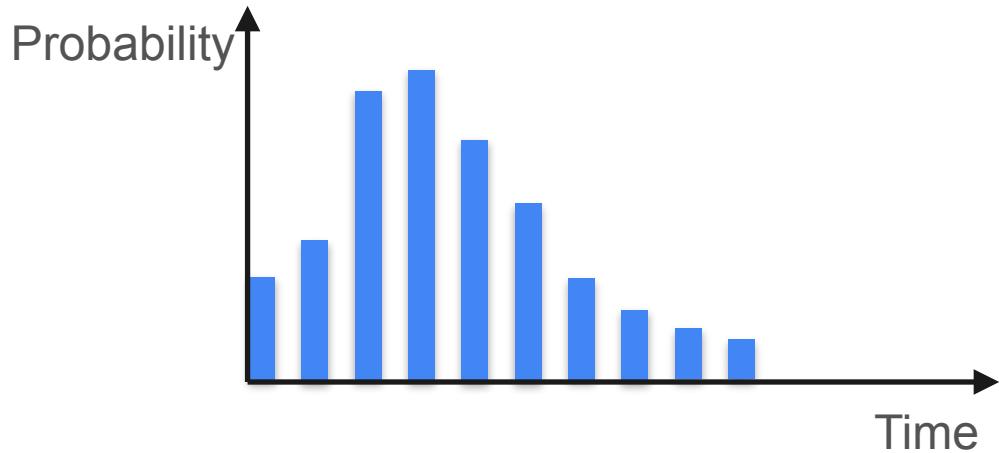
$P(\text{between 4 and 5 mins})$

# From Discrete to Continuous Distributions



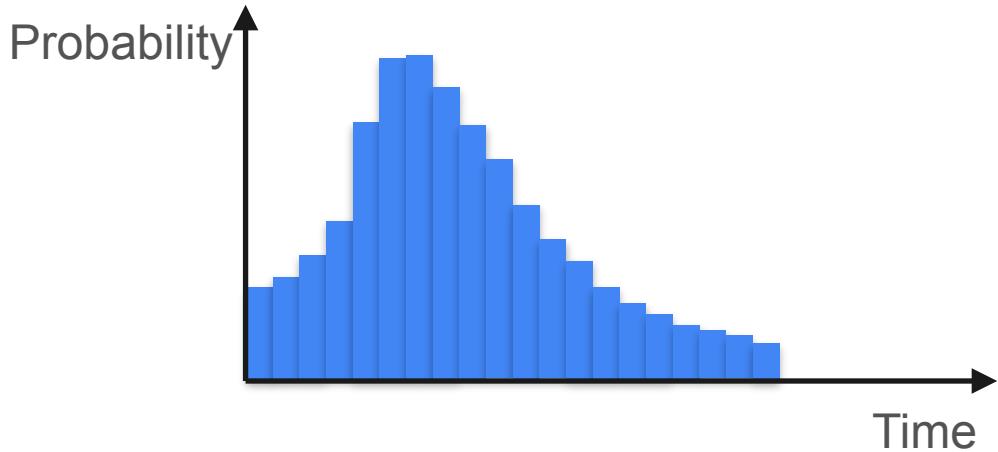
$P(\text{between } 0 \text{ and } 0.5 \text{ mins})$   
 $P(\text{between } 0.5 \text{ and } 1 \text{ mins})$   
 $P(\text{between } 1 \text{ and } 1.5 \text{ mins})$   
⋮  
 $P(\text{between } 3.5 \text{ and } 4 \text{ mins})$   
 $P(\text{between } 4 \text{ and } 4.5 \text{ mins})$   
 $P(\text{between } 4.5 \text{ and } 5 \text{ mins})$

# From Discrete to Continuous Distributions



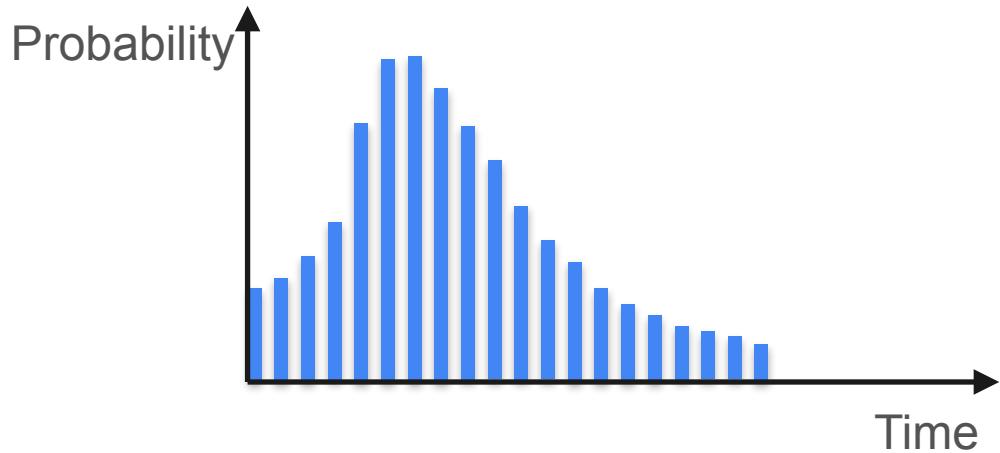
$P(\text{between } 0 \text{ and } 0.5 \text{ mins})$   
 $P(\text{between } 0.5 \text{ and } 1 \text{ mins})$   
 $P(\text{between } 1 \text{ and } 1.5 \text{ mins})$   
⋮  
 $P(\text{between } 3.5 \text{ and } 4 \text{ mins})$   
 $P(\text{between } 4 \text{ and } 4.5 \text{ mins})$   
 $P(\text{between } 4.5 \text{ and } 5 \text{ mins})$

# From Discrete to Continuous Distributions



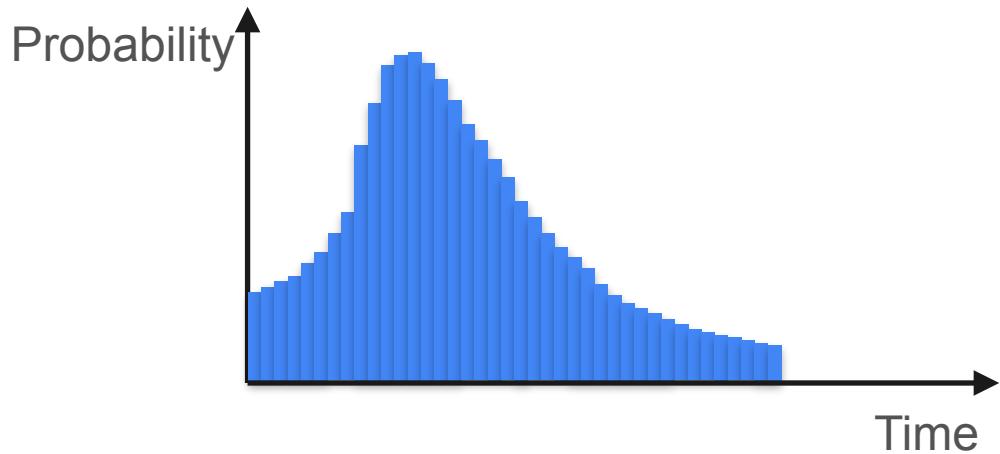
$P(\text{between } 0 \text{ and } 0.25 \text{ mins})$   
 $P(\text{between } 0.25 \text{ and } 0.5 \text{ mins})$   
 $P(\text{between } 0.5 \text{ and } 0.75 \text{ mins})$   
⋮  
 $P(\text{between } 4.25 \text{ and } 4.5 \text{ mins})$   
 $P(\text{between } 4.5 \text{ and } 4.75 \text{ mins})$   
 $P(\text{between } 4.75 \text{ and } 5 \text{ mins})$

# From Discrete to Continuous Distributions



$P(\text{between } 0 \text{ and } 0.25 \text{ mins})$   
 $P(\text{between } 0.25 \text{ and } 0.5 \text{ mins})$   
 $P(\text{between } 0.5 \text{ and } 0.75 \text{ mins})$   
⋮  
 $P(\text{between } 4.25 \text{ and } 4.5 \text{ mins})$   
 $P(\text{between } 4.5 \text{ and } 4.75 \text{ mins})$   
 $P(\text{between } 4.75 \text{ and } 5 \text{ mins})$

# From Discrete to Continuous Distributions

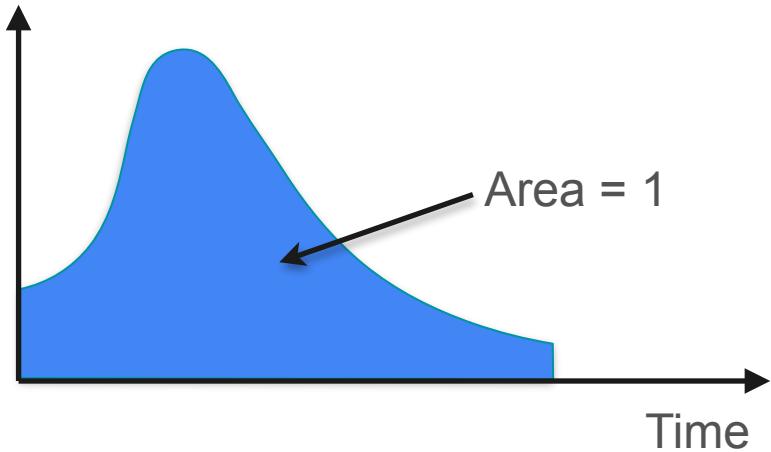


$P(\text{between } 0 \text{ and } 0.125 \text{ mins})$

⋮

$P(\text{between } 4.875 \text{ and } 5 \text{ mins})$

# From Discrete to Continuous Distributions



- Discrete:
  - Sum of heights equals 1
- Continuous:
  - Area under the curve equals 1



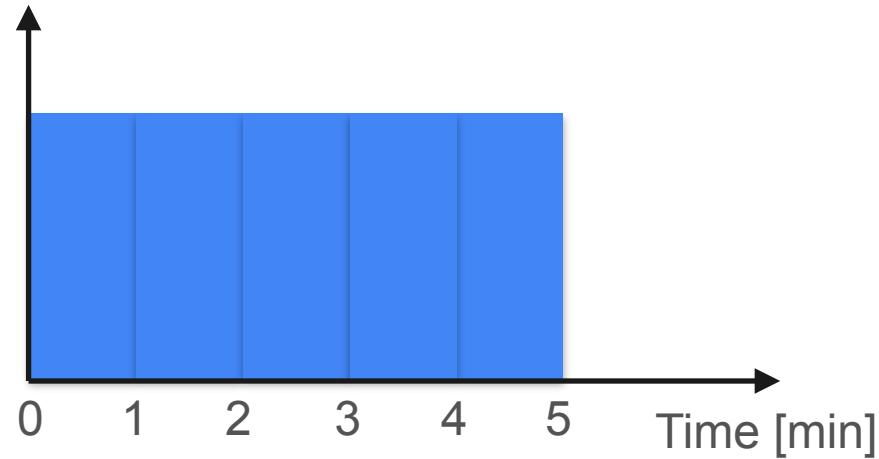
DeepLearning.AI

# Probability Distributions

---

## Probability density function

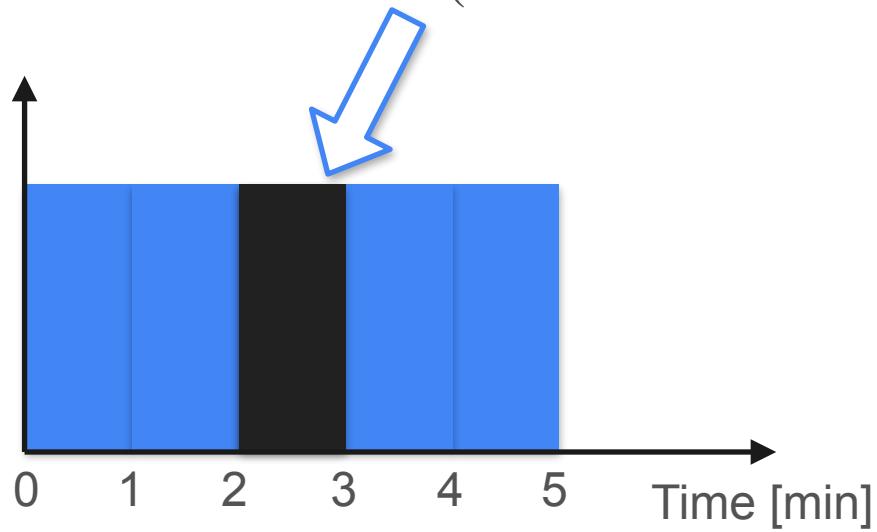
# Probability Density Function



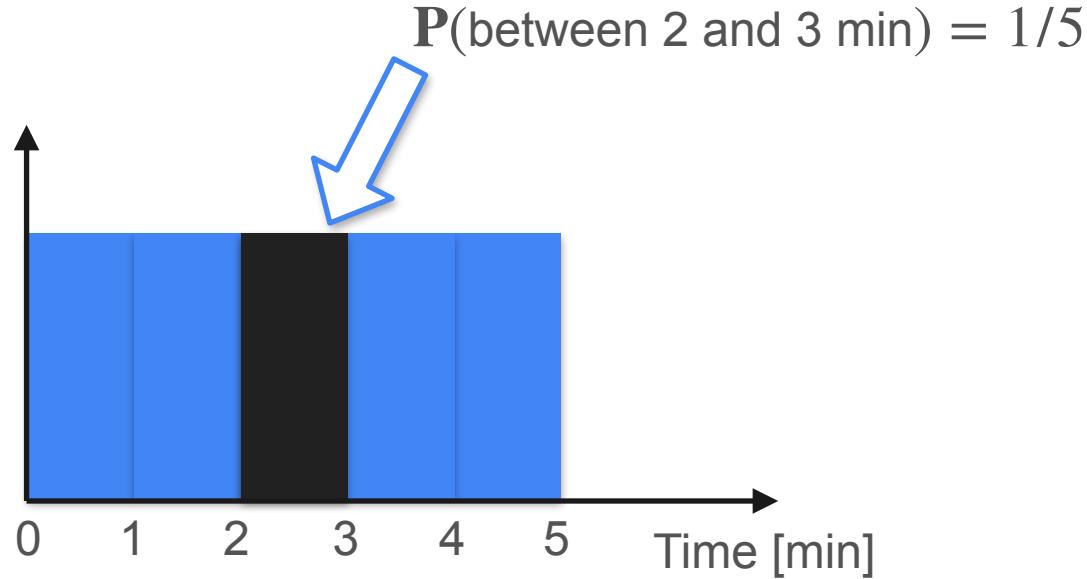
# Probability Density Function



$P(\text{between 2 and 3 min}) = ?$



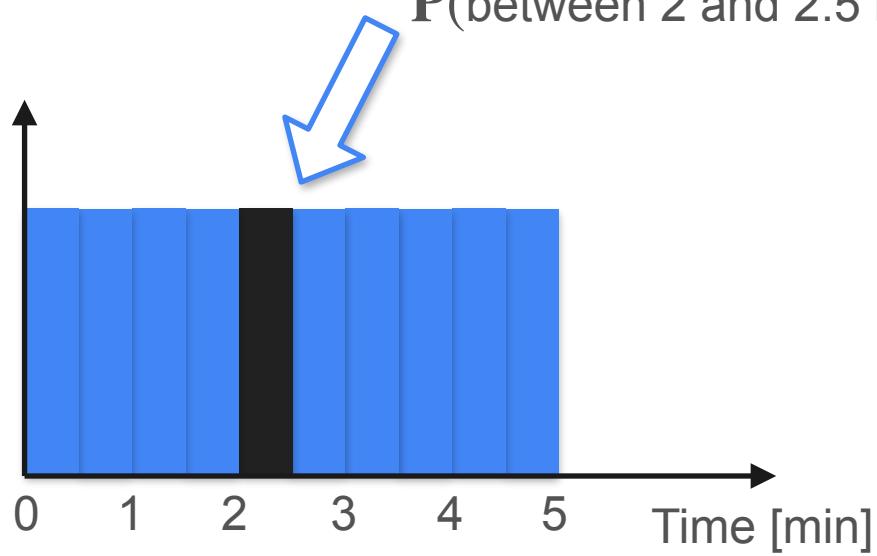
# Probability Density Function



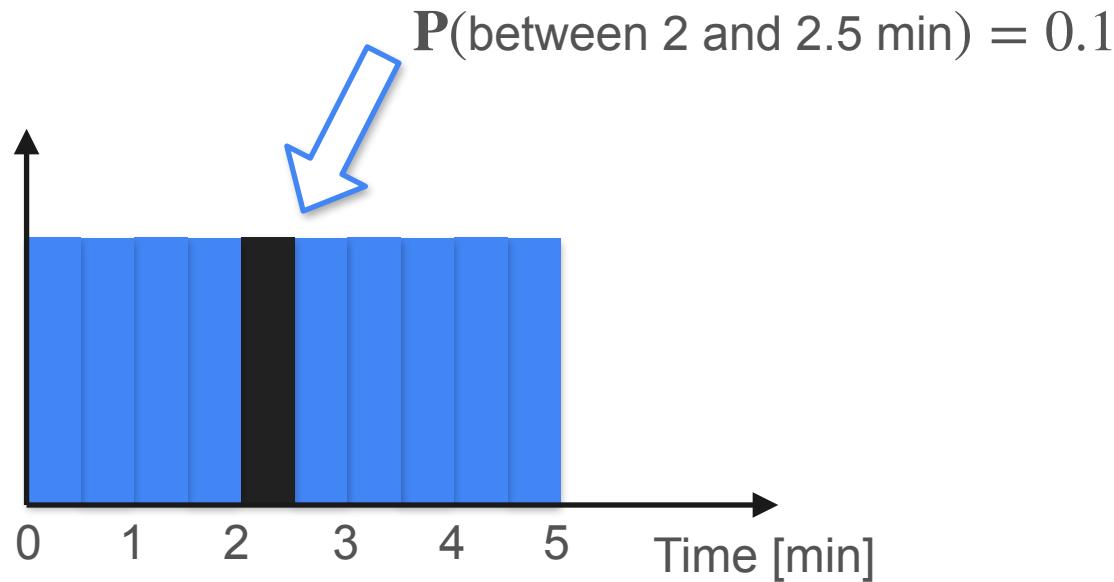
# Probability Density Function



$P(\text{between } 2 \text{ and } 2.5 \text{ min}) = ?$



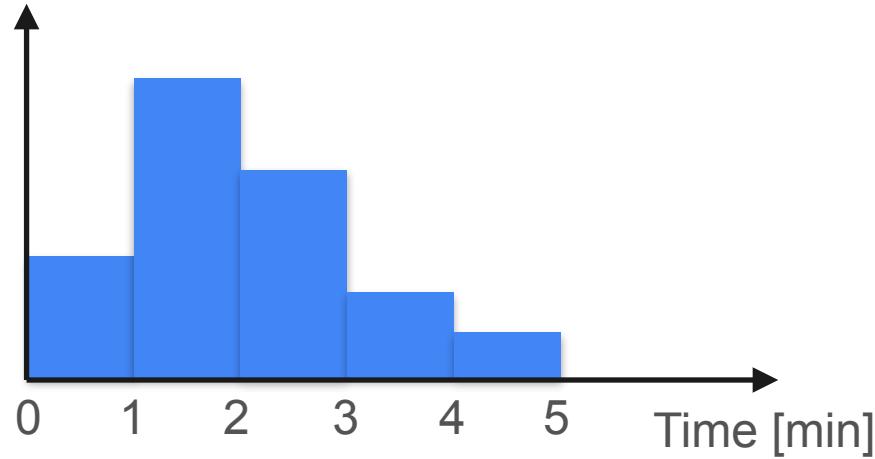
# Probability Density Function



# Probability Density Function



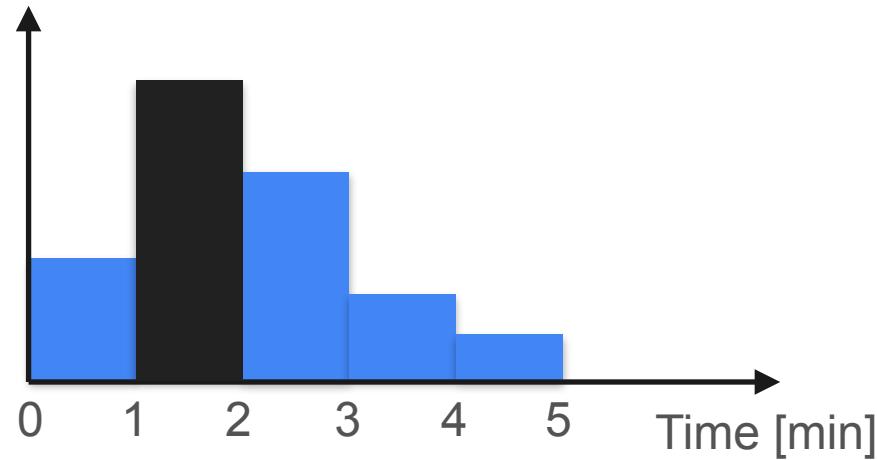
$P(\text{between 1 and 2 min})$



# Probability Density Function



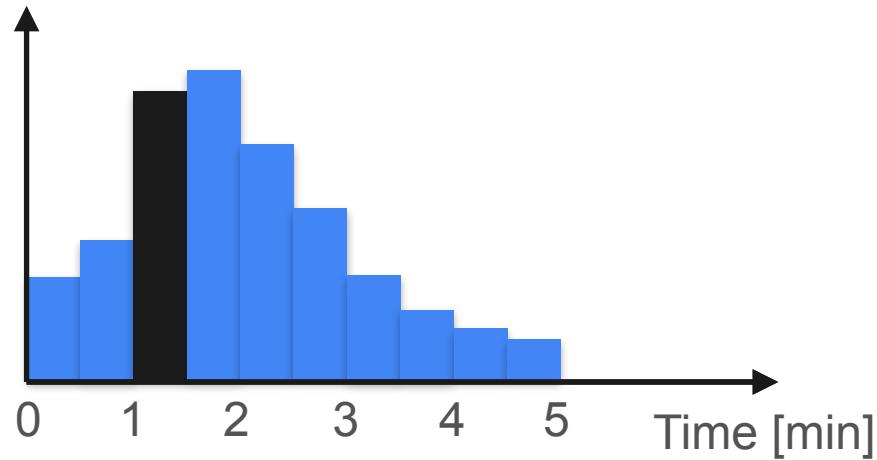
$P(\text{between 1 and 2 min})$



# Probability Density Function



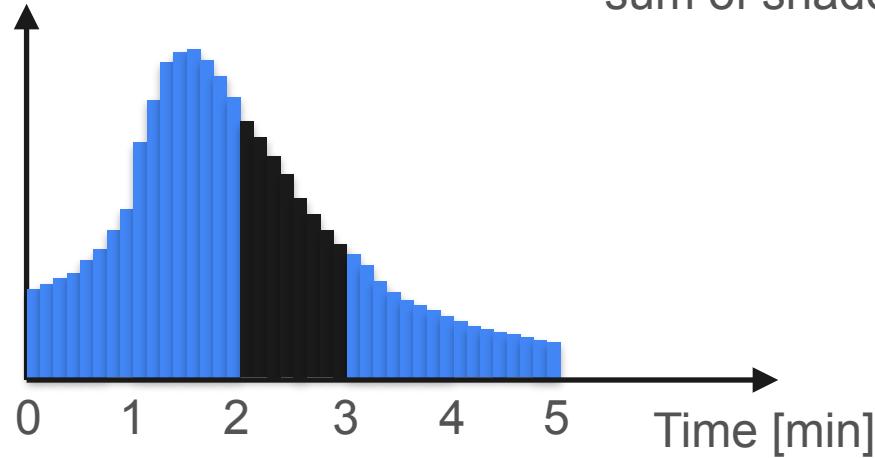
$P(\text{between 1 and 1:30})$



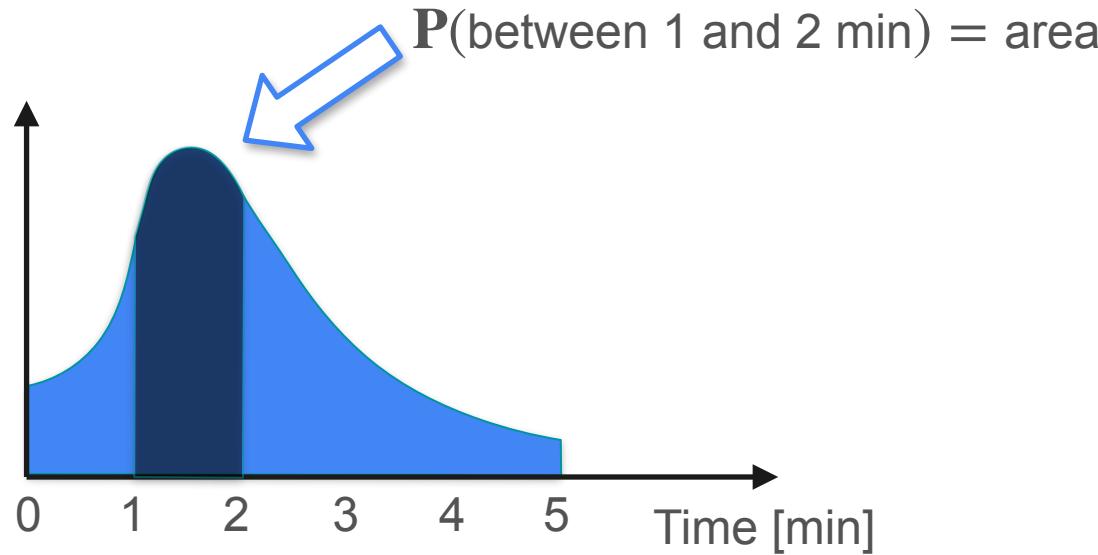
# Probability Density Function



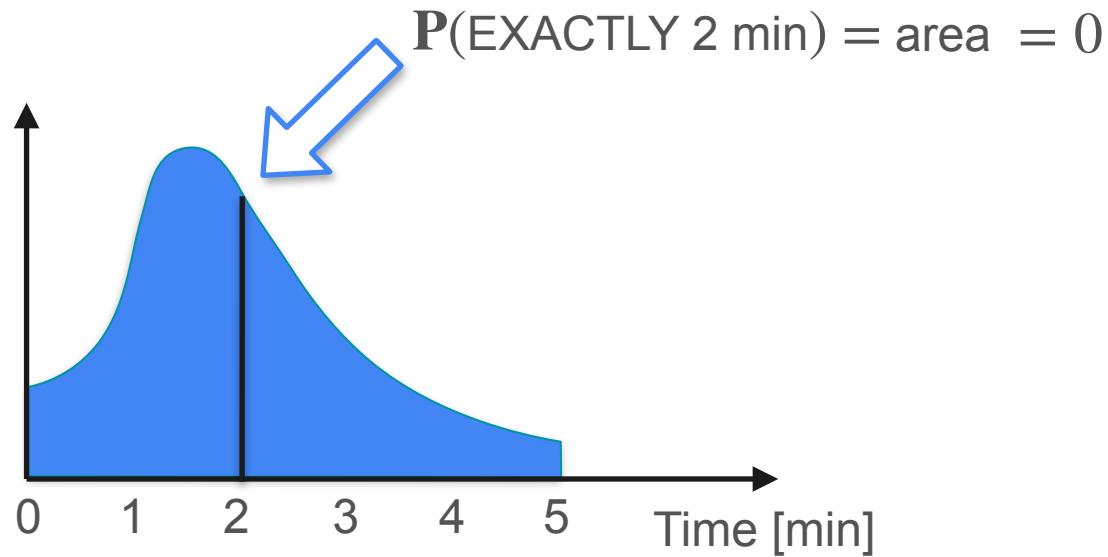
$P(\text{between 2 and 3 min}) =$   
sum of shaded areas



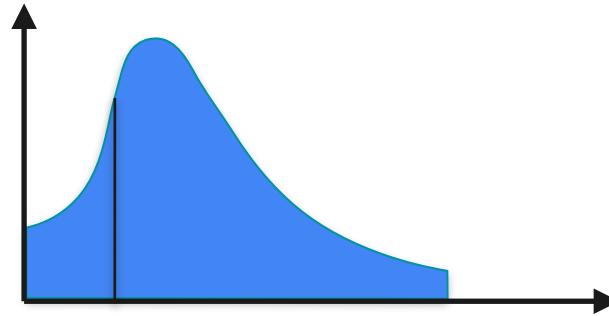
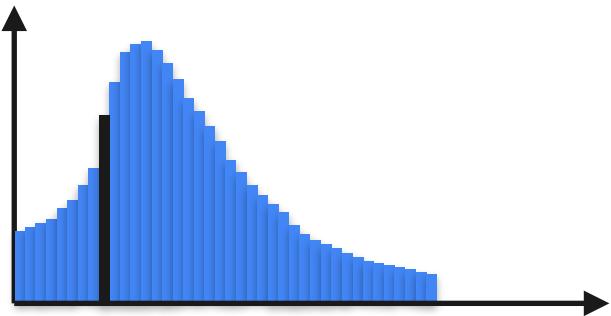
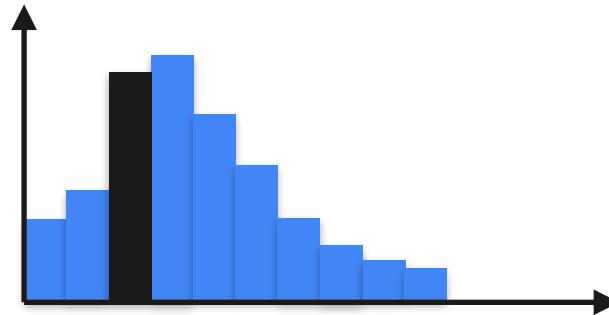
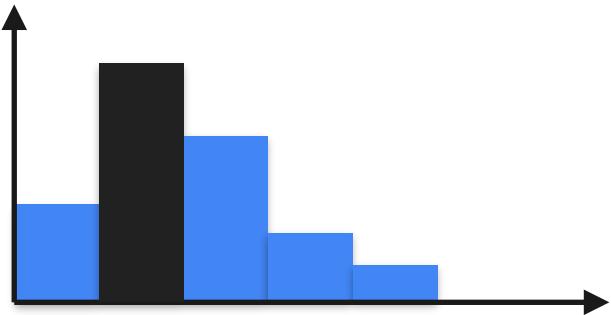
# Probability Density Function



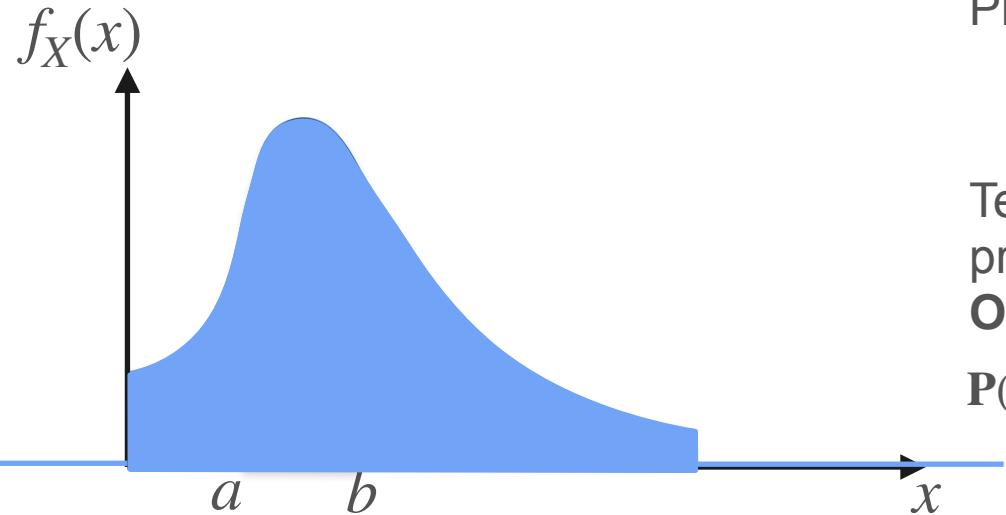
# Probability Density Function



# Probability Density Function



# Probability Density Function: Formal Definition



Probability Density Function (PDF)

$$f_X(x)$$

Tells you the rate you accumulate probability around each point.

**Only defined for continuous variables!**

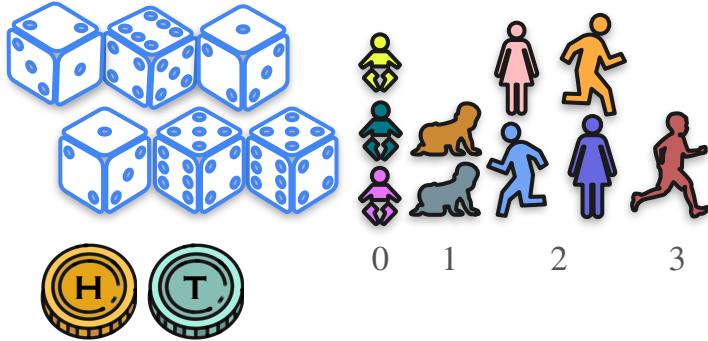
$$\mathbf{P}(a < X < b) = \text{area under } f_X(x)$$

$f_X(x)$  needs to satisfy:

- It is defined for all numbers
- $f_X(x) \geq 0$
- Area under  $f_X(x) = 1$

# Discrete and Continuous Random Variables

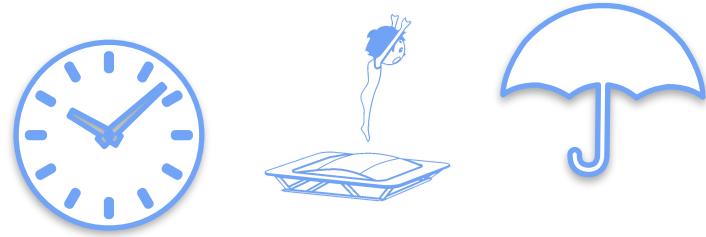
## Discrete random variables



Can take only a **finite** or at most countable number of values

$$\text{PMF: } p_X(x) = \mathbf{P}(X = x)$$

## Continuous random variables



Takes values on an interval  
(infinite possibilities!)

$$\begin{aligned} \text{PDF: } & f_X(x) \\ \mathbf{P}(X = x) &= 0 \quad \forall x \end{aligned}$$



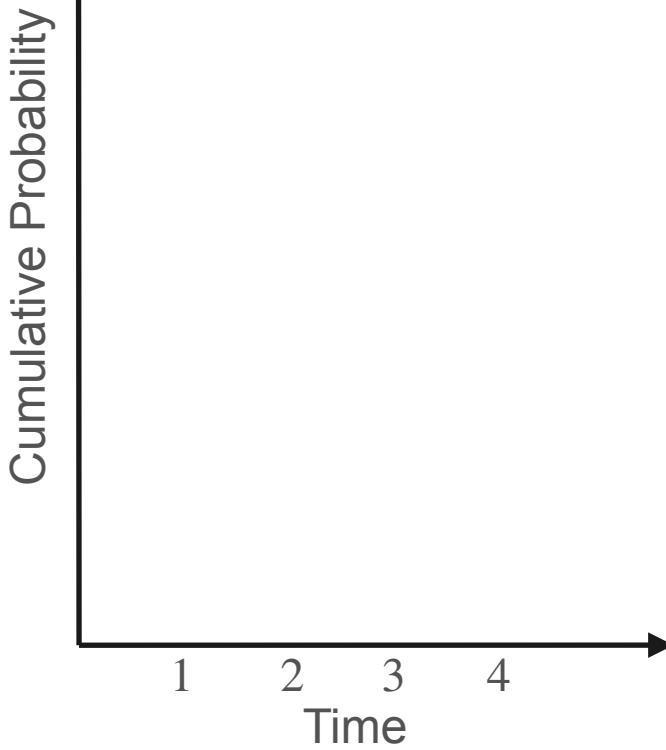
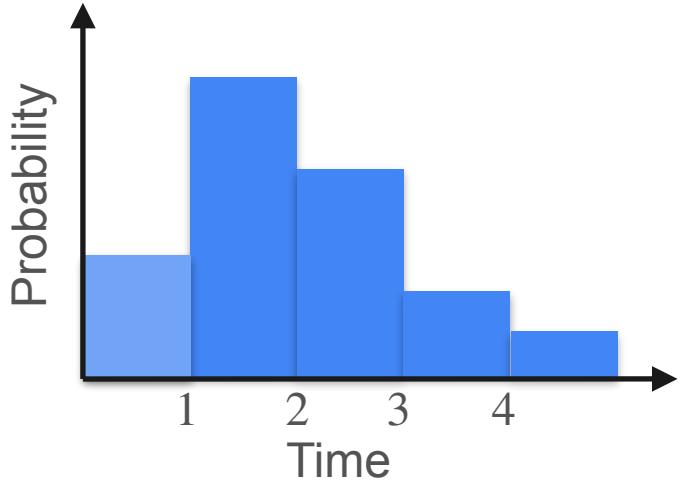
DeepLearning.AI

# Probability Distributions

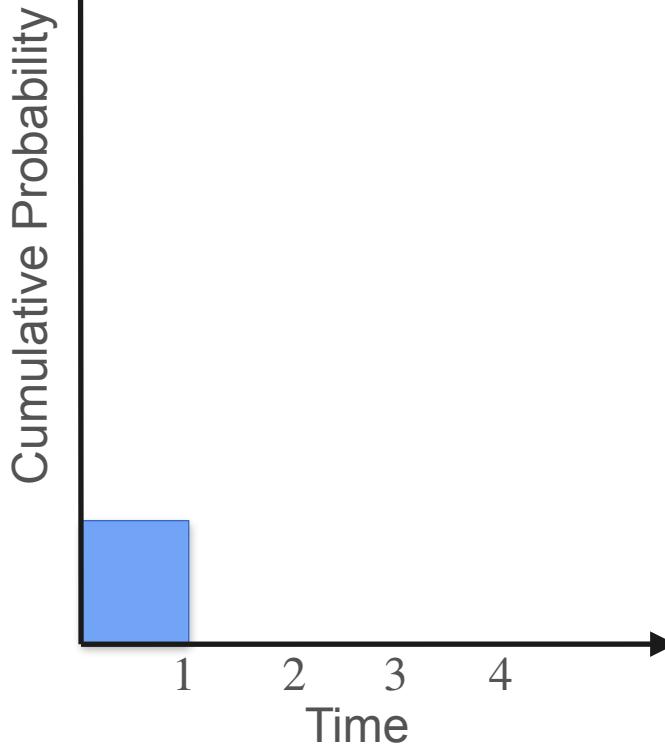
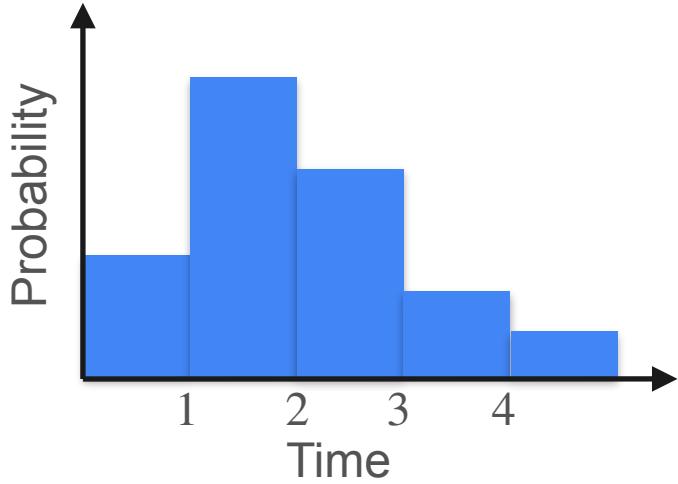
---

## Cumulative Distribution Function

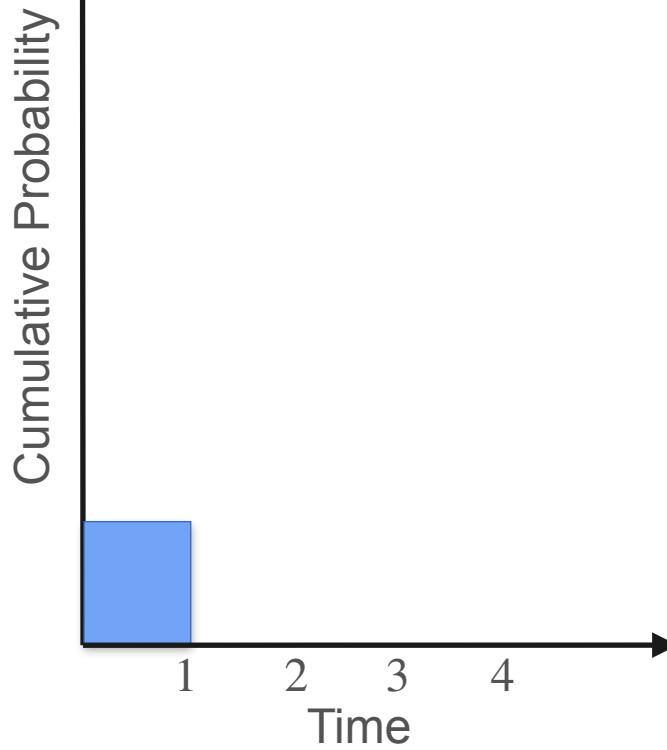
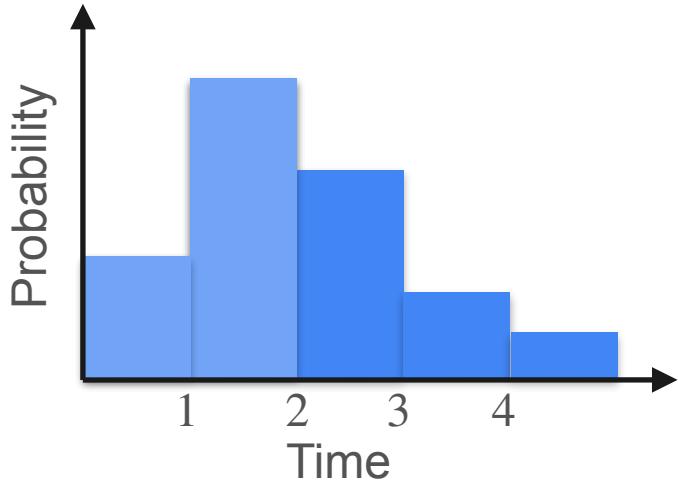
# Cumulative Distribution



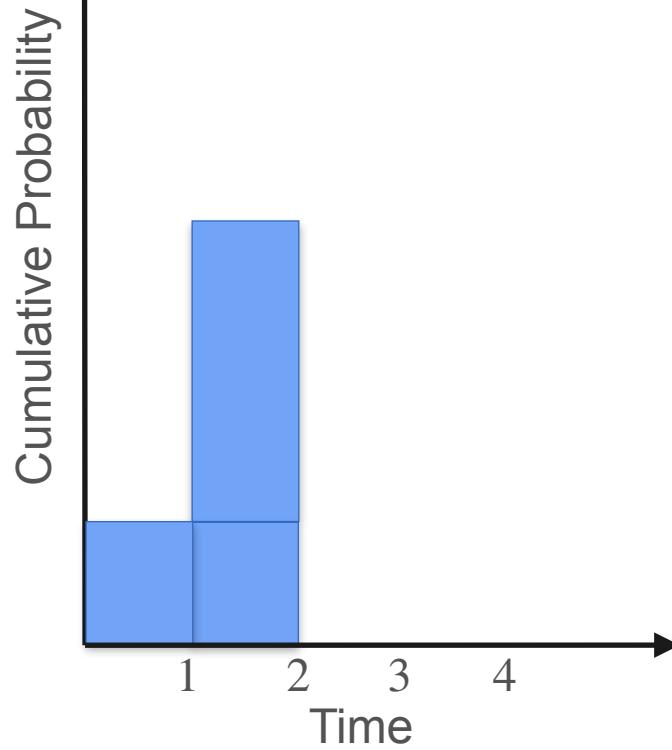
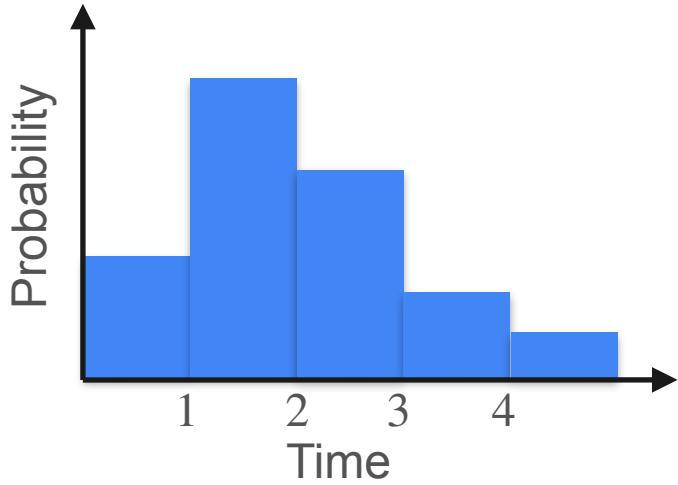
# Cumulative Distribution



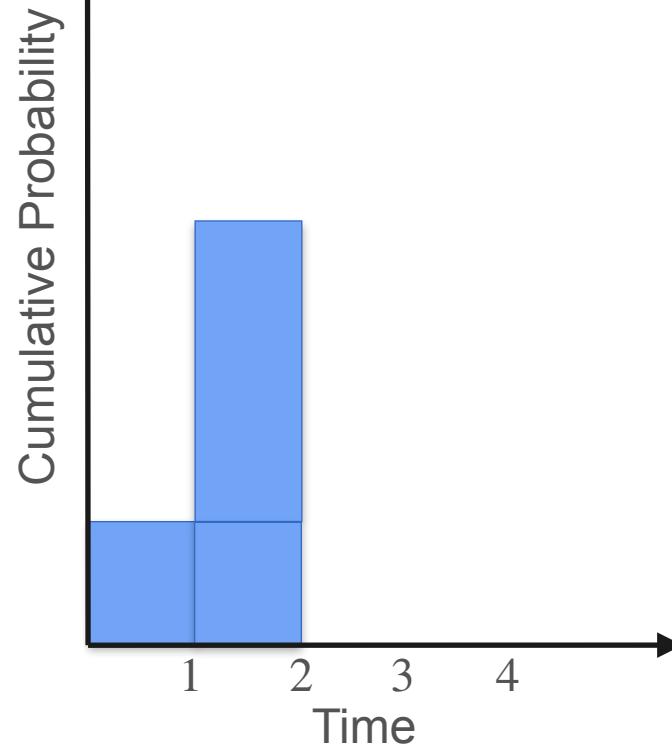
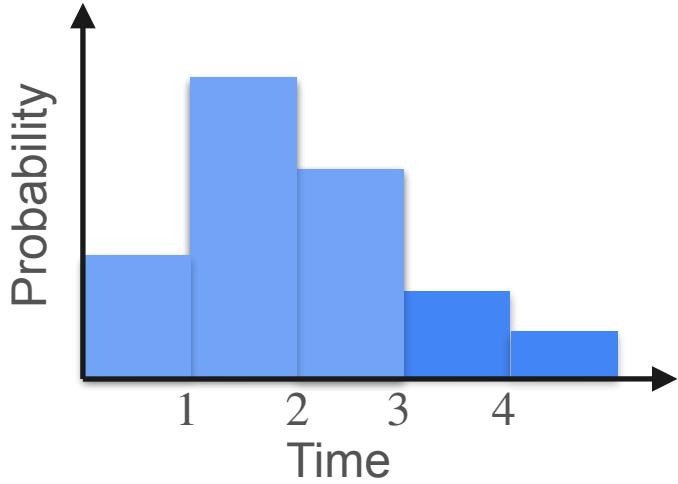
# Cumulative Distribution



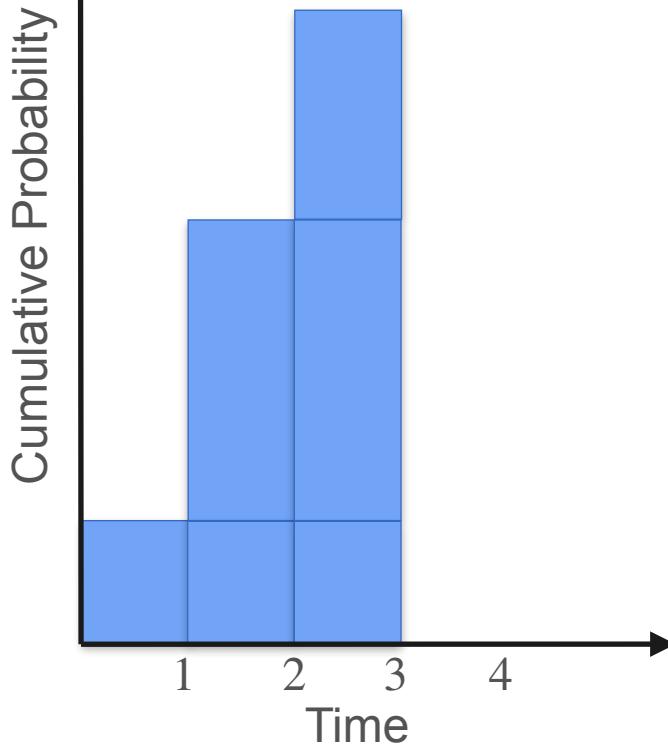
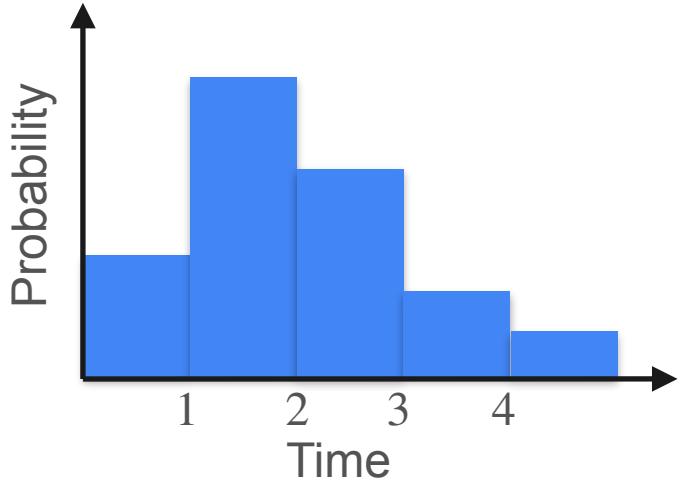
# Cumulative Distribution



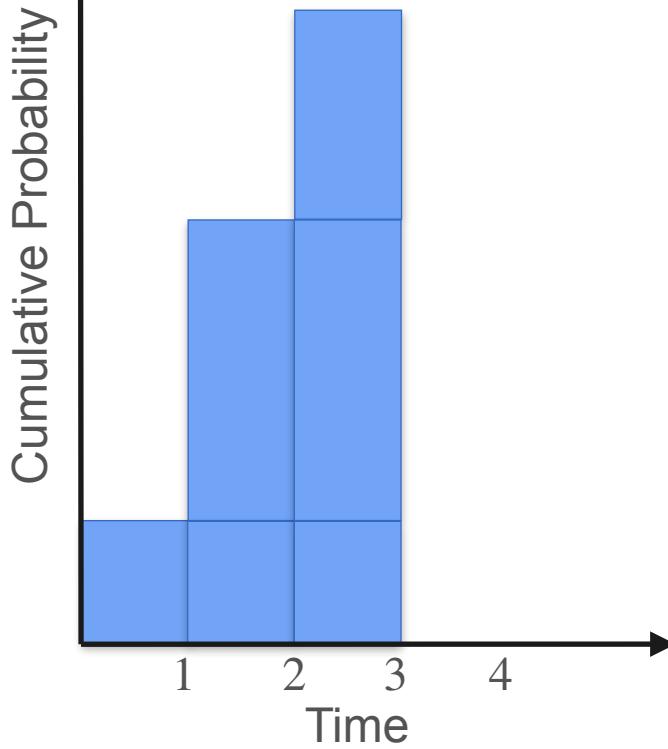
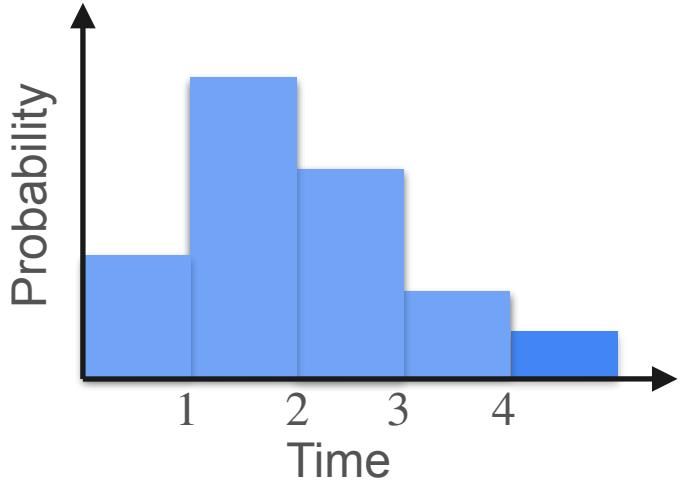
# Cumulative Distribution



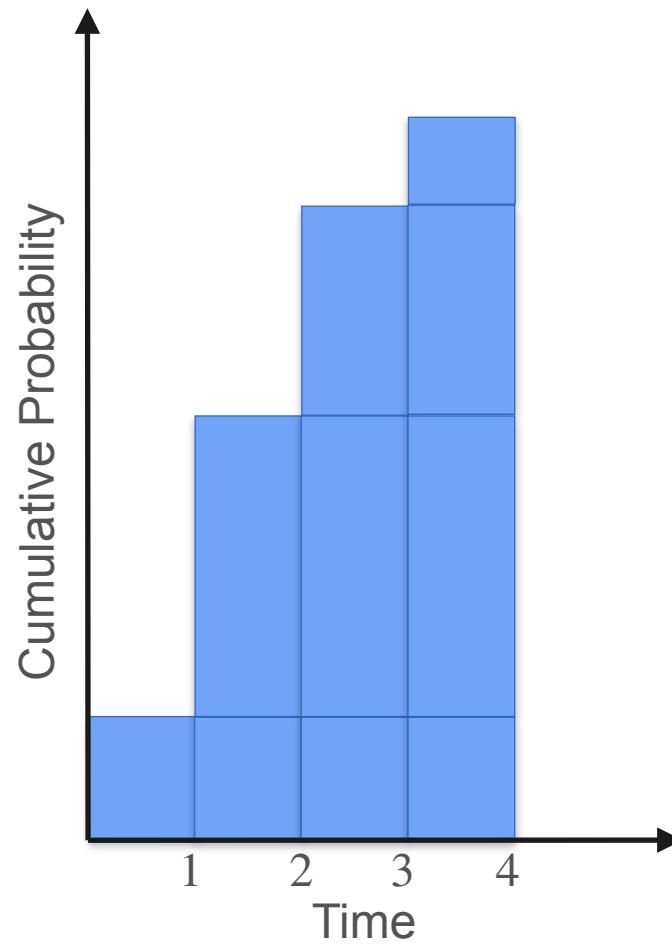
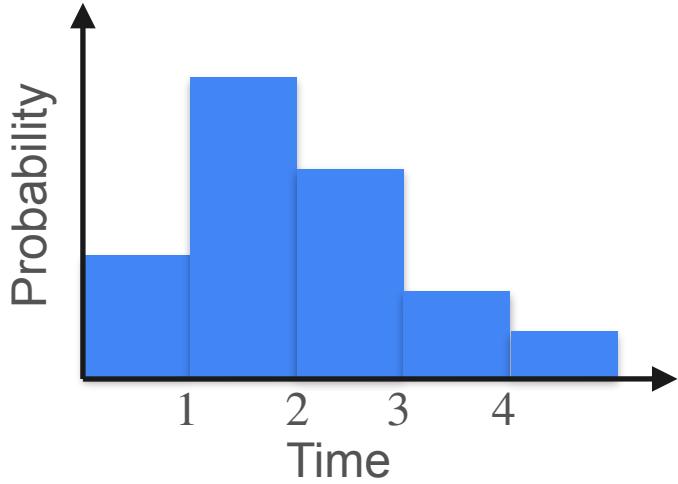
# Cumulative Distribution



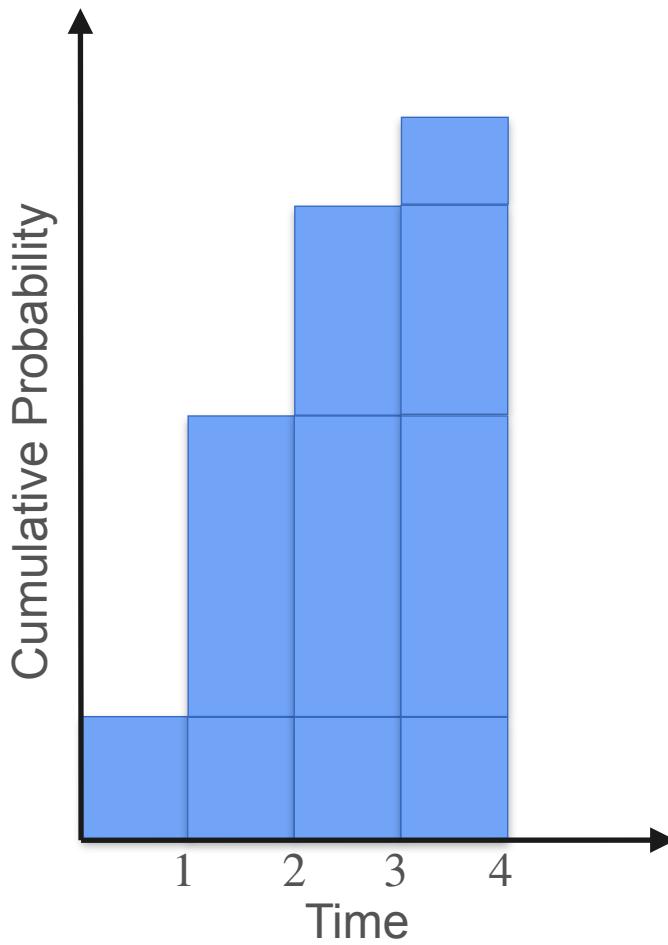
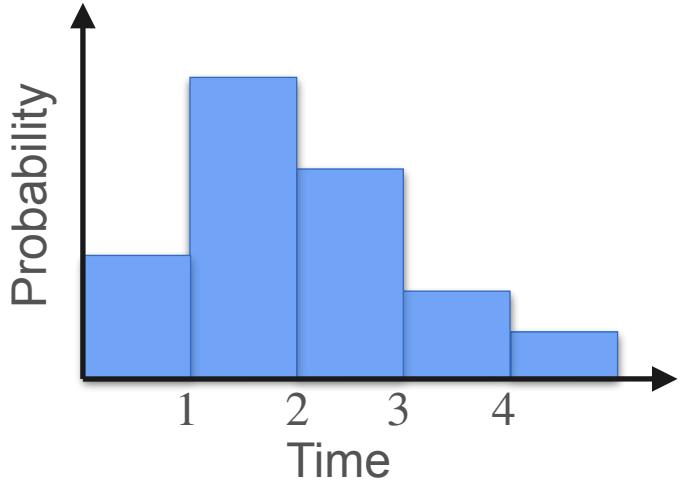
# Cumulative Distribution



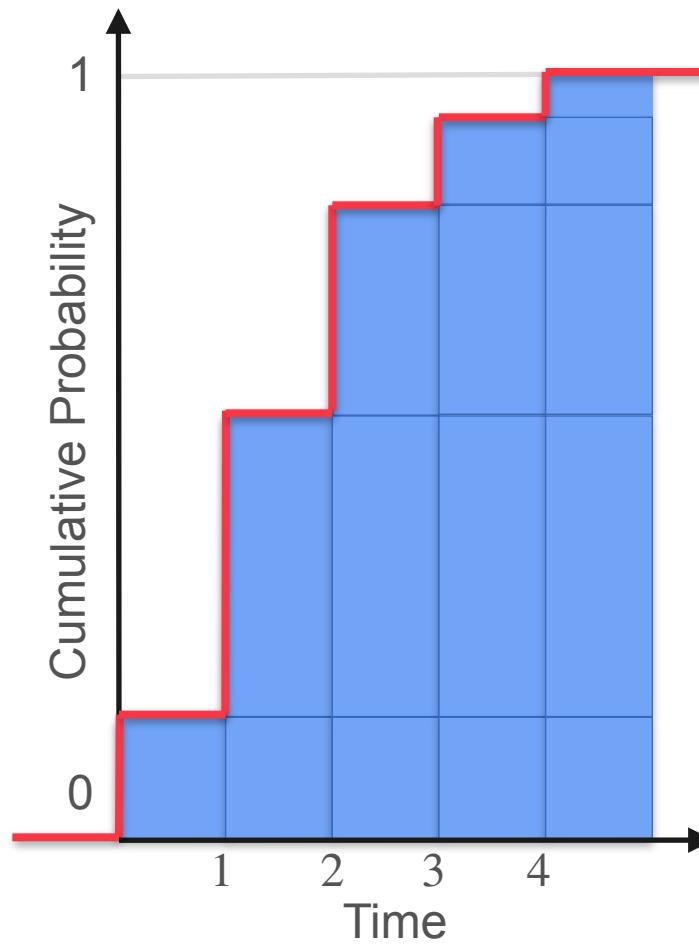
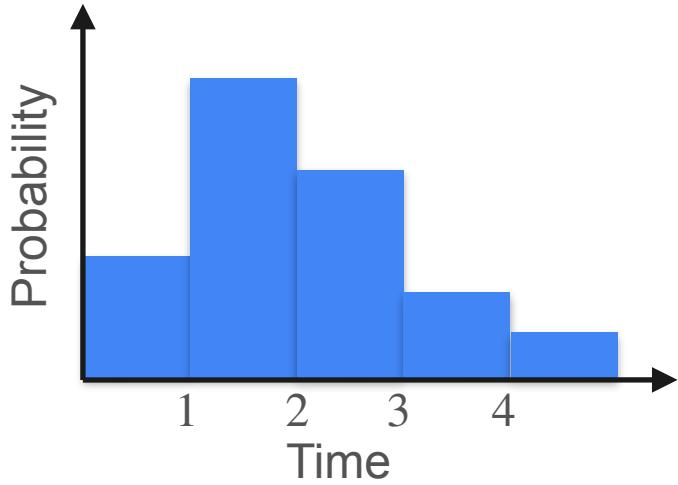
# Cumulative Distribution



# Cumulative Distribution

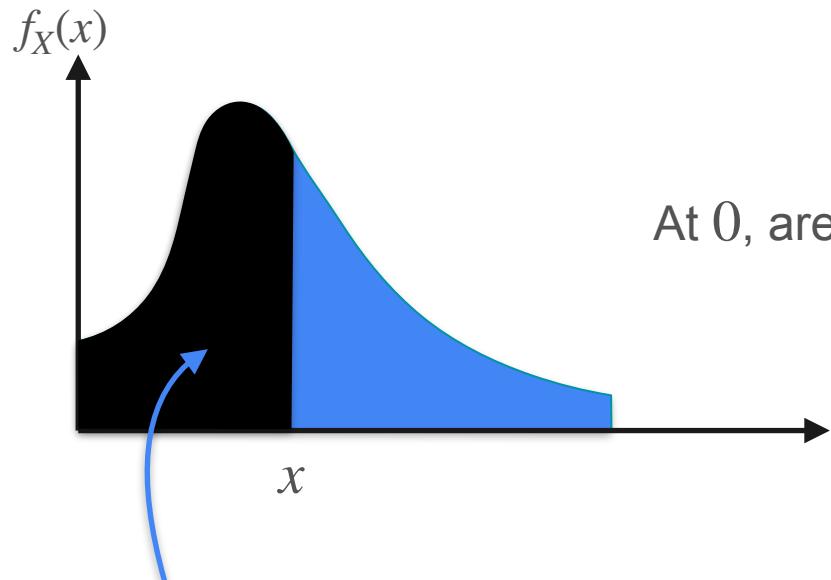


# Cumulative Distribution



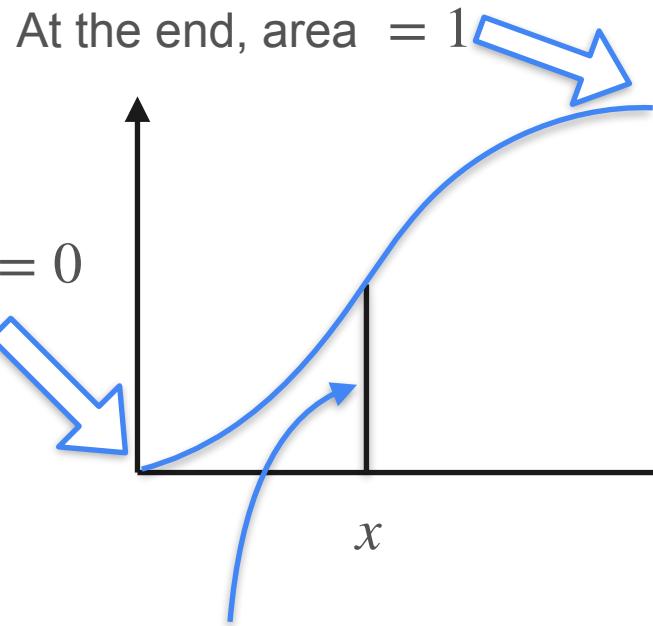
# Cumulative Distribution

CDF: Cumulative distribution function



At 0, area = 0

$P(\text{less than or equal to 2 minutes}) = 0.5$



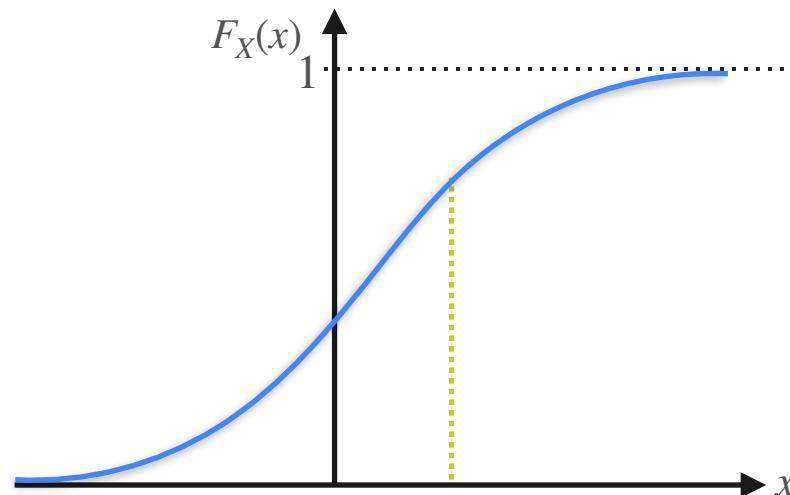
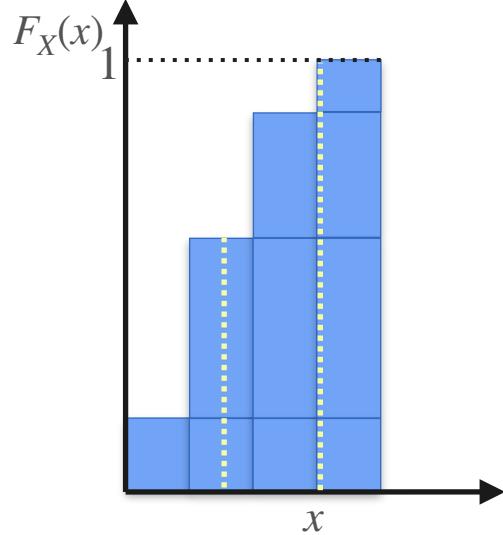
$P(\text{less than or equal to 2 minutes}) = 0.5$

# Cumulative Distribution Function: Formal Definition

The CDF shows how much probability the variable has accumulated until a certain value

That means that

$$F_X(x) = \mathbf{P}(X \leq x) \quad \text{It is defined for every real number}$$

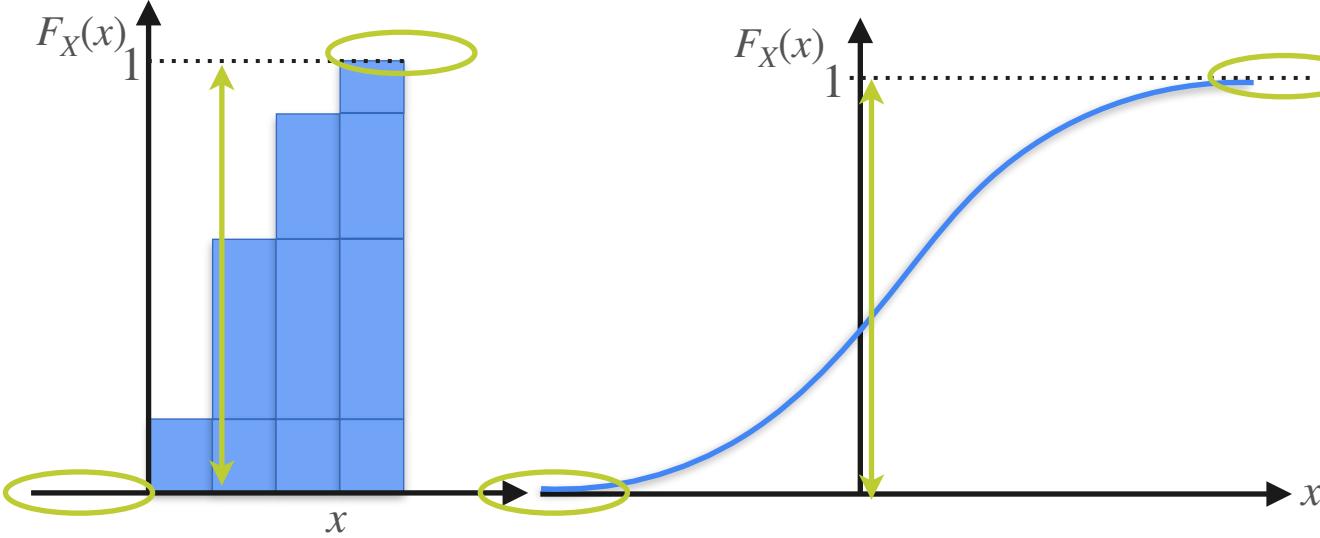


# Cumulative Distribution Function: Formal Definition

The CDF shows how much probability the variable has accumulated until a certain value

That means that

$$F_X(x) = P(X \leq x) \quad \text{It is defined for every real number}$$

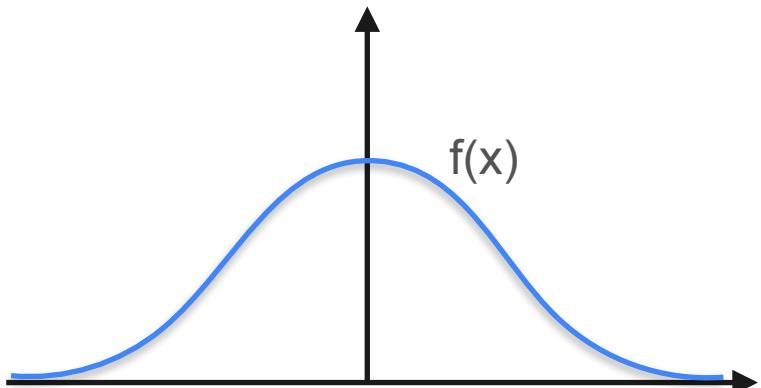


## Properties

- $0 \leq F_X(x) \leq 1$
- Left “endpoint” is 0
- Right “endpoint” is 1
- Never decreases

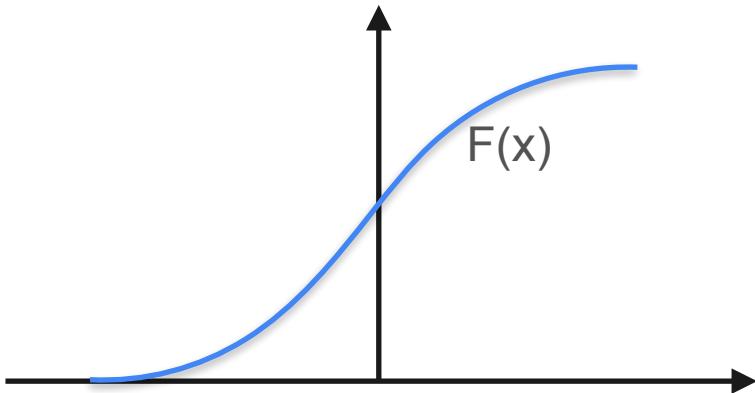
# PDF and CDF Summary

PDF



- area = 1
- Always positive

CDF



- left “endpoint” is 0
- right “endpoint” is 1
- (endpoints can be at infinity)
- Always positive and increasing



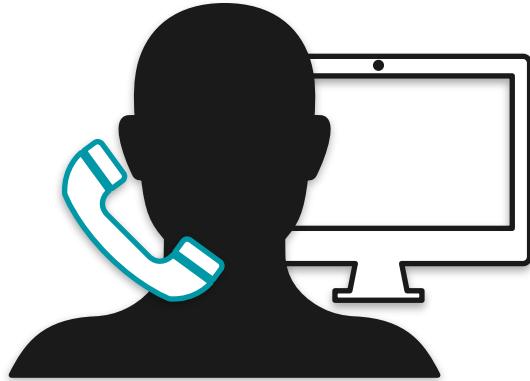
DeepLearning.AI

# Probability Distributions

---

## Uniform Distribution

# Uniform Distribution: Motivation

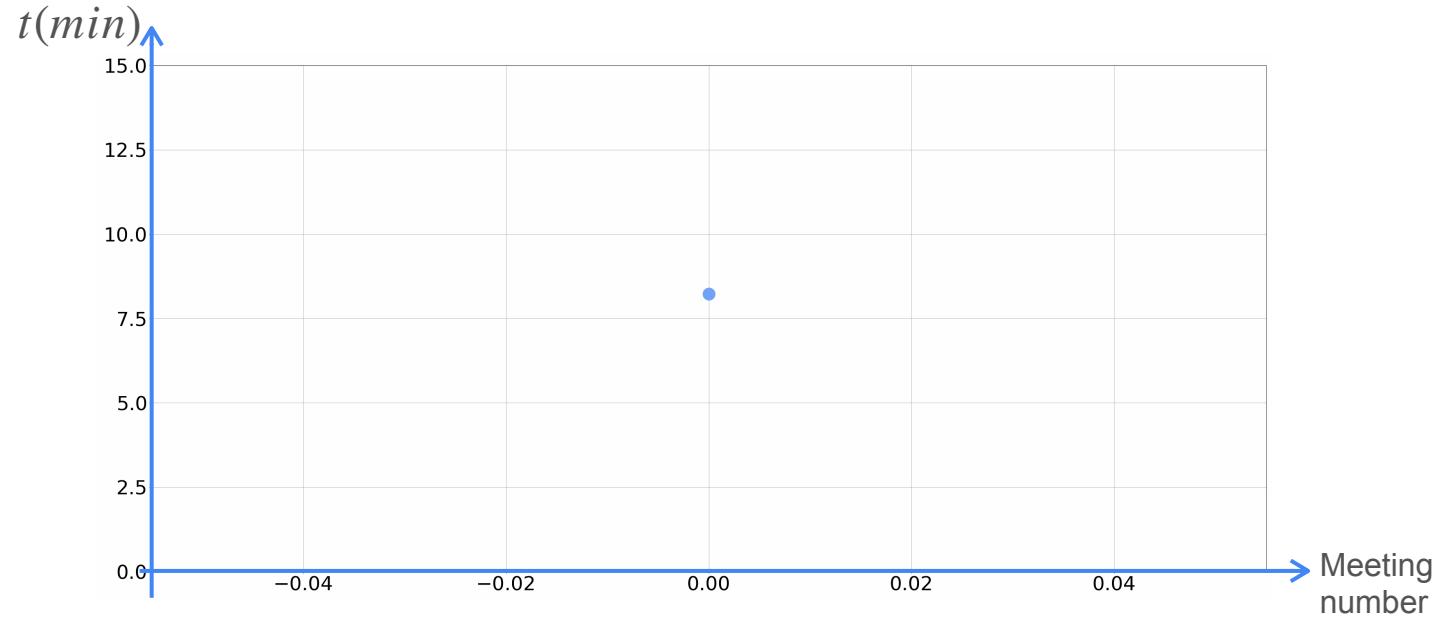


You're calling a tech support line. They can answer any time between zero and 15 minutes and if they don't answer in this time, the line is disconnected.

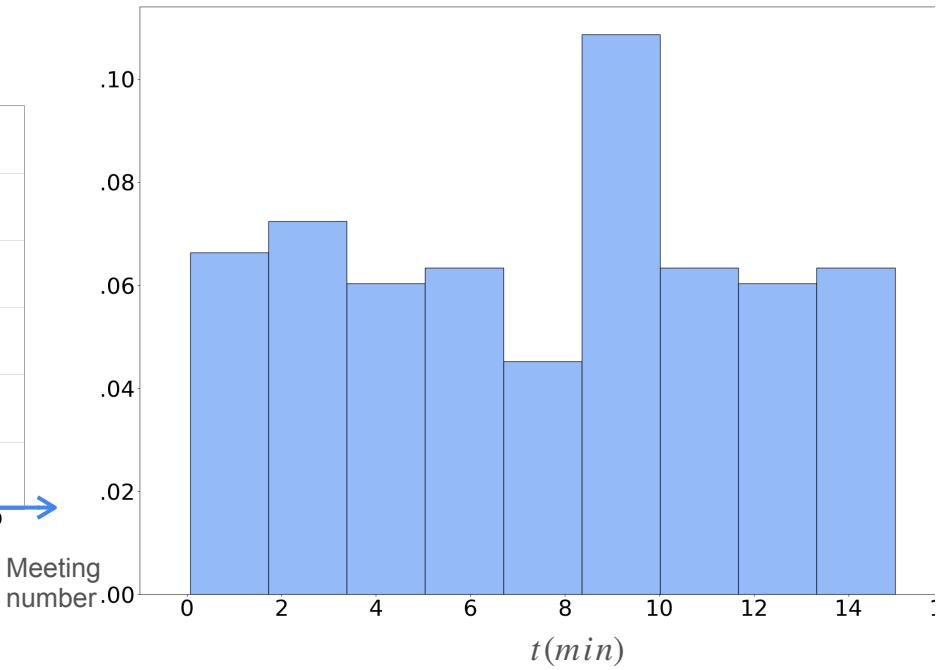
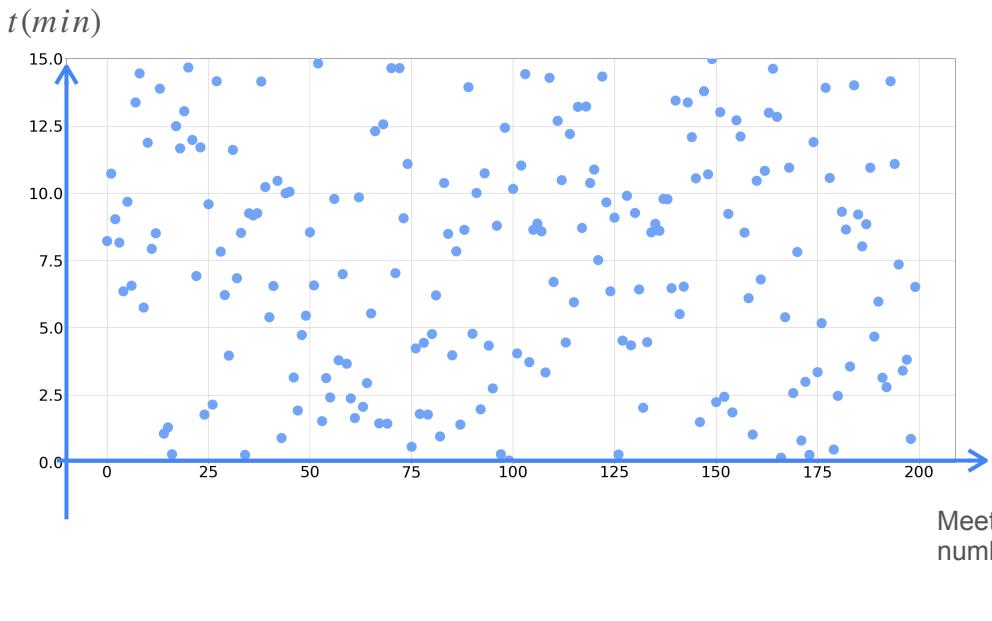
# Uniform Distribution: Motivation



Last 200 times you called them, you took down notes of how long they took to respond



# Uniform Distribution: Motivation



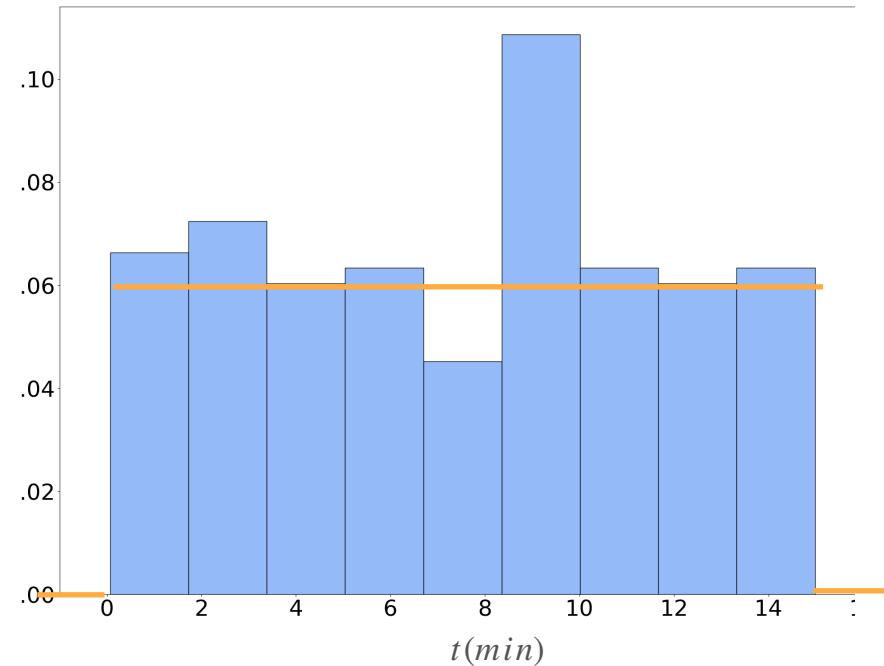
# Uniform Distribution: Motivation

T: time (in minutes) you have to wait

Any value between 0 and 15 minutes must have the same frequency of occurrence.

The pdf must be constant for all values in the interval (0,15)

Which constant?  $\rightarrow 15 \times h = 1 \rightarrow h = \frac{1}{15} = 0.06$



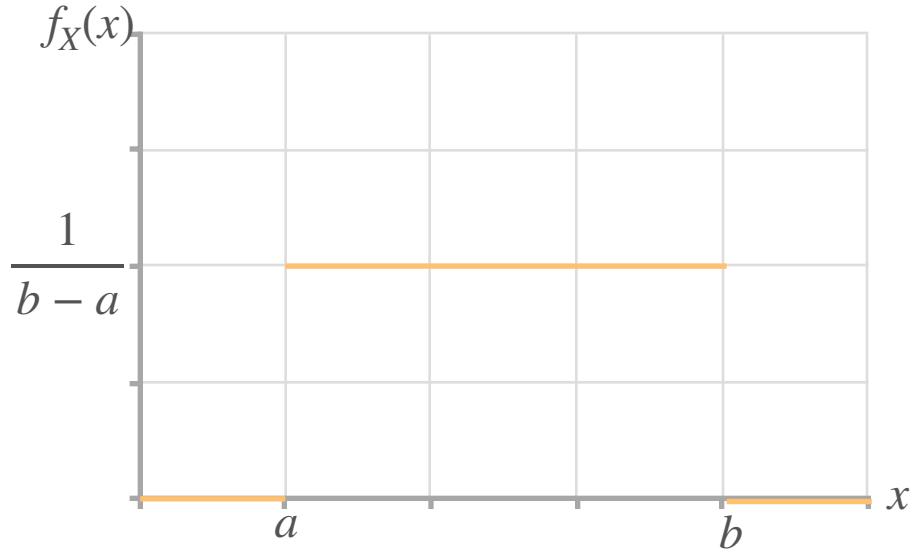
# Uniform Distribution: Model

A continuous random variable can be modeled with a **uniform** distribution if all possible values lie in an interval and have the **same frequency** of occurrence

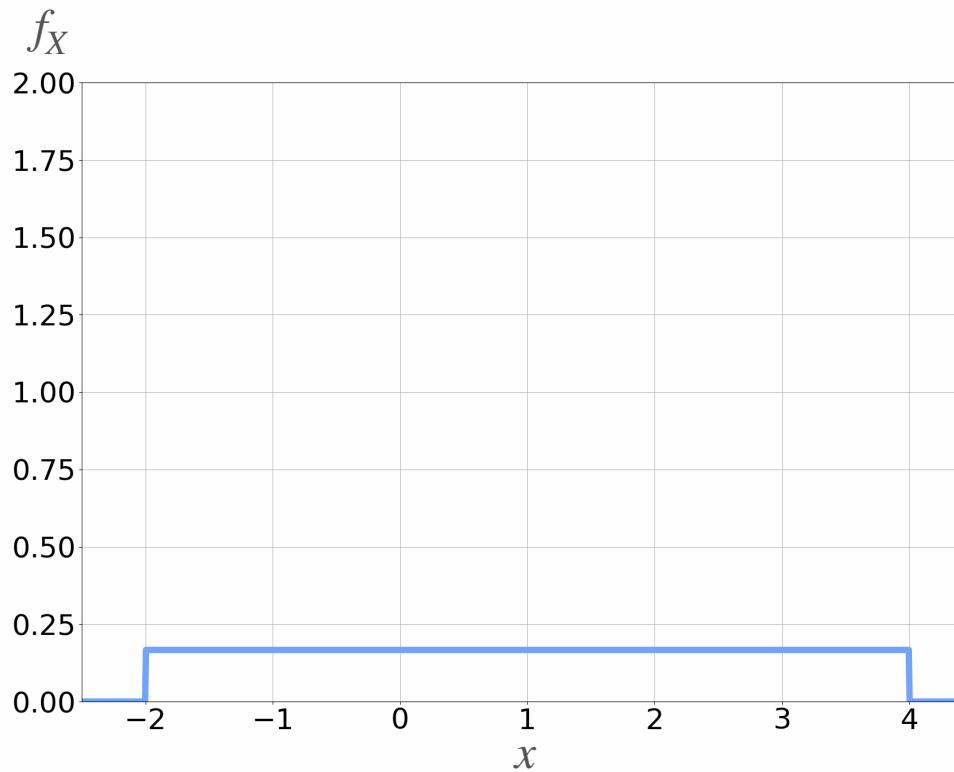
Parameters:

- $a$ : beginning of the interval
- $b$ : end of the interval

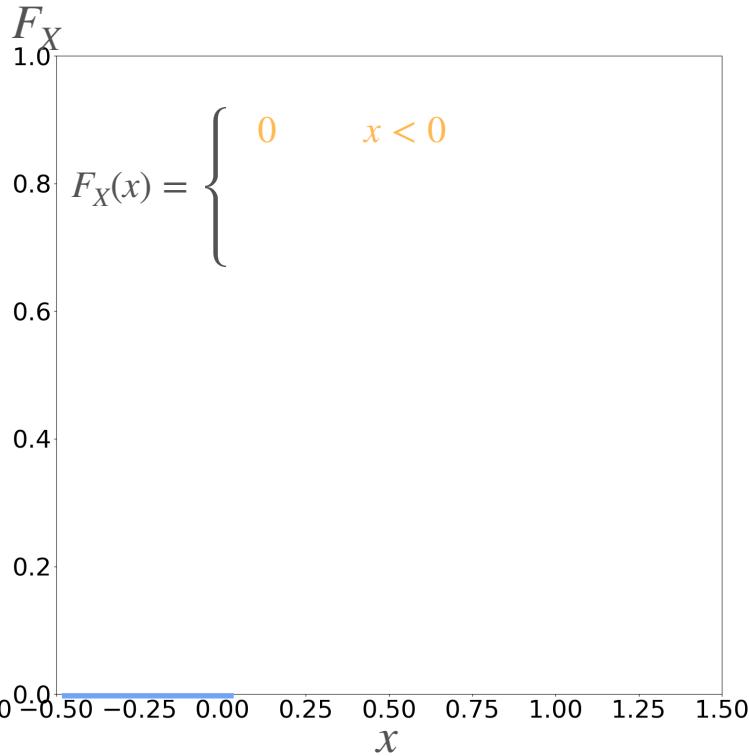
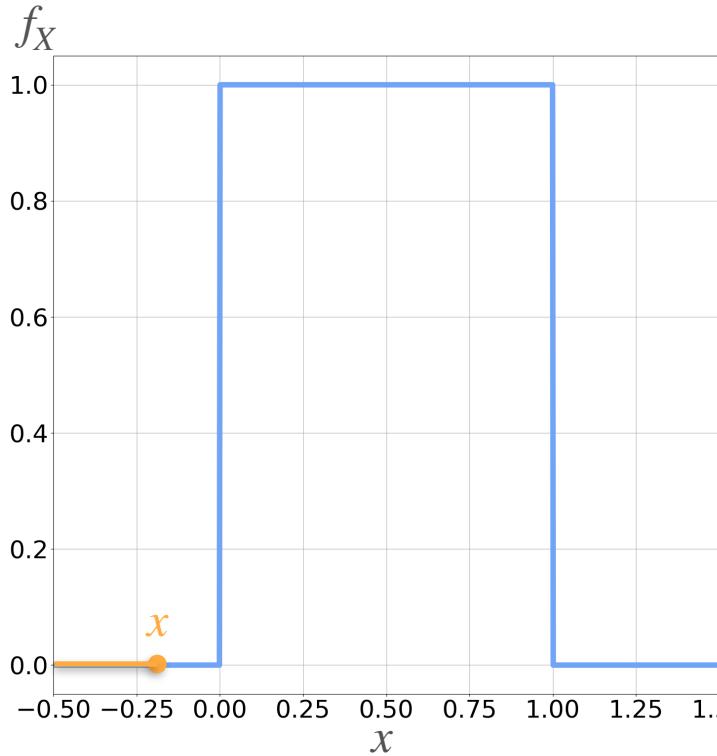
$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x \notin (a, b) \end{cases}$$



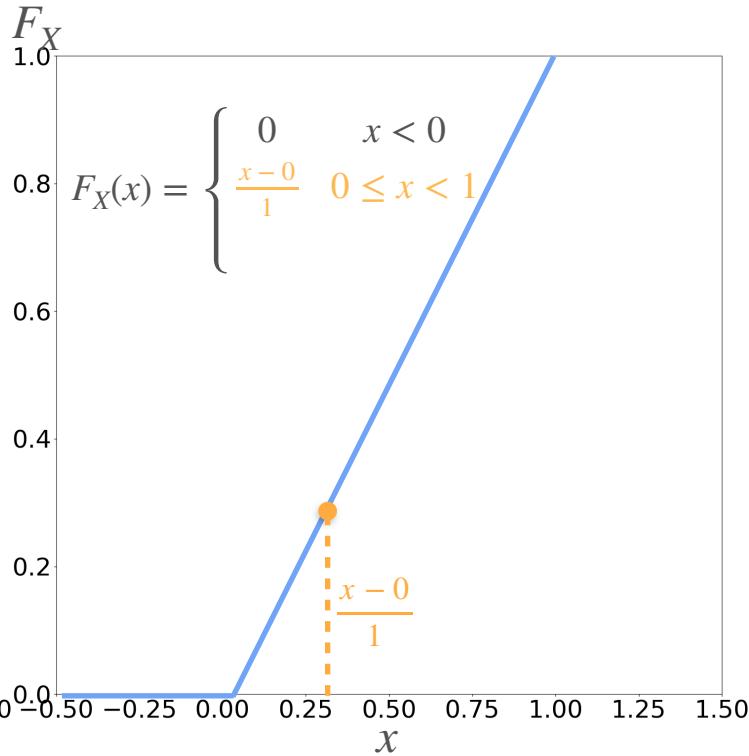
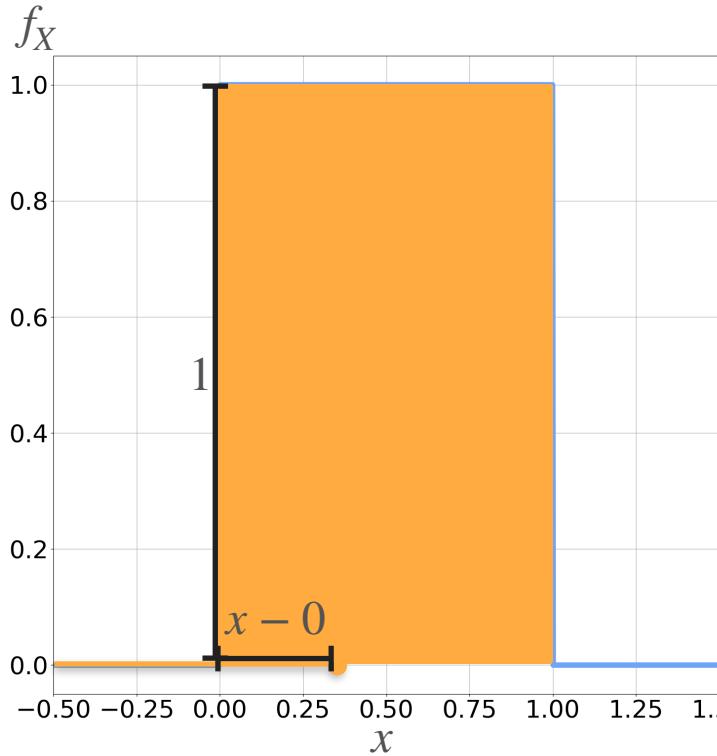
# Uniform Distribution: PDF



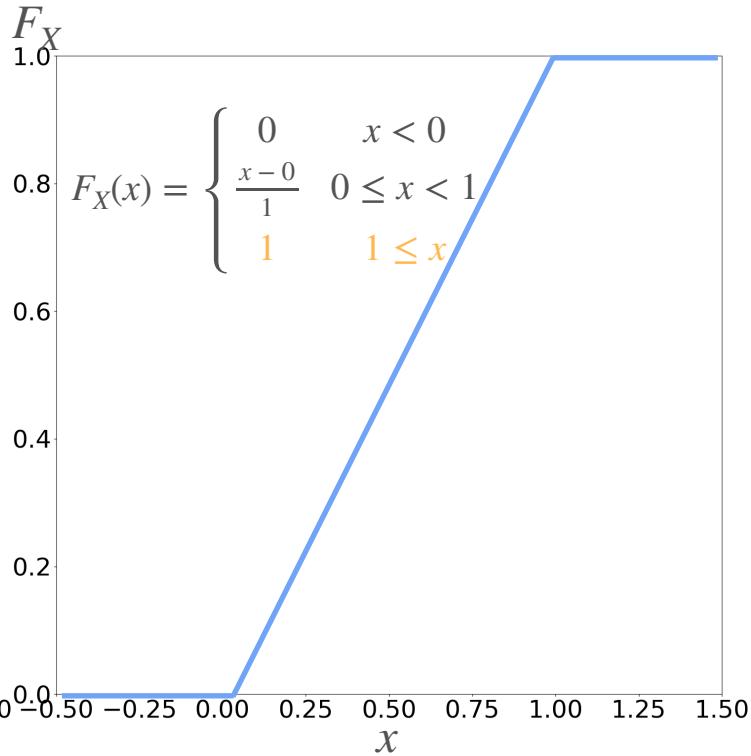
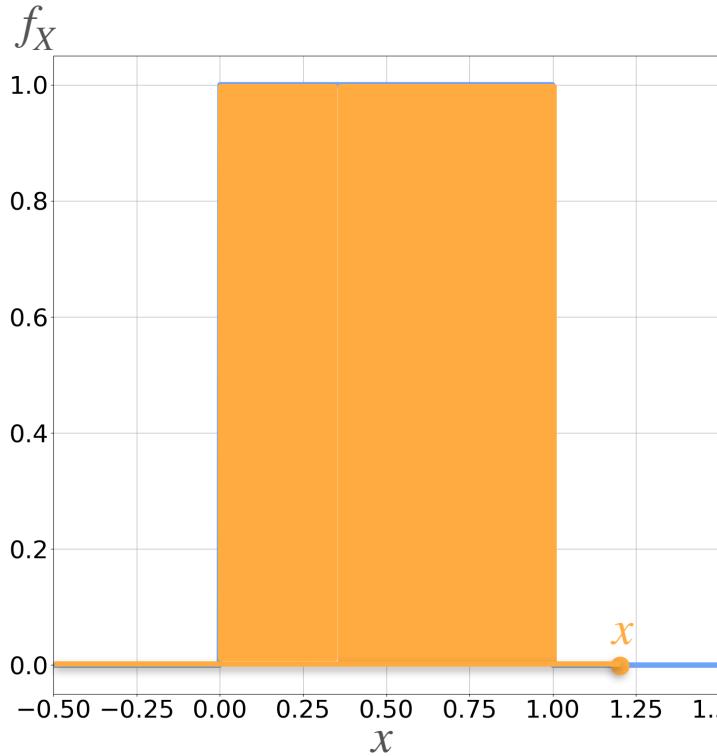
# Uniform Distribution: CDF



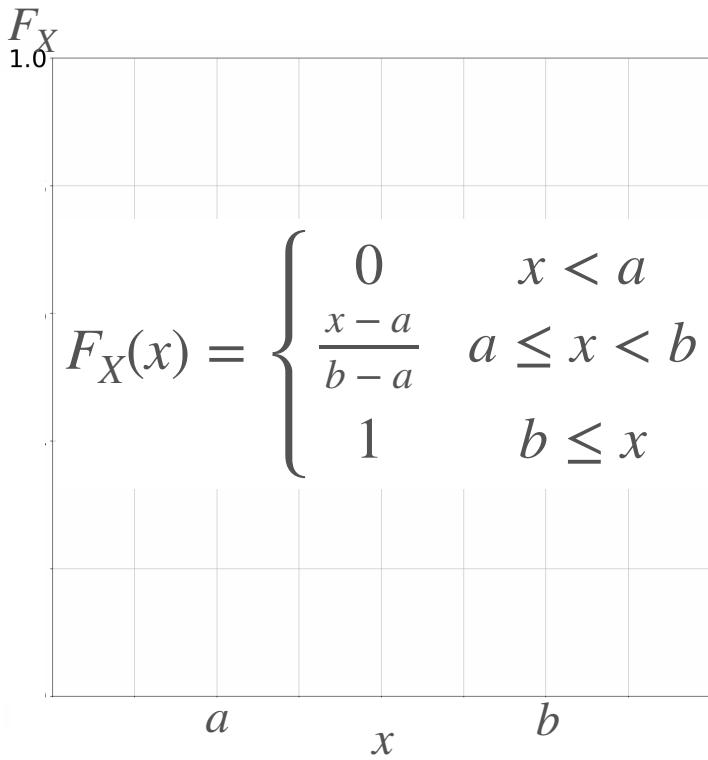
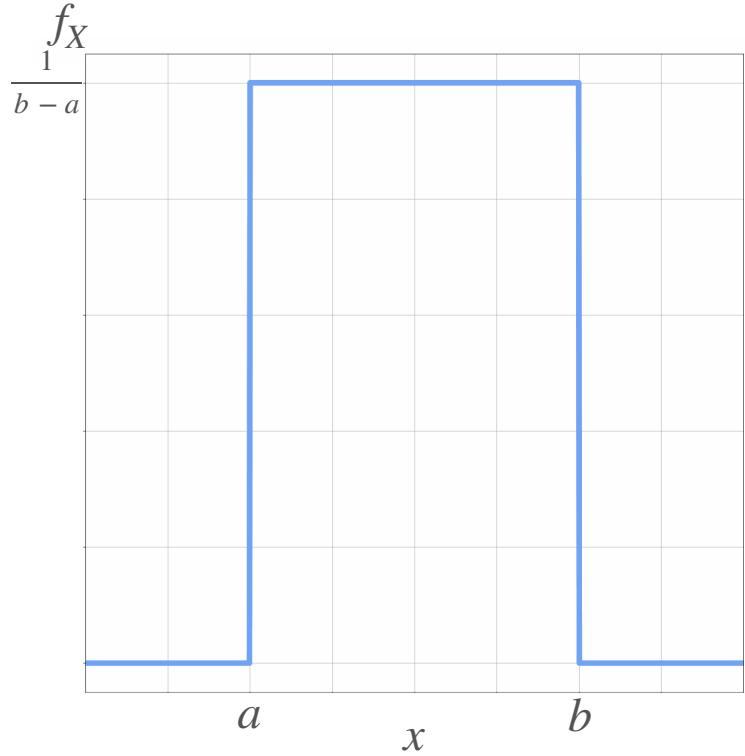
# Uniform Distribution: CDF



# Uniform Distribution: CDF



# Uniform Distribution: CDF





DeepLearning.AI

# Probability Distributions

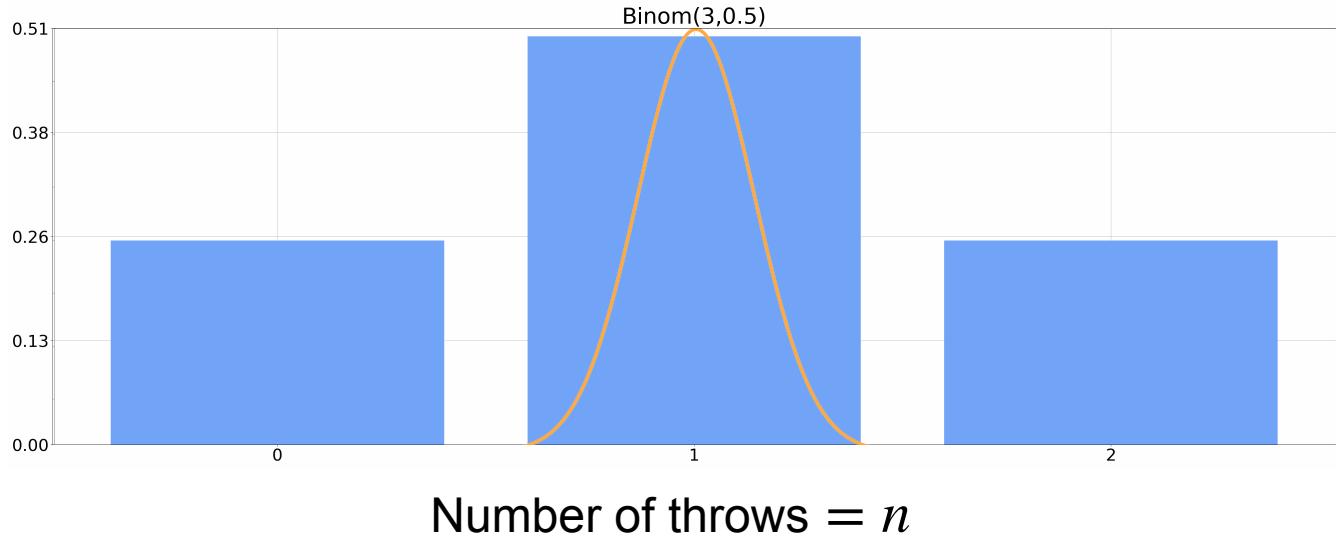
---

## Normal distribution

# Binomial Distribution With Very Large $n$

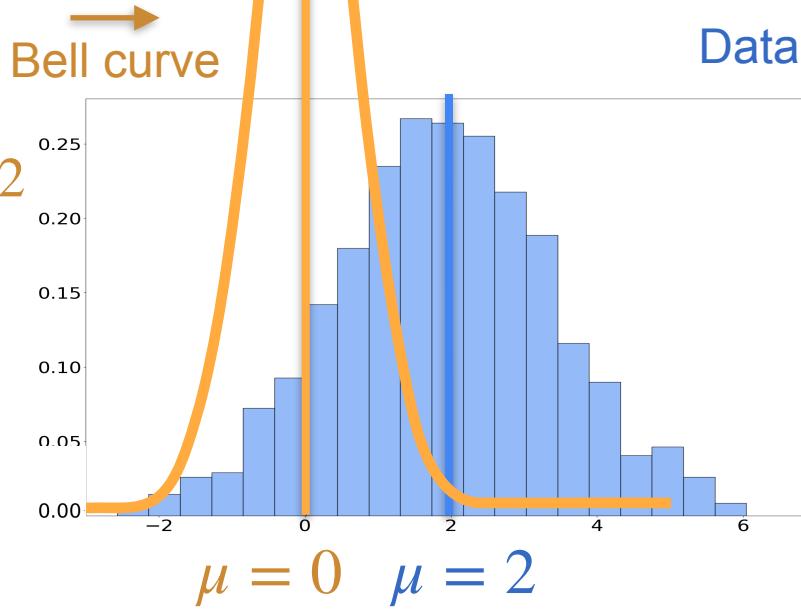


# Binomial Distribution With Very Large $n$

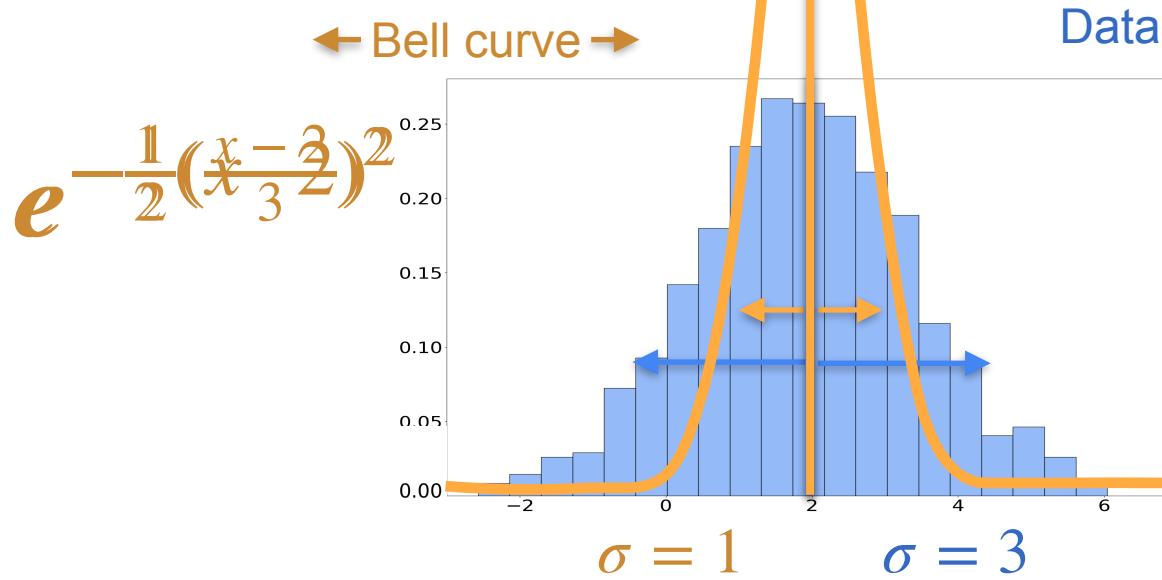


# Bell Shaped Data

$$e^{-\frac{x^2}{2}(x-2)^2}$$

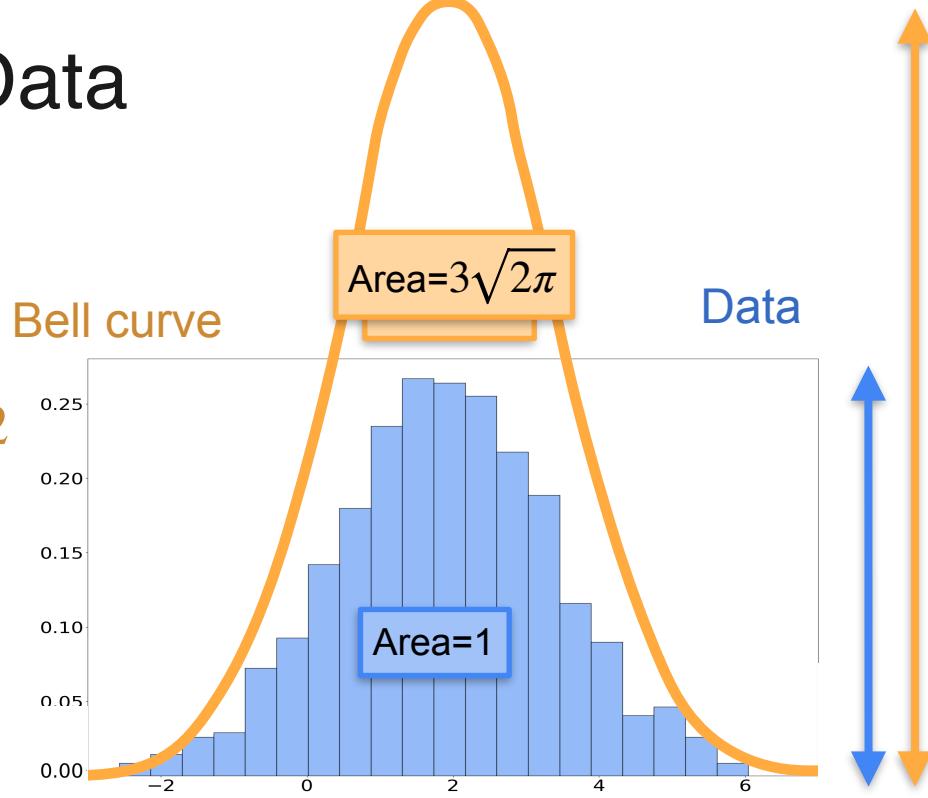


# Bell Shaped Data



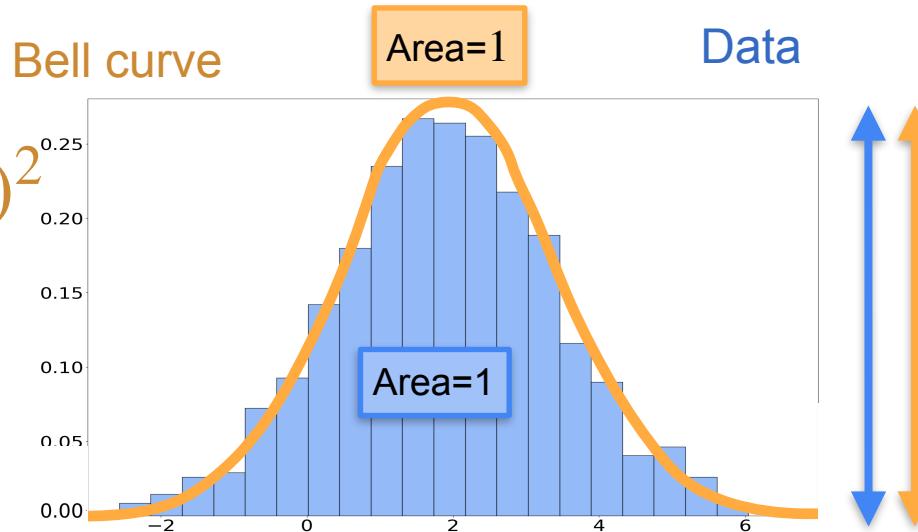
# Bell Shaped Data

$$\frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-2}{3})^2}$$



# Bell Shaped Data

$$\frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-2}{3}\right)^2}$$

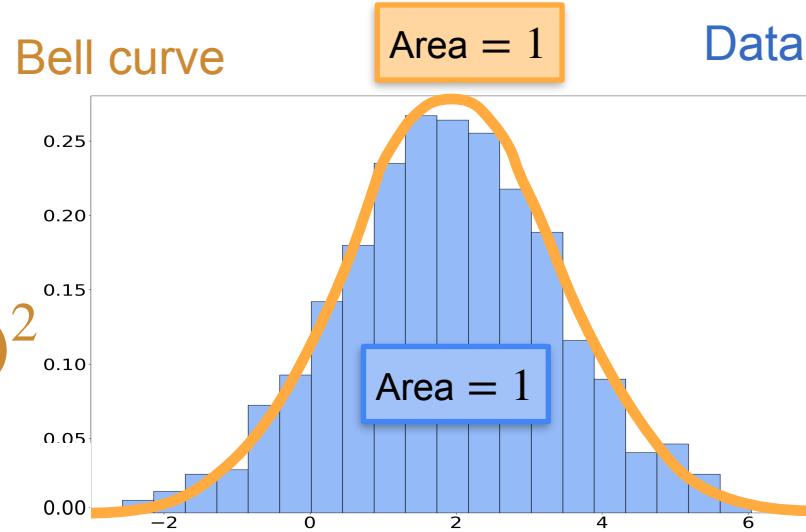


# Bell Shaped Data

Mean =  $\mu$

Standard deviation =  $\sigma$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

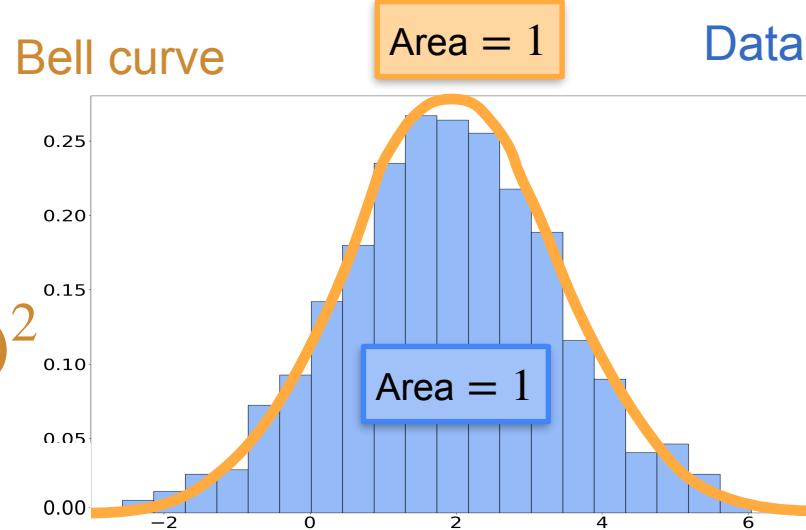


# Bell Shaped Data

Bell curve  
Mean =  $\mu$

Standard deviation =  $\sigma$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

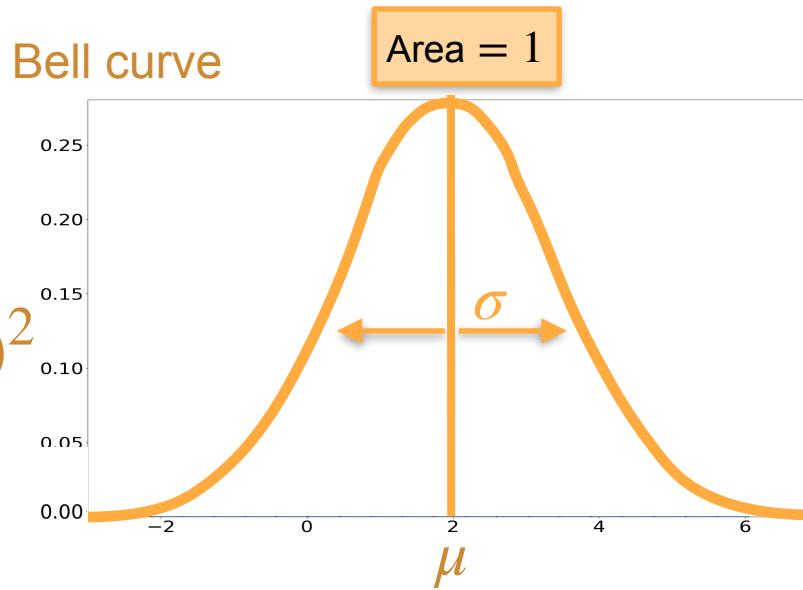


# Normal Distribution

Mean =  $\mu$

Standard deviation =  $\sigma$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



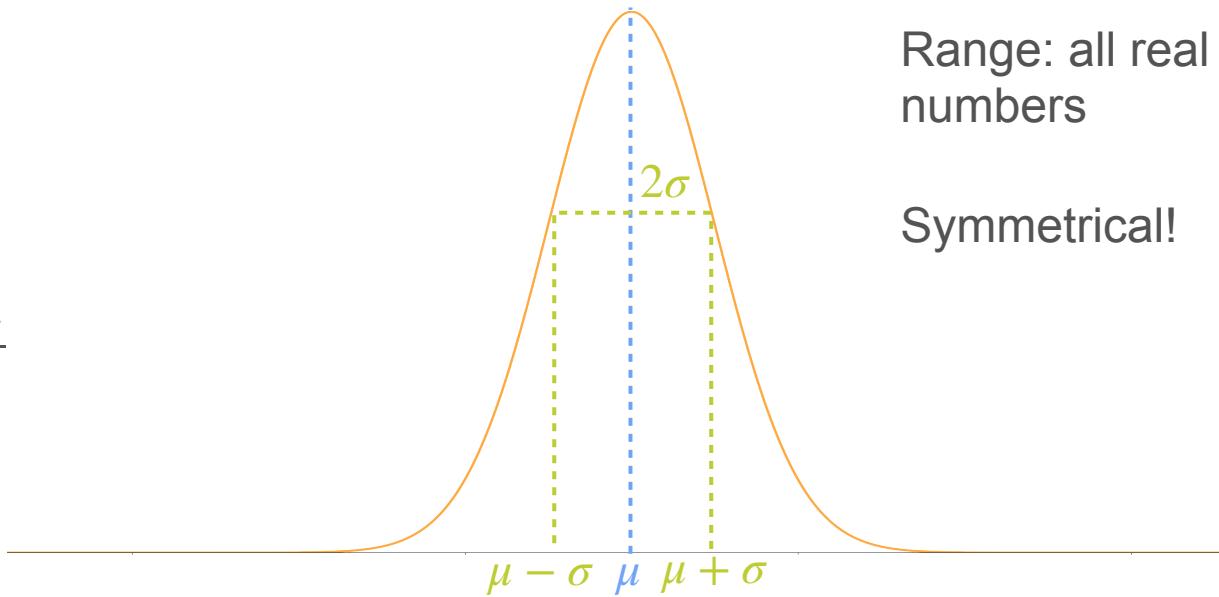
# Normal Distribution

Parameters:

- $\mu$ : center of the bell
- $\sigma$ : spread of the bell

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Scaling constant



Range: all real numbers

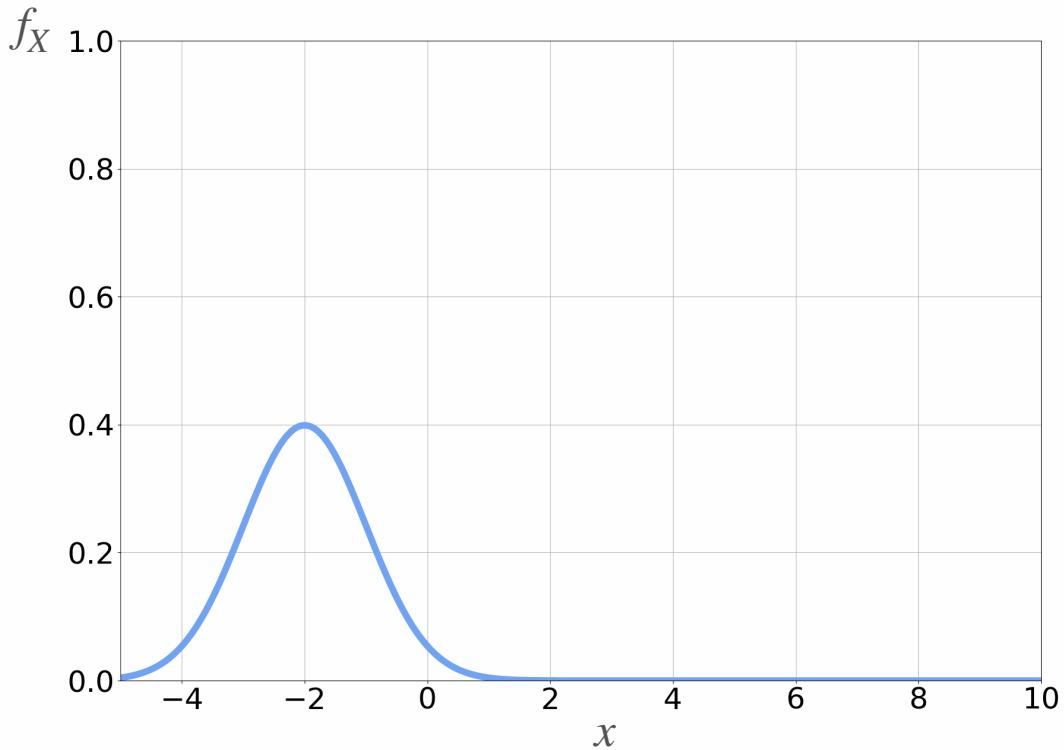
Symmetrical!

# Normal Distribution

Parameters:

- $\mu$ : center of the bell
- $\sigma$ : spread of the bell

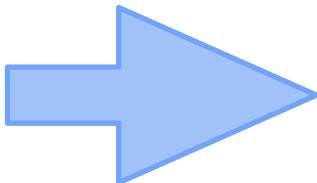
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$



# Normal Distribution - Notation

Parameters:

- $\mu$ : center of the bell
- $\sigma$ : spread of the bell



$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

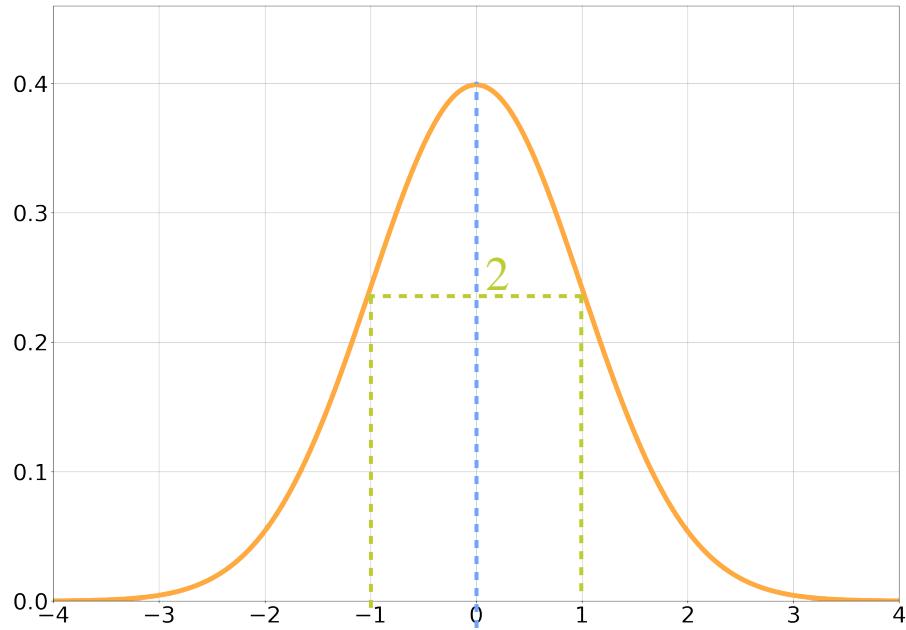
# Standard Normal Distribution

Parameters:

- $\mu$ : 0
- $\sigma$ : 1

$$X \sim \mathcal{N}(0, 1^2)$$

$$\begin{aligned}f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-0)^2}{1^2}} \\&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\end{aligned}$$



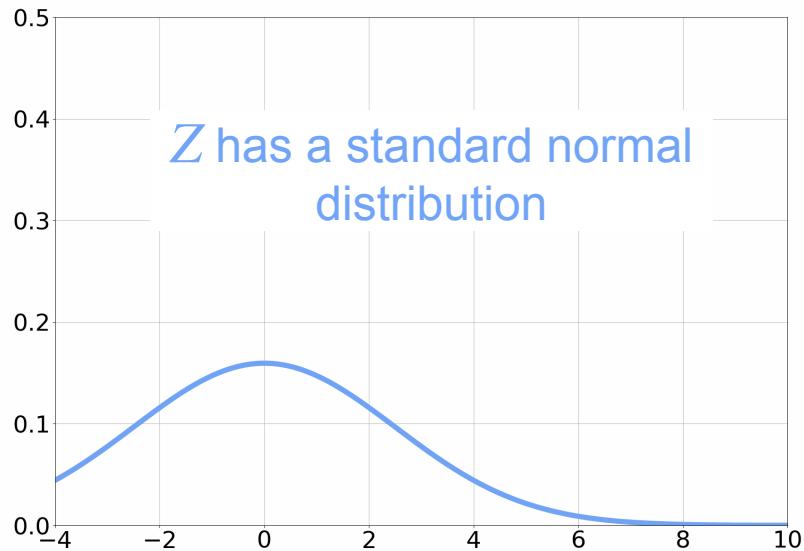
# Standardization

There's a really easy way to convert any normal distribution to the standard one!

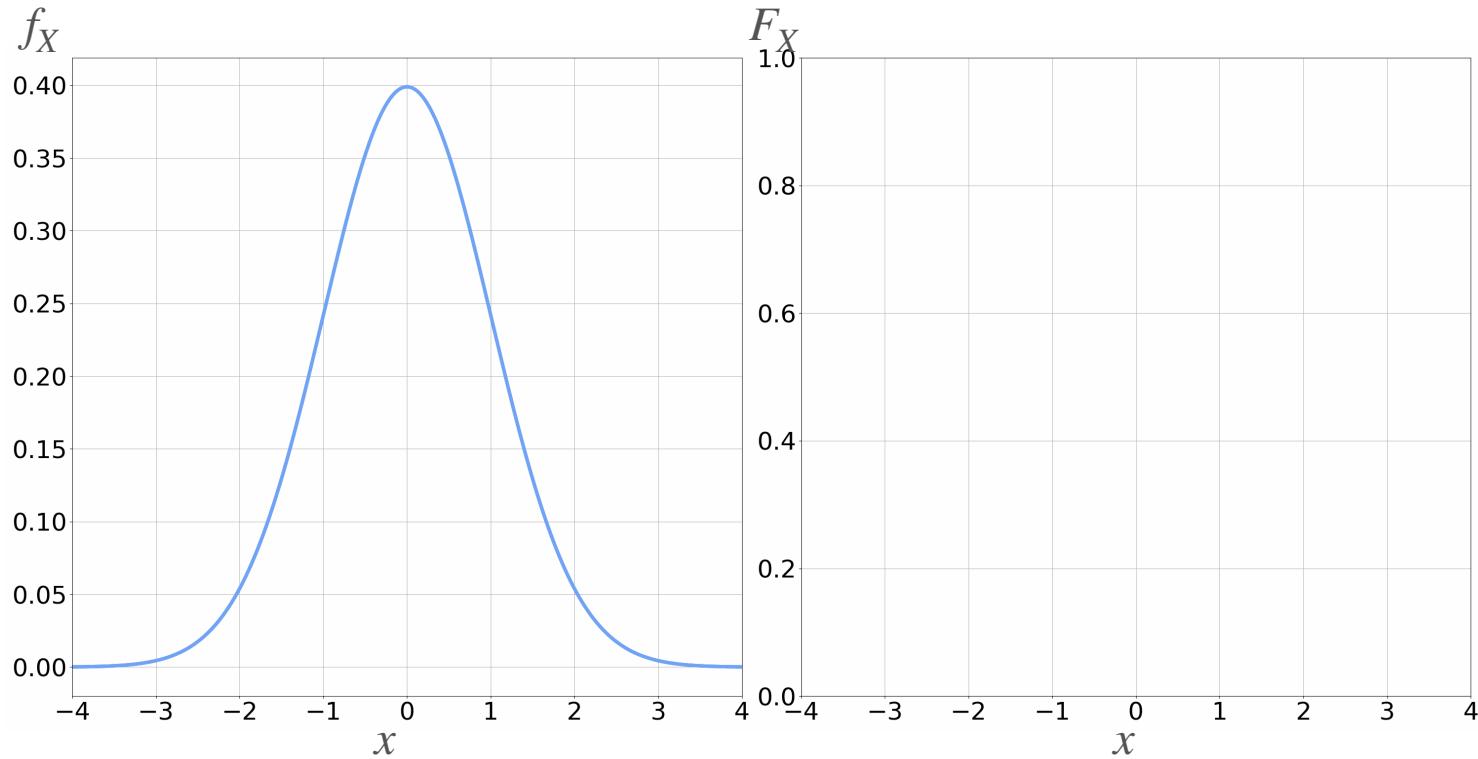
$X$  distributes normally with  
 $\mu = 2, \sigma = 2.5$

$$Z = \frac{X - \mu}{\sigma}$$

Standardization is crucial to compare variables of different magnitudes!

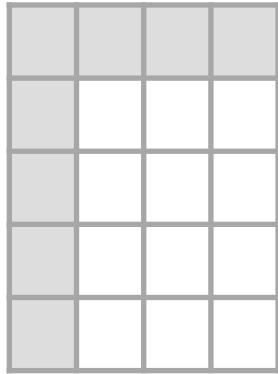


# What Does the CDF Look Like?

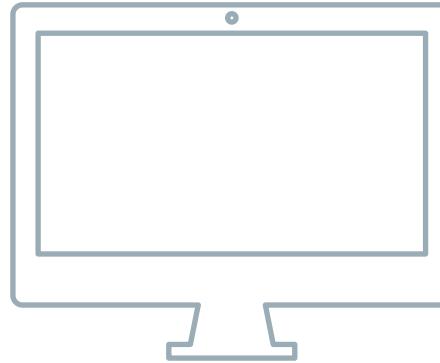


# Computing Probabilities From the PDF

This math can't be done by hand

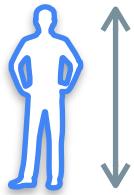


In the old days, people used tables of data



Now, you can use the help of some software to do the approximate area under the curve for you!

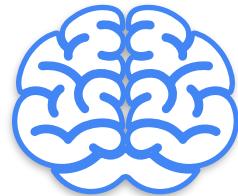
# Normal Distribution: Applications



Height



Weight



IQ



Noise in a  
communication channel

In general, characteristics that are the sum of many independent processes

Many models in ML are designed under the assumption that the variables follow a normal distribution



DeepLearning.AI

# Probability Distributions

---

**(Optional)**  
**Chi-squared distribution**

# Chi-Square Distribution: Motivation



Communication channel

Noise

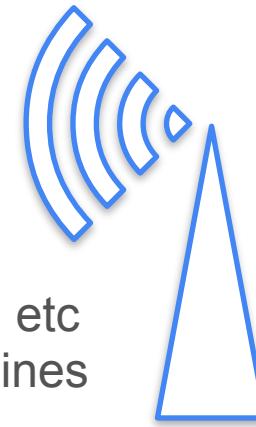
Interference from other devices

Obstructions like walls, trees, etc.

Atmospheric conditions: rain, humidity, etc

Electrical interference, i.e. from power lines

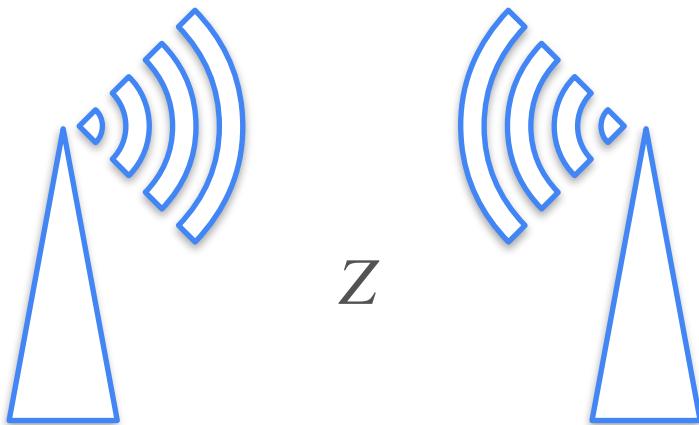
Others



Message sent: 10010

Message received: 10010 +Z

# Chi-Square Distribution: Motivation



The communication channel has noise with a standard normal distribution

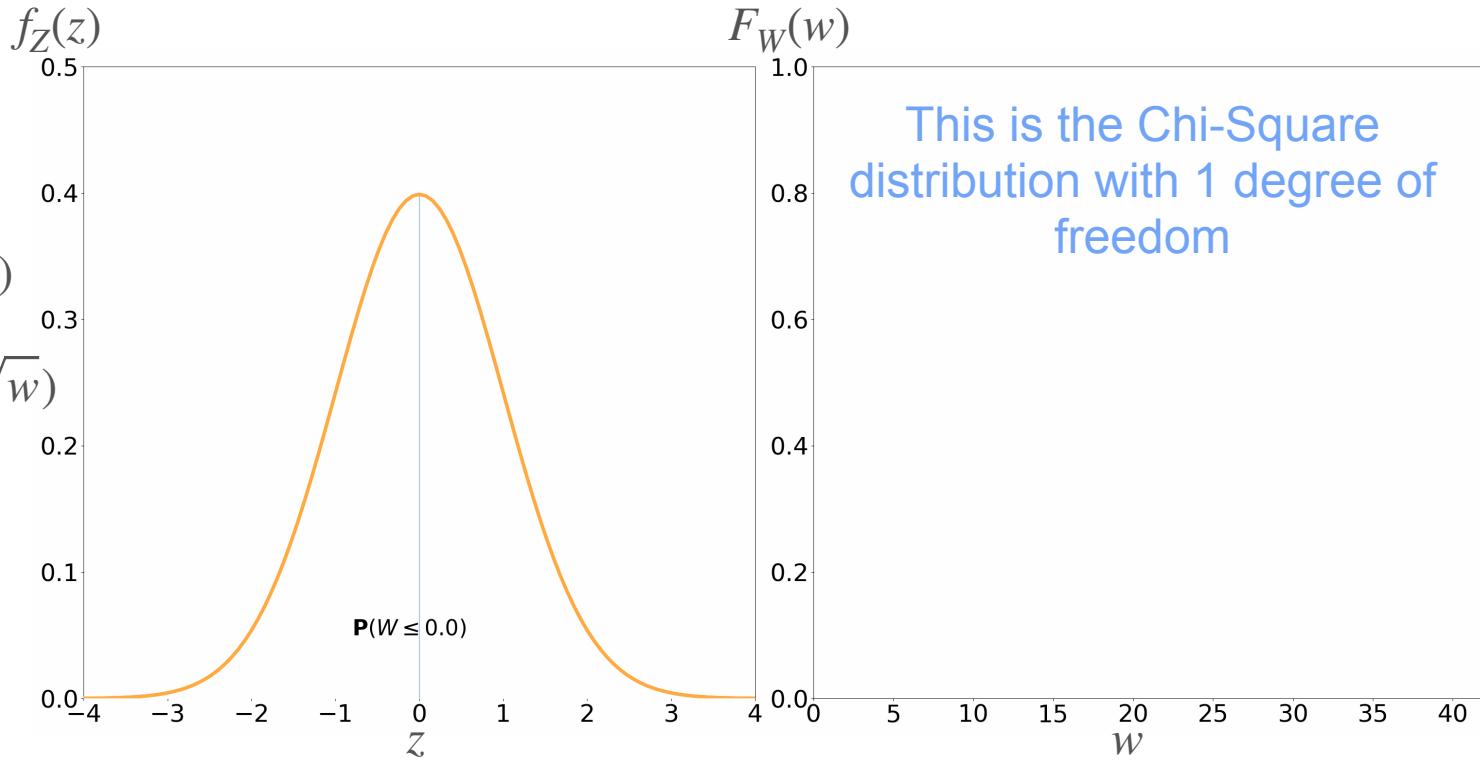
What is the **power** of the noise in the channel?

$$W = Z^2$$

What is the distribution of  $W$ ?

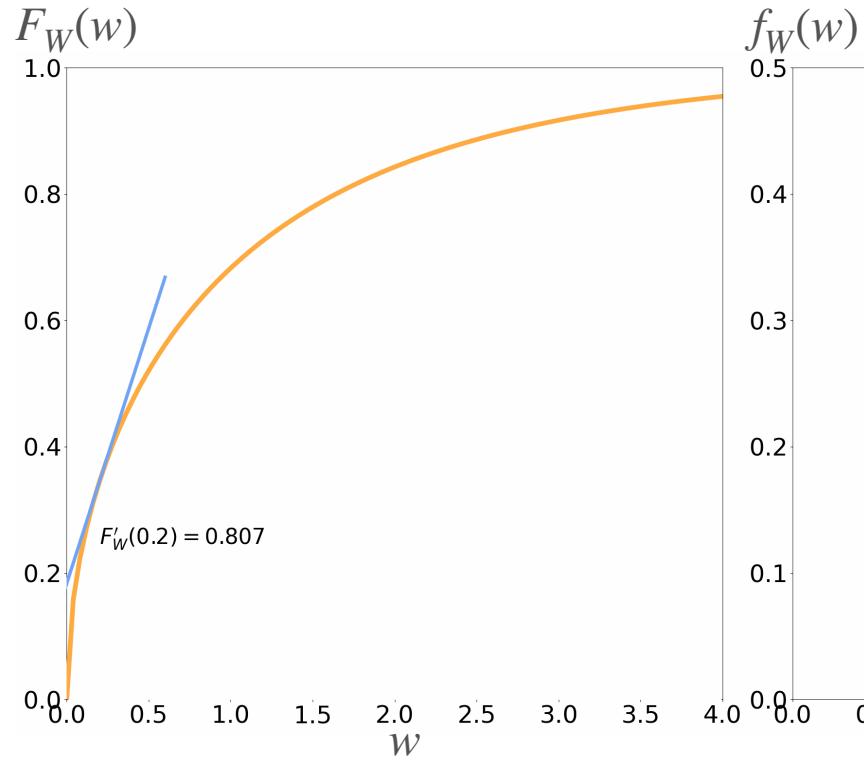
# Chi Square Distribution

$$\begin{aligned}F_W(w) &= \mathbf{P}(W \leq w) \\&= \mathbf{P}(Z^2 \leq w) \\&= \mathbf{P}(|Z| \leq \sqrt{w}) \\&= \mathbf{P}(-\sqrt{w} \leq Z \leq \sqrt{w})\end{aligned}$$



# Chi Square Distribution

$$f_W(w) = F'_W(w)$$



# Chi-Square Distribution

Accumulated power over 2 transmissions?

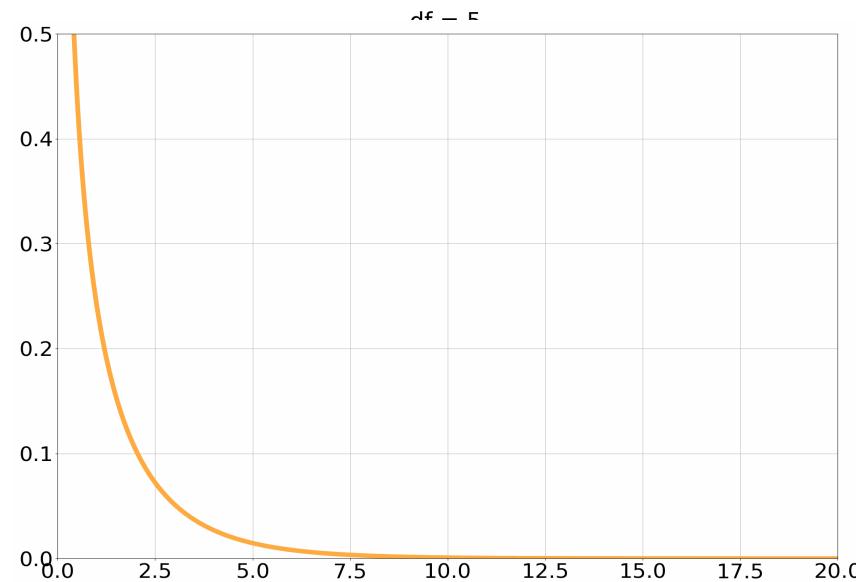
$$W_2 = Z_1^2 + Z_2^2 \quad \text{Chi-Square with 2 df}$$

Accumulated power over 5 transmissions?

$$W_5 = Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2 + Z_5^2 \quad \text{Chi-Square with 5 df}$$

Accumulated power over  $k$  transmissions?

$$W_k = \sum_{i=1}^k Z_i^2 \quad \text{Chi-Square with } k \text{ df}$$





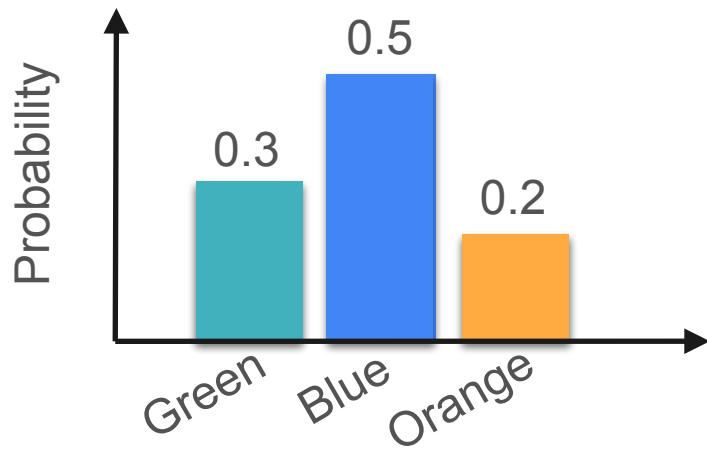
DeepLearning.AI

# Probability Distributions

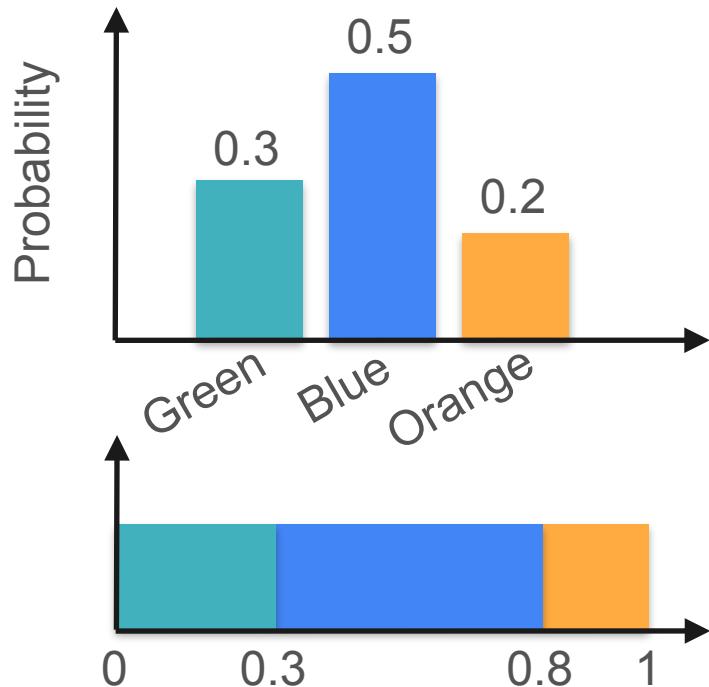
---

## Sampling from a Distribution

# Sampling From a Distribution

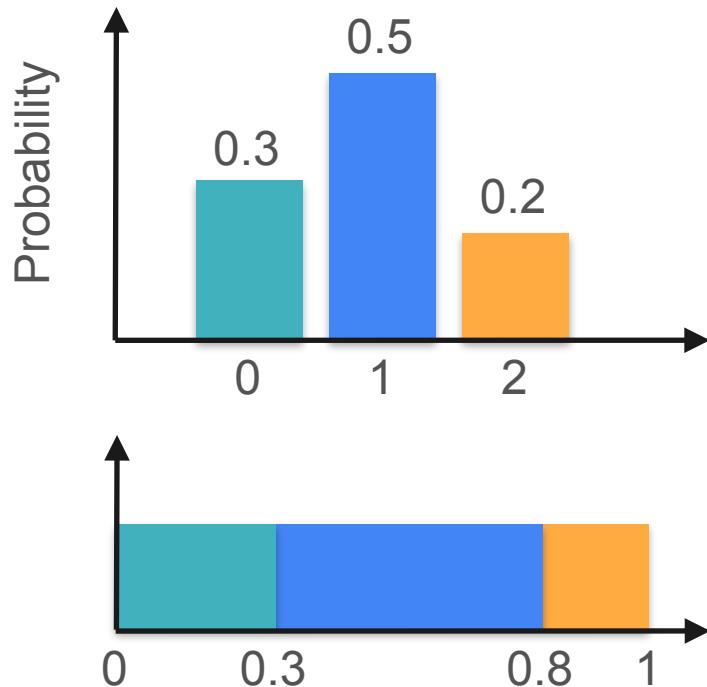


# Sampling From a Distribution



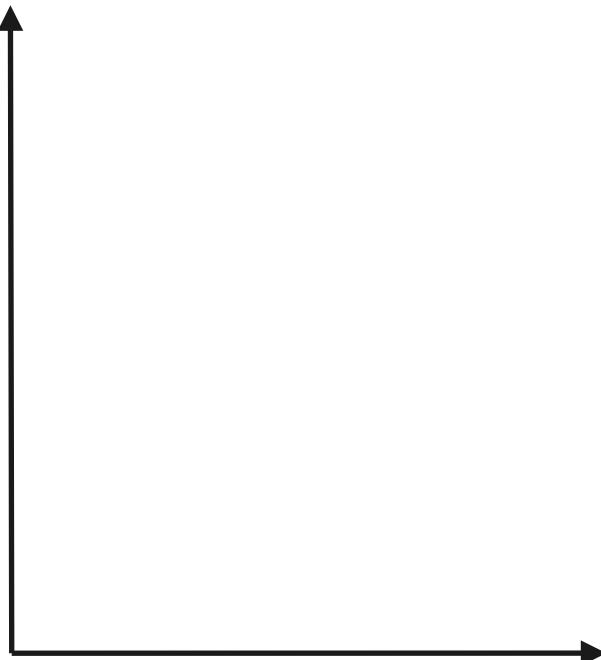
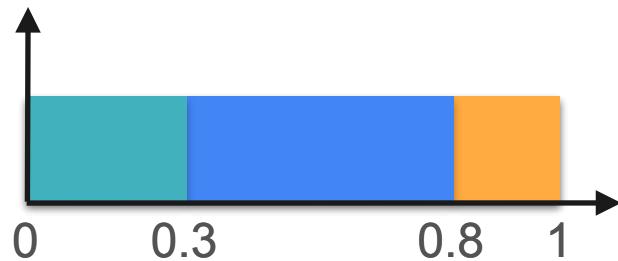
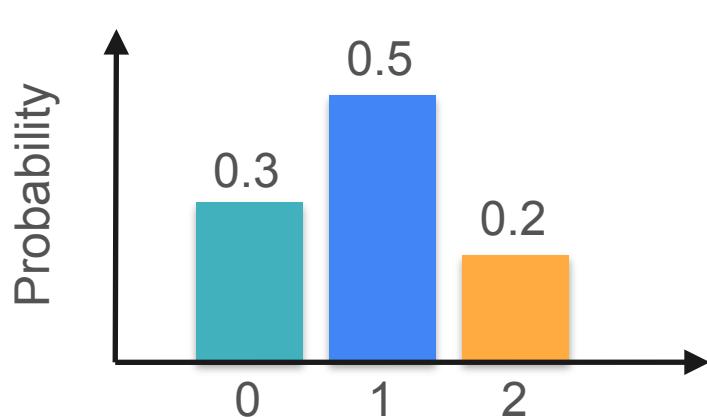
- **Step 1:** generate a random number between 0 and 1
- **Step 2:** find out which interval the number belongs to
  - $[0, 0.3)$
  - $[0.3, 0.8)$
  - $[0.8, 1]$
- **Step 3:** Assign an outcome based on the interval

# Sampling From a Distribution

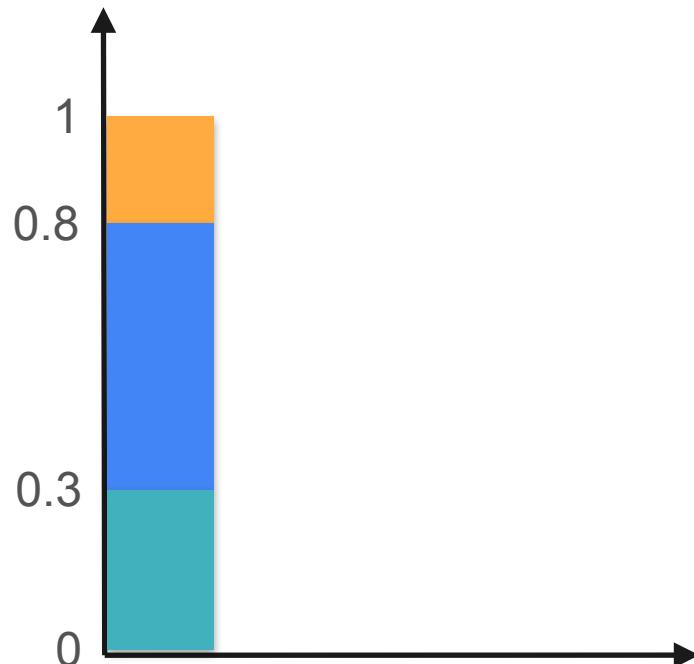
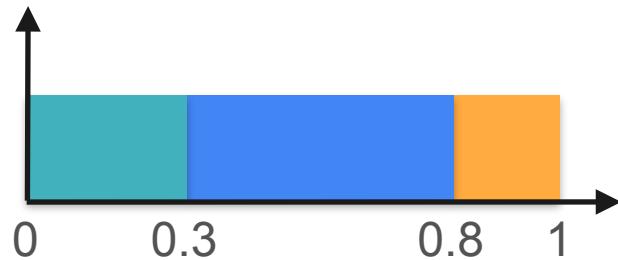
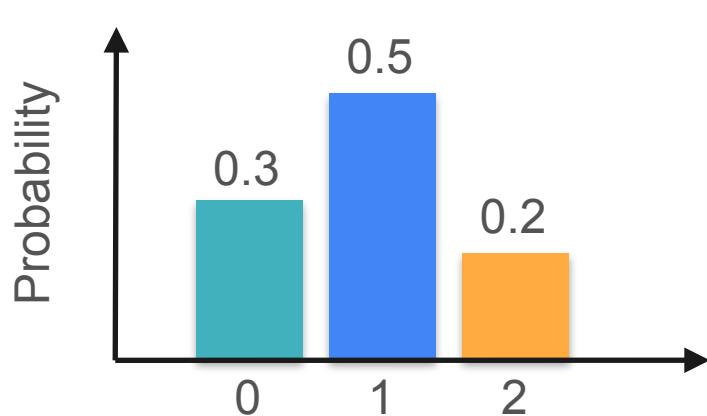


- **Step 1:** generate a random number between 0 and 1
- **Step 2:** find out which interval the number belongs to
  - [0, 0.3)
  - [0.3, 0.8)
  - [0.8, 1]
- **Step 3:** Assign an outcome based on the interval

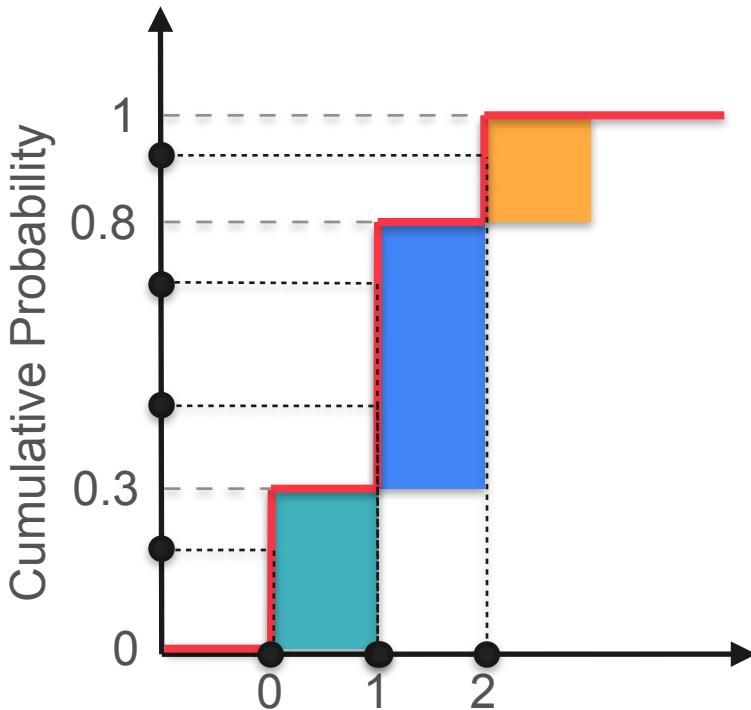
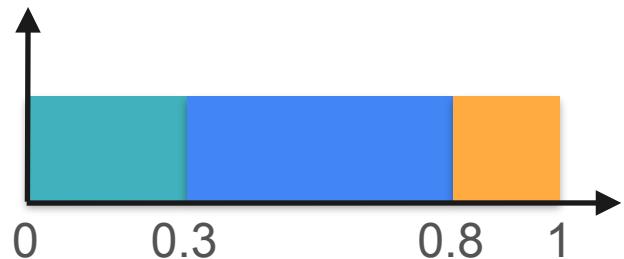
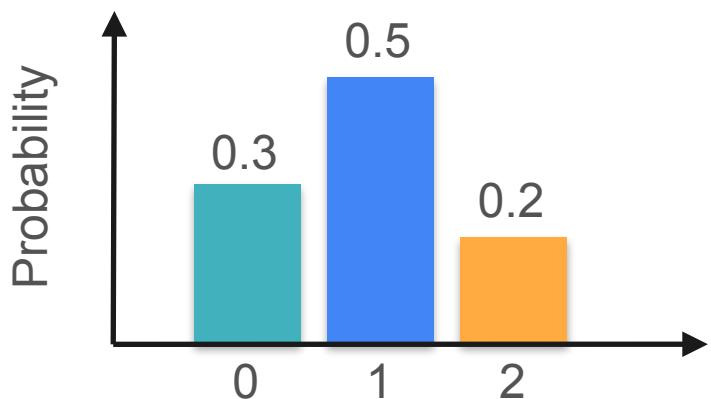
# Sampling From a Distribution



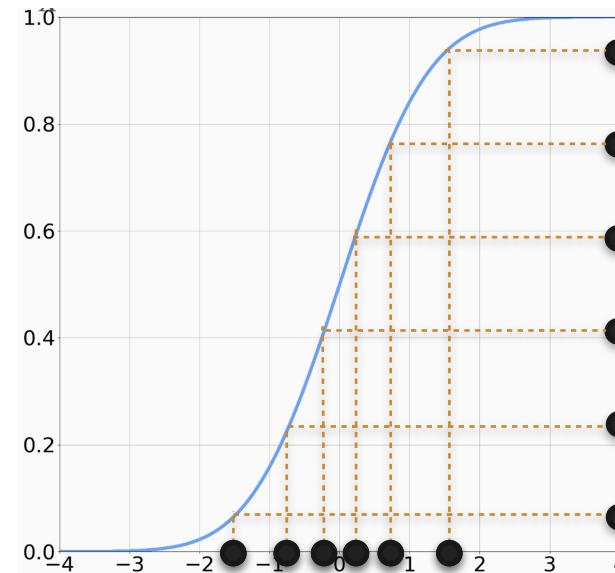
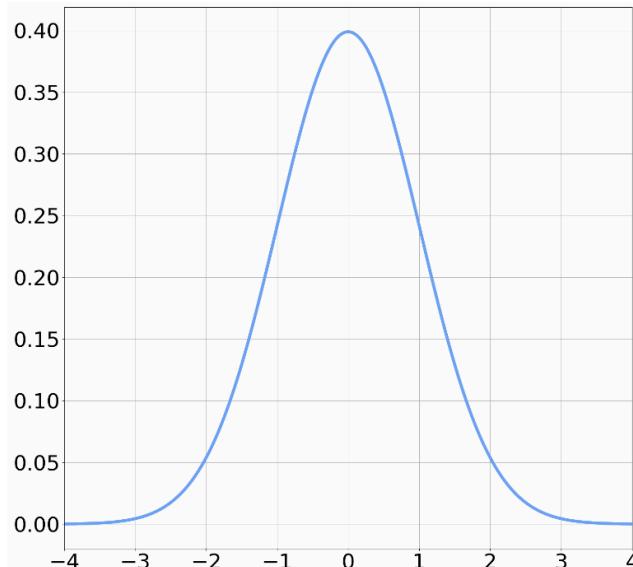
# Sampling From a Distribution



# Sampling From a Distribution



# Sampling From a Normal Distribution





DeepLearning.AI

# Probability Distributions

---

## Conclusion

# Week 1 - Conclusions

Talking head