Copyright Notice

W1 Lesson 1

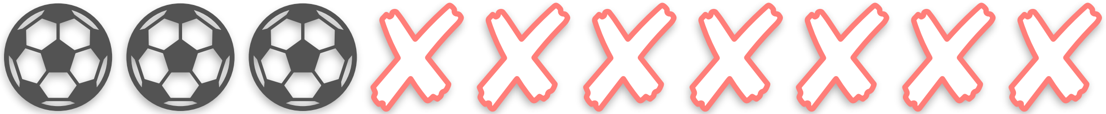
Introduction to Probability

DeepLearning.AI

Introduction to probability

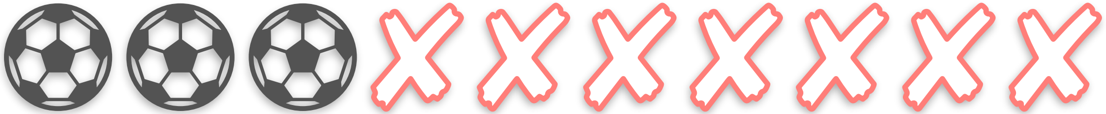
What is Probability?

Introduction to Probability



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

Introduction to Probability

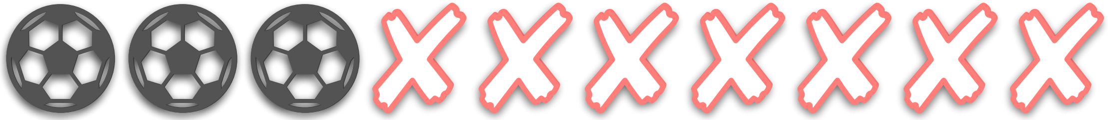


Find the probability that a child picked at random plays soccer.

The probability that a child picked at random plays soccer.

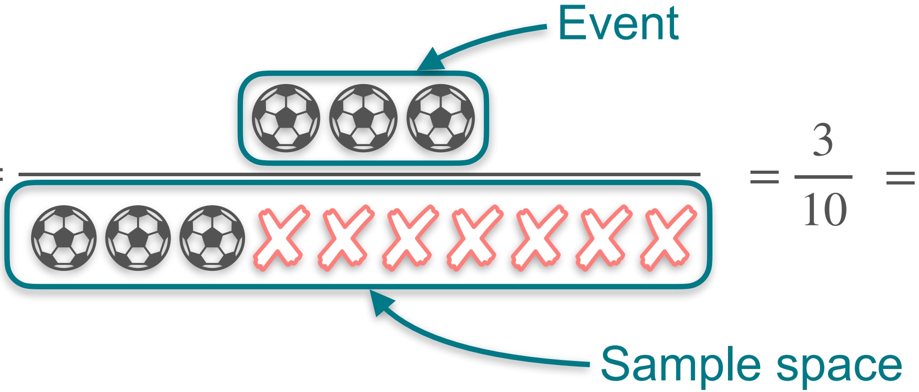
$$\mathbf{P}(\text{soccer})$$

Introduction to Probability



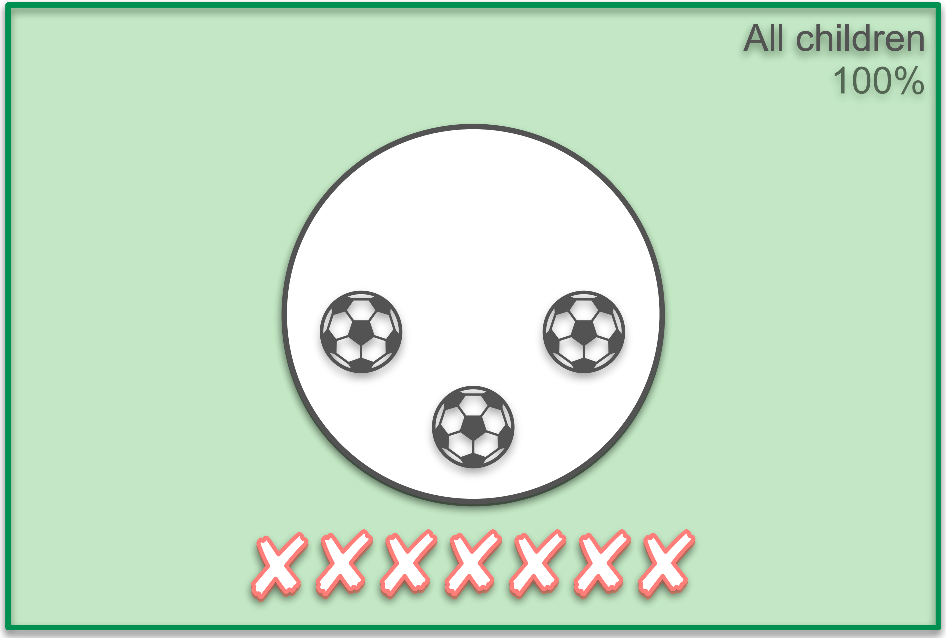
Find the probability that a child picked at random plays soccer.

$$P(\text{soccer}) = \frac{\text{soccer}}{\text{total}} =$$



$$= \frac{3}{10} = 0.3$$

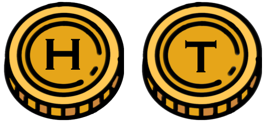
Introduction to Probability: Venn Diagram



Sample Space

Event

Introduction to Probability: Coin Example 1



Experiment

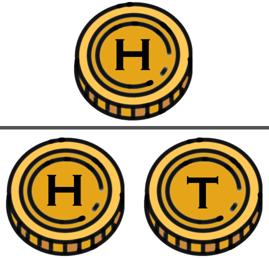
Probability of landing on heads

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

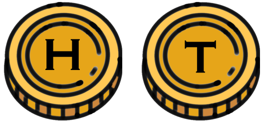
Introduction to Probability: Coin Example 1

50%   50%

$$\mathbf{P}(\text{heads}) = \frac{\text{H}}{\text{H} \quad \text{T}} = \frac{1}{2} = 0.5$$

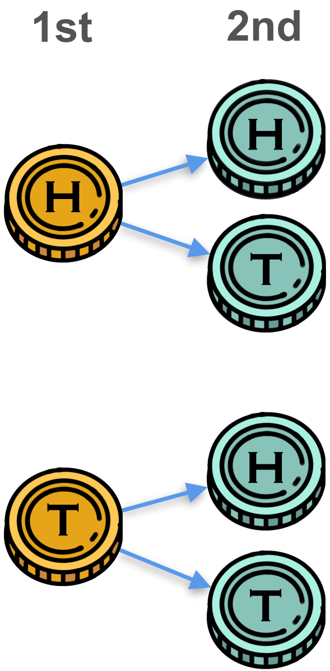


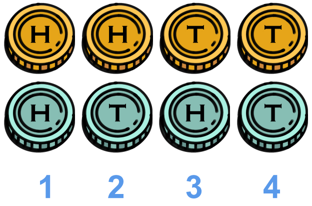
Introduction to Probability: Coin Example 2



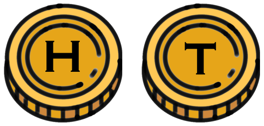
50%   50%

What is the probability of landing on heads twice?

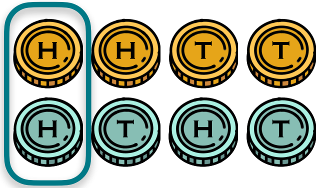




Introduction to Probability: Coin Example 2



50% 50%

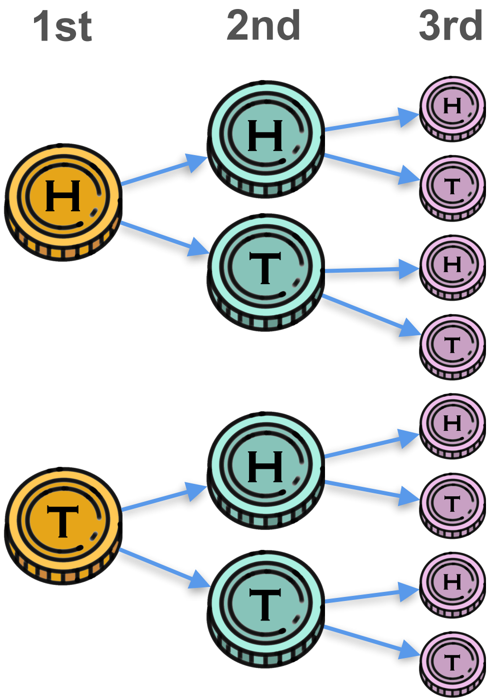


$$\mathbf{P}(HH) = \underline{\hspace{8cm}} = \frac{1}{4} = 0.25$$

Introduction to Probability: Coin Example 3

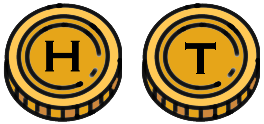
50% 50%

What is the probability of landing on heads 3 times?

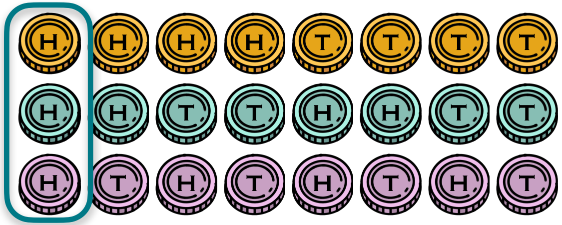


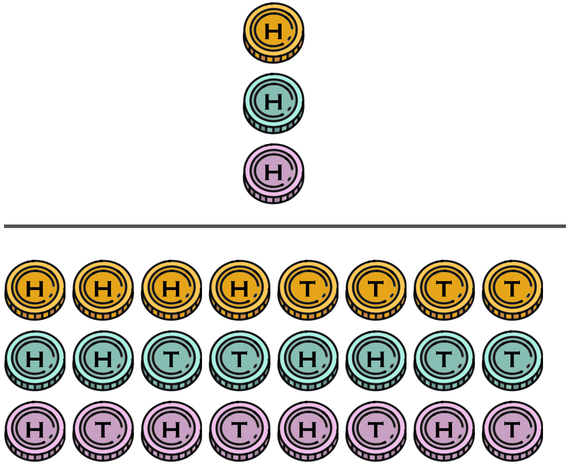
1 2 3 4 5 6 7 8

Introduction to Probability: Coin Example 3



50% 50%



$\mathbf{P}(HHH) =$ 

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

Introduction to Probability

What is Probability? - Dice Example

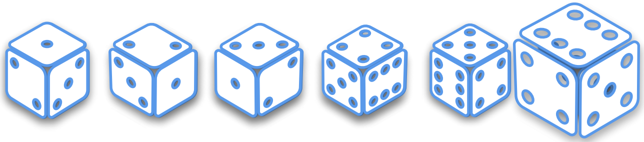
DeepLearning.AI

Introduction to Probability: Dice Example 1

What is the probability of obtaining 6?

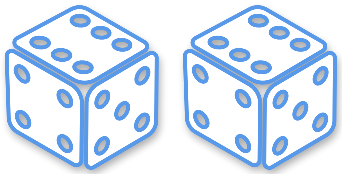
Introduction to Probability: Dice Example 1

What is the probability of obtaining 6?



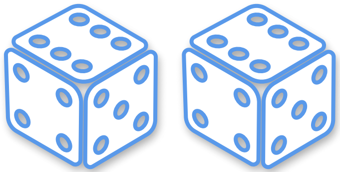
$$\mathbf{P}(6) = \frac{}{} = \frac{1}{6}$$

Introduction to Probability: Dice Example 2



What is the probability of obtaining 6,6?

Introduction to Probability: 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

What is the probability of obtaining 6,6?

$$\mathbf{P}(6,6) = \frac{}{\text{all 36 outcomes}} = \frac{1}{36}$$

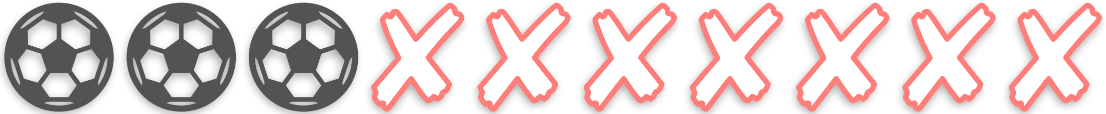
1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

DeepLearning.AI

Introduction to probability

Complement of Probability

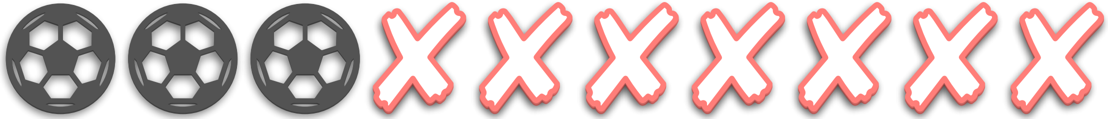
Complement of Probability



30%

What is the probability of a child NOT playing soccer?

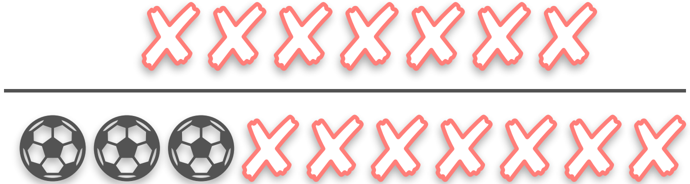
Complement of Probability



30%

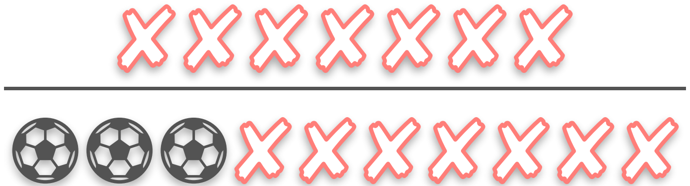
What is the probability of a child NOT playing soccer?

$$\mathbf{P}(\text{not soccer}) = \frac{\text{not soccer}}{\text{total}} =$$

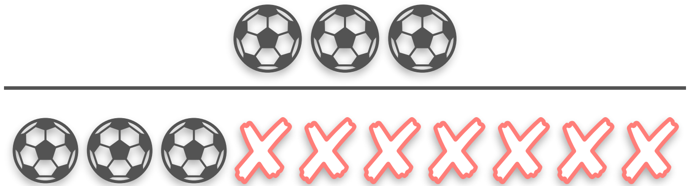


$$= \frac{7}{10} = 0.7$$

Complement of Probability



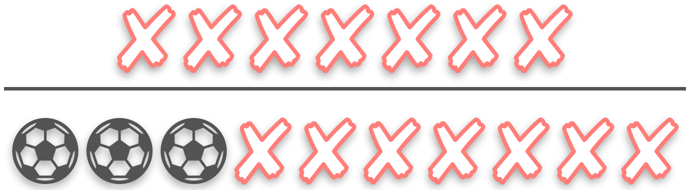
$P(\text{not soccer})$

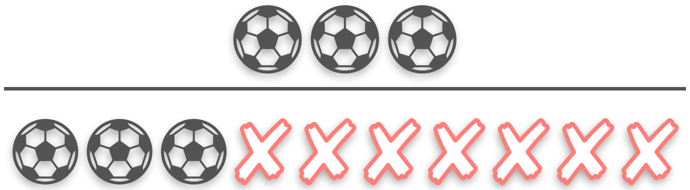


$P(\text{soccer})$

$$0.7 + 0.3 = 1$$

Complement of Probability





$\mathbf{P}(\text{not soccer})$ $\quad$ $\mathbf{P}(\text{soccer})$

$$0.7 = 1 - 0.3$$

Complement Rule

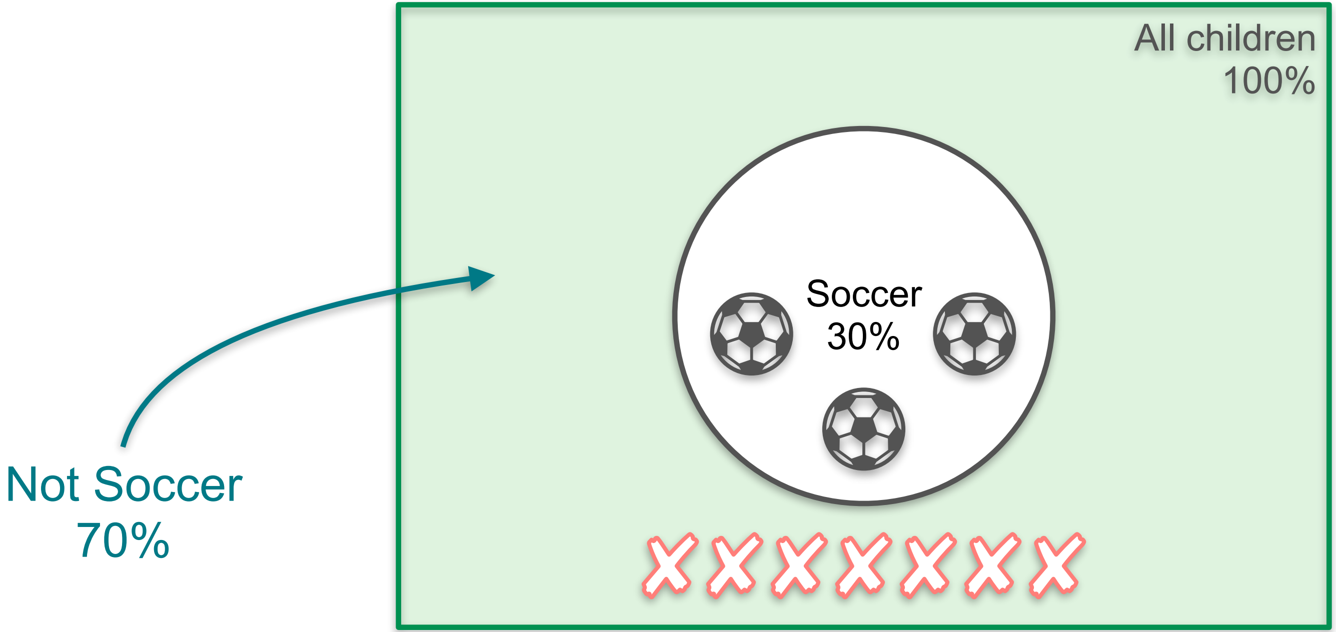
Complement of Probability

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

$$\mathbf{P}(A') = 1 - \mathbf{P}(A)$$

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?

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

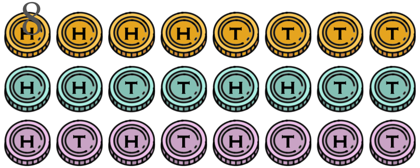
$\mathbf{P}(HHH) =$

Complement of Probability: Coin Example 1

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

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

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



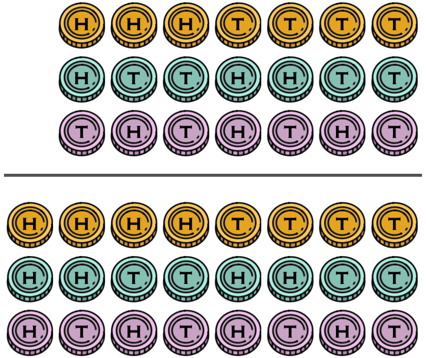
Complement of Probability: Coin Example 1

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

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

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

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

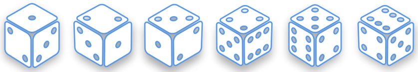


Complement of Probability: Dice Example 1

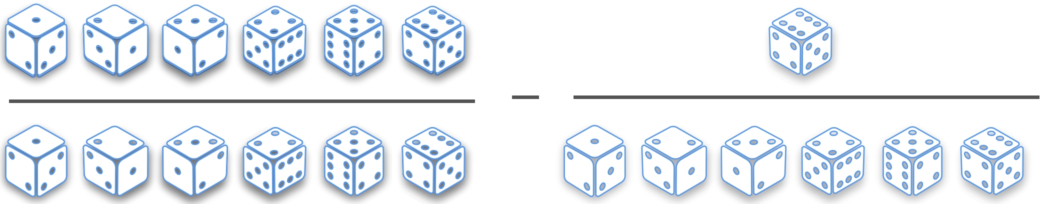
What is the probability of obtaining anything other than 6?

Complement of Probability: Dice Example 1

What is the probability of obtaining anything other than 6?



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



$=$ ______

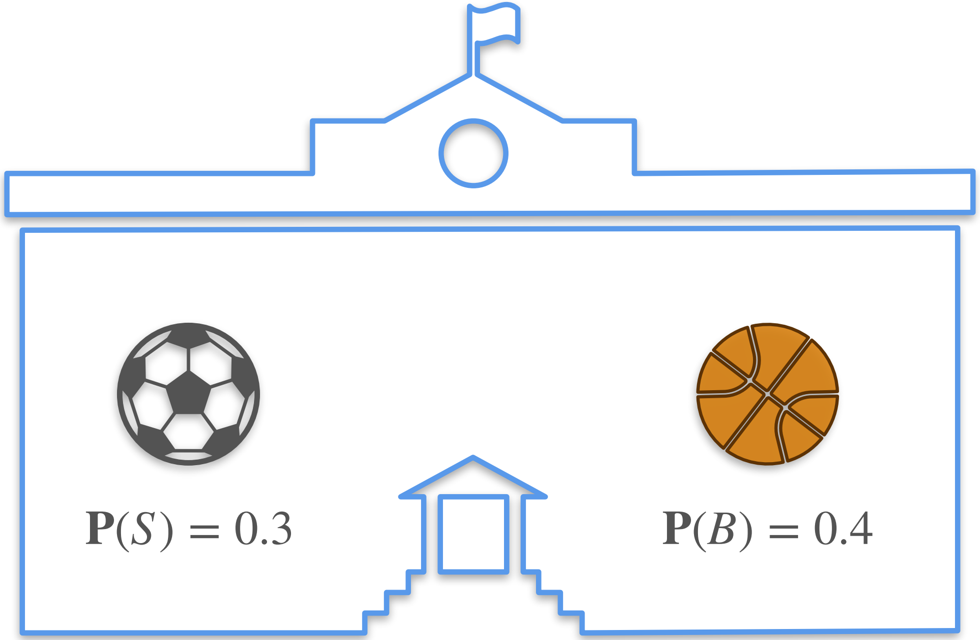
$$= \frac{5}{6}$$

DeepLearning.AI

Introduction to probability

Sum of Probabilities (Disjoint Events)

Sum of Probabilities: Quiz 1

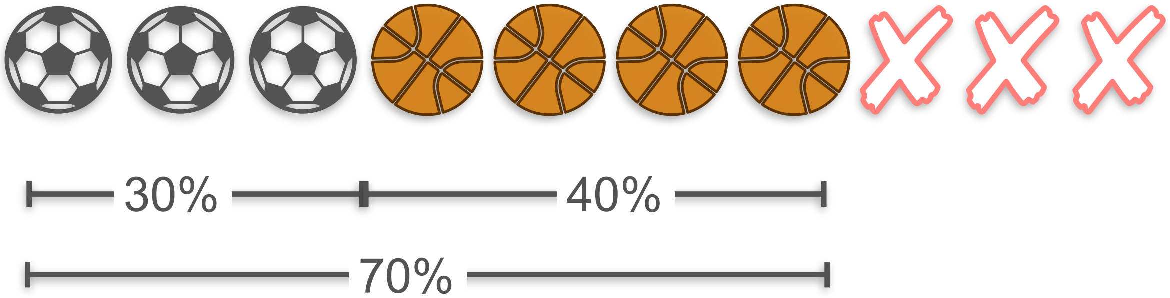


At a school, kids can only play one sport.

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

Hint: What if there were only 10 kids?

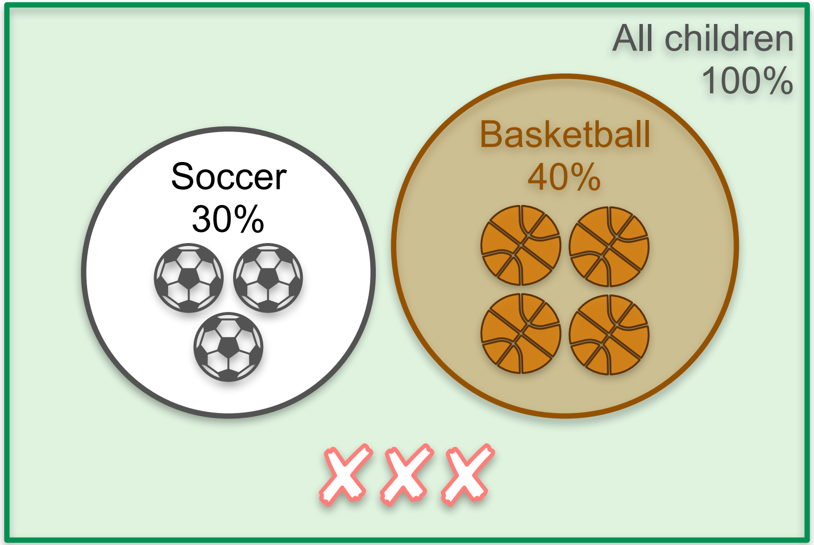
Sum of Probabilities: Quiz 1 Solution



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

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

Sum of Probabilities: Quiz 1 Solution



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

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$$

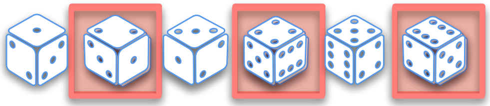
Sum of Probabilities: Dice Example 1

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

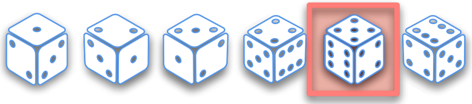
Sum of Probabilities: Dice Example 1

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

A



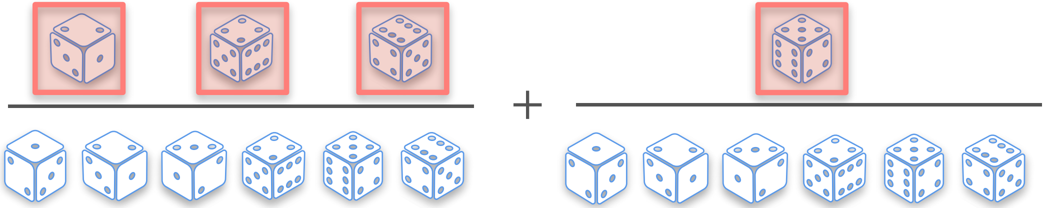
B



______ + ______

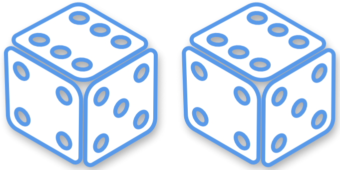
Sum of Probabilities: Dice Example 1

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



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

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
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

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

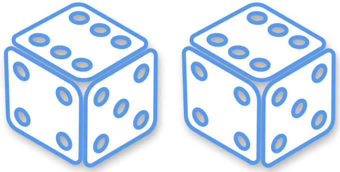
$\mathbf{P}(\text{sum of 7 or sum of 10})$

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

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

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

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
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

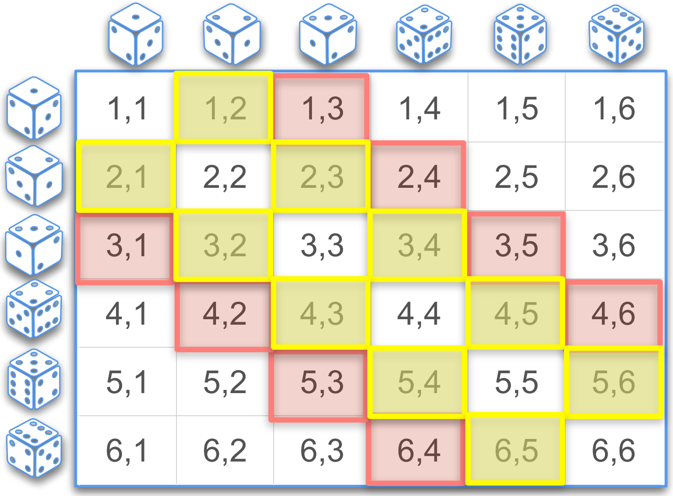
diff = 1

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

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
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 3

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

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

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

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

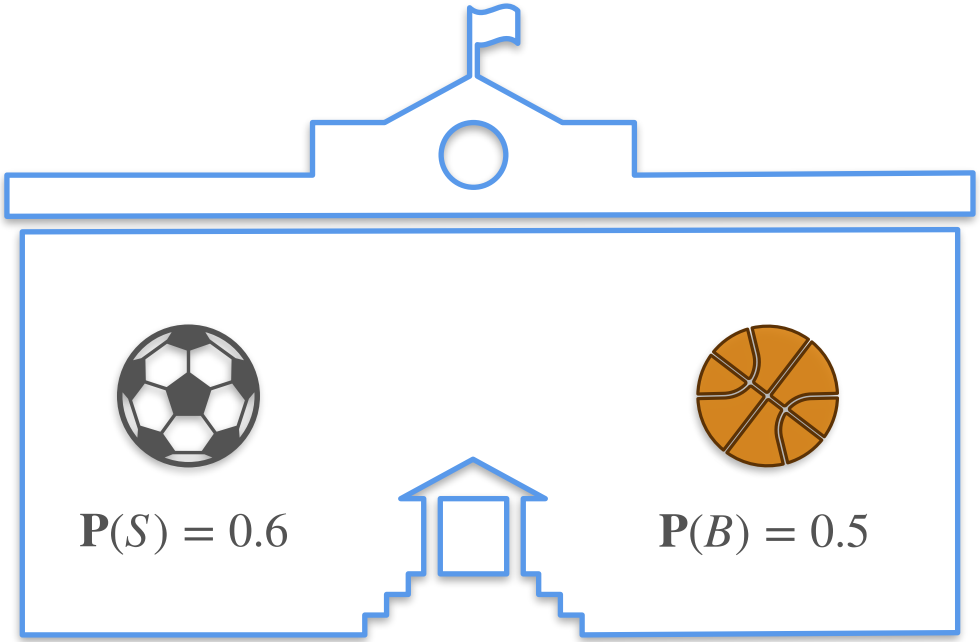
$$= \frac{18}{36}$$

DeepLearning.AI

Introduction to probability

Sum of Probabilities (Joint Events)

Sum of Probabilities (Joint Events): Quiz 1

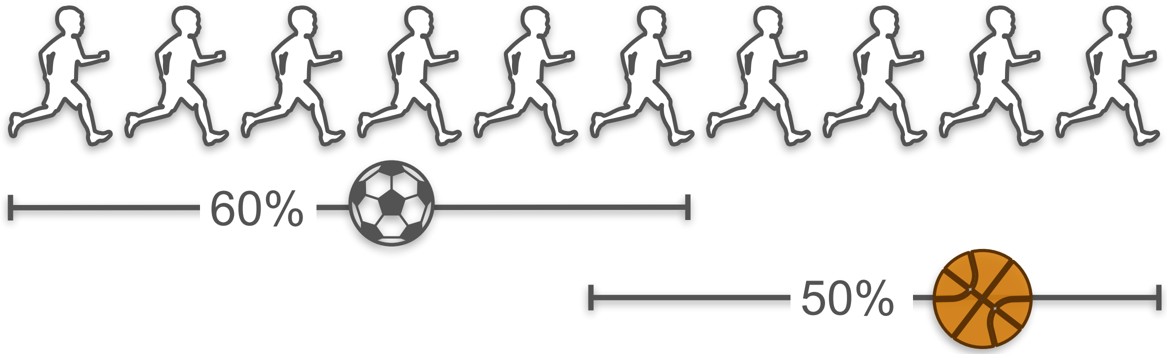


At a school, kids can play as many sports as they want.

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

Hint: What if there were only 10 kids?

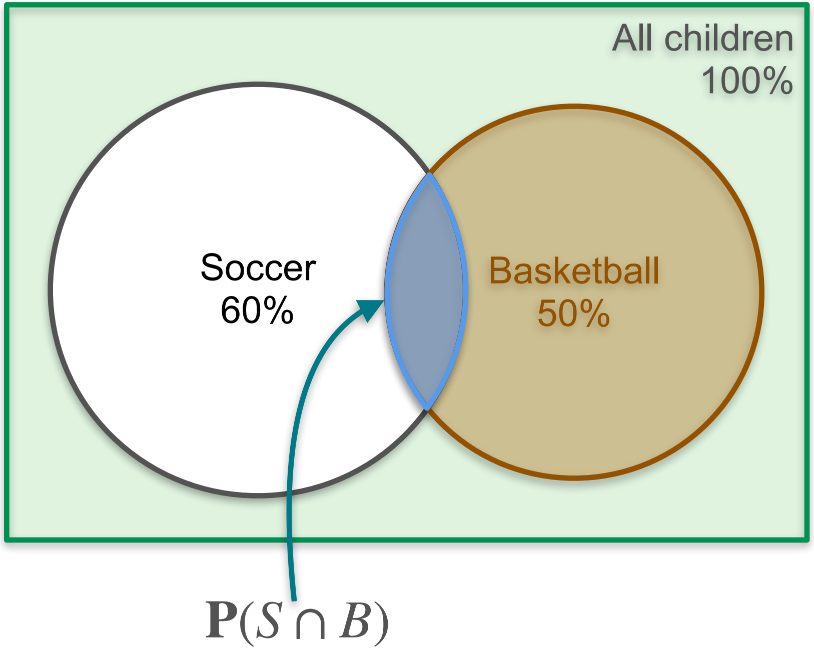
Sum of Probabilities (Joint Events): Quiz 1 Solution

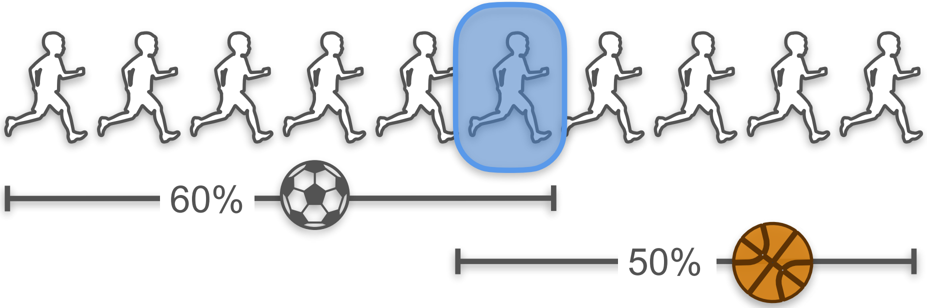


$\mathbf{P}(\text{soccer} \cup \text{basketball}) = ?$

We don't know how many children play multiple sports

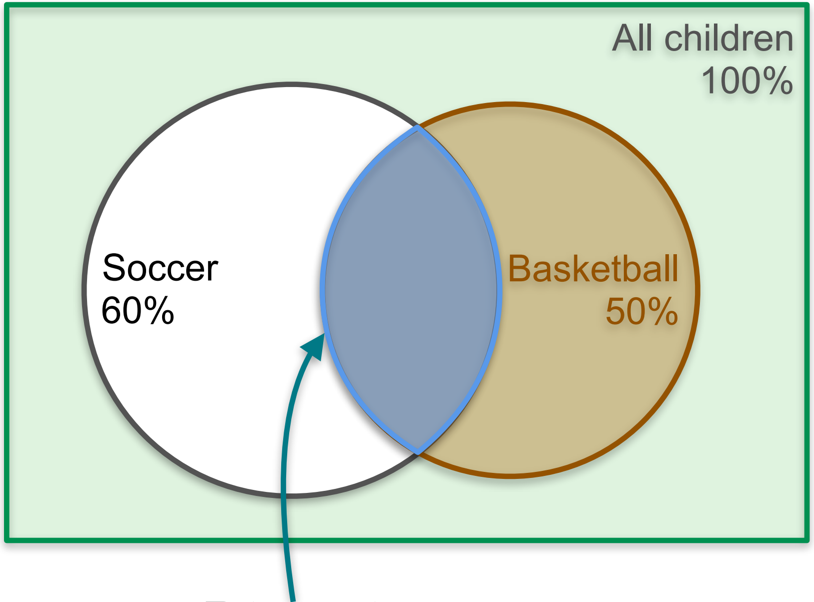
Sum of Probabilities (Joint Events): Quiz 1 Solution



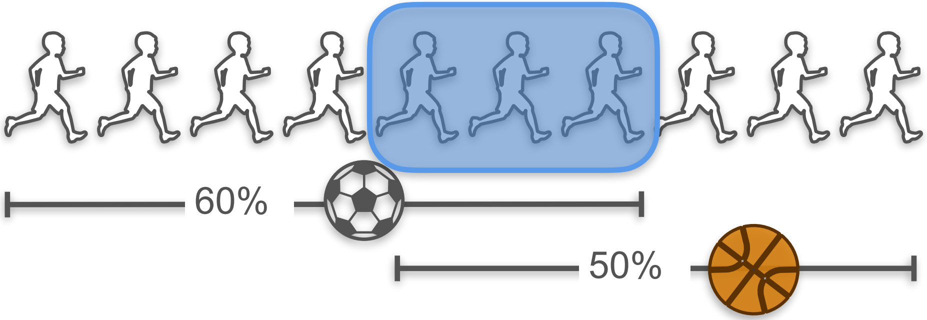


$$\mathbf{P}(S \cup B) = \mathbf{P}(S) + \mathbf{P}(B)$$

Sum of Probabilities (Joint Events): Quiz 1 Solution

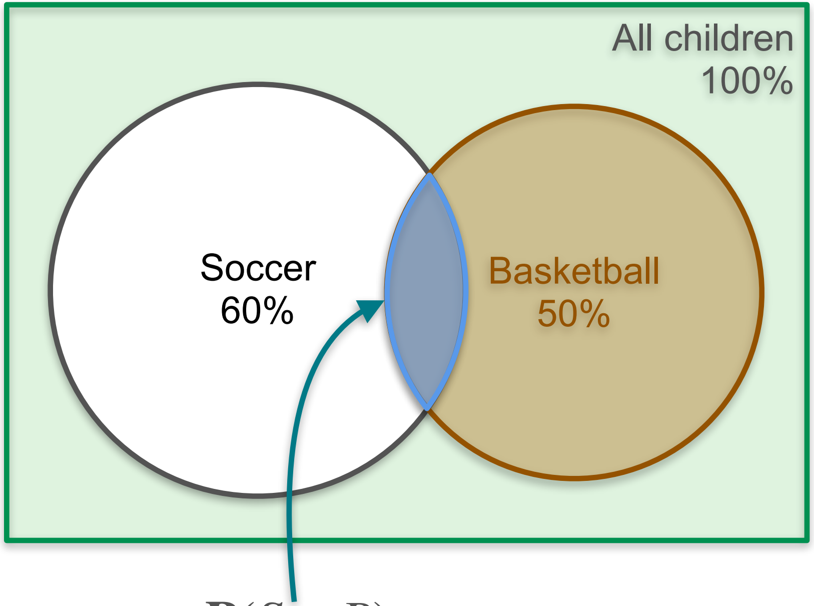


$\mathbf{P}(S \cap B)$

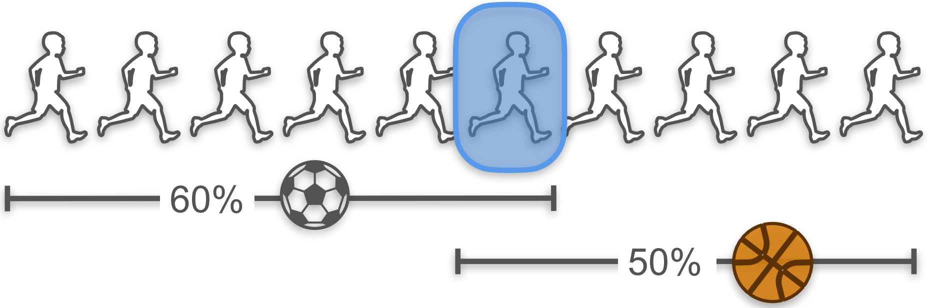


$$\mathbf{P}(S \cup B) = \mathbf{P}(S) + \mathbf{P}(B)$$

Sum of Probabilities (Joint Events): Quiz 1 Solution

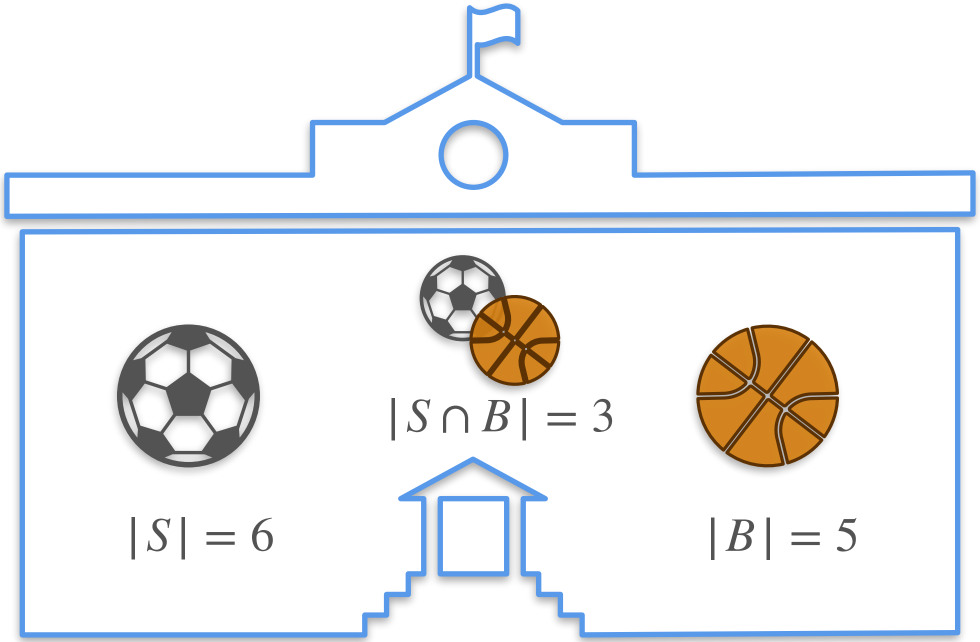


$\mathbf{P}(S \cap B)$



$$\mathbf{P}(S \cup B) = \mathbf{P}(S) + \mathbf{P}(B)$$

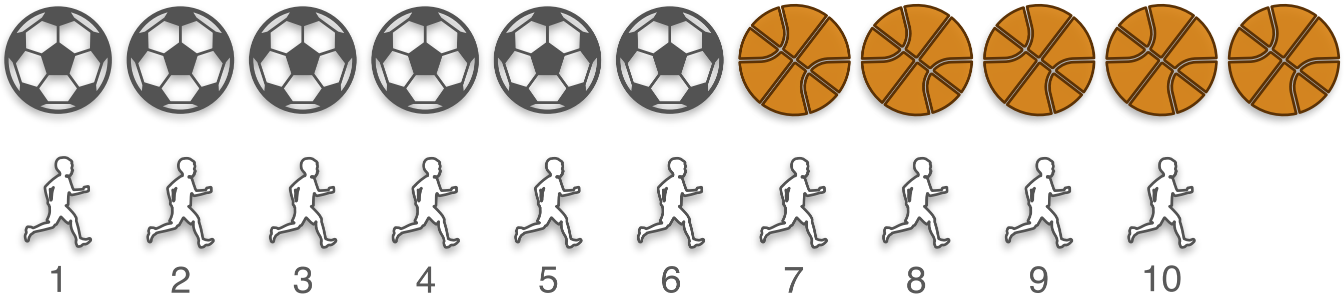
Sum of Probabilities (Joint Events): Quiz 2



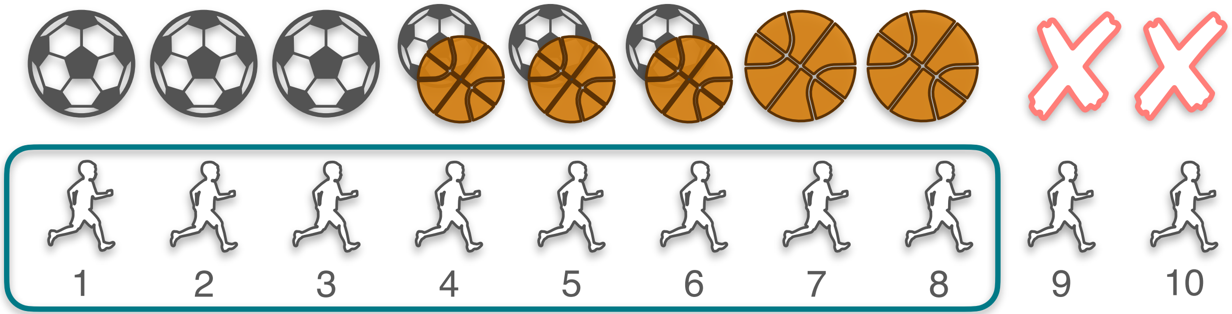
How many kids play soccer or basketball?

Hint: What if there were only 10 kids?

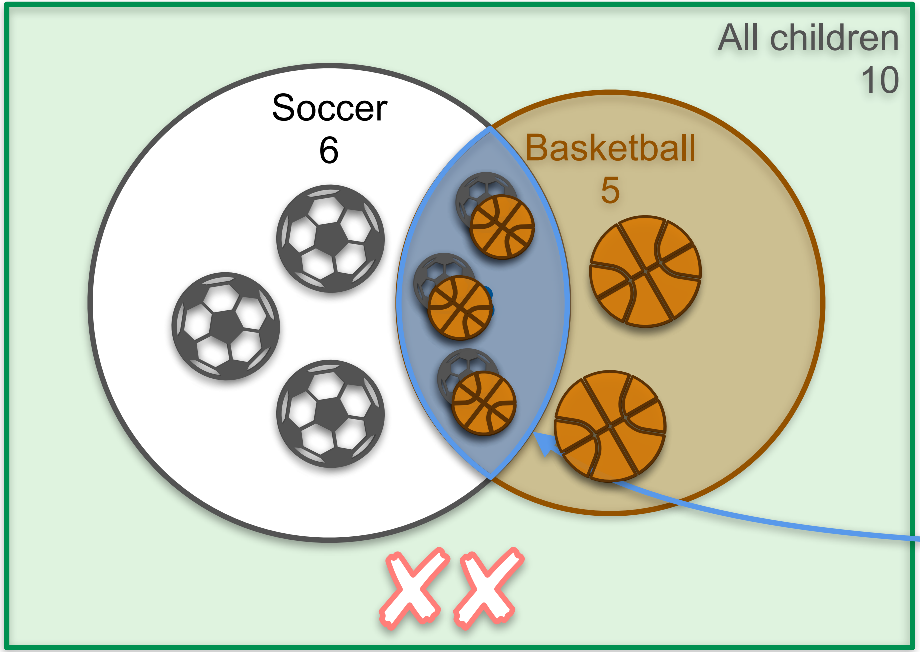
Sum of Probabilities (Joint Events): Quiz 2 Solution



Sum of Probabilities (Joint Events): Quiz 2 Solution



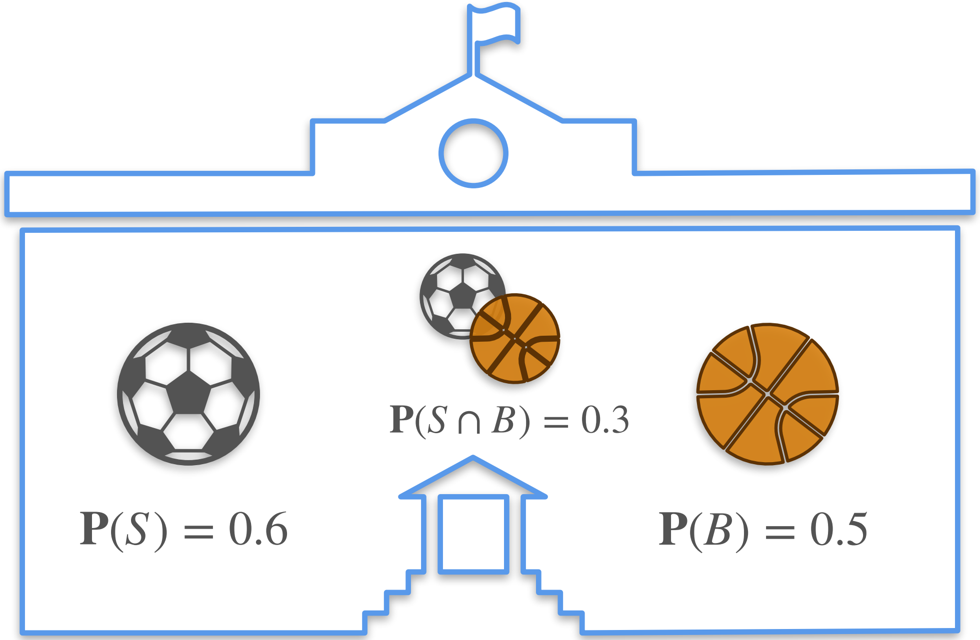
Sum of Probabilities (Joint Events): Venn Diagram



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

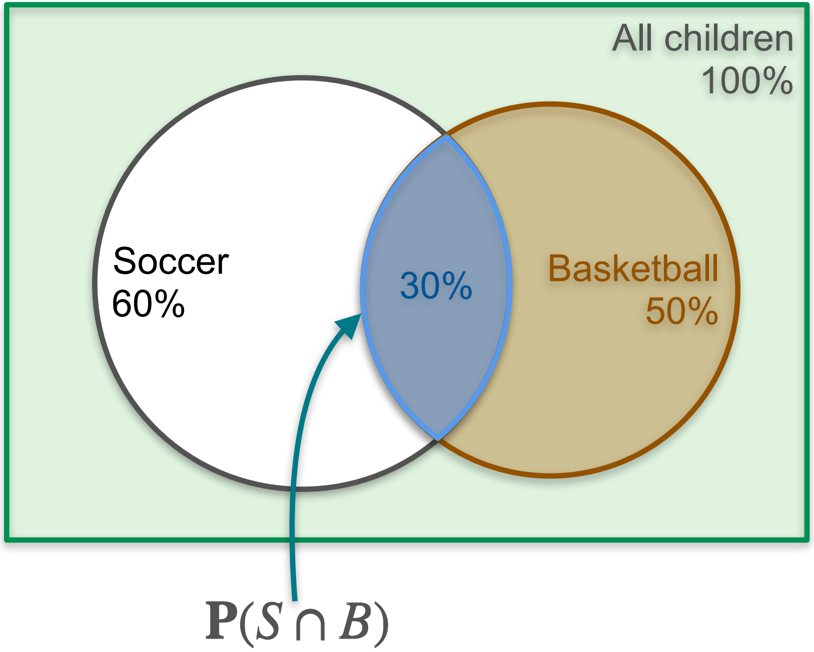
Soccer and Basketball

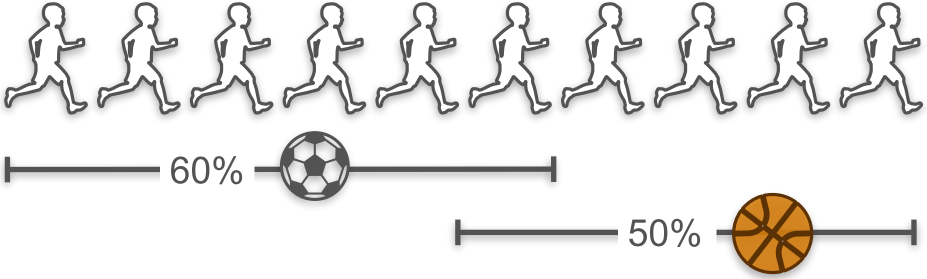
Sum of Probabilities (Joint Events): Quiz 3



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

Sum of Probabilities (Joint Events): Quiz 3 Solution





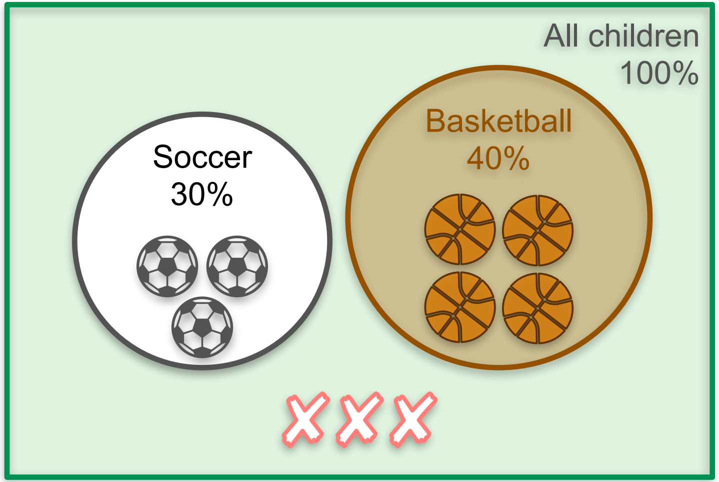
$$\mathbf{P}(S \cup B) = \mathbf{P}(S) + \mathbf{P}(B) - \mathbf{P}(S \cap B)$$

$$= 0.6 + 0.5 - 0.3$$

$$= 0.8$$

Disjoint Events vs Joint Events

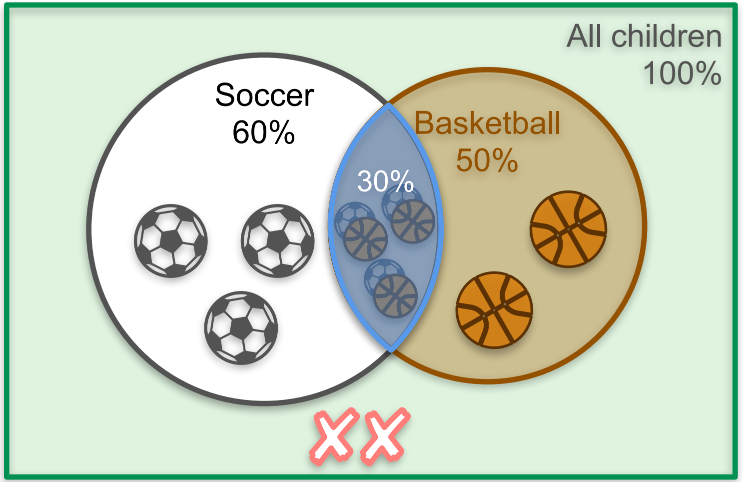
Disjoint



Mutually exclusive

$$\mathbf{P}(S \cup B) = \mathbf{P}(S) + \mathbf{P}(B)$$

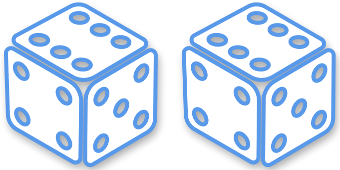
Joint



Non-mutually exclusive

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

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
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

diff = 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

Sum of Probabilities (Joint Events): Dice Example 1

A or B

sum = 7 or diff = 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

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

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

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

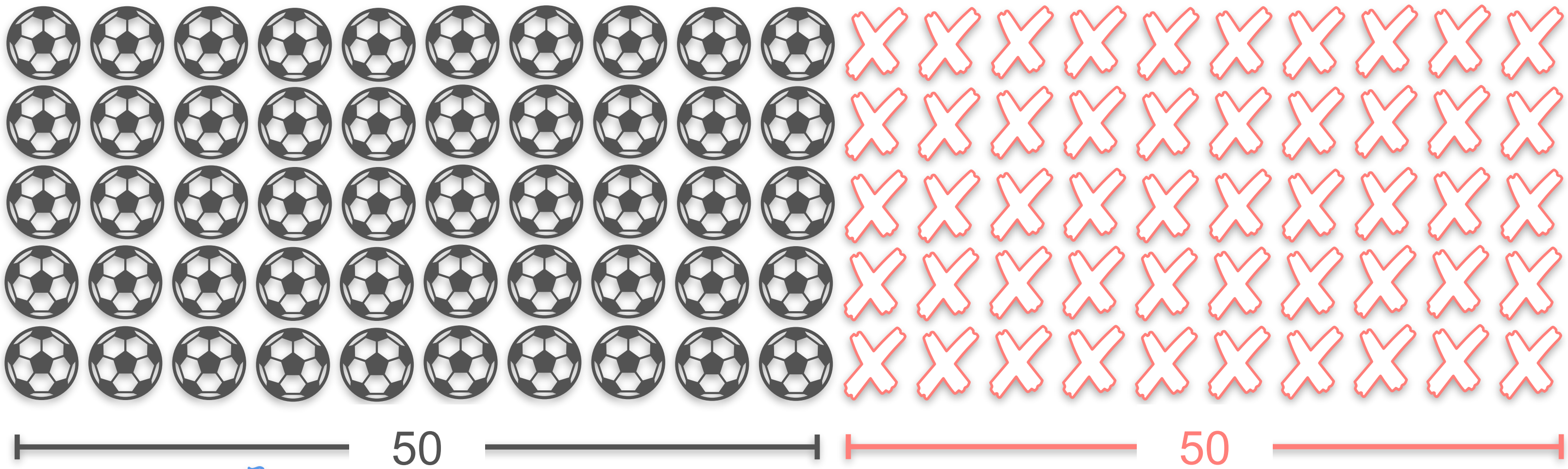
$$= \frac{6}{36} + \frac{10}{36} - \frac{2}{36}$$

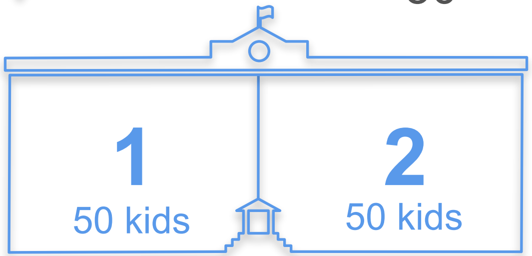
$$= \frac{14}{36}$$

Introduction to probability

Independence

Independence: Quiz 1

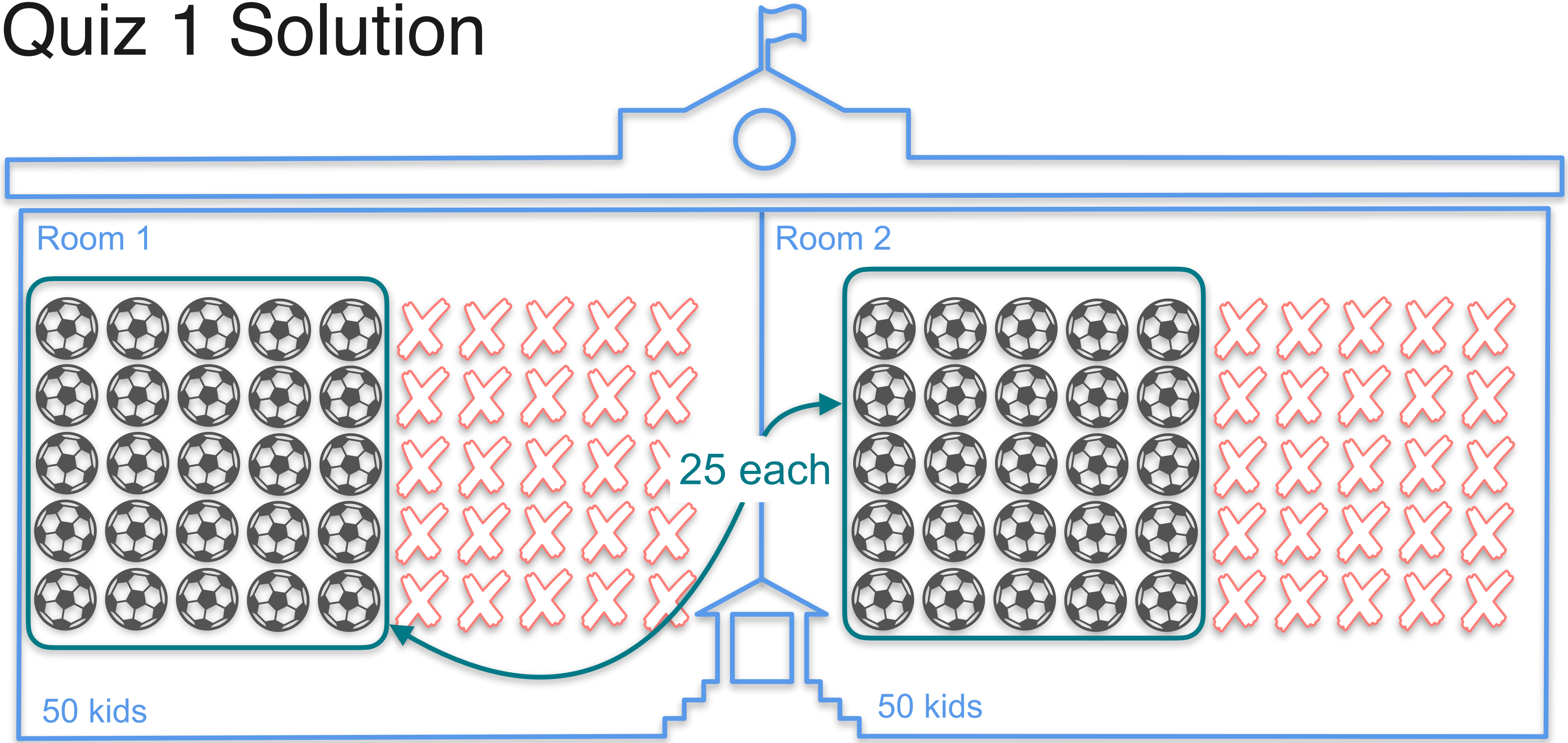




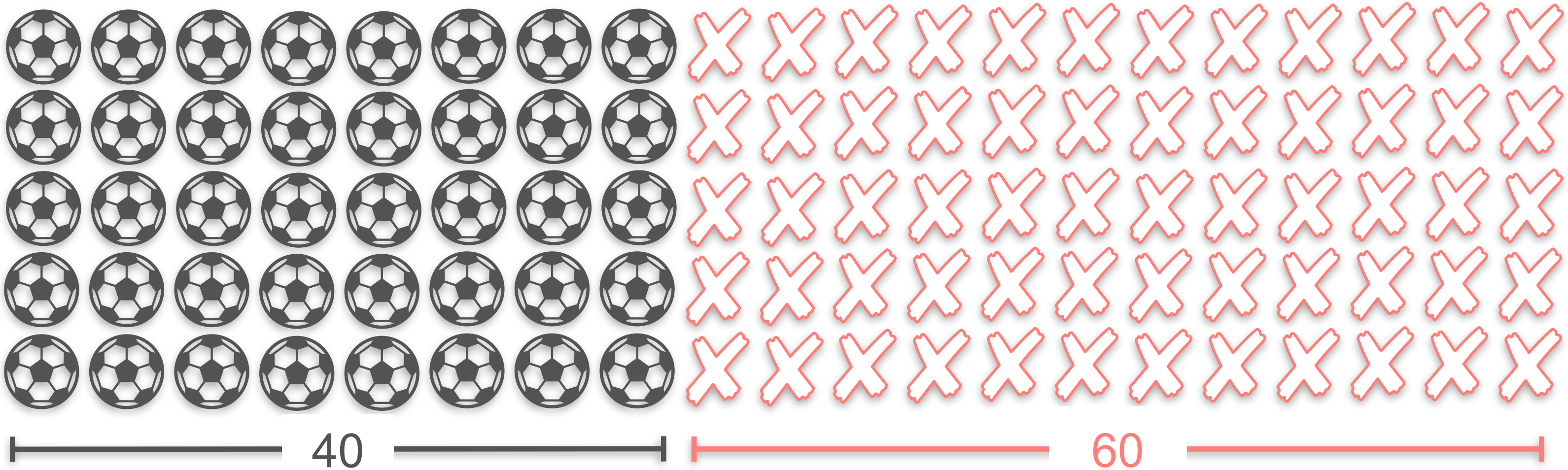
100 kids

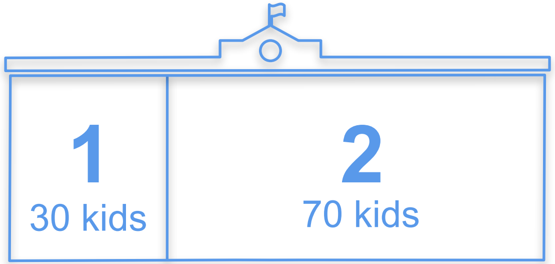
How many kids in room 1 like soccer?

Quiz 1 Solution



Independence: Quiz 2

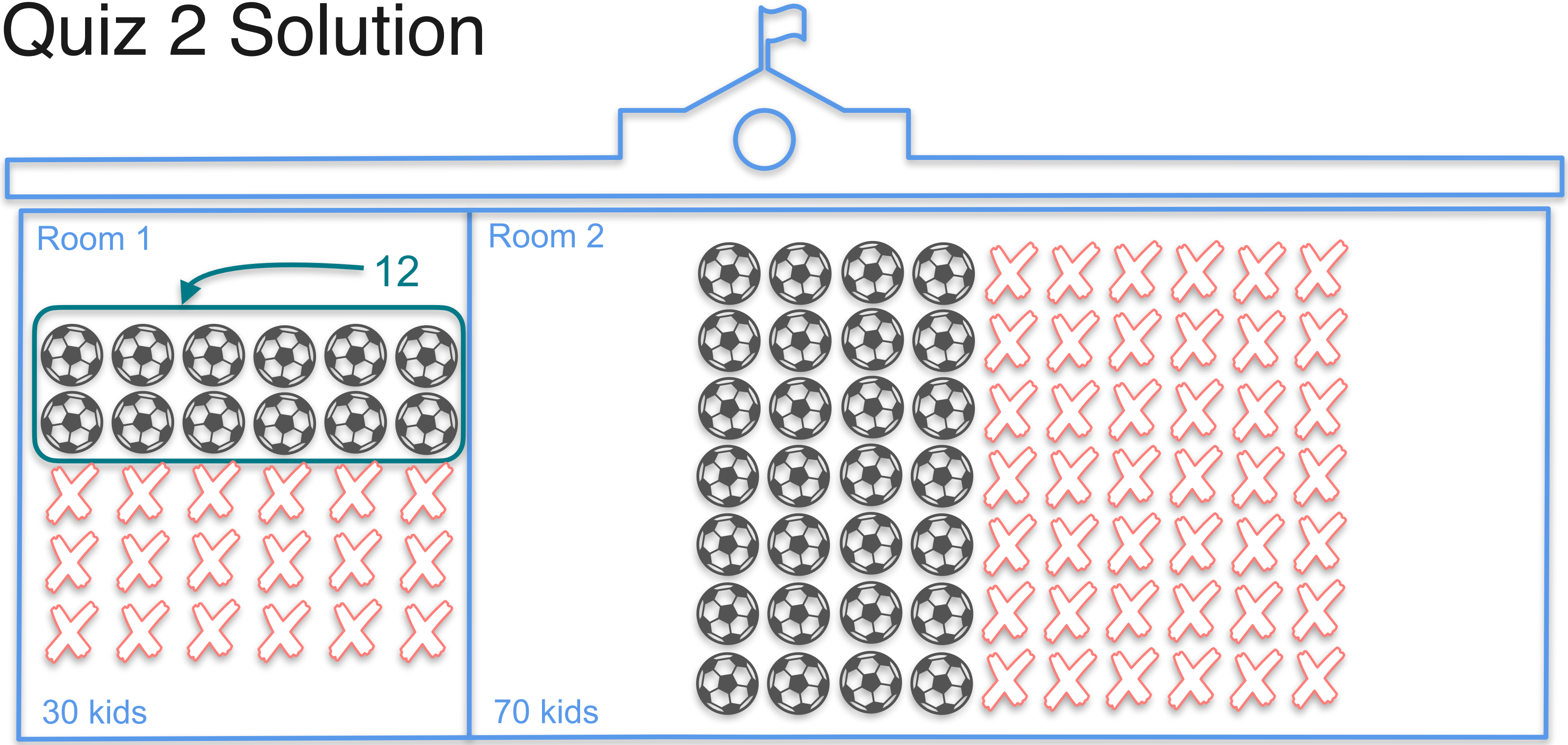




100 kids

How many kids in the smaller room like to play soccer?

Quiz 2 Solution



DeepLearning.AI

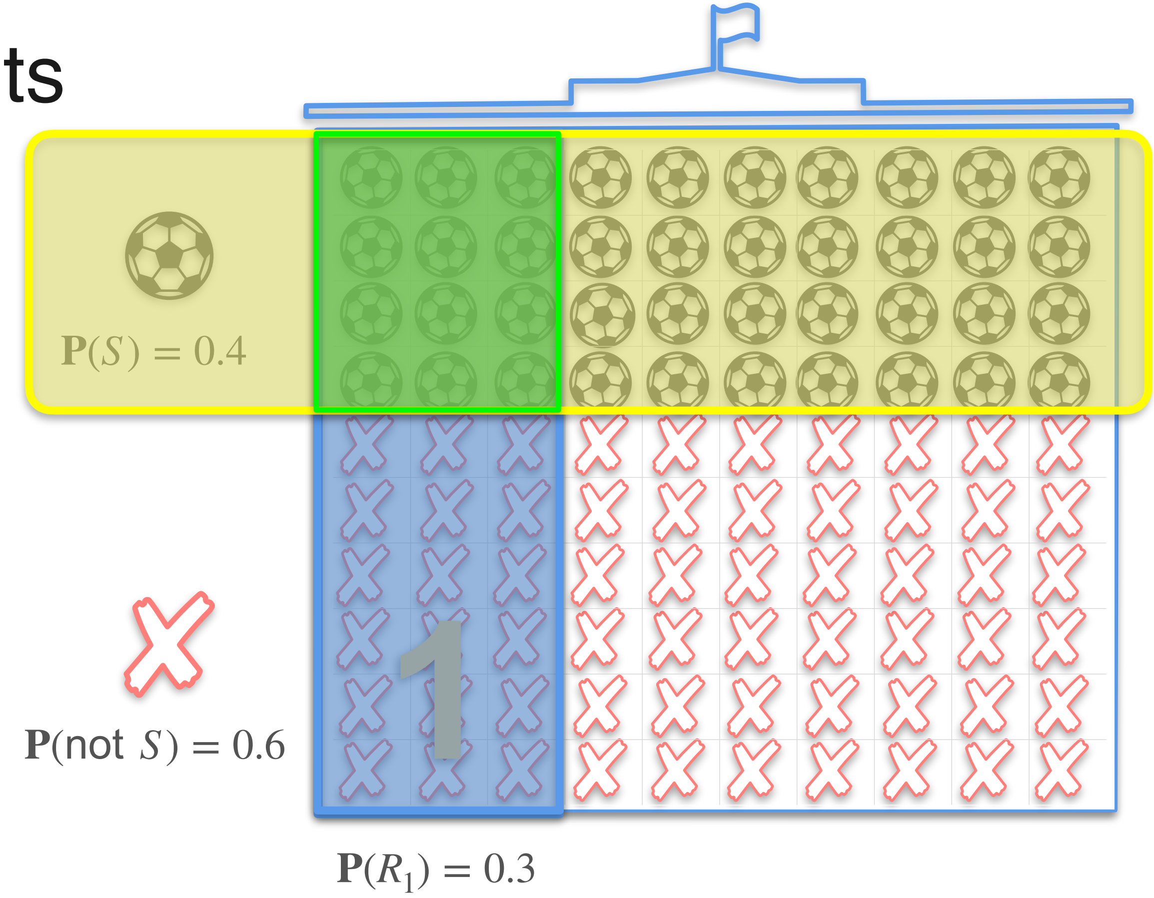
Independent Events

P(Soccer and Room 1)

$$\mathbf{P}(S \cap R_1) = \mathbf{P}(S) \bullet \mathbf{P}(R_1)$$

$$= 0.4 \bullet 0.3$$

$$= 0.12$$



Product Rule (for Independent Events)

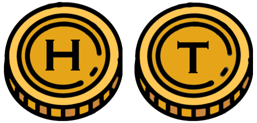
$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

Independent Events: Coin Example 1

50% 50%

What is the probability of landing on heads five times?

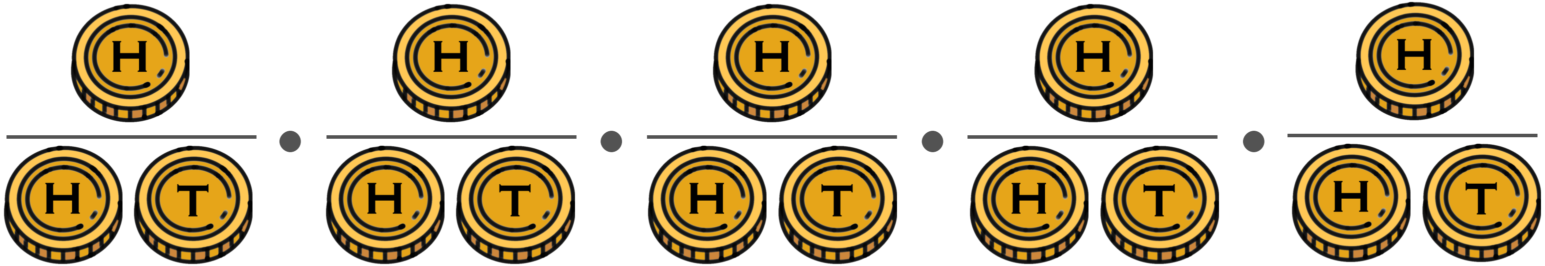
Independent Events: Coin Example 1



50% 50%

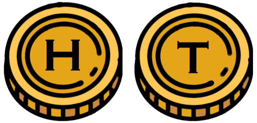
What is the probability of landing on heads five times?

P(5 heads) =



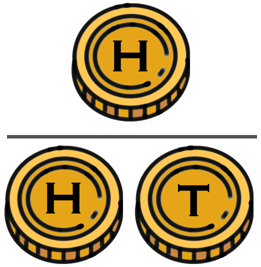
$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$

Independent Events: Coin Example 2



50%   50%

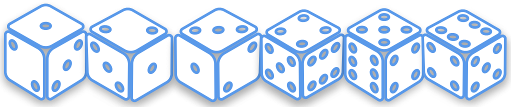
What is the probability of landing on heads five times?

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

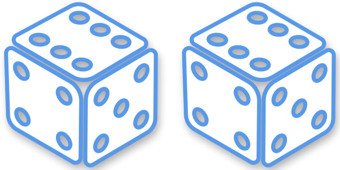
$$\left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

Independent Events: Dice Example 1

$$\mathbf{P}(6) = \frac{\text{[die]}}{\text{[six dice]}} = \frac{1}{6}$$



Independent Events: Dice Example 1



2 dice

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

$$\mathbf{P}(6,6) = \underline{\hspace{6cm}} = \frac{1}{36}$$

Independent Events: Dice Example 1

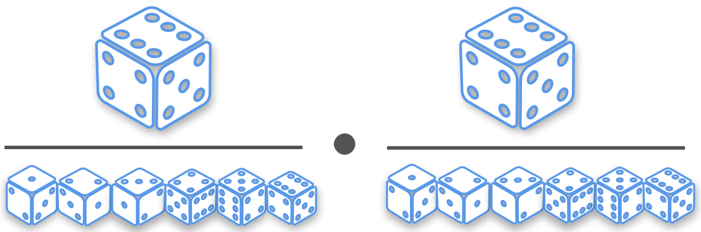
$\frac{1}{6}$

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

$\frac{1}{6}$

2 dice

$\mathbf{P}(6,6) =$

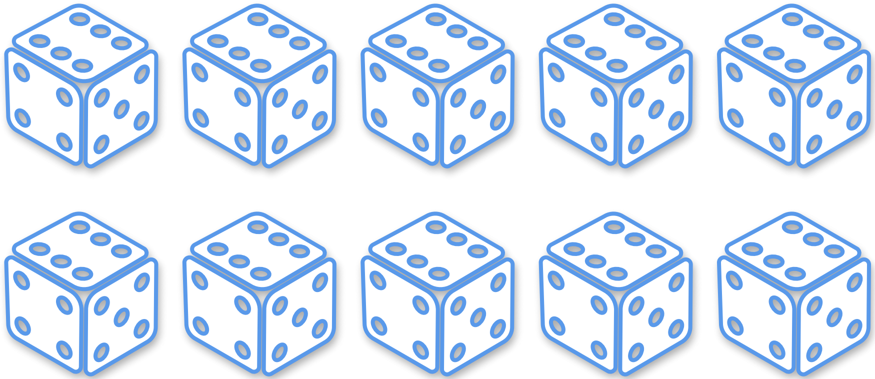


$$\frac{1}{6} \bullet \frac{1}{6}$$

$$= \left(\frac{1}{6}\right) \quad = \frac{1}{36}$$

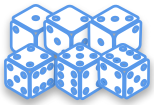
Independent Events: Dice Example 2

What is the probability of getting 10 sixes?



Independent Events: Dice Example 2

What is the probability of getting 10 sixes?

$\mathbf{P}$(10 sixes) = ( / )^{10}

$$= \left(\frac{1}{6}\right)^{10}$$

DeepLearning.AI

Introduction to probability

Birthday problem

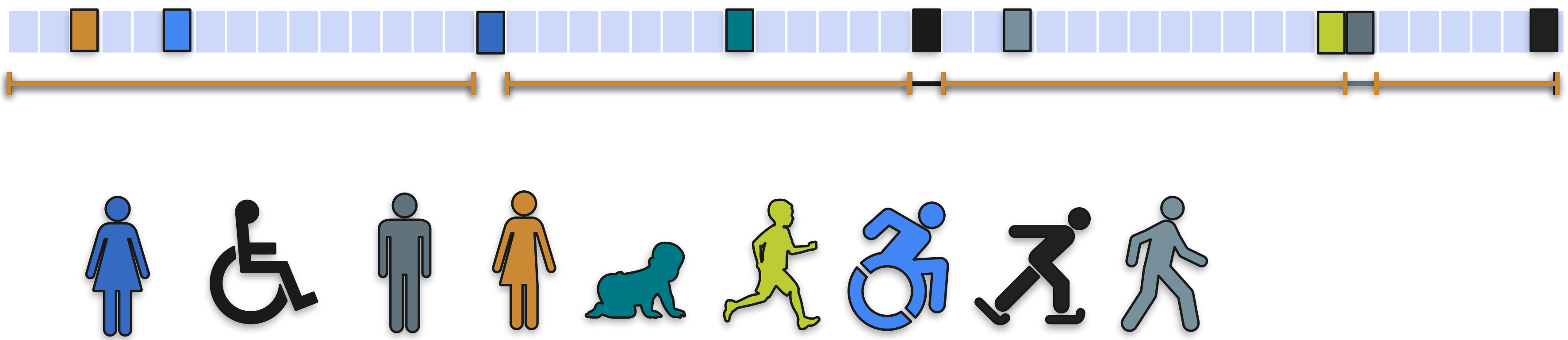
Quiz

- 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.

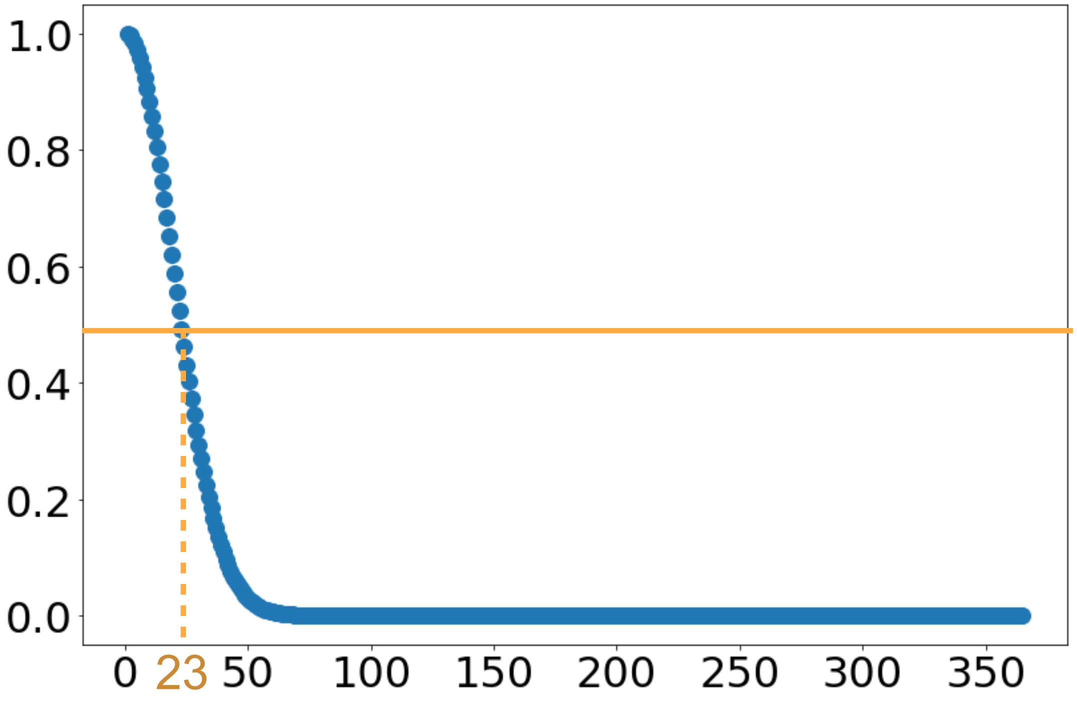
Probability That Everyone Has a Different Birthday



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

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

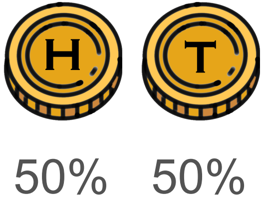


DeepLearning.AI

Introduction to Probability

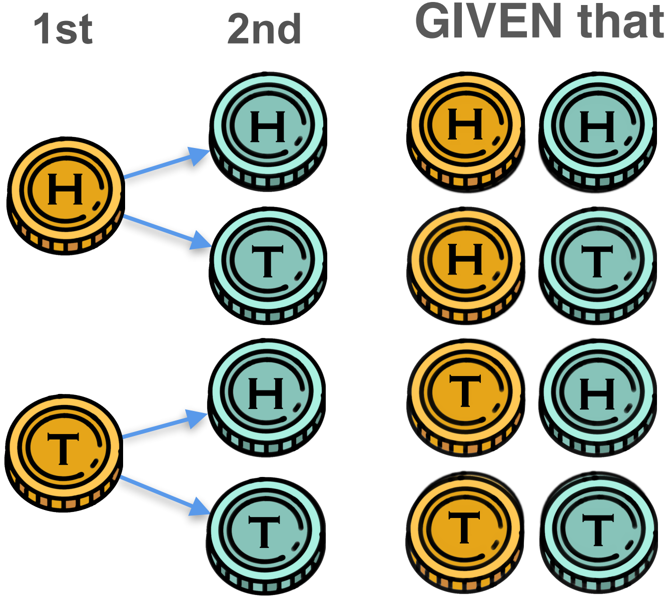
Conditional probability

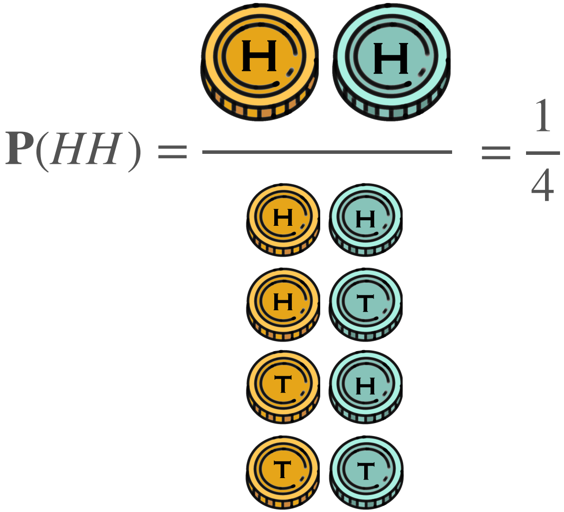
Conditional Probability: Coin Example 1



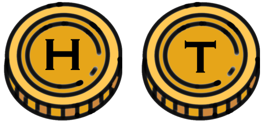
What is the probability of landing on heads twice?

GIVEN that the first one is heads





Conditional Probability: Coin Example 1

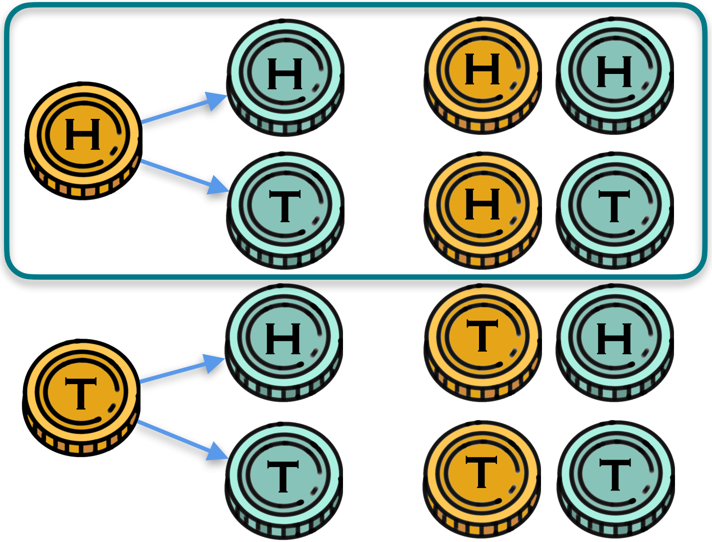


50% 50%

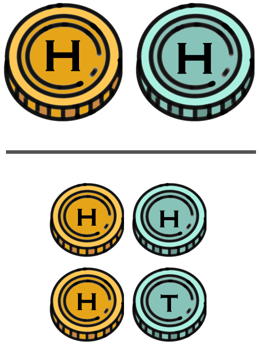
What is the probability of landing on heads twice?

GIVEN that the first one is heads

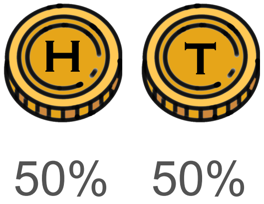
1st 2nd



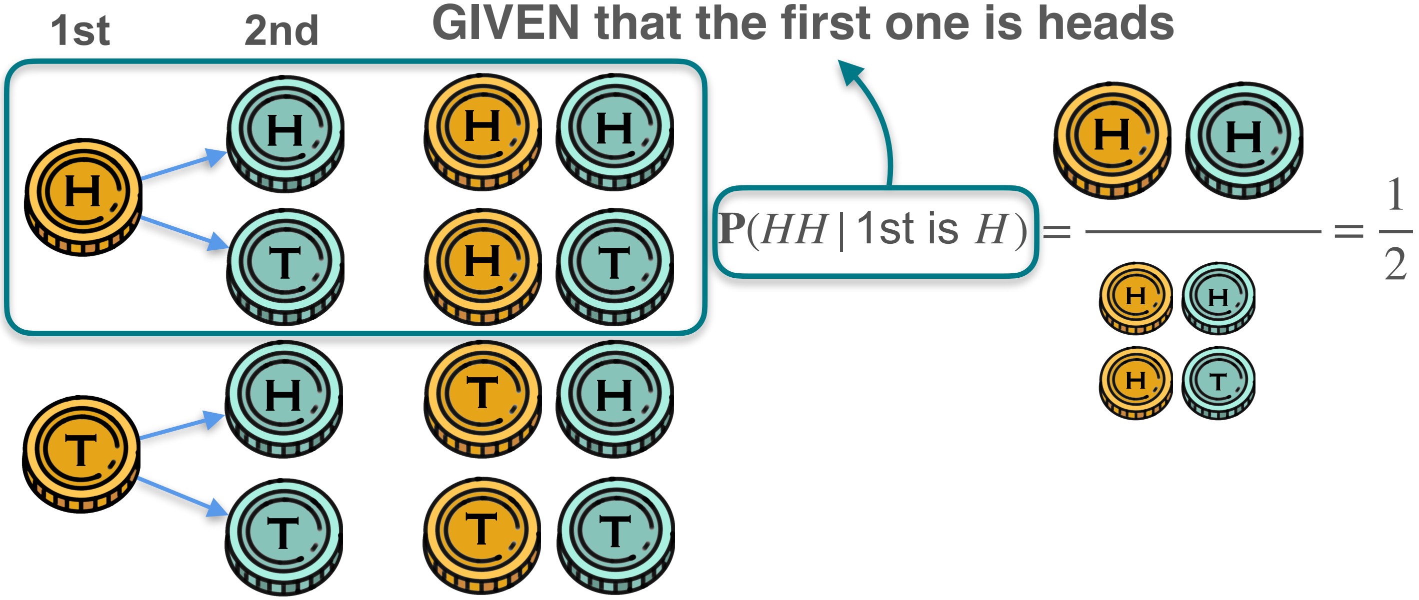
$$\mathbf{P}(HH) = \frac{\text{[HH]}}{\text{[HH, HT]}} = \frac{1}{4}$$



Conditional Probability: Coin Example 1



What is the probability of landing on heads twice?



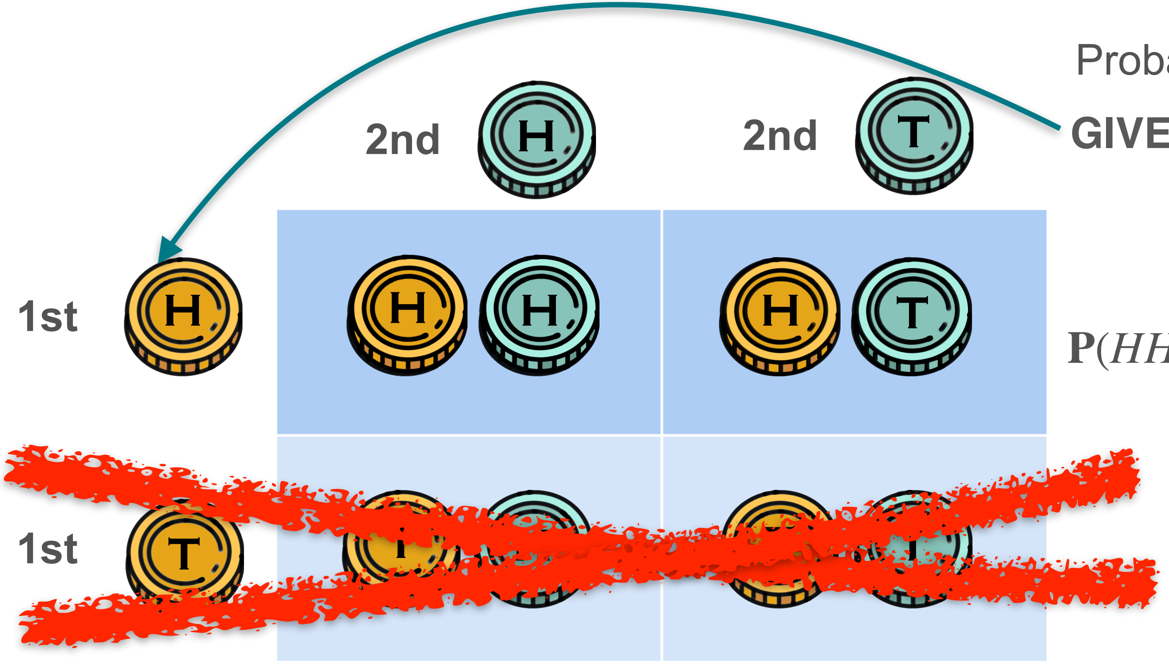
Conditional Probability: Coin Example 1

What is the probability of landing on heads twice?

GIVEN that the first one is heads

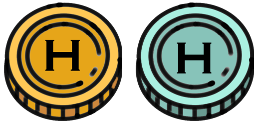
$$\mathbf{P}(HH \mid \boxed{\text{1st is } H})$$

Conditional Probability: Coin Example 1



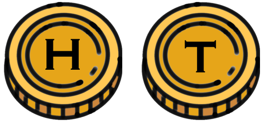
Probability of landing on heads twice

GIVEN that the first one is heads



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

Conditional Probability: Coin Example 2

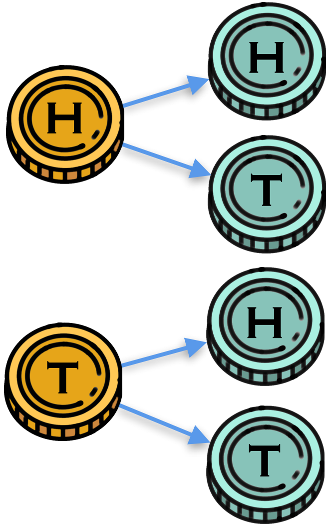


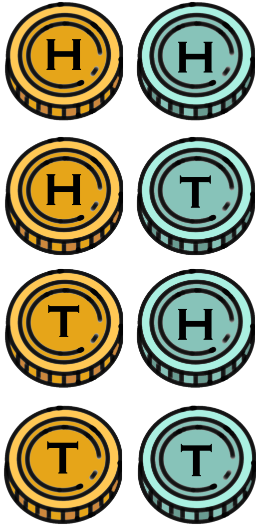
50% 50%

What is the probability of landing on heads twice?

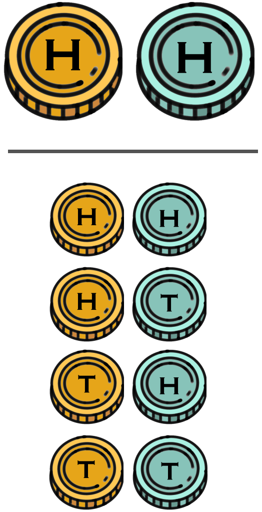
1st **2nd**

GIVEN that the first one is tails



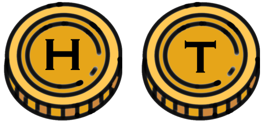


$\mathbf{P}(HH) =$



$= \frac{1}{4}$

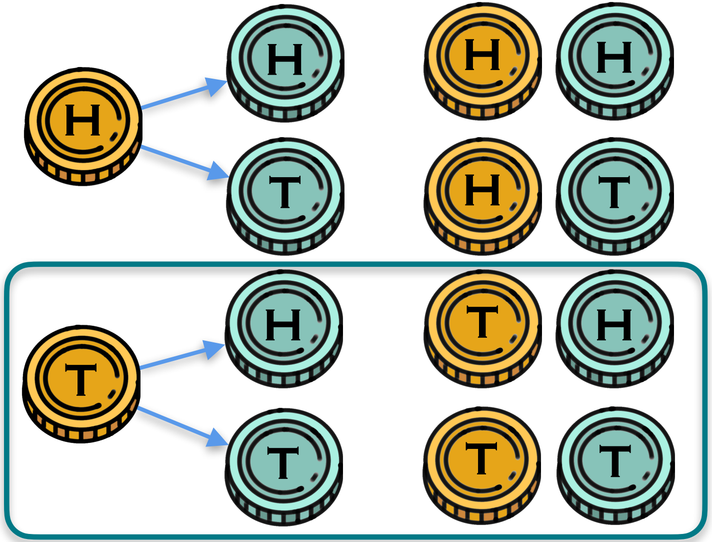
Conditional Probability: Coin Example 2

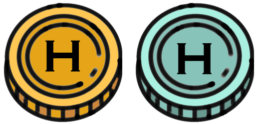


50% 50%

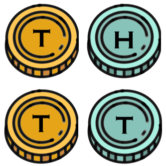
What is the probability of landing on heads twice?

1st **2nd** **GIVEN that the first one is tails**

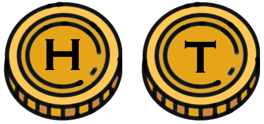




$$\mathbf{P}(HH) = \frac{}{} = \frac{1}{4}$$



Conditional Probability: Coin Example 2

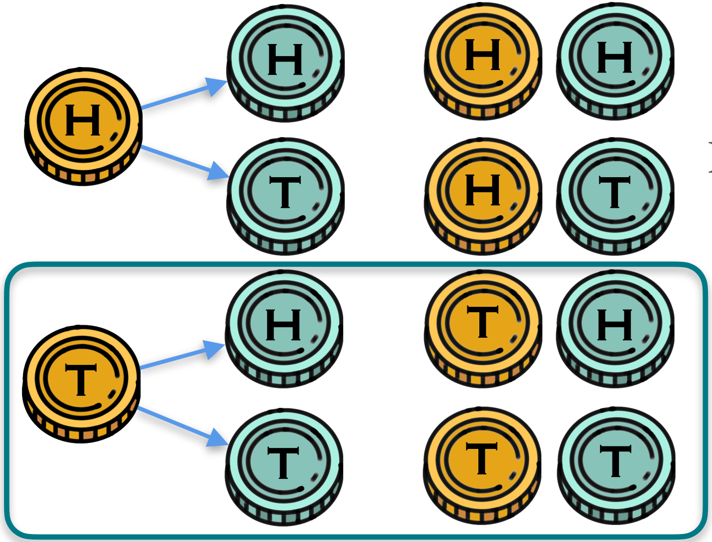


50% 50%

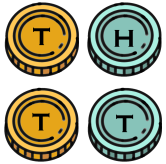
What is the probability of landing on heads twice?

GIVEN that the first one is tails

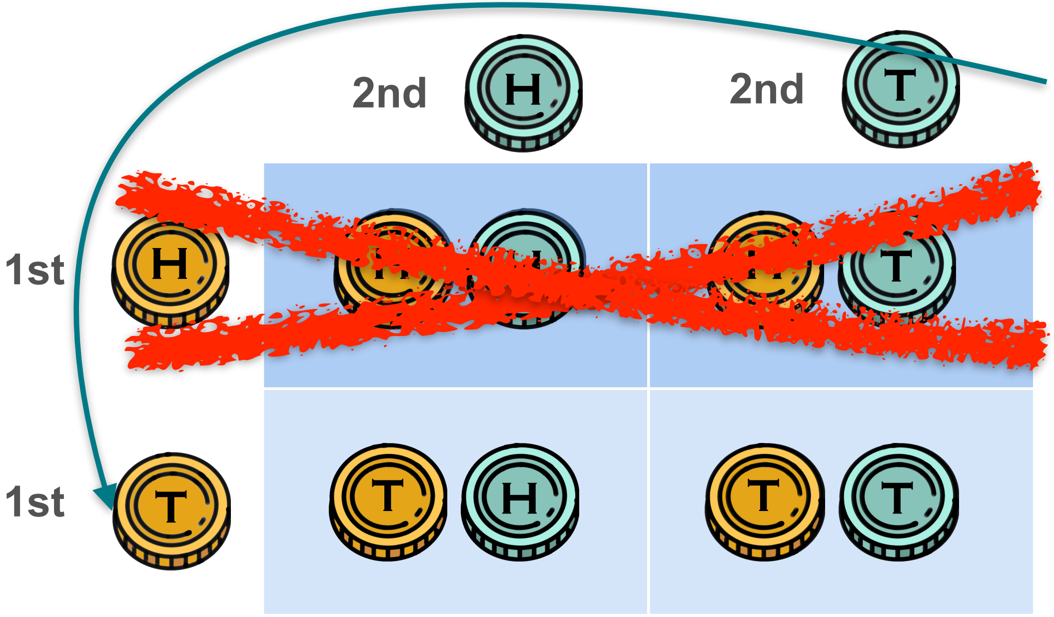
1st **2nd**



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



Conditional Probability: Coin Example 2



Probability of landing on heads twice

GIVEN that the first one is tails

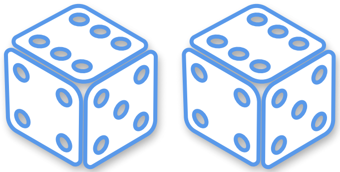
$$\mathbf{P}(HH \mid \text{1st is } T) = \frac{\quad}{\quad} = 0$$

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



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	$\boxed{6,4}$	6,5	6,6

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

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

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

Conditional Probability: Dice Example 3

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

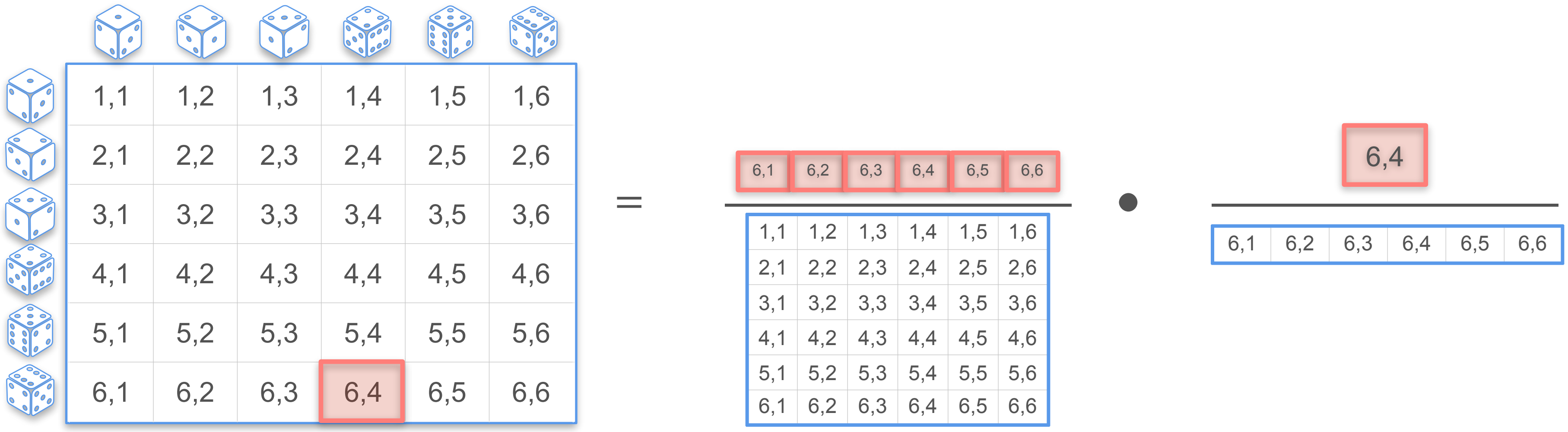
$\mathbf{P}(\text{1st is 6} \cap \text{sum} = 10) =$

$\mathbf{P}(\text{1st is 6}) \quad \bullet \quad \mathbf{P}(\text{sum} = 10 \mid \text{1st 6})$

$\underline{\qquad\qquad} \quad \bullet \quad \underline{\qquad\qquad}$

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

Conditional Probability: Dice Example 3



$\mathbf{P}$(1st is 6 ∩ sum = 10) = $\mathbf{P}$(1st is 6) • $\mathbf{P}$(sum = 10 | 1st 6)

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

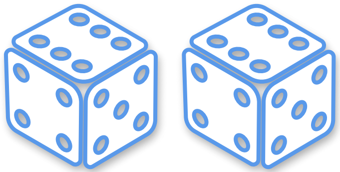
The General Product Rule

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

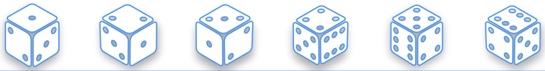
When independent

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

Conditional Probability: Dice Example 1



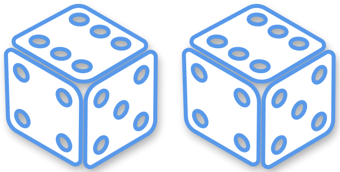
What is the probability that the sum is 10?



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

$$\mathbf{P}(\text{sum} = 10) = \underline{\hspace{6cm}} = \frac{3}{36}$$

Conditional Probability: Dice Example 1



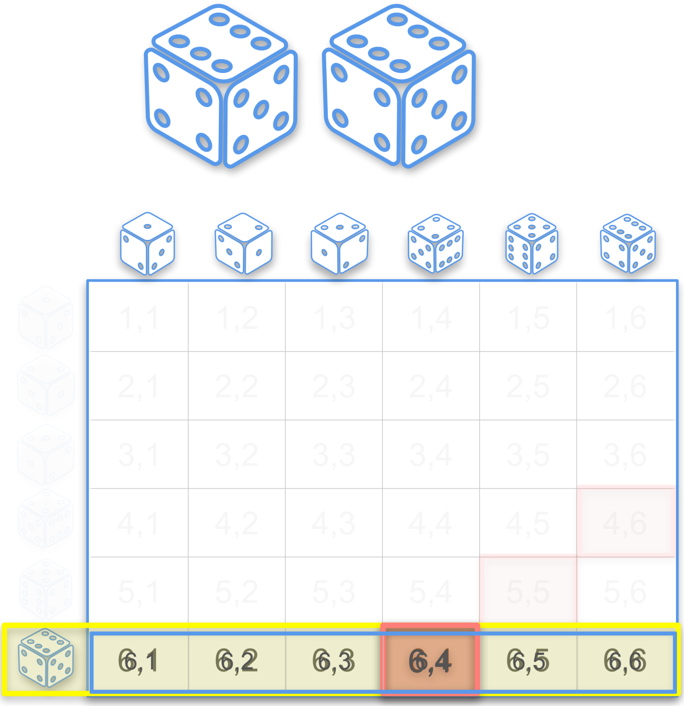
What is the probability that the sum is 10?

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

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

Conditional Probability: Dice Example 1



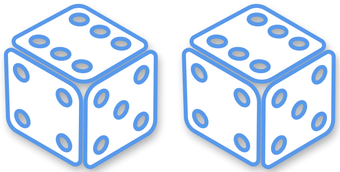
What is the probability that the sum is 10?

GIVEN that the first one is 6

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

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

Conditional Probability: Dice Example 2



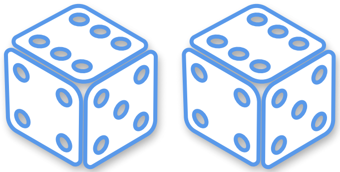
1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	$\boxed{4,6}$
5,1	5,2	5,3	5,4	$\boxed{5,5}$	5,6
6,1	6,2	6,3	$\boxed{6,4}$	6,5	6,6

What is the probability that the sum is 10?

$$\mathbf{P}(\text{sum} = 10) = \frac{\boxed{6,4}\ \boxed{5,5}\ \boxed{4,6}}{\text{all 36 outcomes}} = \frac{3}{36}$$

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

Conditional Probability: Dice Example 2



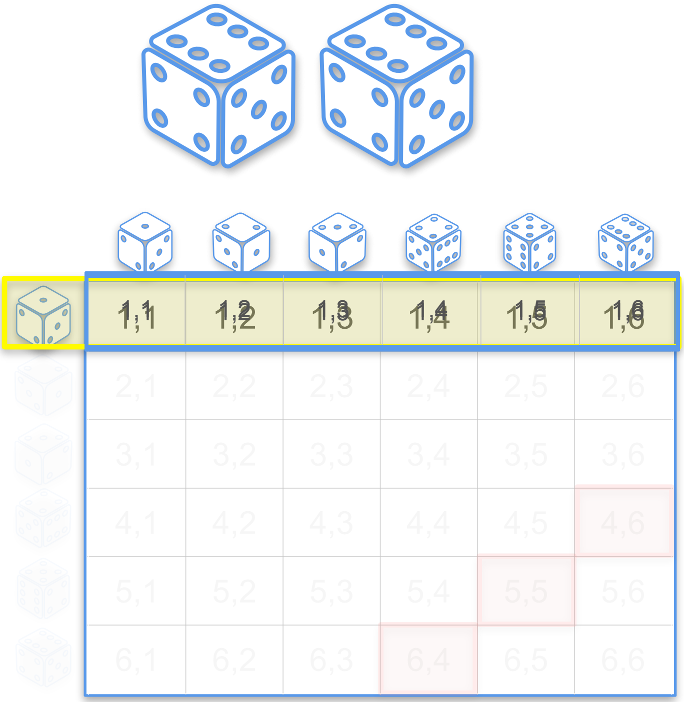
	1,1	1,2	1,3	1,4	1,5	1,6
	2,1	2,2	2,3	2,4	2,5	2,6
	3,1	3,2	3,3	3,4	3,5	3,6
	4,1	4,2	4,3	4,4	4,5	4,6
	5,1	5,2	5,3	5,4	5,5	5,6
	6,1	6,2	6,3	6,4	6,5	6,6

What is the probability that the sum is 10?

GIVEN that the first one is 1

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

Conditional Probability: Dice Example 2



What is the probability that the sum is 10?

GIVEN that the first one is 1

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

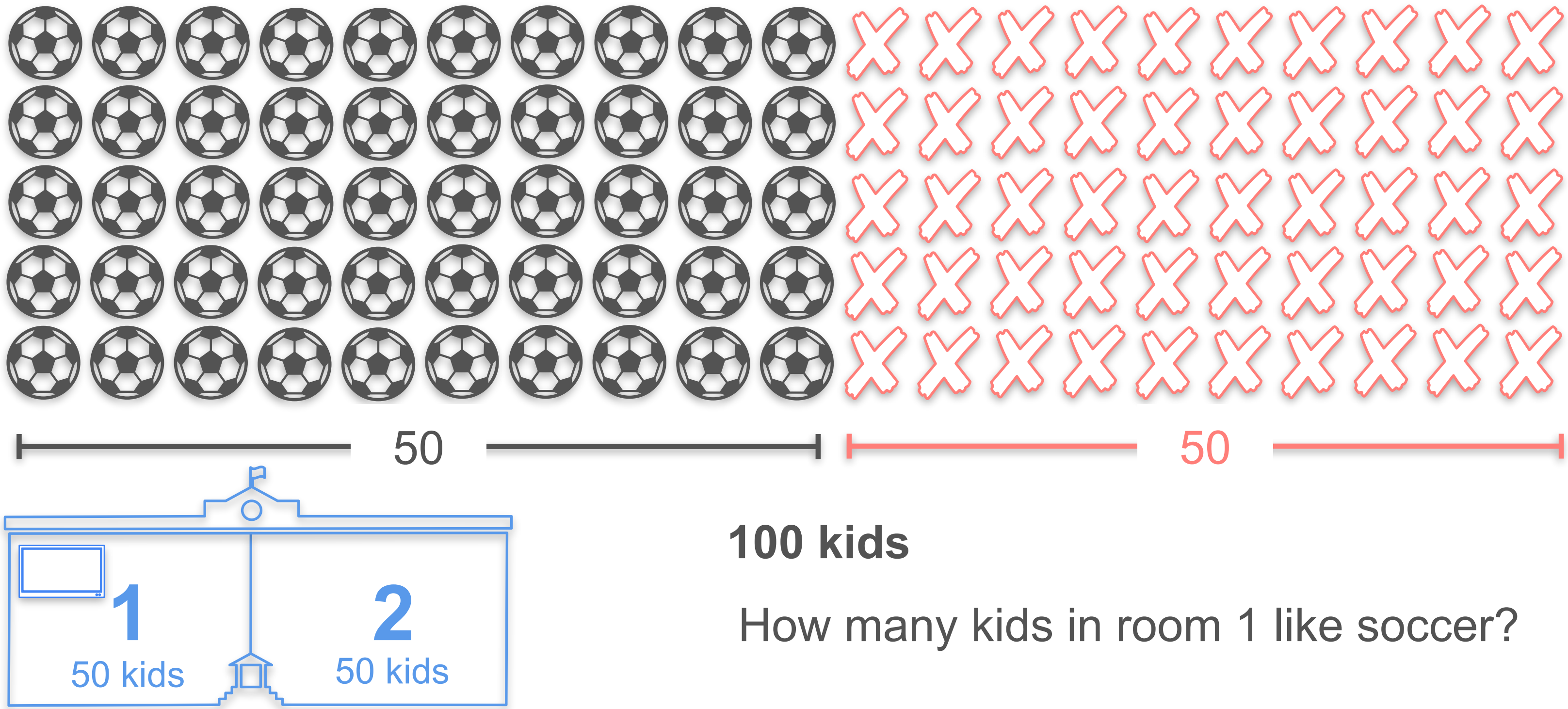
$$= 0$$

DeepLearning.AI

Introduction to Probability

Conditional probability - Part 2

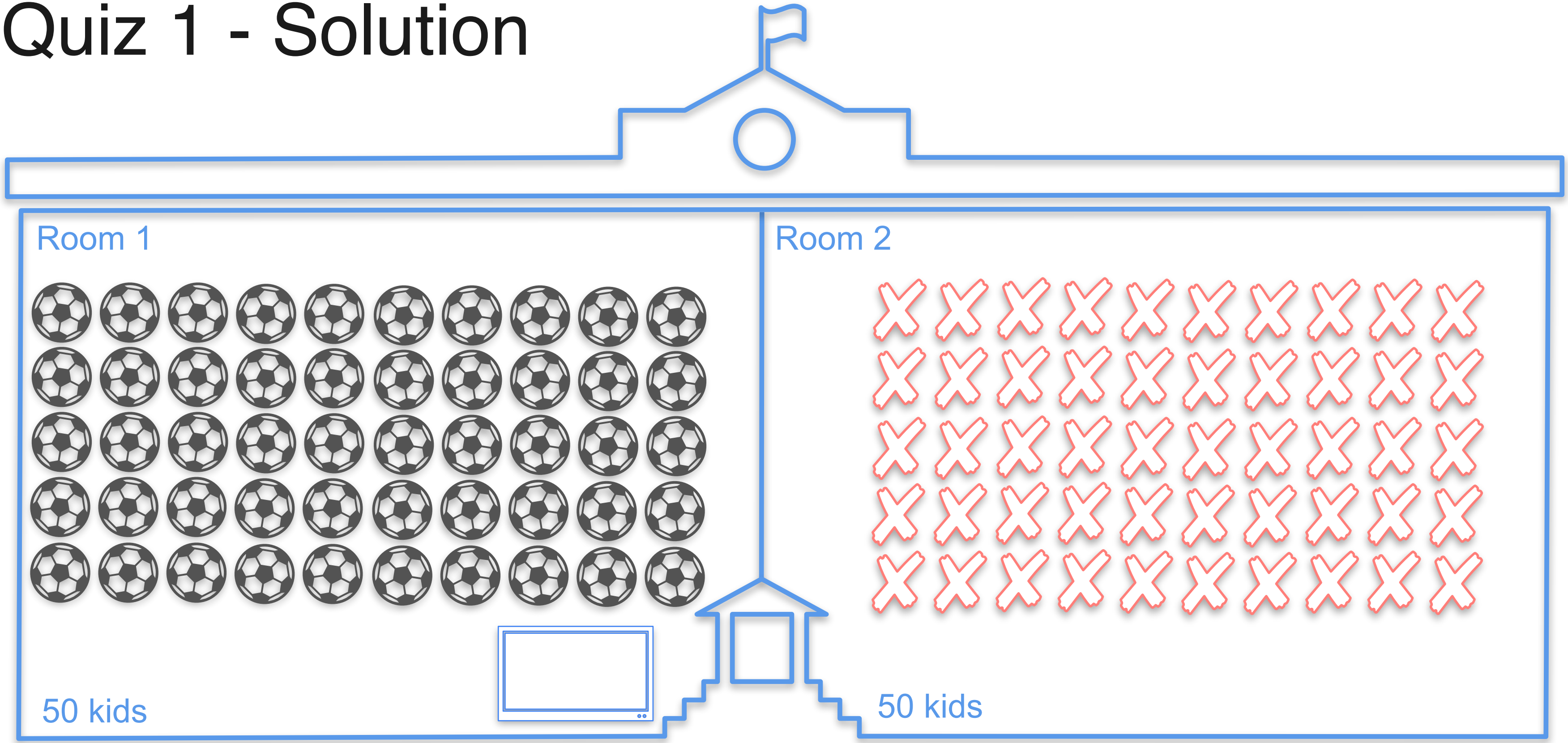
Quiz 1



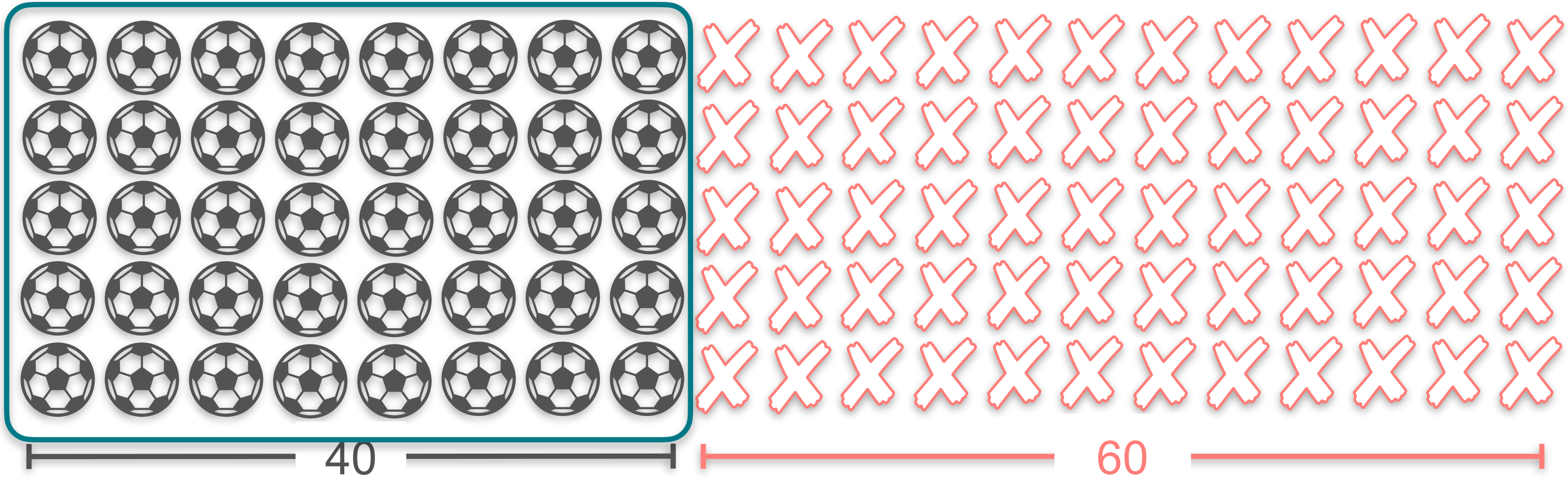
100 kids

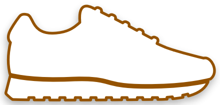
How many kids in room 1 like soccer?

Quiz 1 - Solution



Quiz 2





80%

100 kids

How many kids play soccer and wear running shoes

Quiz 2 - Solution

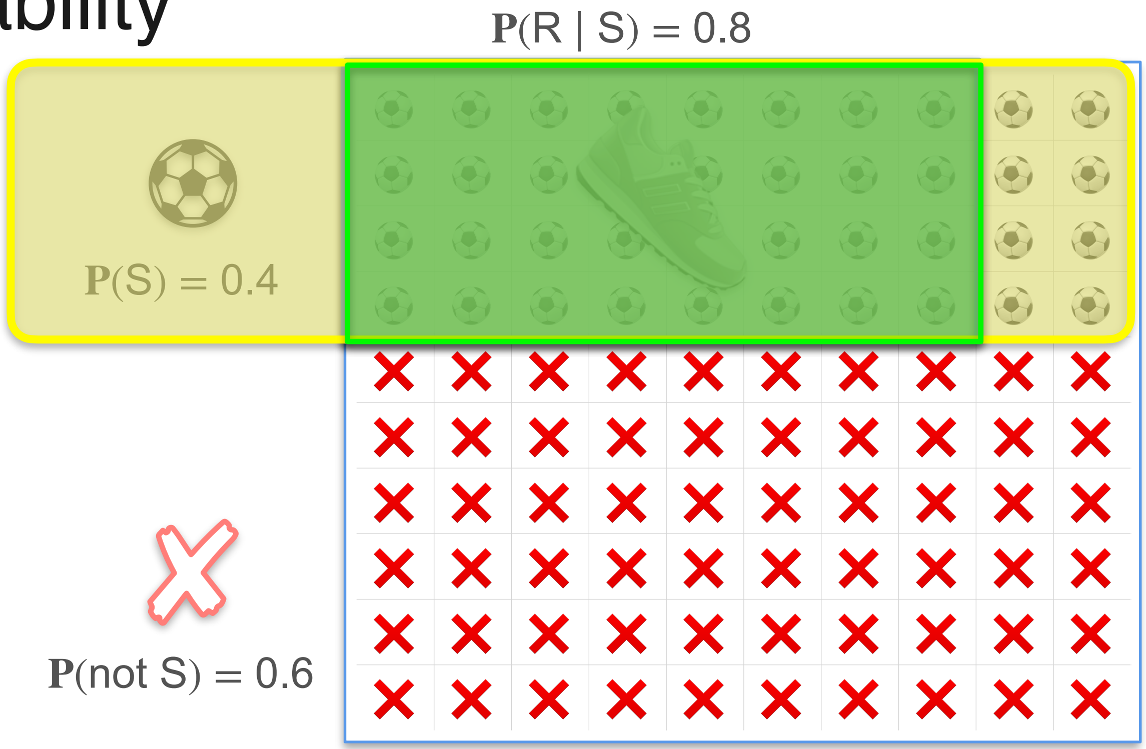
Conditional Probability

P(Soccer and Running shoes)

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

$$= 0.4 \bullet 0.8$$

$$= 0.32$$



Conditional Probability

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

$\mathbf{P}(S \cap R) = 0.32$

$\mathbf{P}(S) = 0.4$

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

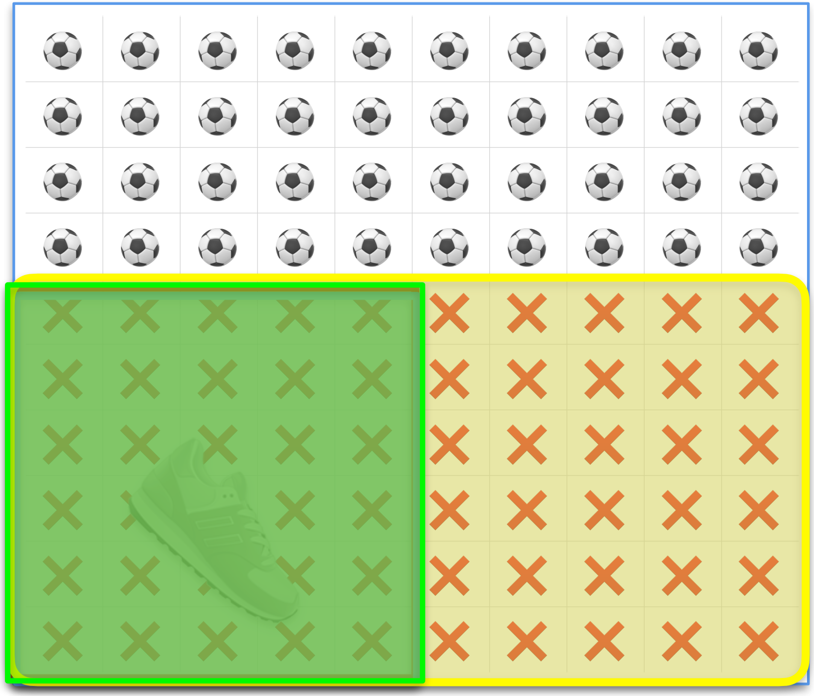
$\mathbf{P}(\text{not } S \cap R) = \mathbf{P}(\text{not } S) \bullet \mathbf{P}(R \mid \text{not } S)$

$= 0.6 \bullet 0.5$

$= 0.3$

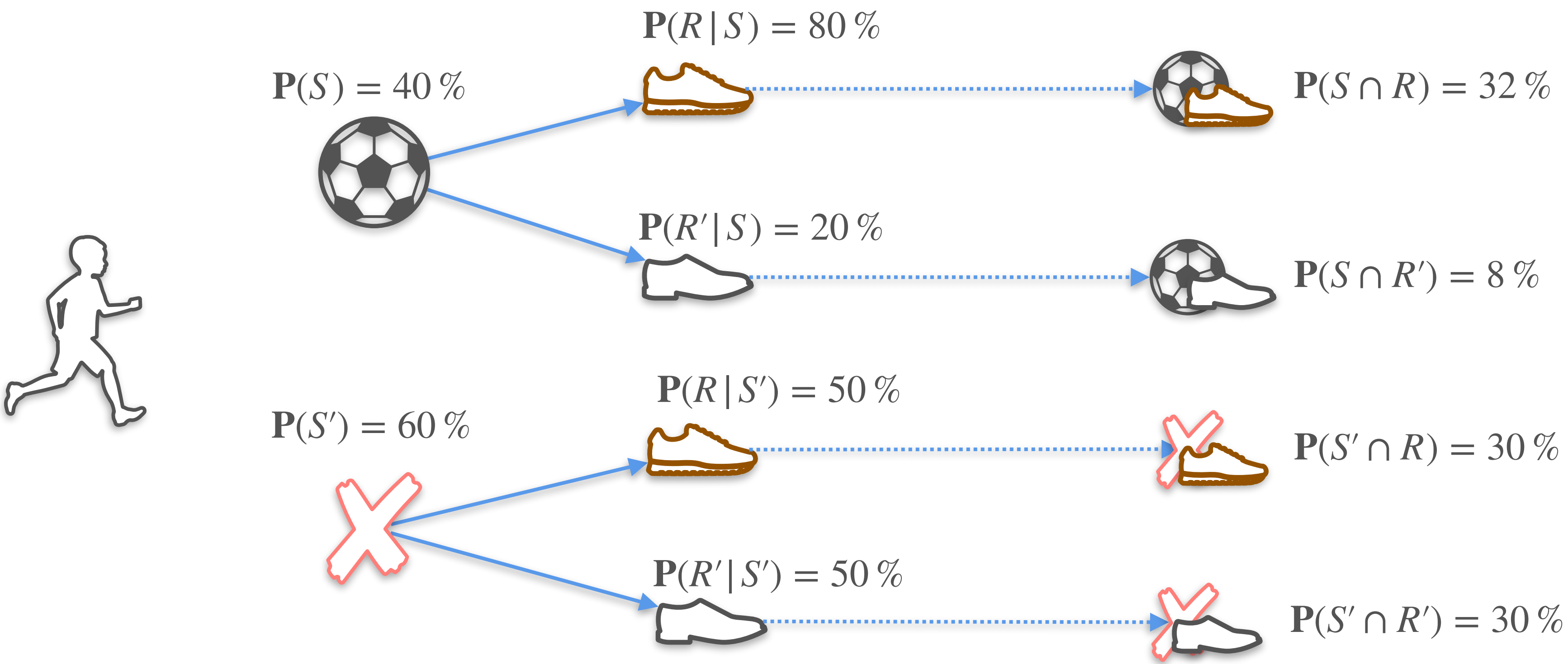
$\mathbf{P}(\text{not } S) = 0.6$

$\mathbf{P}(R \mid S) = 0.8$

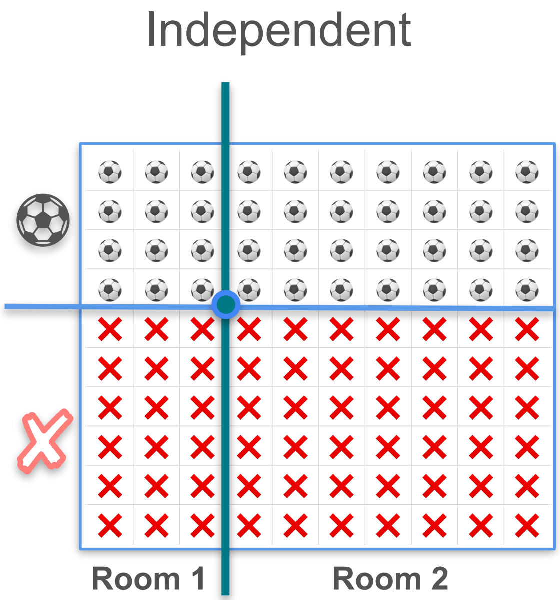


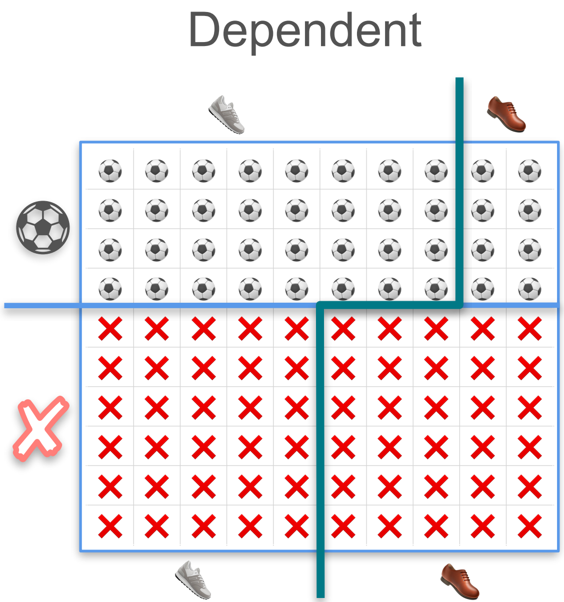
$\mathbf{P}(R \mid \text{not } S) = 0.5$

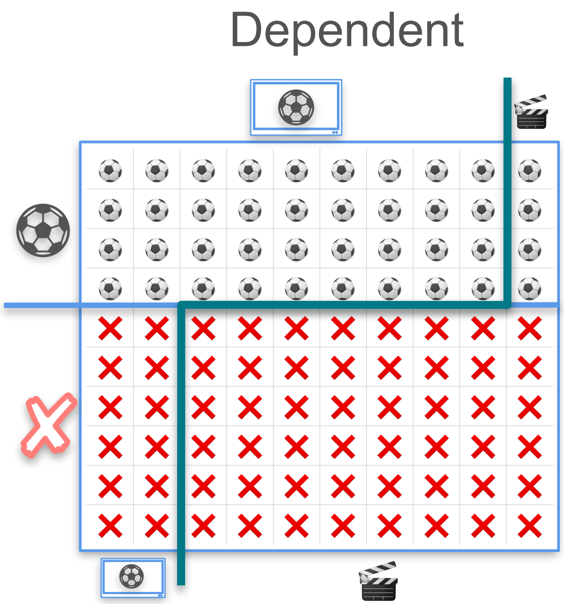
Conditional Probability



Independent vs Dependent Events





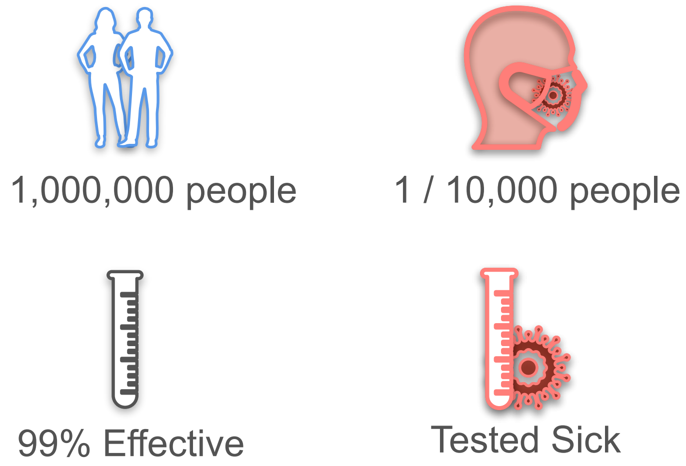


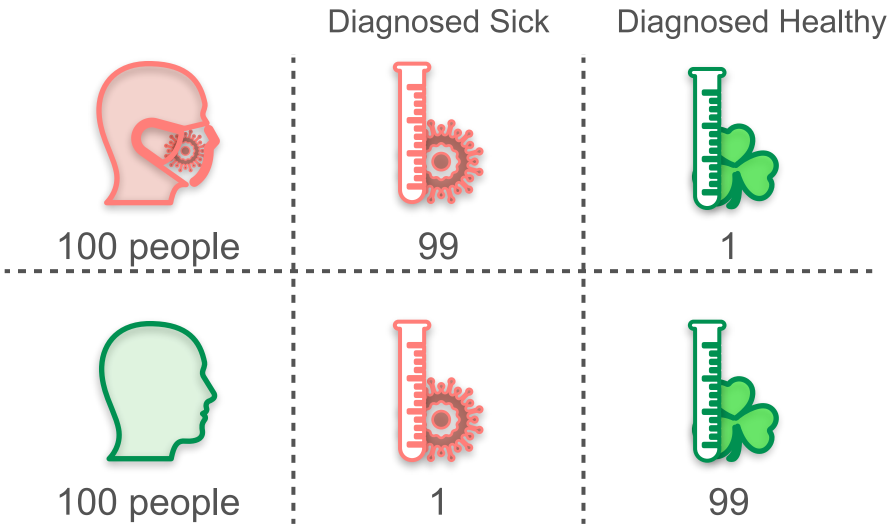
DeepLearning.AI

Introduction to probability

Bayes Theorem - Intuition

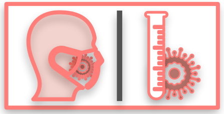
Bayes Theorem: Intuition





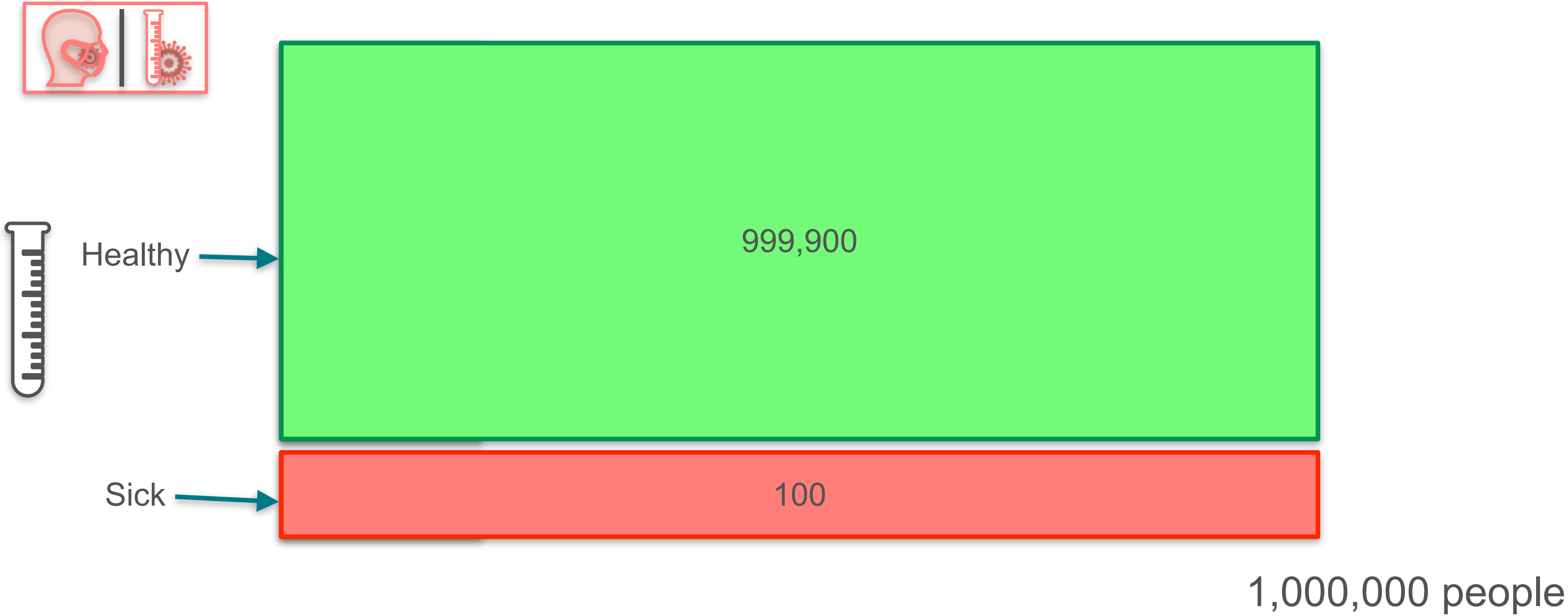
What's the probability that **you are sick GIVEN that you tested sick**?

Bayes Theorem: Intuition

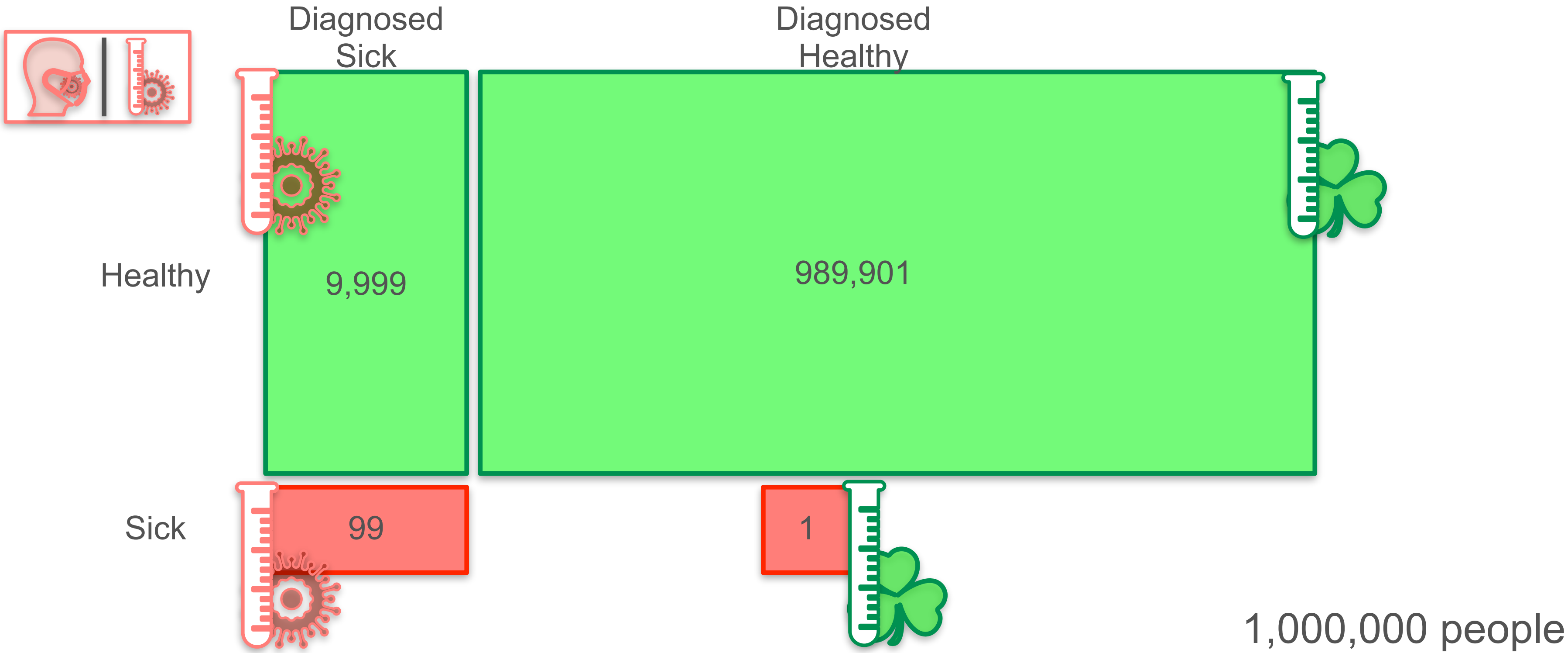


1,000,000 people

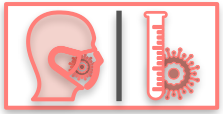
Bayes Theorem: Intuition



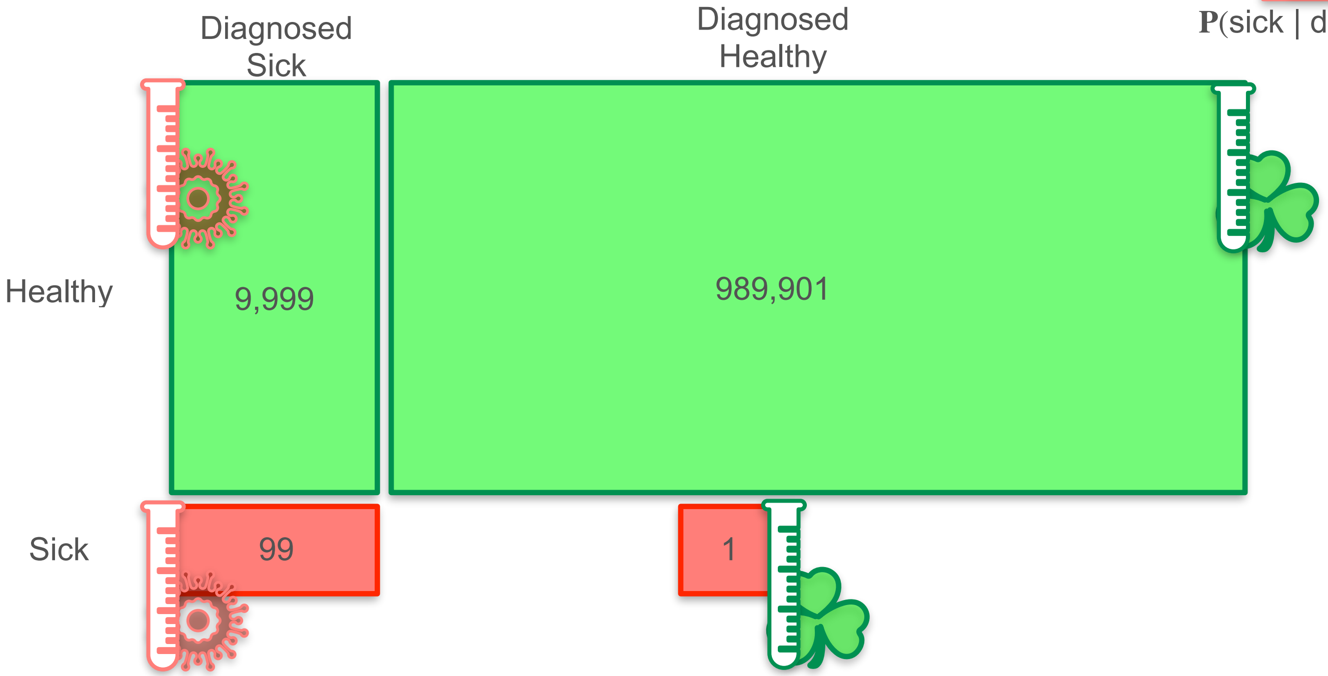
Bayes Theorem: Intuition



Bayes Theorem: Intuition

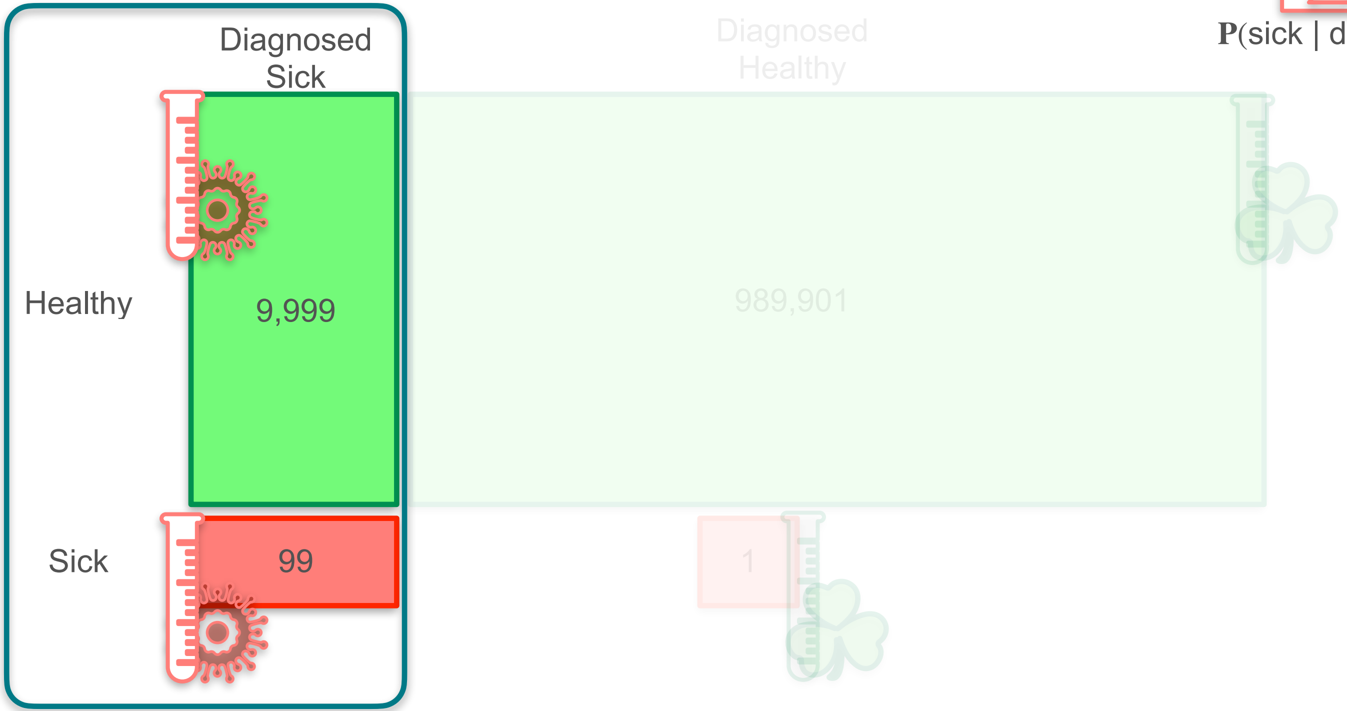


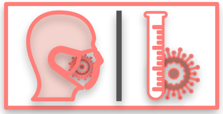
$\mathbf{P}(\text{sick} \mid \text{diagnosed sick}) = \underline{\qquad\qquad\qquad}$



1,000,000 people

Bayes Theorem: Intuition

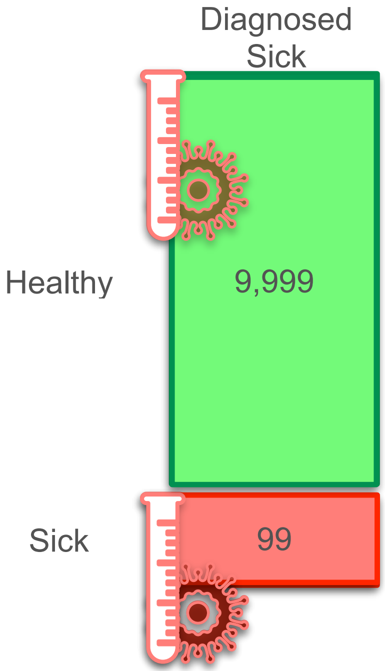


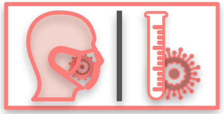


$\mathbf{P}(\text{sick} \mid \text{diagnosed sick}) = \underline{\hspace{4cm}}$

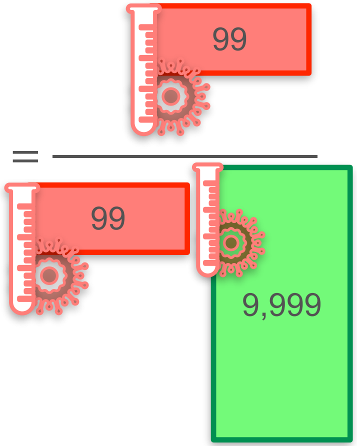
1,000,000 people

Bayes Theorem: Intuition





P(sick | diagnosed sick) =

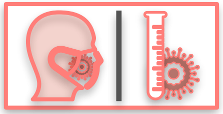


$$= \frac{99}{99 + 9999}$$

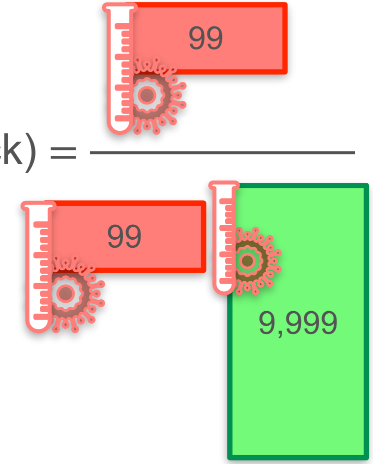
$$= \frac{99}{10098} = 0.0098$$

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

Bayes Theorem: Intuition



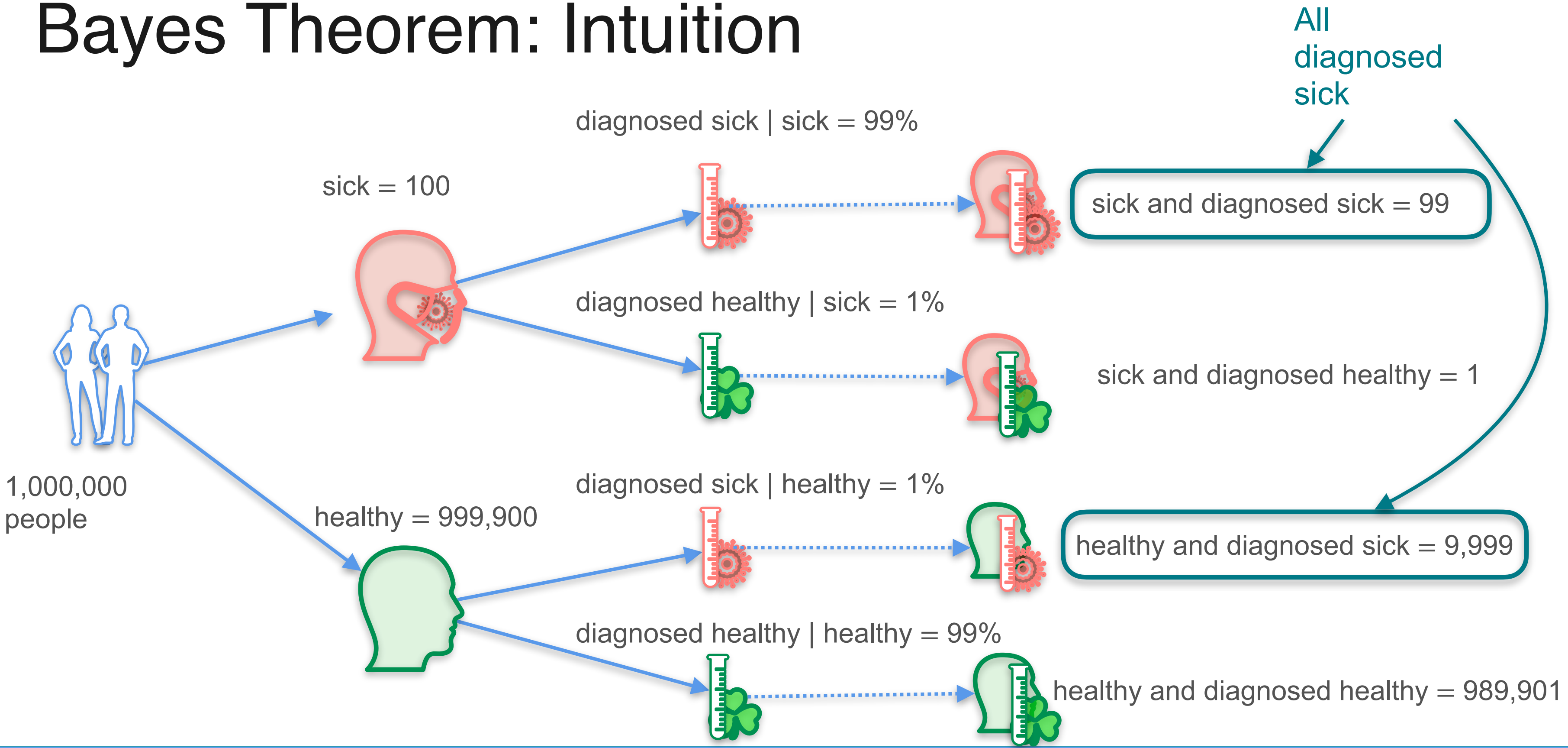
$\mathbf{P}(\text{sick} \mid \text{diagnosed sick}) =$



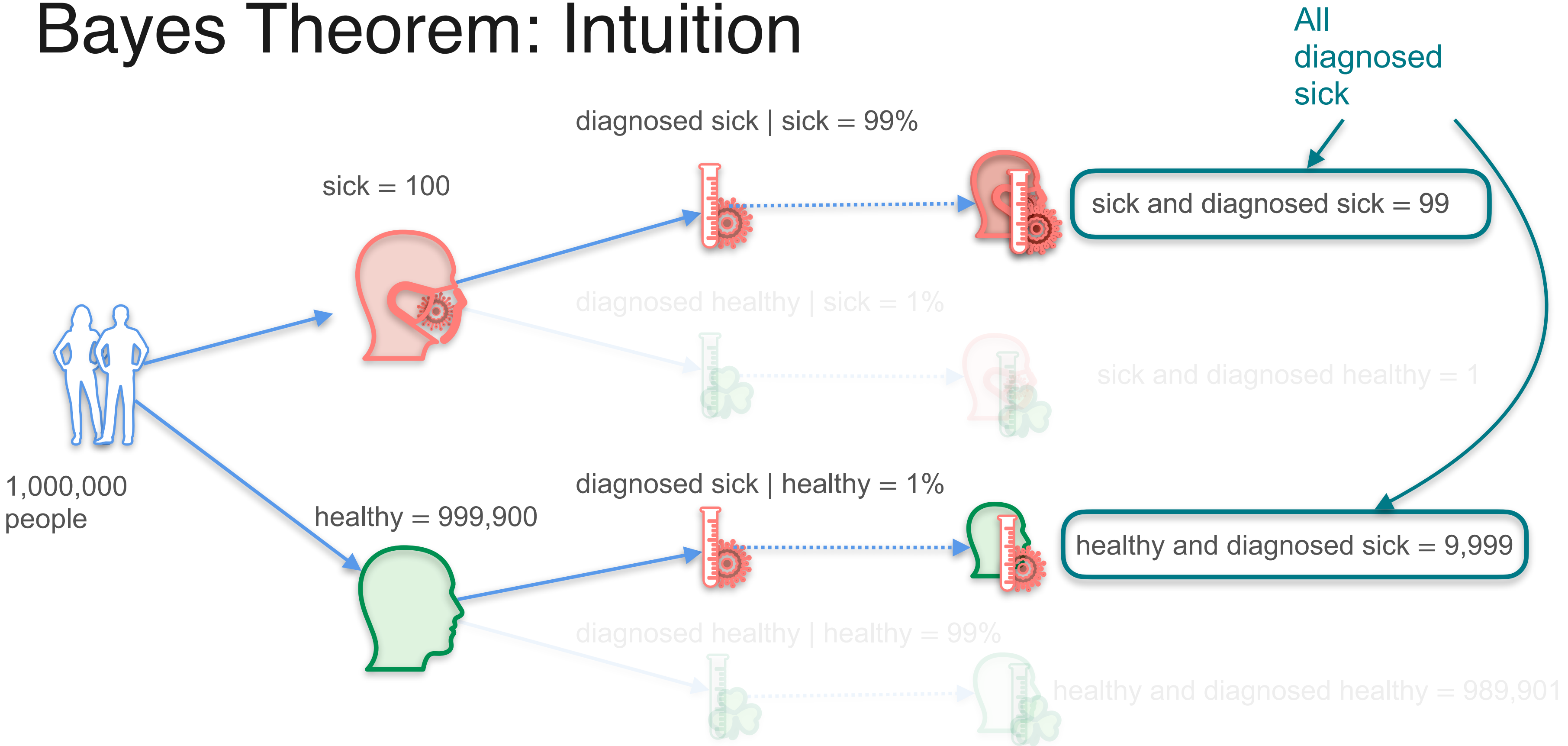
$$= \frac{\text{sick and diagnosed sick}}{\text{sick and diagnosed sick} + \text{healthy and diagnosed sick}}$$

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

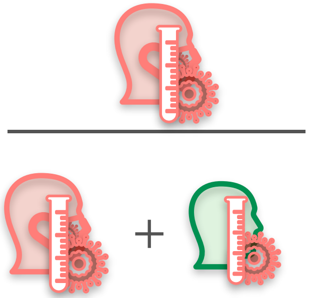
Bayes Theorem: Intuition



Bayes Theorem: Intuition



Bayes Theorem: Intuition

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

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

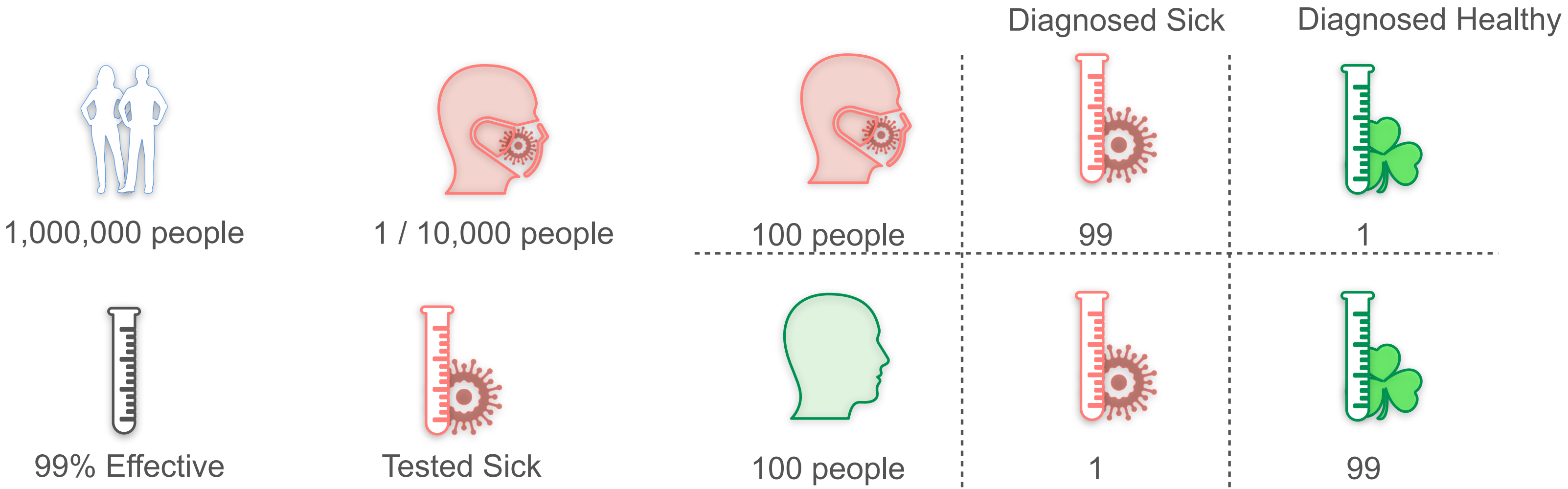
$$\mathbf{P}(\text{sick} \mid \text{diagnosed sick}) = \frac{99}{10098} = 0.0098$$

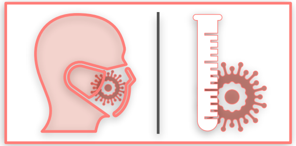
DeepLearning.AI

Introduction to probability

Bayes Theorem - Mathematical Formula

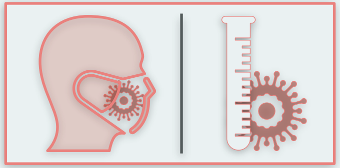
Bayes Theorem: Formula





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

Bayes Theorem: Formula



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

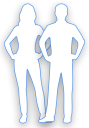
$\mathbf{P}(\text{sick}) = 0.01\%$

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

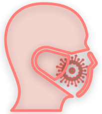
$\mathbf{P}(\text{diagnosed sick} \mid \text{sick}) = 99\%$

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

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



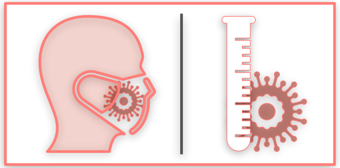
1,000,000



1 / 10,000

99% Effective

Bayes Theorem: Formula



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

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

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

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

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

A: sick

B: diagnosed sick

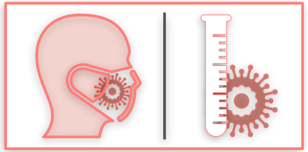
$\mathbf{P}(A \mid B) = ?$

From Conditional Probability

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

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

Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

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

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

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

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

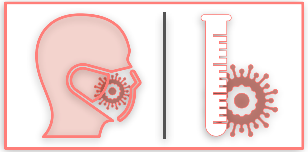
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(\text{sick and diagnosed sick})}{\mathbf{P}(\text{diagnosed sick})}$$

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

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

BAYES THEOREM FORMULA CAN HELP

Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

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

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

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

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

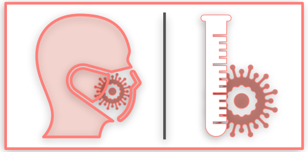
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(\text{sick and diagnosed sick}) = ?}{\mathbf{P}(\text{diagnosed sick}) = ?}$$

From Conditional Probability

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

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

Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

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

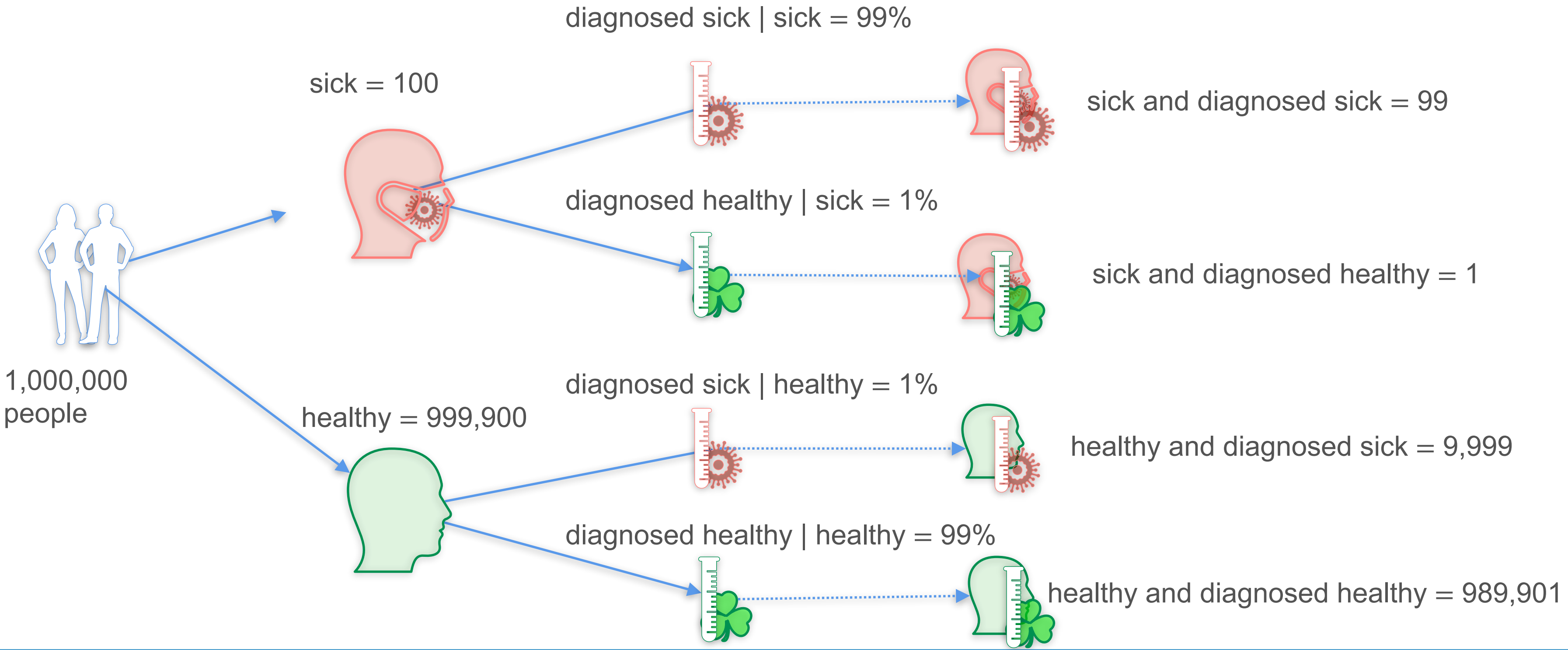
$\mathbf{P}(\text{diagnosed sick} \mid \text{sick}) = 99\%$

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

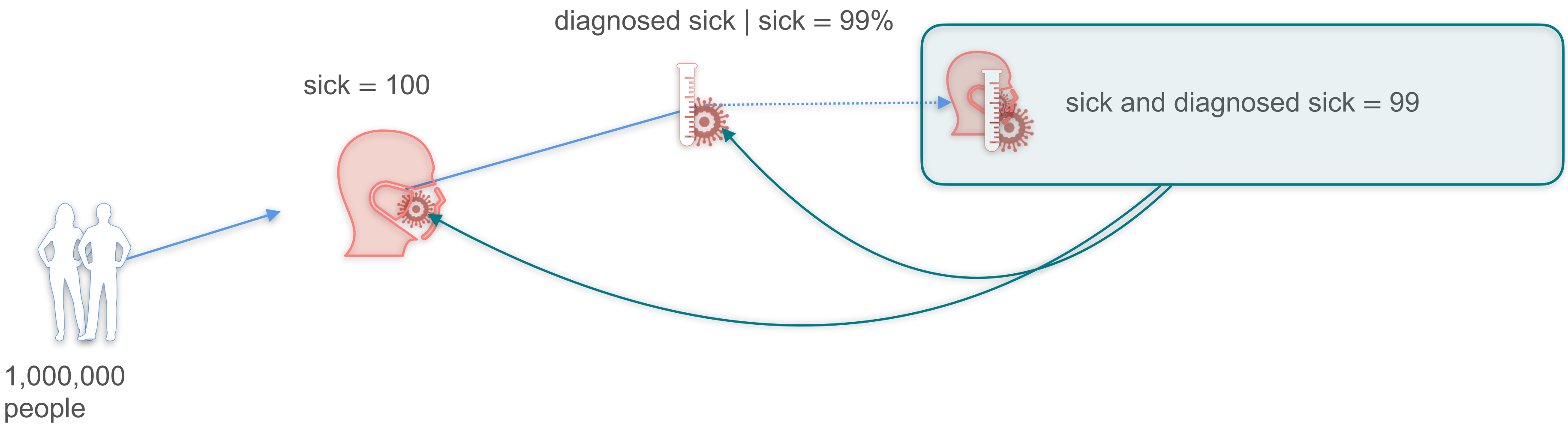
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

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

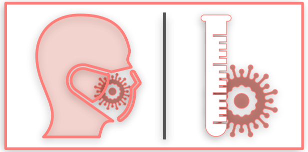
Bayes Theorem: Formula



Bayes Theorem: Formula



Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

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

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

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

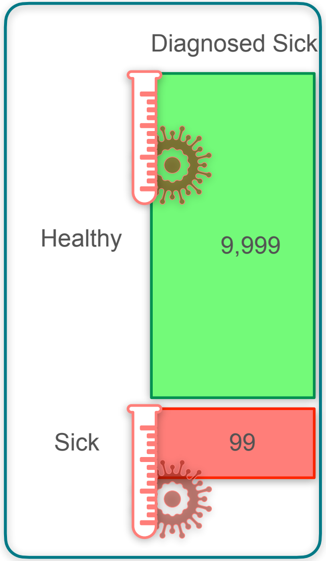
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

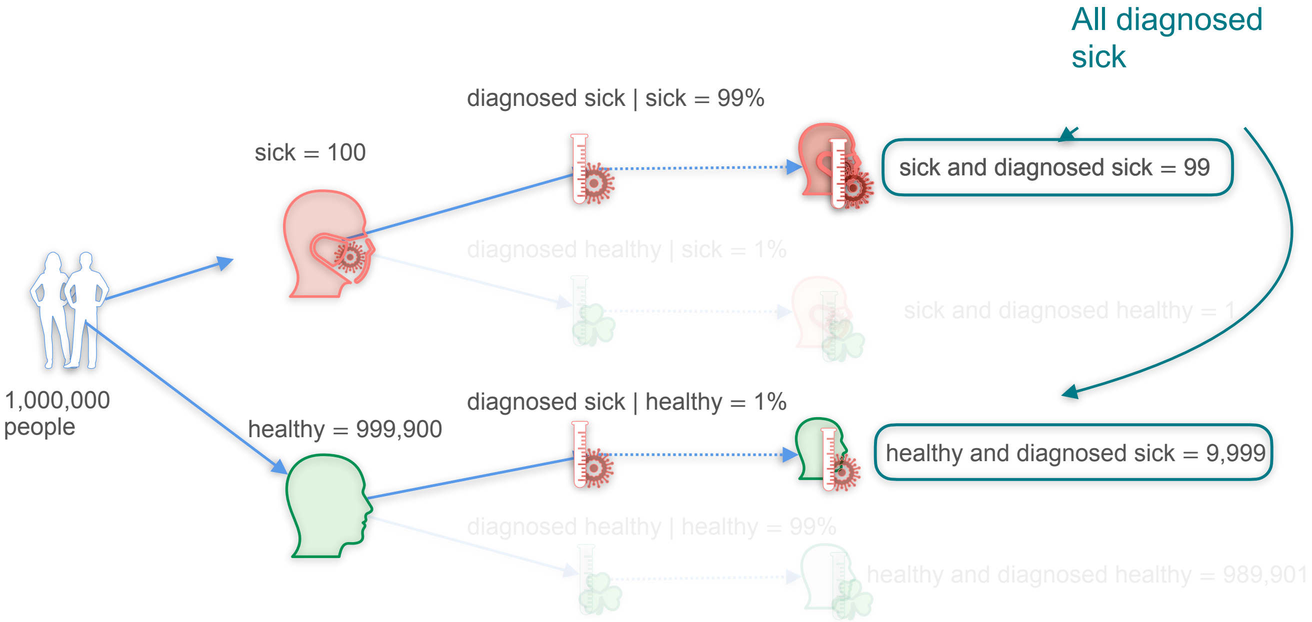
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick})}{\mathbf{P}(\text{diagnosed sick}) = ?}$$

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

Bayes Theorem: Formula

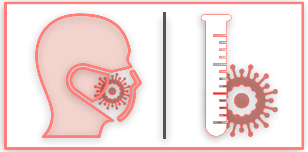
$\mathbf{P}(\text{diagnosed sick}) =$





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

Bayes Theorem: Formula



A: sick

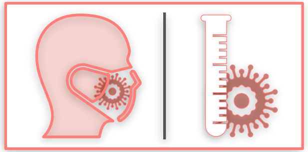
B: diagnosed sick

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

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

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

Bayes Theorem: Formula



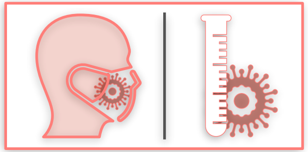
A: sick

B: diagnosed sick

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

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick})}{\mathbf{P}(\text{sick} \cap \text{diagnosed sick}) + \mathbf{P}(\text{not sick} \cap \text{diagnosed sick})}$$

Bayes Theorem: Formula



A: sick

B: diagnosed sick

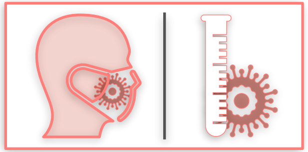
$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

$$\mathbf{P}(\text{A} \mid \text{B}) = \frac{\mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick})}{\mathbf{P}(\text{sick} \cap \text{diagnosed sick}) + \mathbf{P}(\text{not sick} \cap \text{diagnosed sick})}$$

$$\begin{aligned}\mathbf{P}(\text{sick} \cap \text{diagnosed sick}) &= \mathbf{P}(A \cap B) \\ &= \mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick})\end{aligned}$$

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

Bayes Theorem: Formula



A: sick

B: diagnosed sick

$$\mathbf{P}(A \mid B) = \frac{\boxed{\mathbf{P}(A \cap B)}}{\boxed{\mathbf{P}(B)}} \quad ?$$

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick})}{\mathbf{P}(\text{sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{sick}) + \mathbf{P}(\text{not sick}) \bullet \mathbf{P}(\text{diagnosed sick} \mid \text{not sick})}$$

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

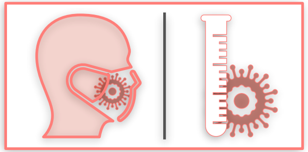
$\mathbf{P}(\text{not sick}) = 99.99\%$

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

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

BAYES THEOREM FORMULA

Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

$\mathbf{P}(A)$ = 0.01%

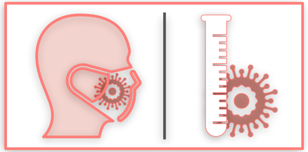
$\mathbf{P}(A')$ = 99.99%

$\mathbf{P}(B \mid A)$ = 99%

$\mathbf{P}(B \mid A')$ = 1%

BAYES THEOREM FORMULA

Bayes Theorem: Formula



A: sick

B: diagnosed sick

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

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

$\mathbf{P}(A)$ = 0.01%

$\mathbf{P}(A')$ = 99.99%

$\mathbf{P}(B \mid A)$ = 99%

$\mathbf{P}(B \mid A')$ = 1%

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

$$\mathbf{P}(A \mid B) = 0.0098$$

BAYES THEOREM FORMULA

DeepLearning.AI

Introduction to probability

Bayes Theorem - Spam Example

Bayes Theorem: Spam Example

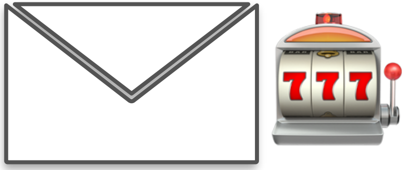
Bayes Theorem: Spam Example

20 spam

80 not spam (ham)

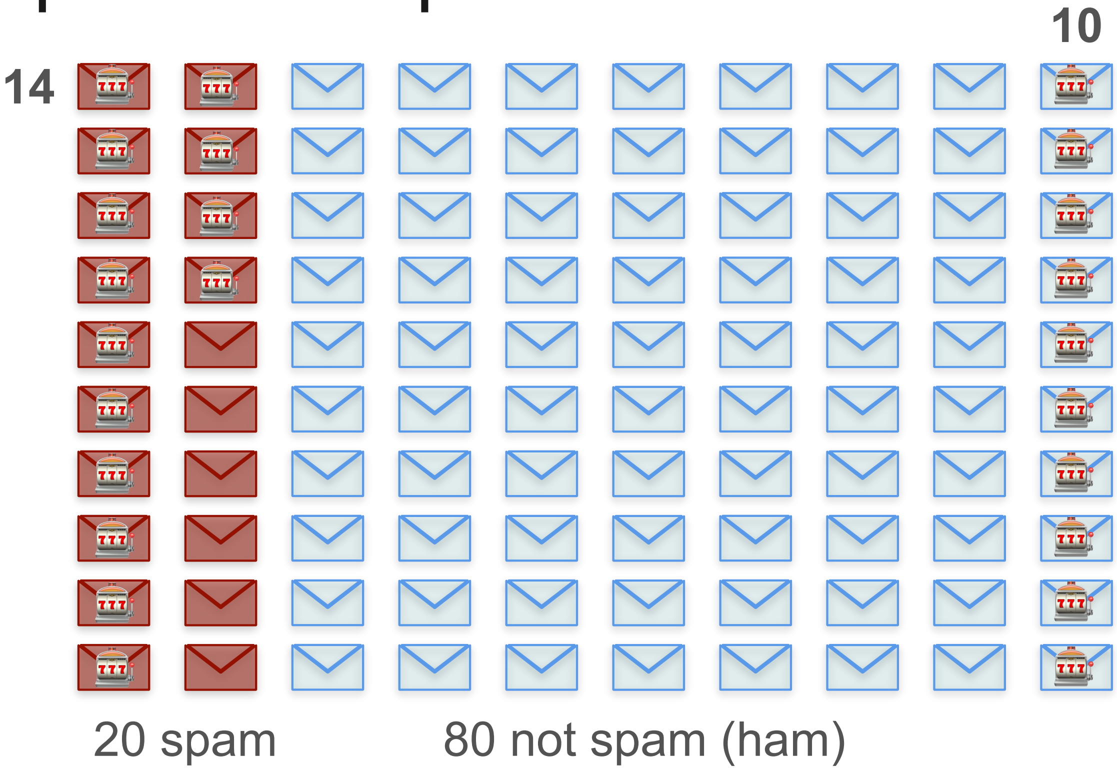
“lottery”

Bayes Theorem: Spam Example

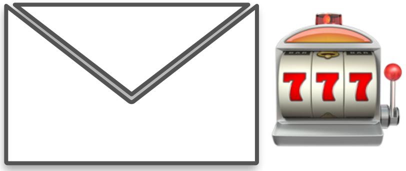


What is the probability that an email containing lottery is a spam?

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

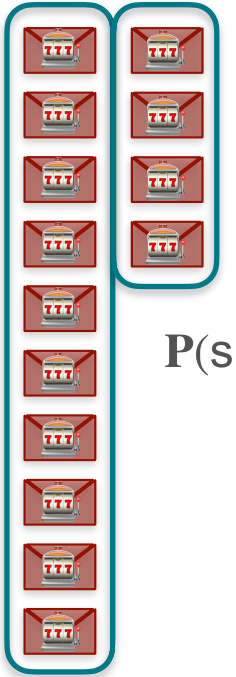


Bayes Theorem: Spam Example (Intuition Solution)



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

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



14

24 emails

containing lottery

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

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

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

10

Bayes Theorem: Spam Example (Formula Solution)

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

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

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

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

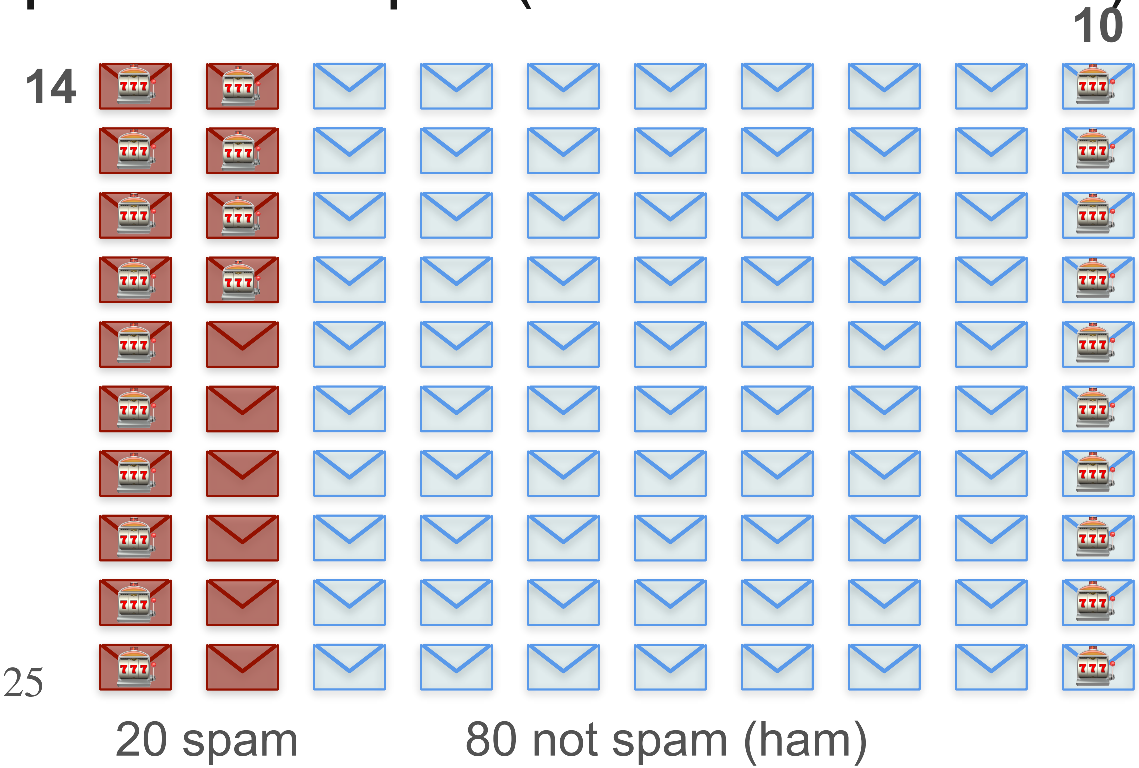
Bayes Theorem: Spam Example (Formula Solution)

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

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

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

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



Bayes Theorem: Spam Example (Formula Solution)

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

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

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

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

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

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

DeepLearning.AI

Introduction to probability

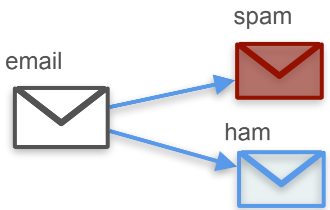
Bayes Theorem - Prior and Posterior

Bayes Theorem

PRIOR	EVENT	POSTERIOR
$\mathbf{P}(A)$	E	$\mathbf{P}(A \mid E)$

Prior and Posterior

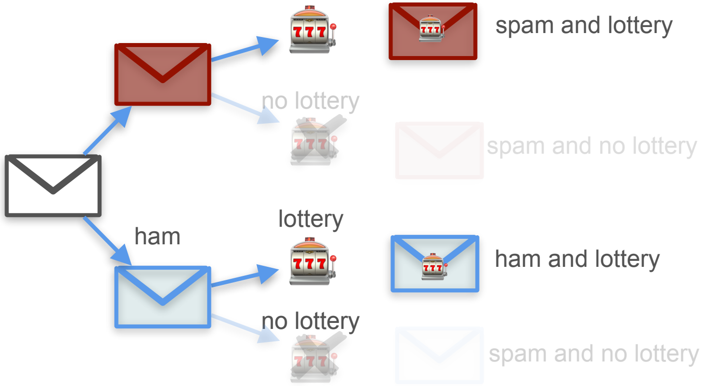
PRIOR



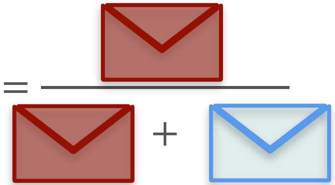
EVENT

Email contains lottery

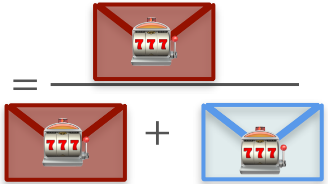
POSTERIOR



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



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



Prior and Posterior

PRIOR

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

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

EVENT

?

1st dice is 6

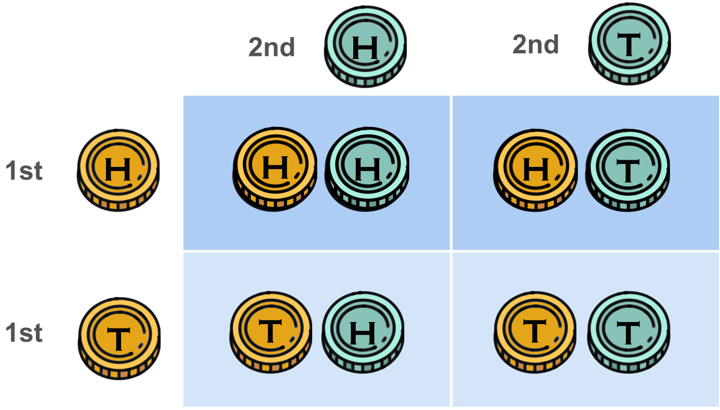
POSTERIOR

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

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

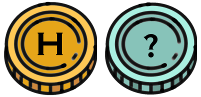
Prior and Posterior

PRIOR



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

EVENT



1st coin is H

POSTERIOR

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

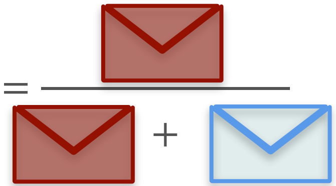
DeepLearning.AI

Introduction to probability

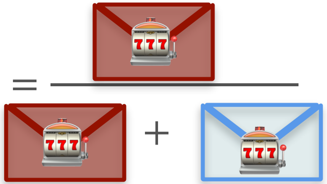
Bayes Theorem - The Naive Bayes Model

What About 2 Events?

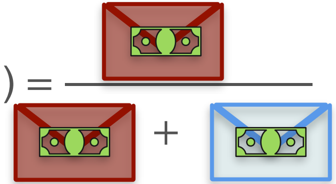
PRIOR | **EVENT** | **POSTERIOR**

P(spam) = 

Email contains ‘lottery’

P(spam | lottery) = 

Email contains ‘winning'

P(spam | winning) = 

Email contains ‘lottery’ and ‘winning’

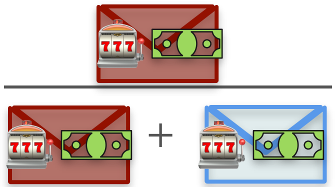
?

What About 2 Events?

EVENT

Email contains ‘lottery’ and ‘winning’

POSTERIOR

$\mathbf{P}(\text{spam} \mid \text{lottery \& winning}) =$ 

$$\frac{\text{\# Spam emails with 'lottery' and 'winning'}}{\text{\# Spam emails}}$$

?

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

What About More Than 2 Events?

EVENT | **POSTERIOR**

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

$$\frac{\cancel{\text{\# Spam emails with } w_1, \ldots, w_{100}}\ 0}{\text{\#Spam emails}}$$

?

$$\mathbf{P}(\text{spam} \mid w_1, \ldots, w_{100}) = \frac{\mathbf{P}(\text{spam})\ \mathbf{P}(w_1, \ldots, w_{100} \mid \text{spam})}{\mathbf{P}(\text{spam}) \cdot \mathbf{P}(w_1, \ldots, w_{100} \mid \text{spam}) + \mathbf{P}(\text{ham}) \cdot \mathbf{P}(w_1, \ldots, w_{100} \mid \text{ham})}$$

Is There a Quicker Way To Estimate the Probability?

Naive assumption

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

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

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

Is There a Quicker Way To Estimate the Probability?

Naive assumption

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

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

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

Is There a Quicker Way To Estimate the Probability?

Naive assumption

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

$$\mathbf{P}(\text{spam} \mid w_1, \ldots, w_n) = \frac{\mathbf{P}(\text{spam}) \cdot \mathbf{P}(w_1 \mid \text{spam}) \cdots \mathbf{P}(w_n \mid \text{spam})}{\mathbf{P}(\text{spam}) \cdot \mathbf{P}(w_1 \mid \text{spam}) \cdots \mathbf{P}(w_n \mid \text{spam}) + \mathbf{P}(\text{ham}) \cdot \mathbf{P}(w_1 \mid \text{ham}) \cdots \mathbf{P}(w_n \mid \text{ham})}$$

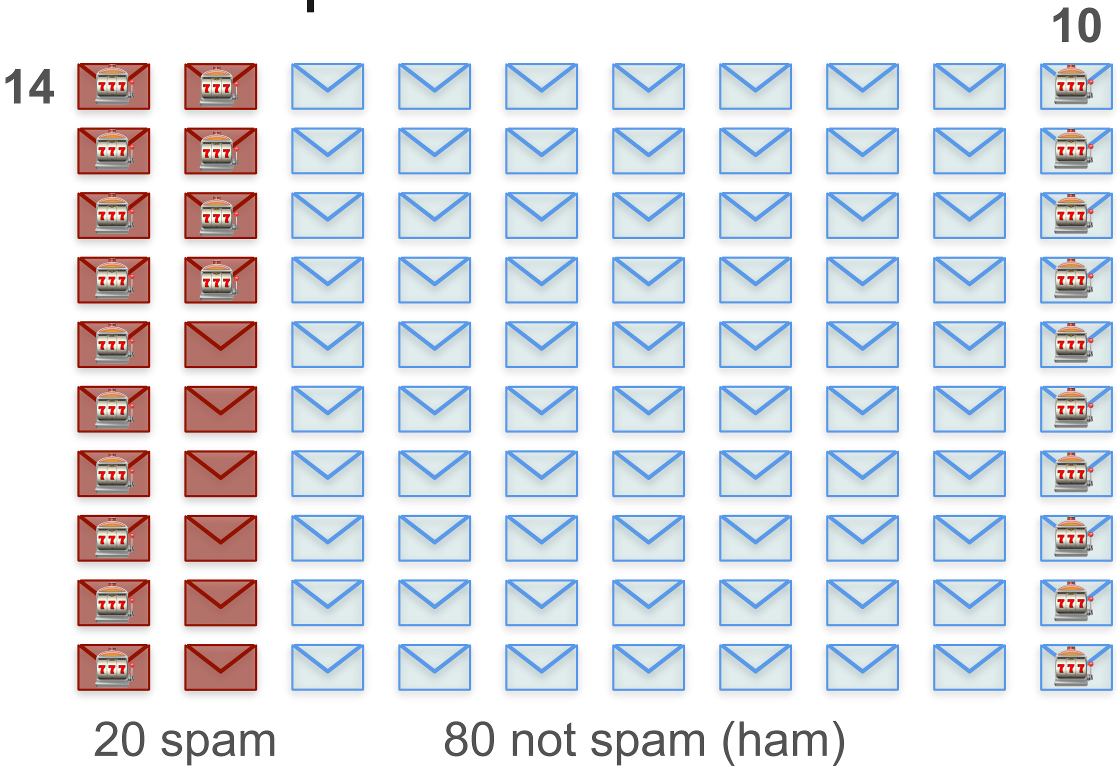
Naive Bayes: Spam Example

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

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

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

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



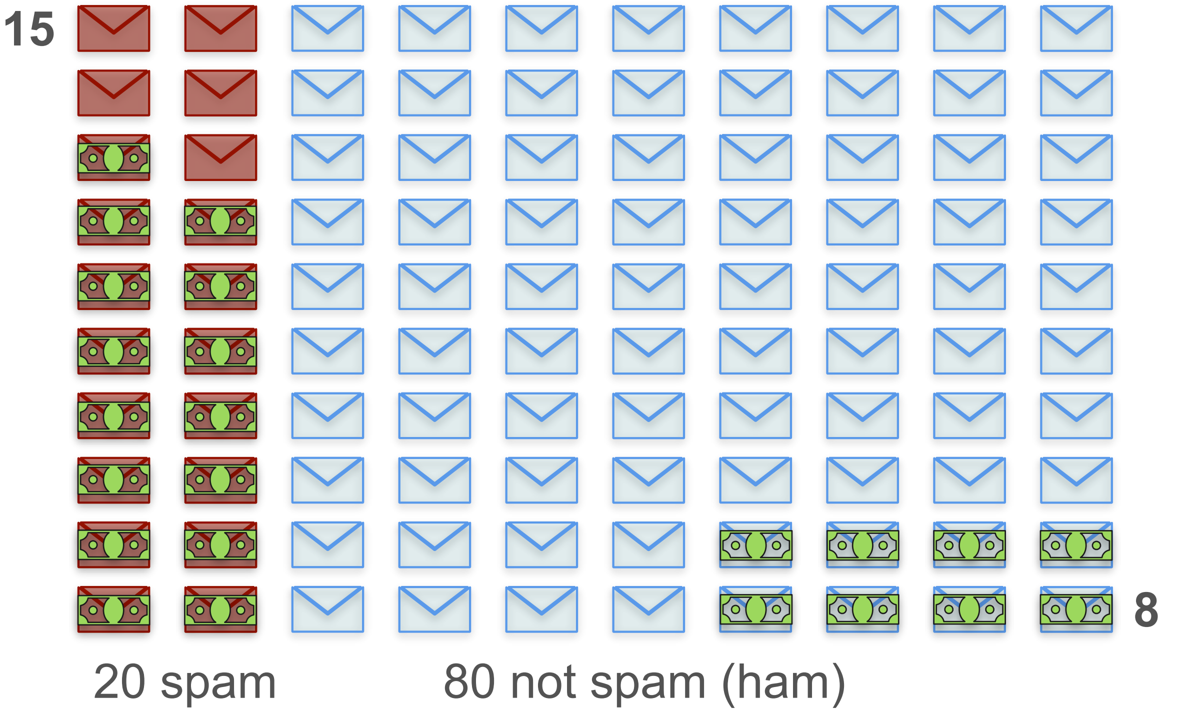
Naive Bayes: Spam Example

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

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

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

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



Naive Bayes: Spam Example

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

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

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

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

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

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

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

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

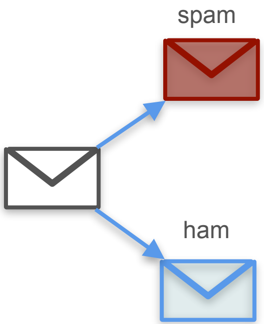
DeepLearning.AI

Introduction to probability

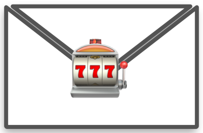
Probability in Machine Learning

Bayes Theorem

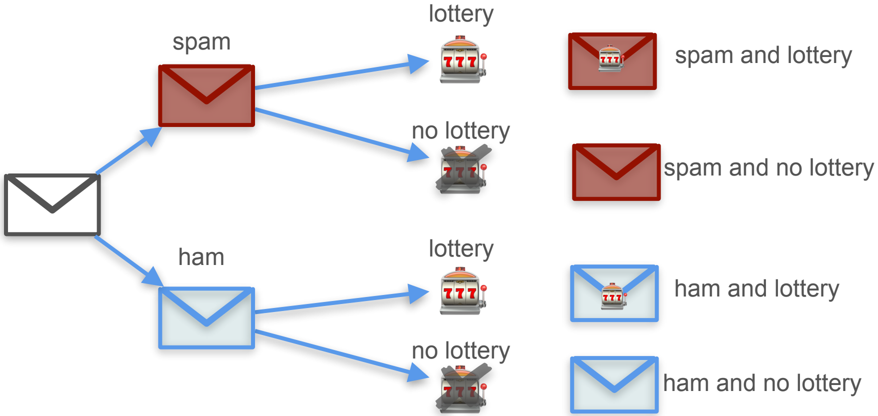
PRIOR

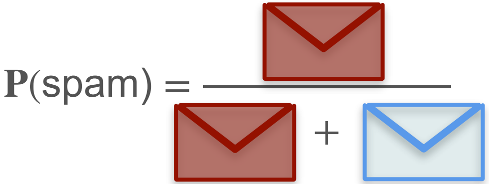


EVENT



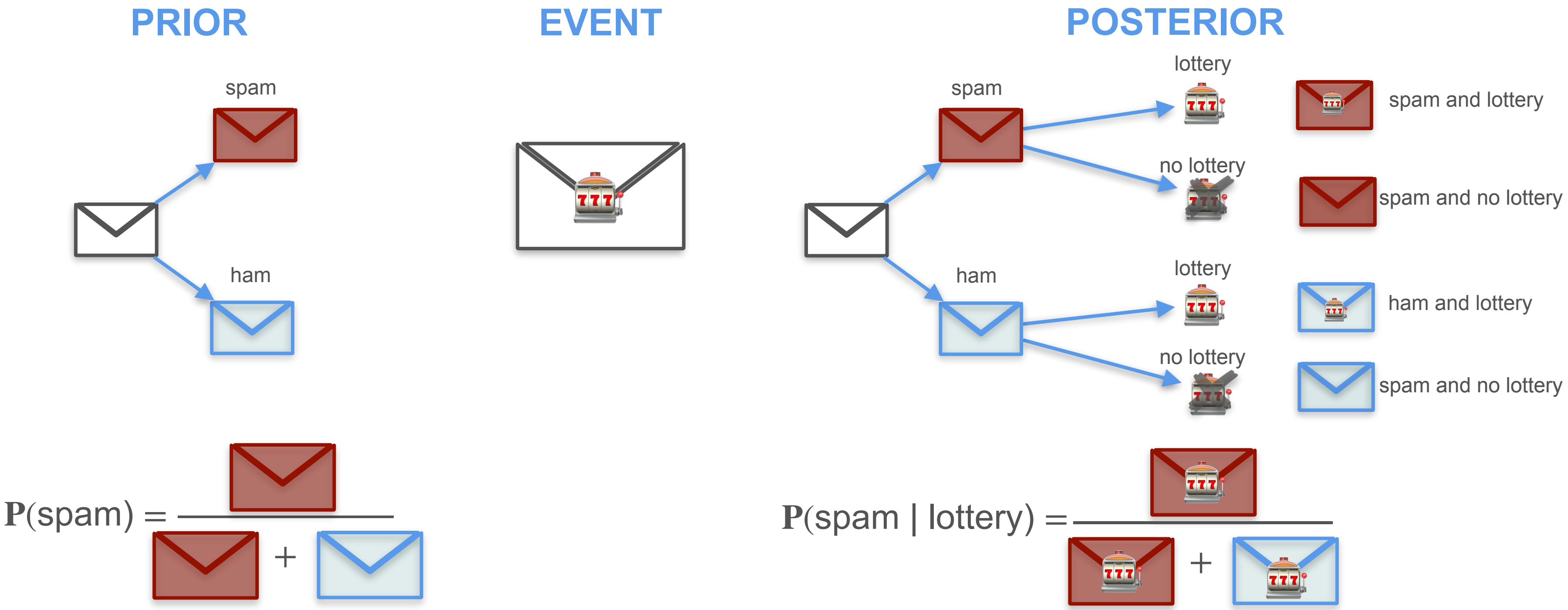
POSTERIOR





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

Bayes Theorem



Example Problem

Image recognition

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

Example Problem

Classification

Patient 1	
Age	29
Gender	Female
Height	169 cm
Weight	62 kg
Smoker	No
...	...
Heart rate	63
Blood pressure	120 90

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

Example Problem

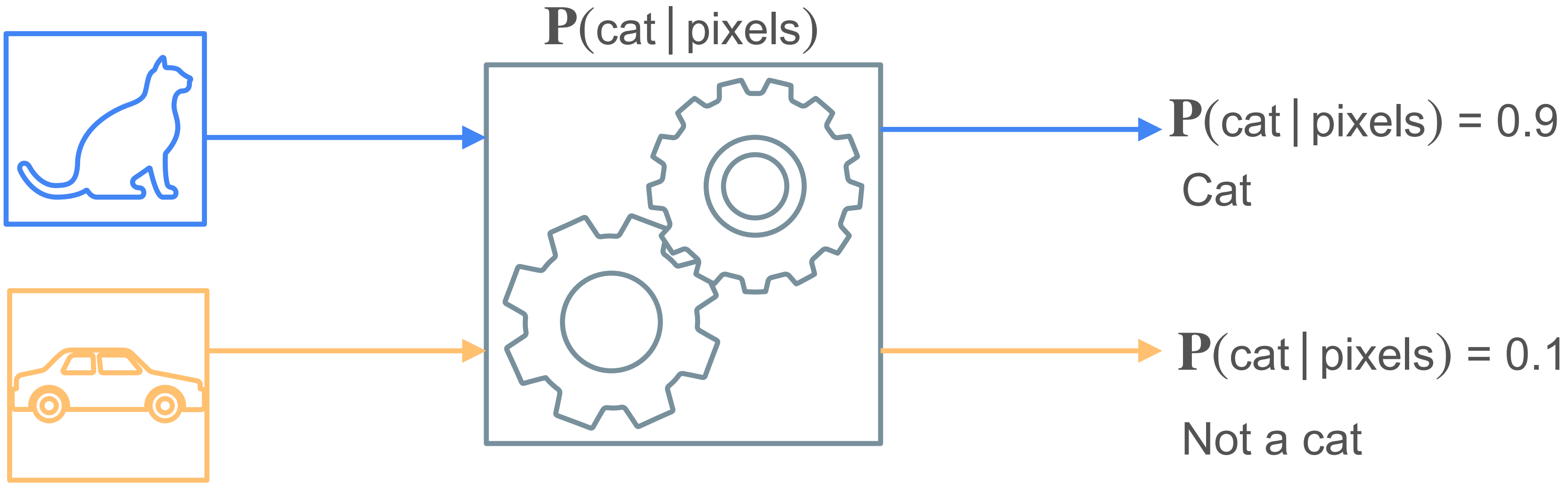
Sentiment analysis

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

Matsuo Bashō

- Is this a happy sentence?
- Calculate P(happy I 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 $\mathbf{P}(\text{face} \mid \text{pixels})$ is high.

Image generated by a StyleGAN

W1 Lesson 2

Probability Distributions

DeepLearning.AI

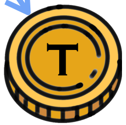
DeepLearning.AI

Probability Distributions

Random Variables

From Events to Random Variables

Flip a coin



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

$X =$ Number of heads

Flip a coin

1 heads $\quad$ 0 heads

$X = 1$ $\quad$ $X = 0$

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

X is a random variable

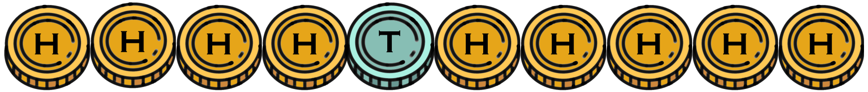
From Events to Random Variables

$X =$ Number of heads in 10 coin tosses

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

$X = 10$ $\quad 0.5^{10}$

$X = 9$ $\quad 0.5^{9}0.5$



$X = 9$ $\quad 0.5^{9}0.5$

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

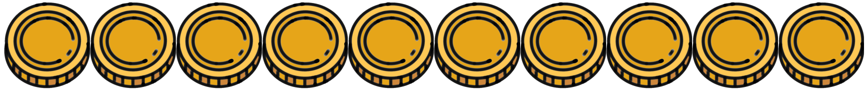
$\mathbf{P}(X = 1)?$

$\dots$

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

From Events to Random Variables

$X =$ Number of heads in 10 coin tosses



$X = 5$

$X = 6$

$X = 5$

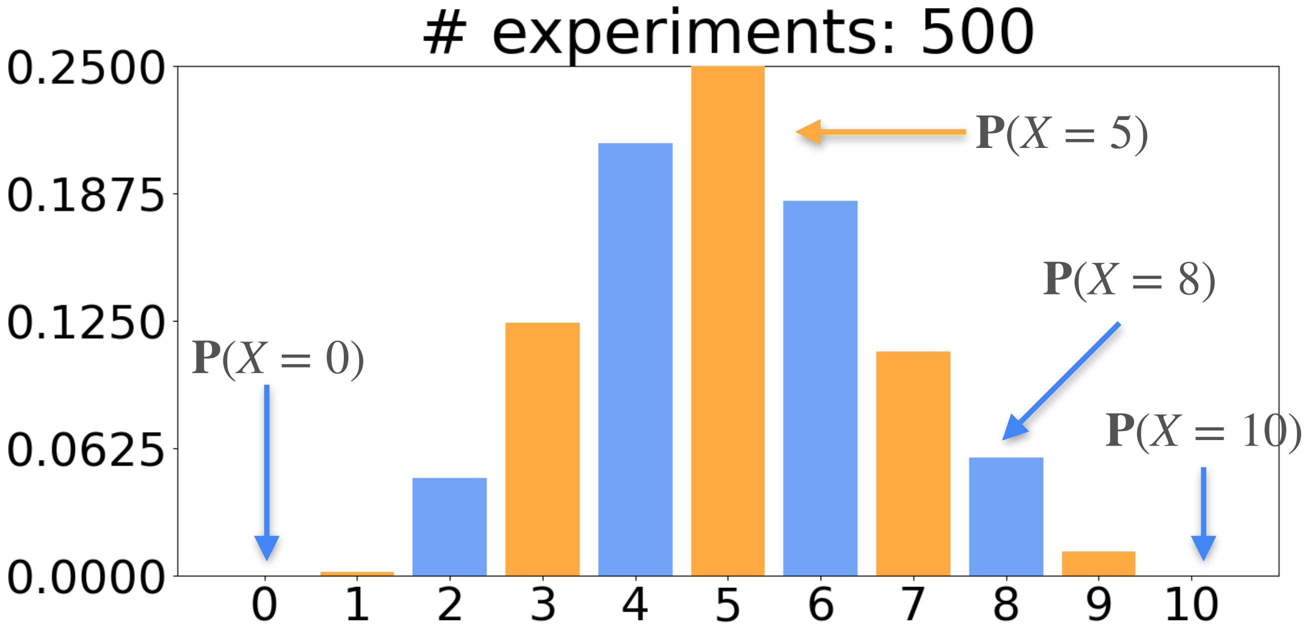
Possible outcomes:

0	4	8
1	5	9
2	6	10
3	7	

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

Repeat 500 times

Flipping a Fair Coin 500 Times

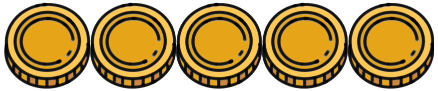


$\mathbf{P}(H) = 0.5$

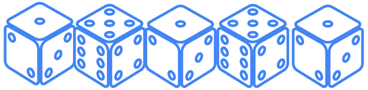
Repeat 500 times

Why Random Variables?

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



$X =$ Number of heads



$X =$ Number of 1's

$X =$ Number of sick patients

?

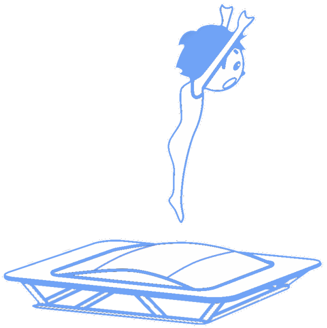
$\mathbf{P}(X = 1) = 0.5$

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

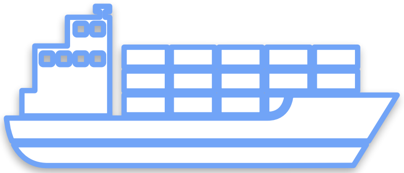
$\mathbf{P}(X = 3.14159) = 0.3$

Other Random Variables

Wait time until the next bus arrives



Height of an gymnast's jump

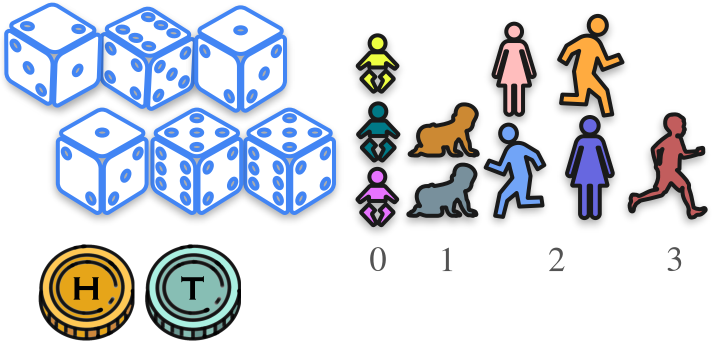


Number of defective products in a shipment

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

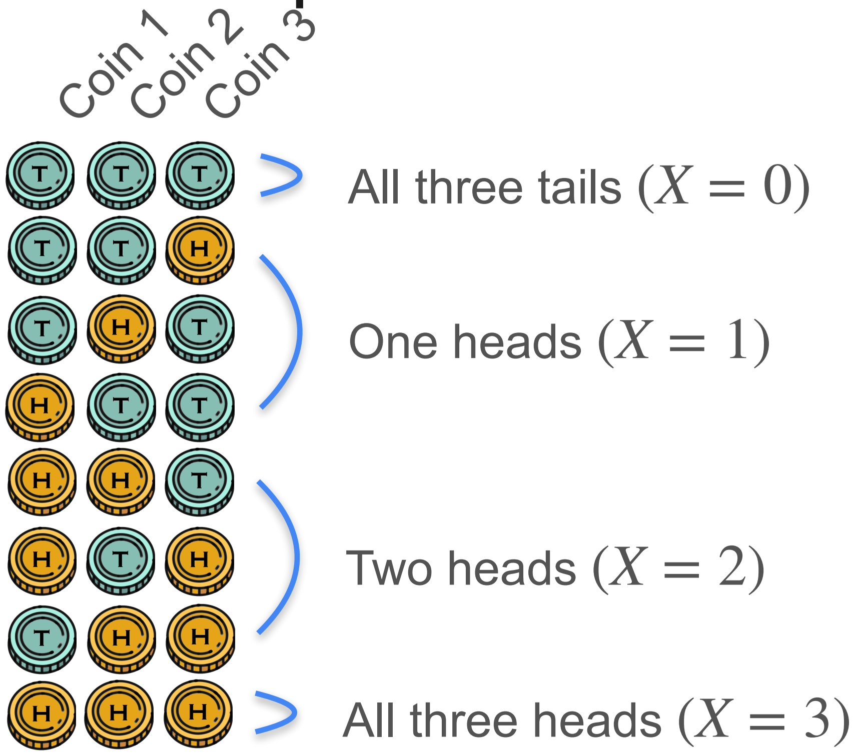
DeepLearning.AI

Probability Distributions

Probability Distributions (Discrete)

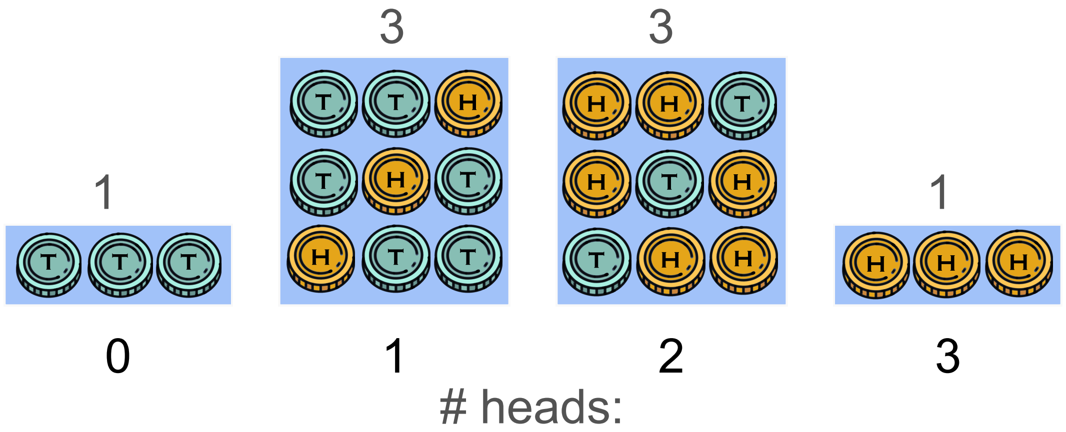
Discrete Distributions: Flip Three Coins

X_1: number of heads in **3** coin tosses



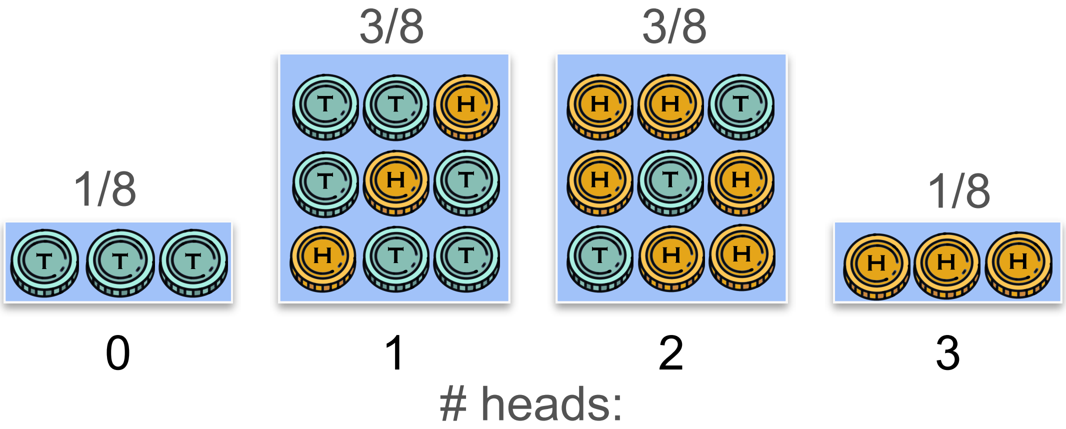
Discrete Distributions: Flip Three Coins

X_1: number of heads in **3** coin tosses



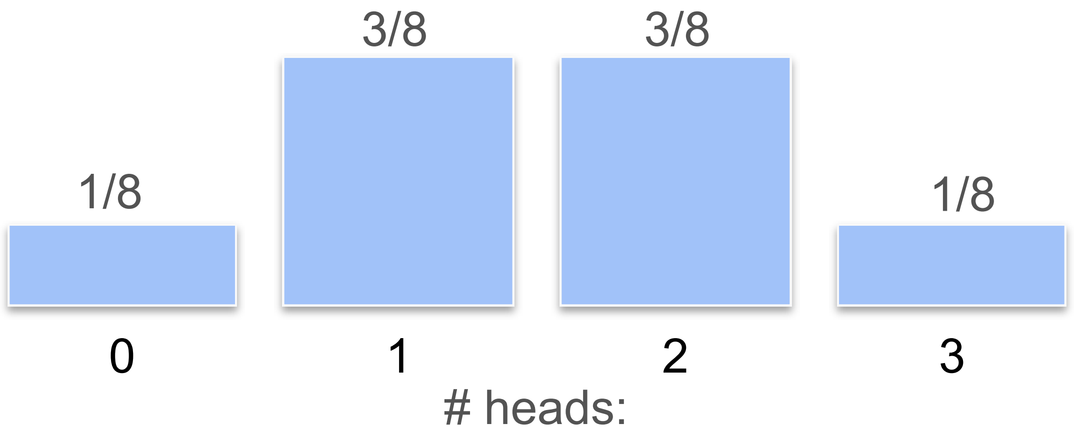
Discrete Distributions: Flip Three Coins

X_1: number of heads in **3** coin tosses



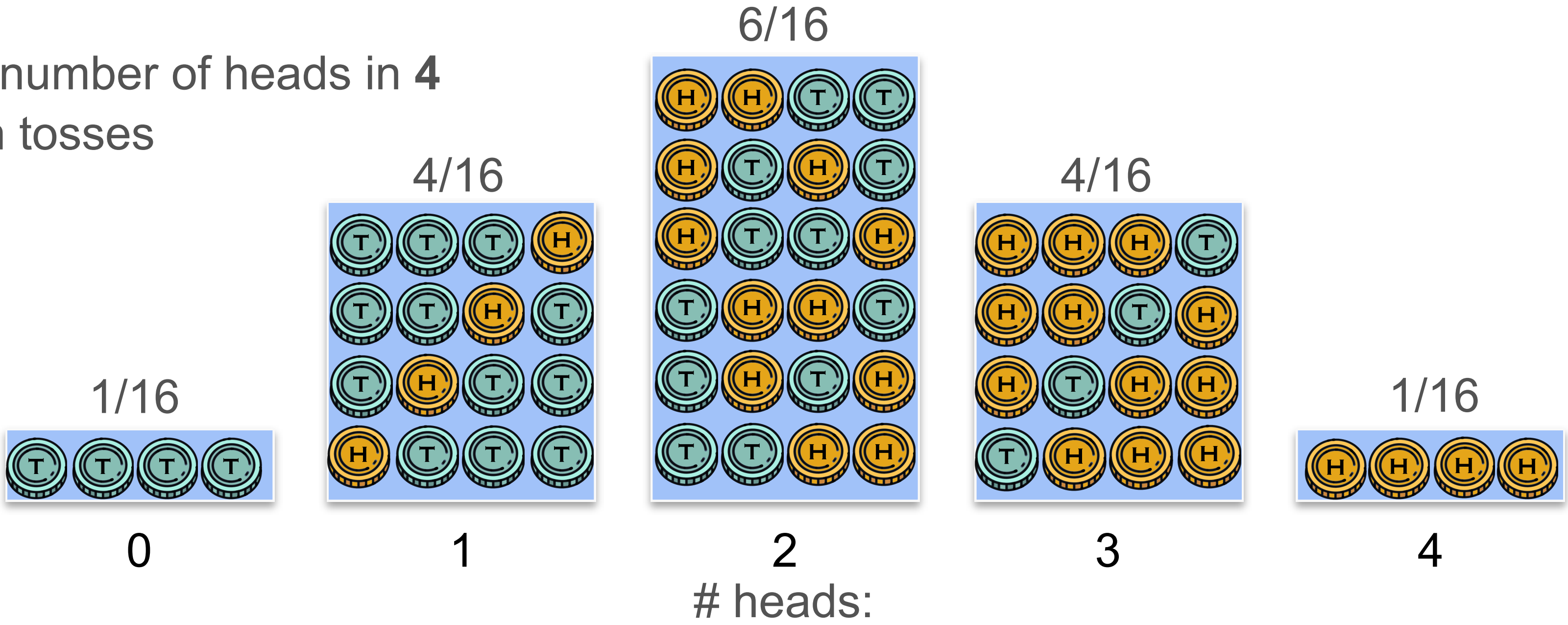
Discrete Distributions: Flip Three Coins

X_1: number of heads in **3** coin tosses



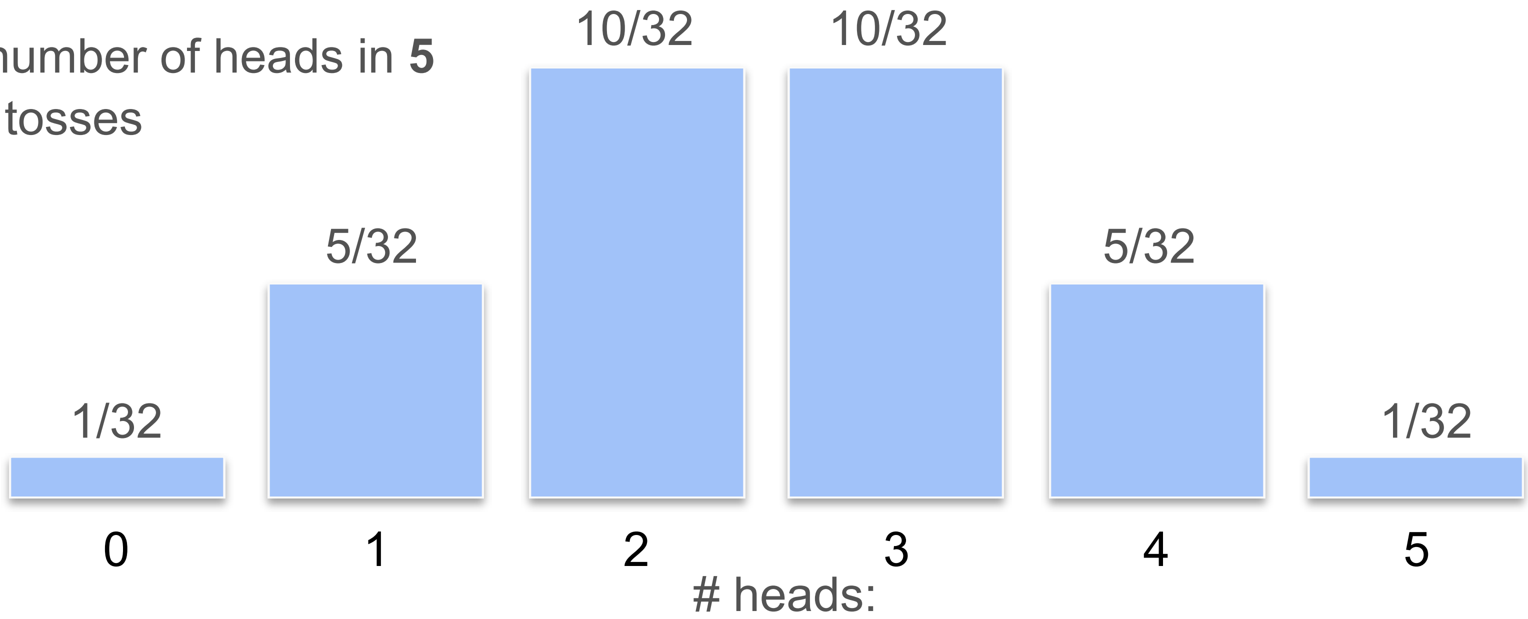
Discrete Distributions: Flip Four Coins

X_2: number of heads in **4** coin tosses



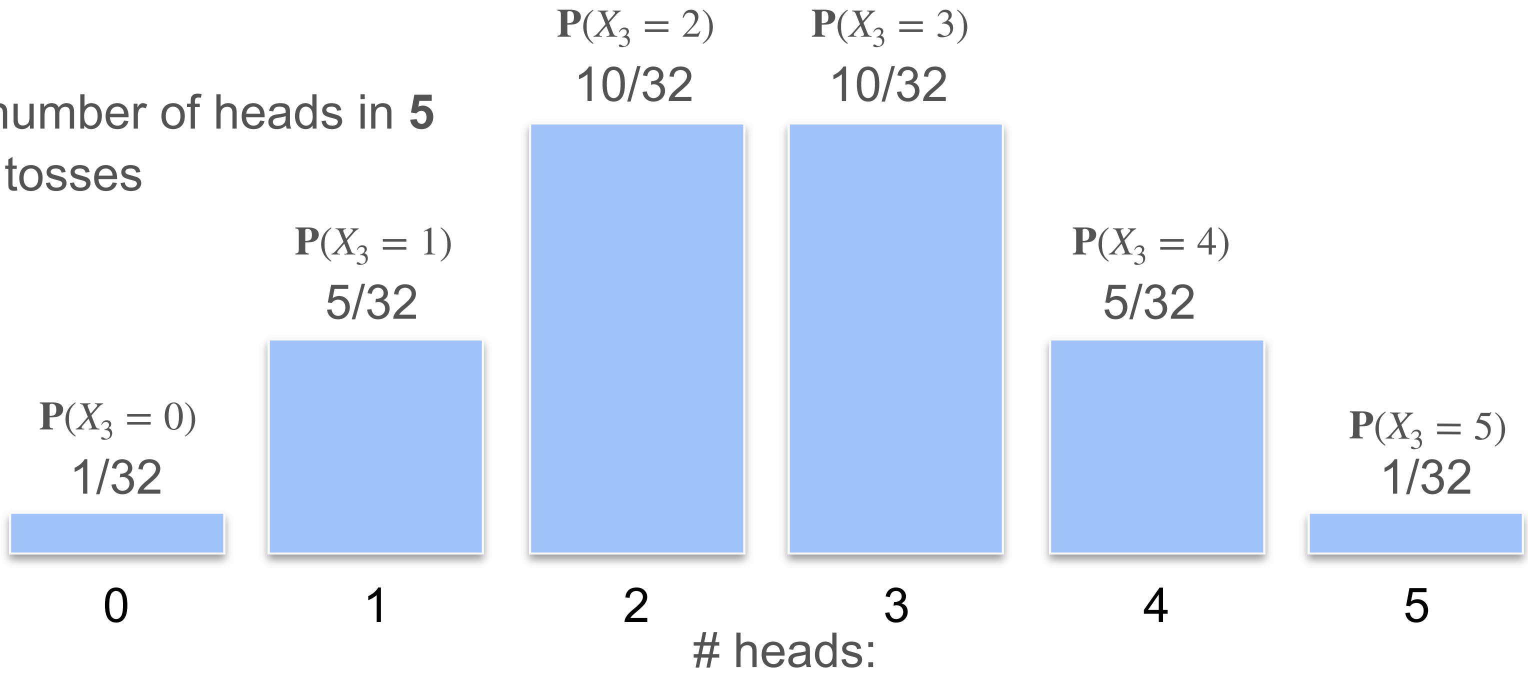
Discrete Distributions: Flip Five Coins

X_3: number of heads in **5** coin tosses



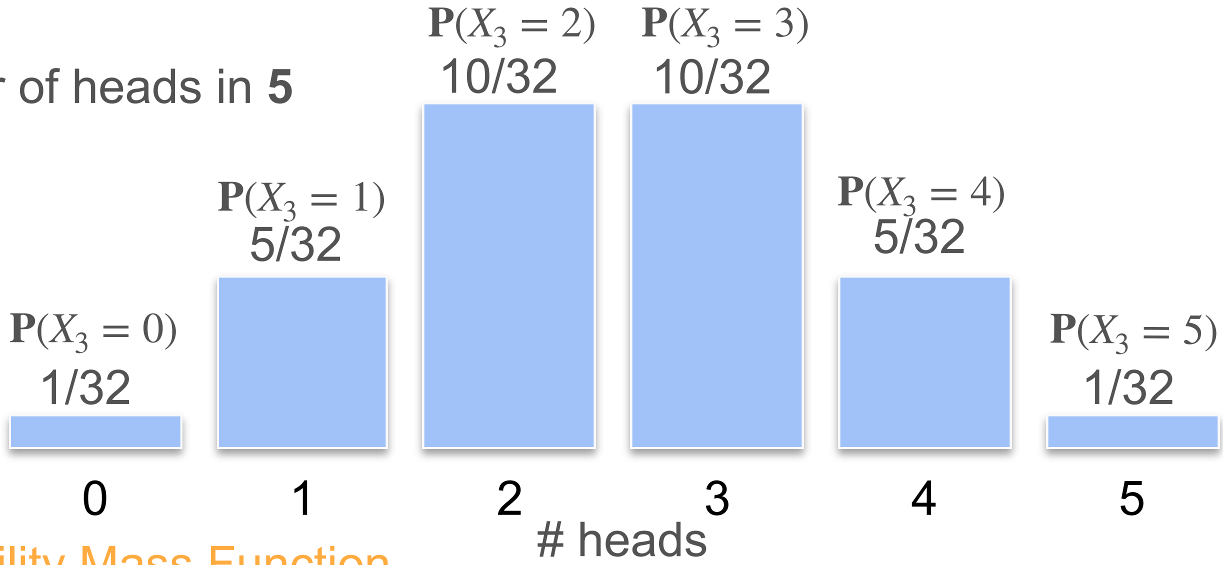
Discrete Distributions: Flip Five Coins

X_3: number of heads in **5** coin tosses



Discrete Distributions: Flip Five Coins

X_3: number of heads in **5** coin tosses

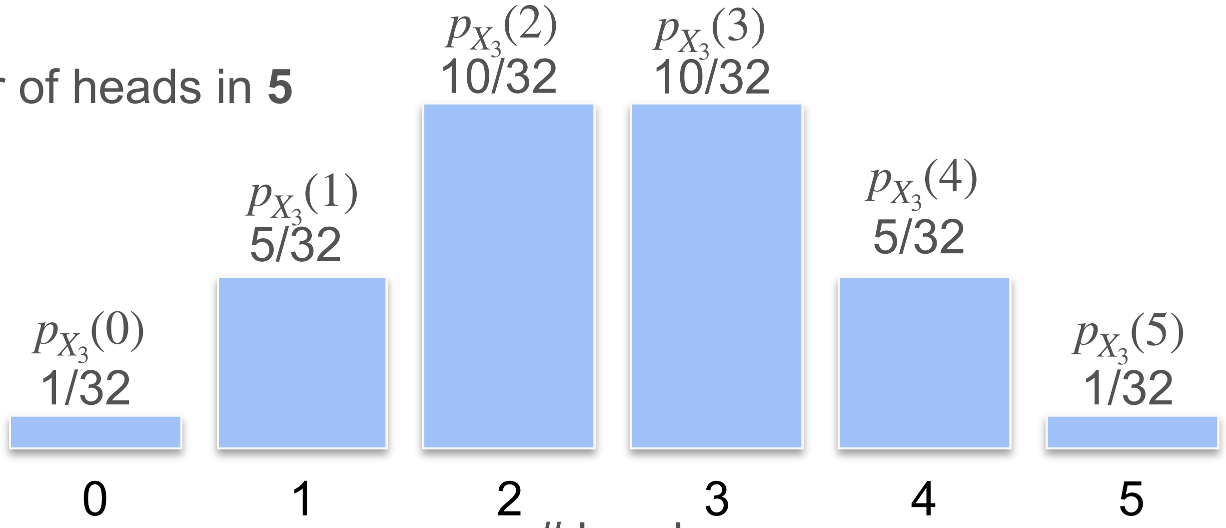


Probability Mass Function (PMF)

$$p_X(x) = \mathbf{P}(X_3 = x), \quad x = 0,1,2,3,4,5$$

Discrete Distributions: Flip Five Coins

X_3: number of heads in **5** coin tosses

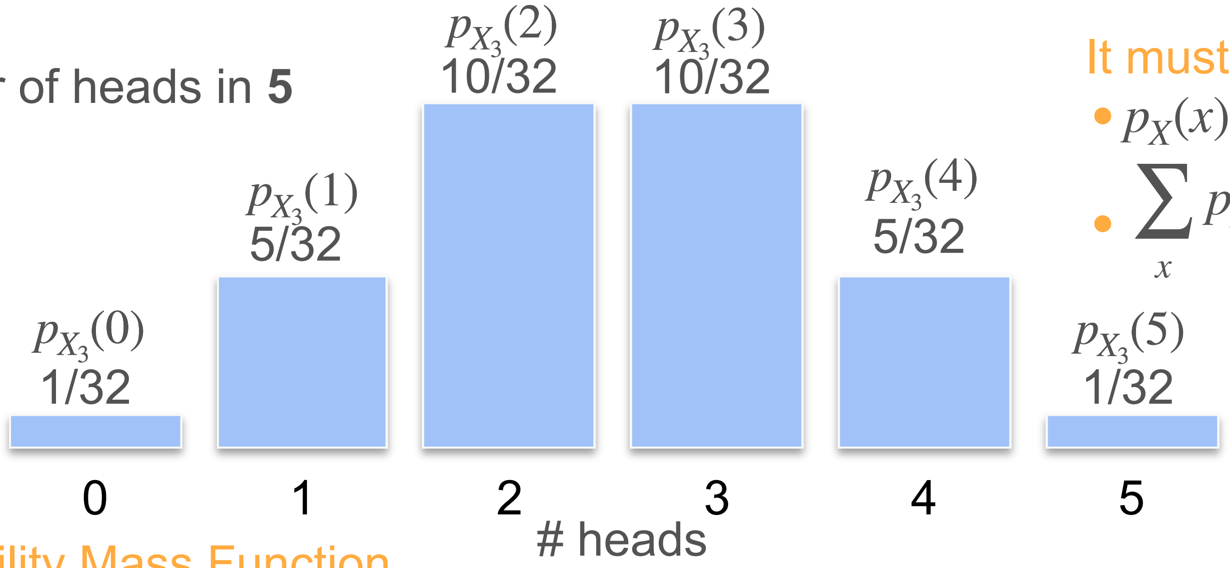


Probability Mass Function (PMF)

$$p_X(x) = \mathbf{P}(X_3 = x), \quad x = 0,1,2,3,4,5$$

Discrete Distributions: Flip Five Coins

X_3: number of heads in **5** coin tosses



It must satisfy:

- $p_X(x) \geq 0$
- $\sum_x p_X(x) = 1$

Probability Mass Function (PMF)

$$p_X(x) = \mathbf{P}(X_4 = x), \quad x = 0,1,2,3,4,5$$

$+ \quad + \quad + \quad + \quad +$

$$= \frac{1}{32} + \frac{5}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} + \frac{1}{32} = 1$$

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

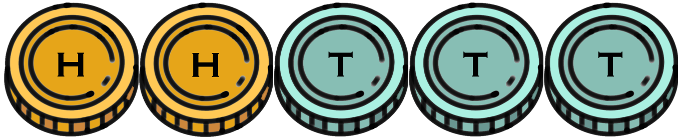
DeepLearning.AI

Probability Distributions

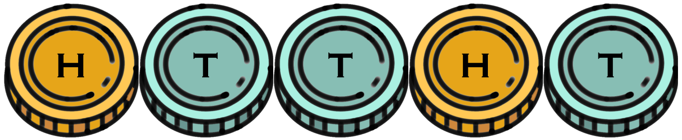
Binomial Distribution

Binomial Distribution: Example

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



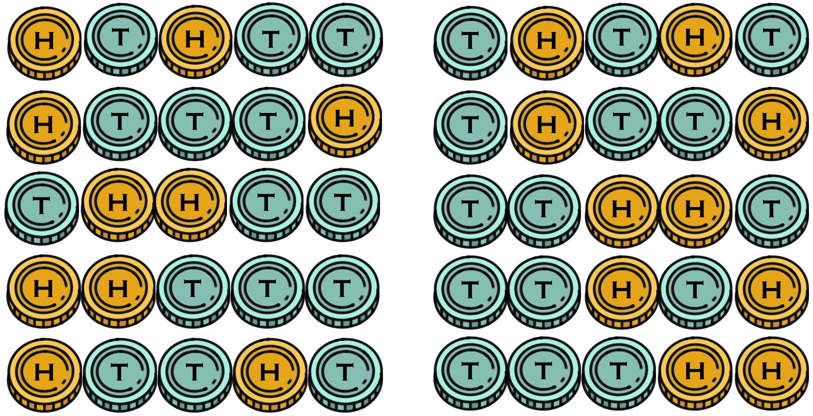
$$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \quad = \frac{1}{32}$$



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

Binomial Distribution: Example

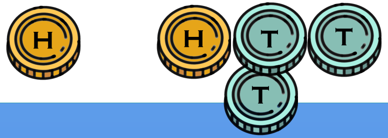
What is the probability that if I flip 5 coins, 2 of them land in heads?



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

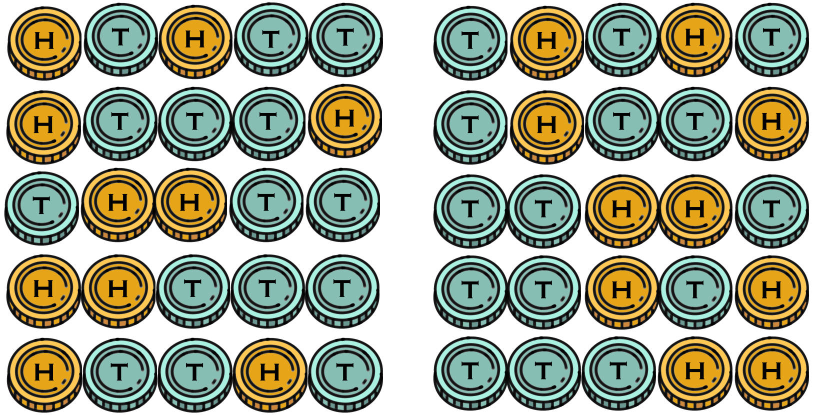
Number of ways you can order 5 coins

Number of H



Binomial Distribution: Example

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



Number of ways you can order 5 coins

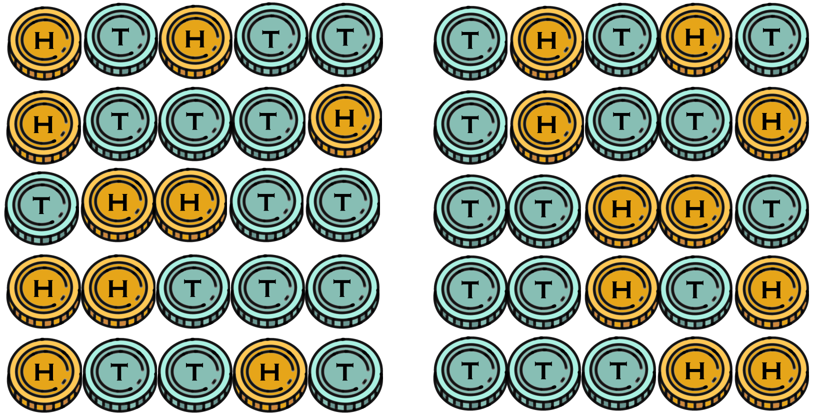
$$10 = \frac{5!}{2!(5-2)!}$$

Number of H

Number of T

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

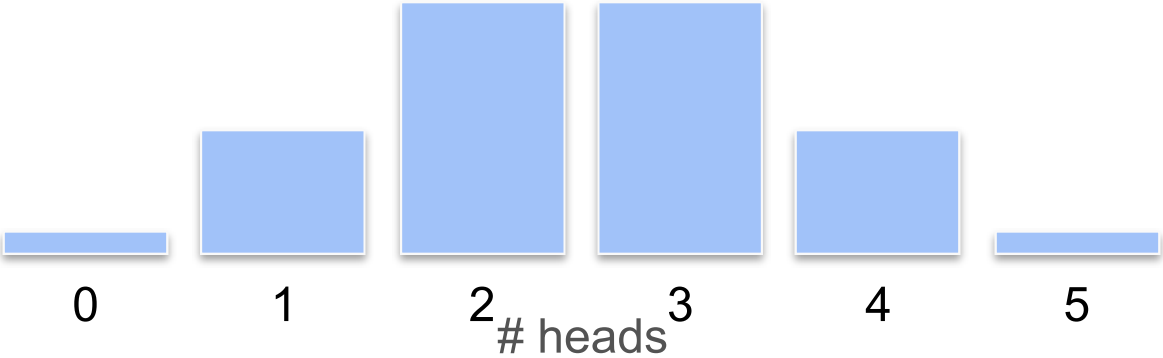
In general:

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

Property:

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

The PMF with a fair coin is symmetrical



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

Probability of seeing $5 - x$ tails

All the possible orders

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

Probability of seeing x heads

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**

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

Number of flips

$\mathbf{P}(H)$

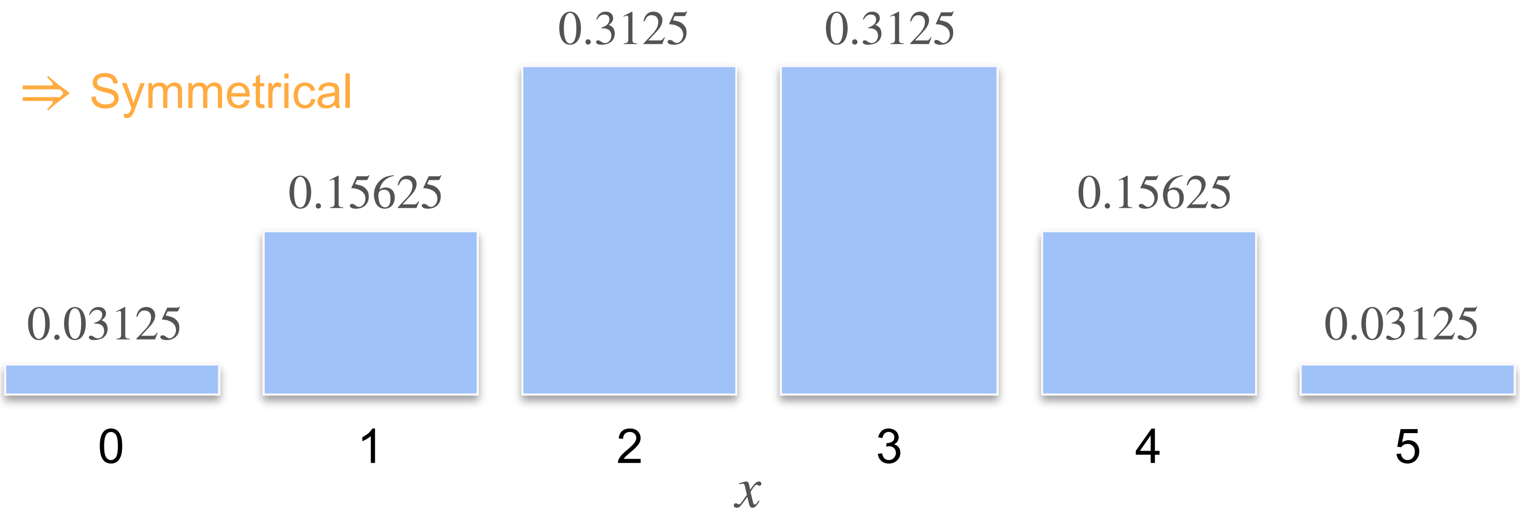
DeepLearning.AI

Binomial Distribution

$$\boxed{\begin{matrix} n = 5 \\ p = 0.5 \end{matrix}}$$

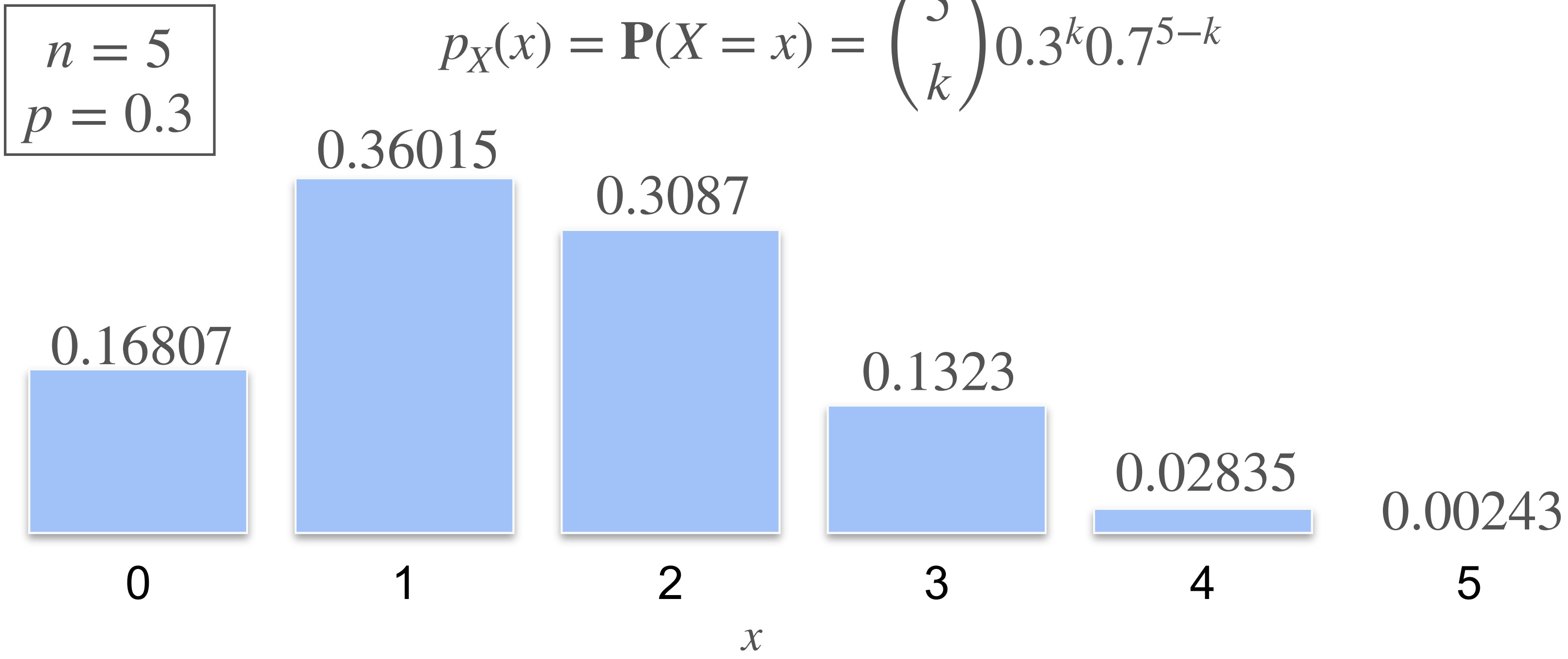
$$p_X(x) = \mathbf{P}(X = x) = \binom{5}{k} 0.5^k 0.5^{5-k}$$

$p = 0.5 \Rightarrow$ Symmetrical



Binomial Distribution

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



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

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

$$X \sim Binomial(5,p)$$

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

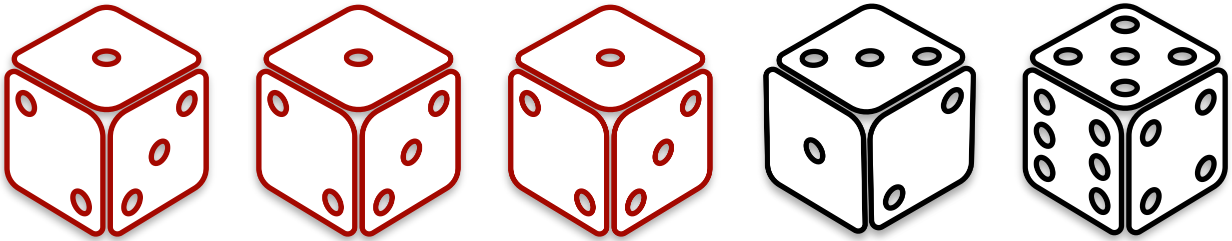
$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,2,\ldots,n$$

$$X \sim Binomial(n, p)$$

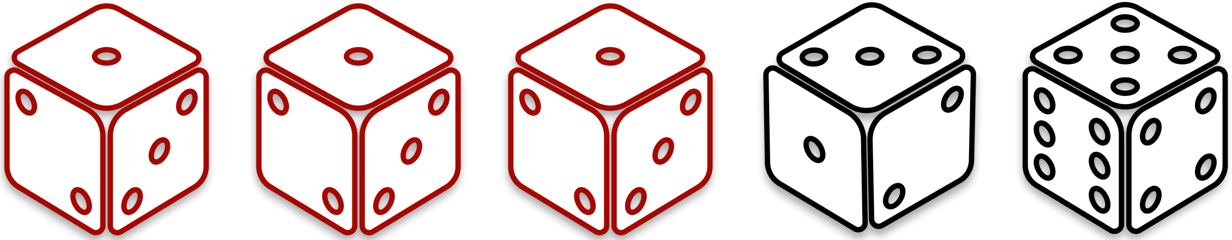
n and p are called the **parameters** of the binomial distribution

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!

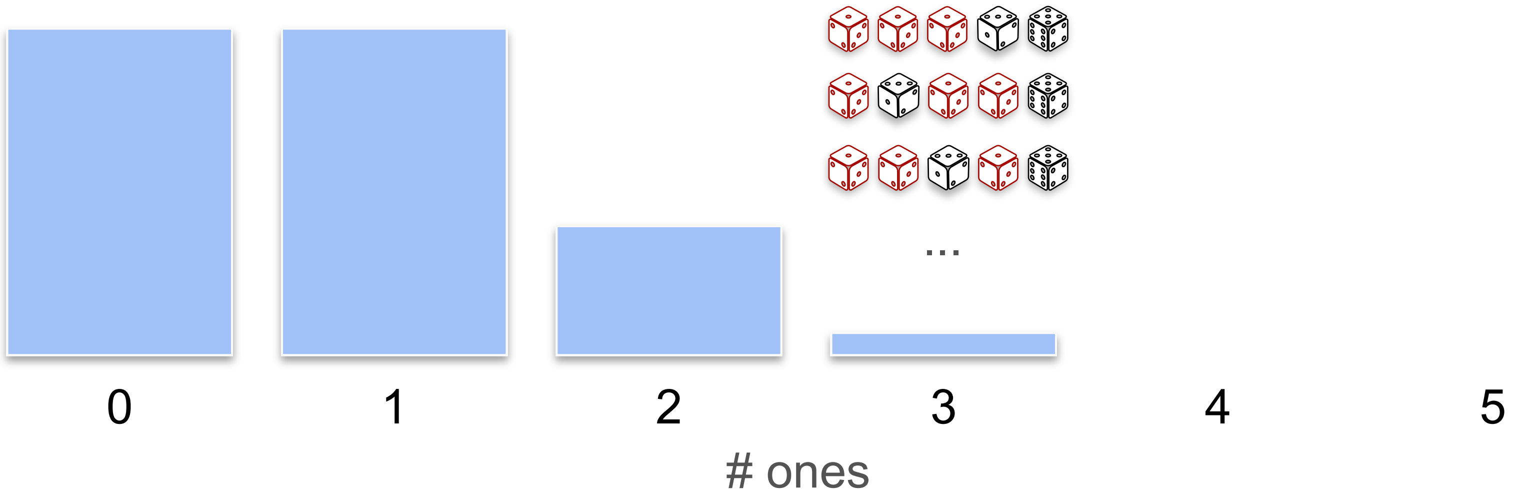


one	one	one	not one	not one
$p = \frac{1}{6}$	$p = \frac{1}{6}$	$p = \frac{1}{6}$	$p = \frac{5}{6}$	$p = \frac{5}{6}$

heads	heads	heads	not heads	not heads
$p = \frac{1}{6}$	$p = \frac{1}{6}$	$p = \frac{1}{6}$	$p = \frac{5}{6}$	$p = \frac{5}{6}$

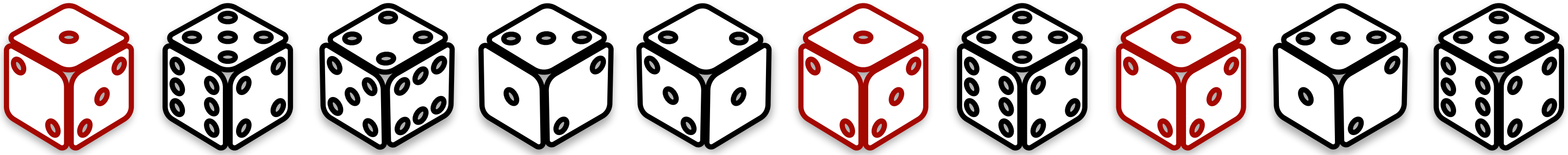
Binomial Distribution: Dice Is a Biased Coin!

n = 5
p = 0.1666



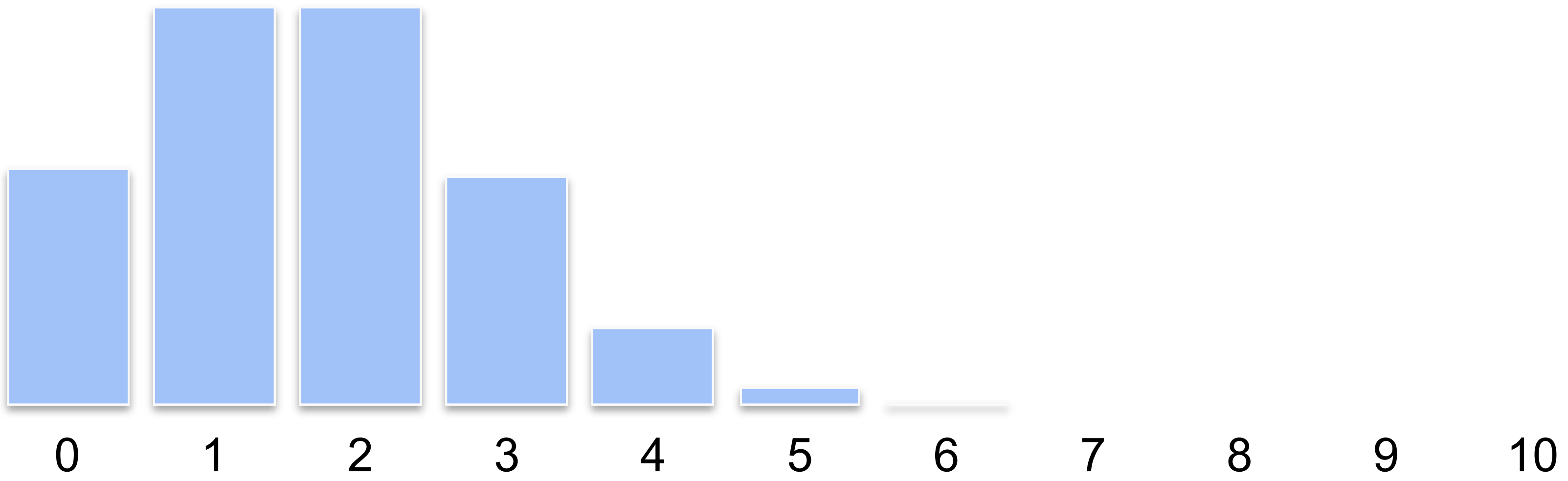
Binomial Distribution: Quiz

- 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

$$n = 10$$
$$p = 0.1666$$



DeepLearning.AI

Probability Distributions

(Optional)
Bionomial Coefficient

Binomial Coefficient

Pick 1st number Pick 2nd number Pick 3rd number Pick k-th number

n options $\quad n-1$ options $\quad n-2$ options $\quad \dots \quad n-(k-1)$options

Ordered sets of length k

$1,2,3,4,\dots,k$

$2,1,3,4,\dots,k$

$5,1,3,k,\dots,2$

$n,1,3,4,\dots,2$

$\dots$

Some of them repeat

$$n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-(k-1))$$

Binomial Coefficient

	Pick 1st	Pick 2nd	Pick 3rd	Pick 4th
	4 options	3 options	2 options	1 option

1,2,3,4

1,2,4,3

1,3,2,4

1,3,4,2

...

4,3,2,1

$$4 \cdot 3 \cdot 2 \cdot 1 = 4!$$

For five numbers:

$$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$$

General solution: $k!$

Binomial Coefficient

Pick 1st number | Pick 2nd number | Pick 3rd number | Pick k-th number

n options → $n-1$ options → $n-2$ options → $\dots$ → $n-(k-1)$options

Ordered sets of length k → Unordered sets of length k

$1,2,3,4,\dots,k$

$2,1,3,4,\dots,k$

$5,1,3,k,\dots,2$

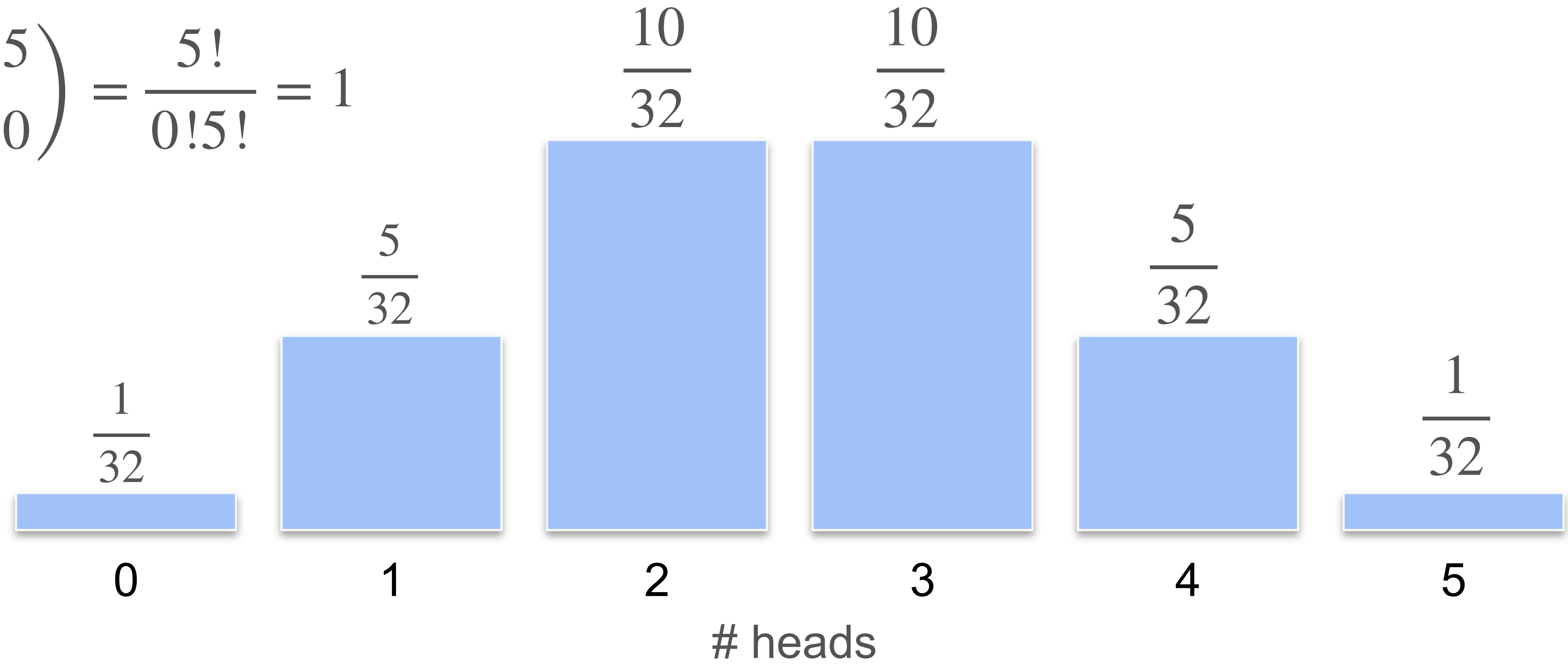
$n,1,3,4,\dots,2$

$\dots$

$$\frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-(k-1))}{k!}=\frac{n!}{(n-k)!\,k!}=\binom{n}{k}$$

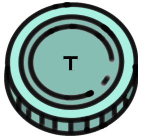
Binomial Distribution: Fair Coins

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



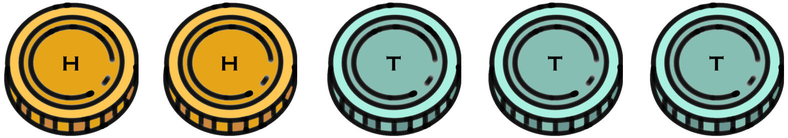
Binomial Distribution: Fair Coins

50 %



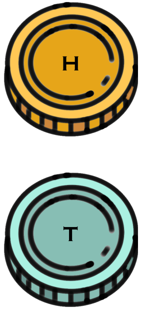
50 %

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



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

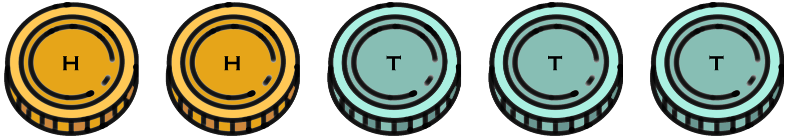
Binomial Distribution: Biased Coins



30 %

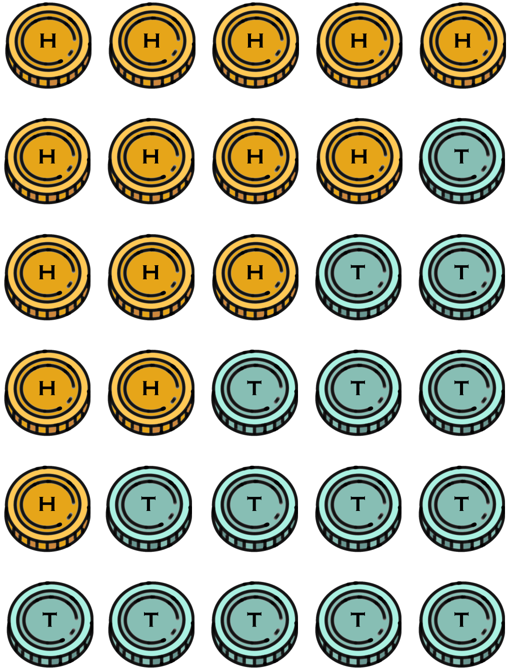
70 %

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



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

Binomial Distribution: Biased Coins



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

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

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

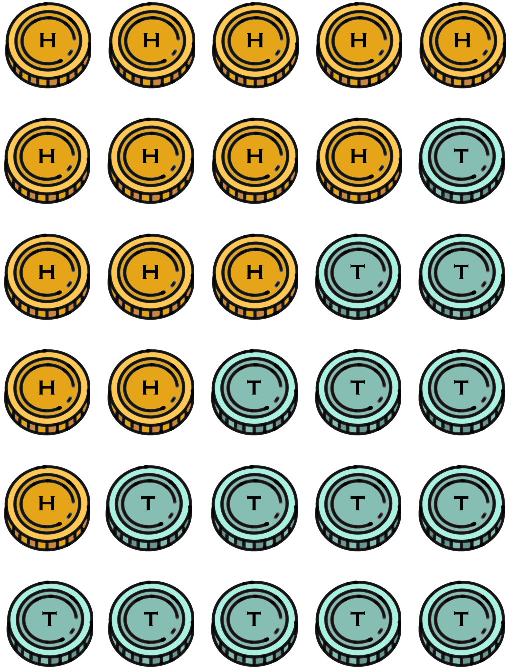
$$0.3 \cdot 0.3 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.3^2 \cdot 0.7^3$$

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

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

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

Binomial Distribution: Biased Coins



$0.3^5 \cdot 0.7^0$

$0.3^4 \cdot 0.7^1$

$0.3^3 \cdot 0.7^2$

$0.3^2 \cdot 0.7^3$

$0.3^1 \cdot 0.7^4$

$0.3^0 \cdot 0.7^5$

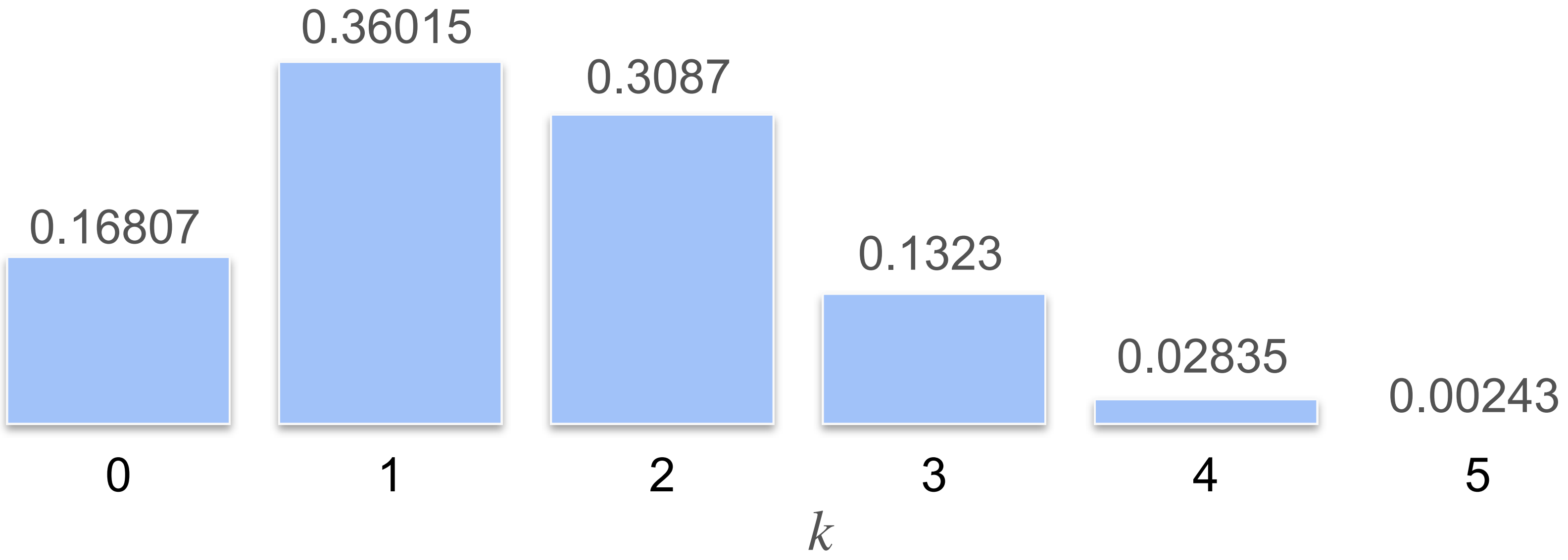
$$= 0.3^k \cdot 0.7^{n-k} \Rightarrow \binom{n}{k} 0.3^k \cdot 0.7^{n-k}$$

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

DeepLearning.AI

Probability Distributions

Bernoulli Distribution

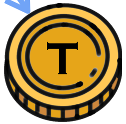
Bernoulli Distribution

$X =$ Number of heads

Flip a coin

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

Success

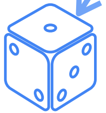


$\mathbf{P}(X = 0) = 0.5$

Failure

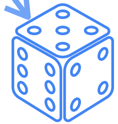
$X =$ Number of 1's

Throw a dice



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

Success

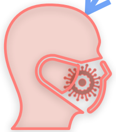


$\mathbf{P}(X = 0) = \frac{5}{6}$

Failure

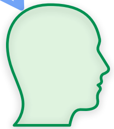
$X =$ Number of sick patients

Flip a coin



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

Success



$\mathbf{P}(X = 0) = 1 - p$

Failure

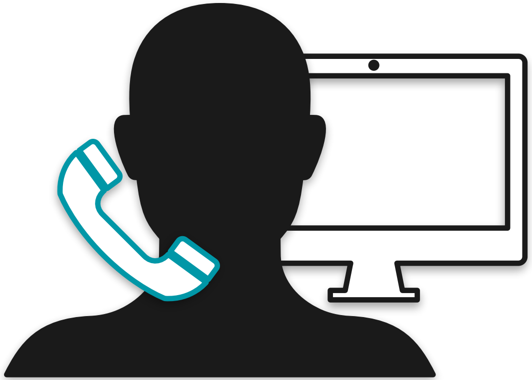
$X \sim Bernoulli(p)$ ← p is the parameter of the Bernoulli distribution

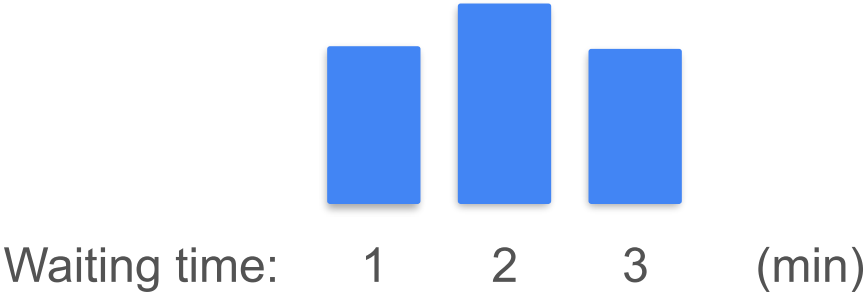
DeepLearning.AI

Probability Distributions

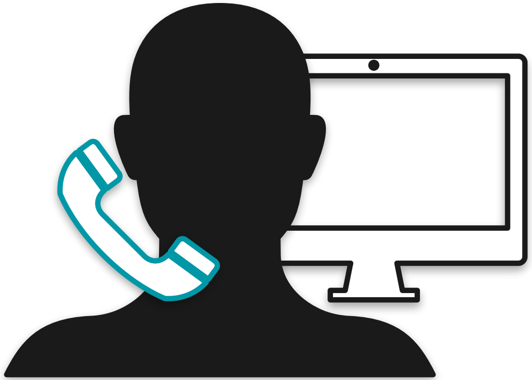
Probability Distributions (Continuous)

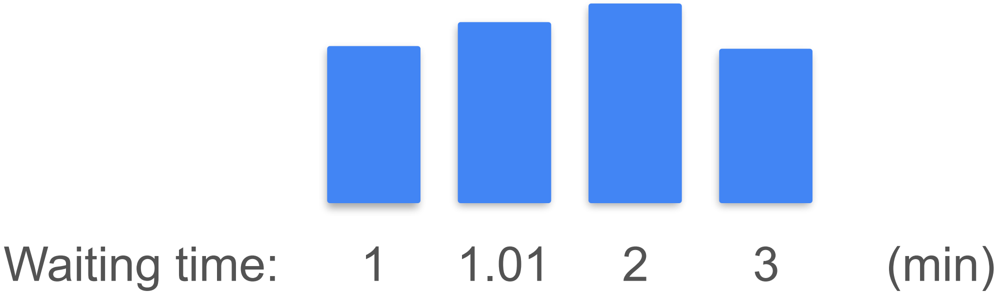
From Discrete to Continuous Distributions



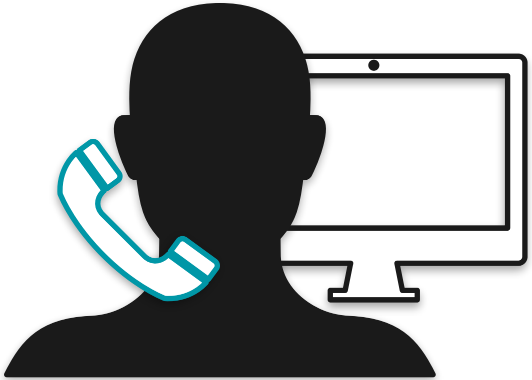


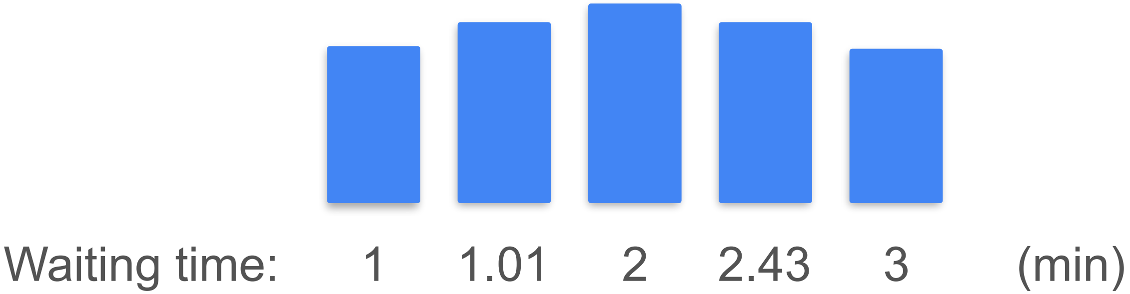
From Discrete to Continuous Distributions



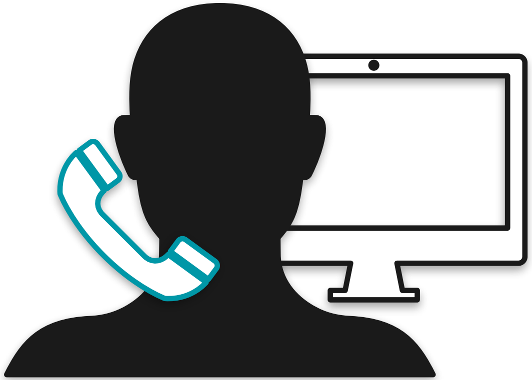


From Discrete to Continuous Distributions

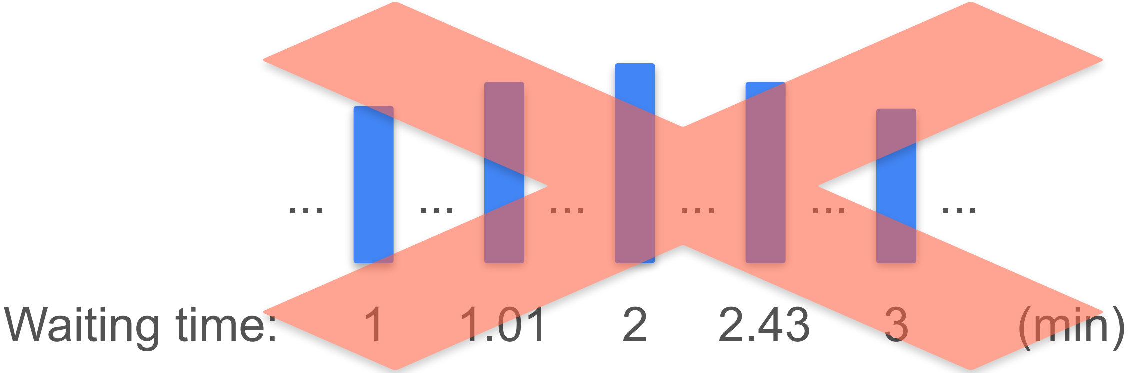




From Discrete to Continuous Distributions



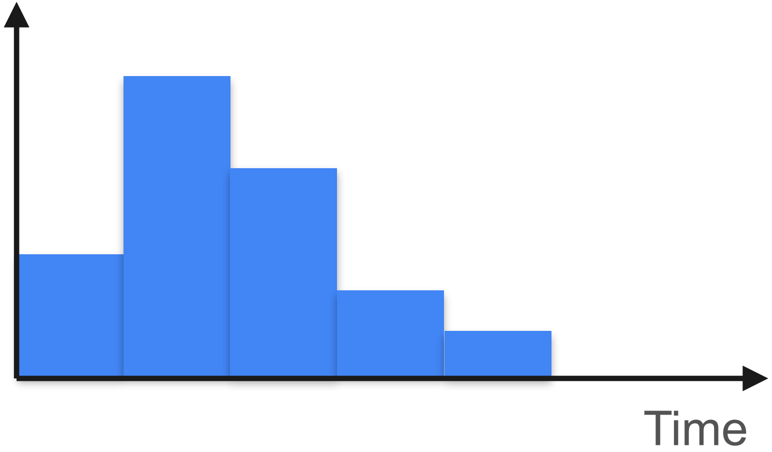
This is a continuous distribution!



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

Answer: ZERO

From Discrete to Continuous Distributions



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

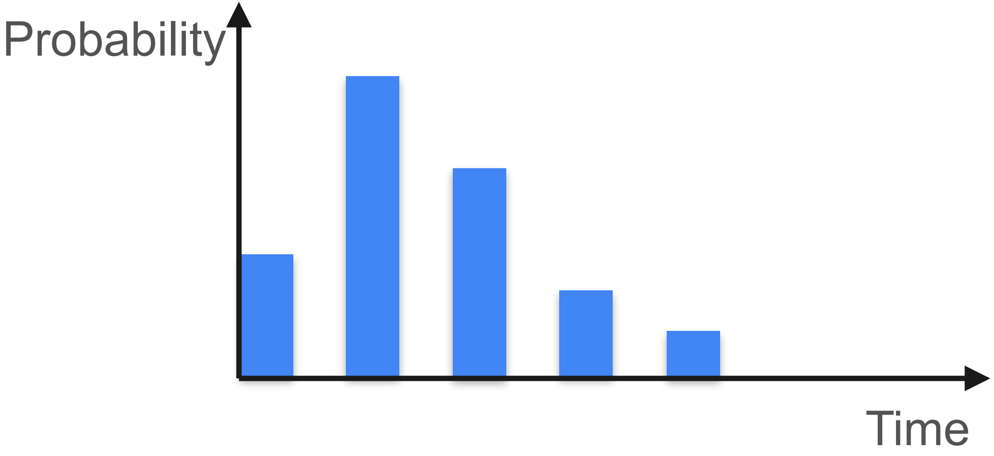
$\mathbf{P}$(between 1 and 2 mins)

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

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

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

From Discrete to Continuous Distributions



P(between 0 and 1 mins)

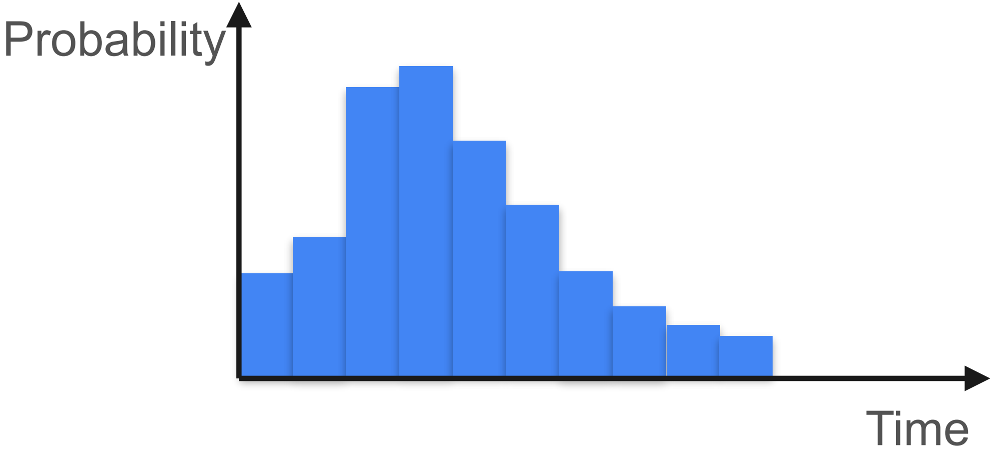
P(between 1 and 2 mins)

P(between 2 and 3 mins)

P(between 3 and 4 mins)

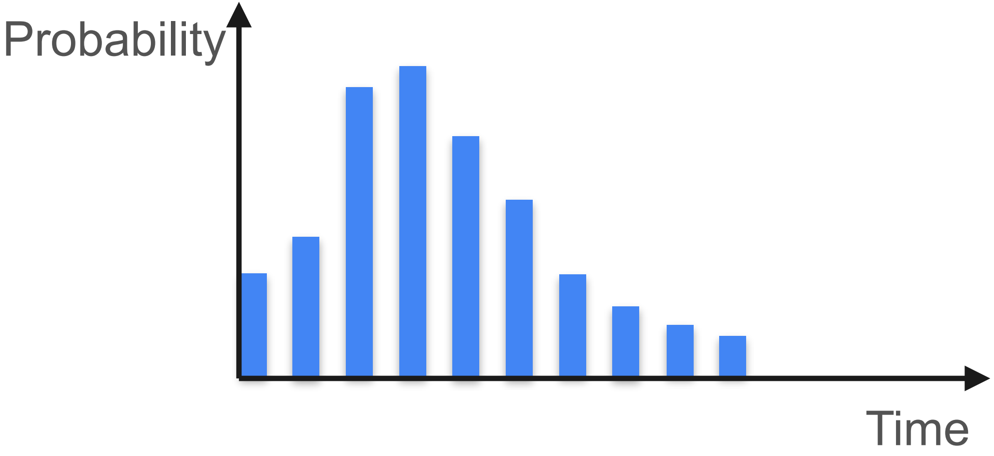
P(between 4 and 5 mins)

From Discrete to Continuous Distributions



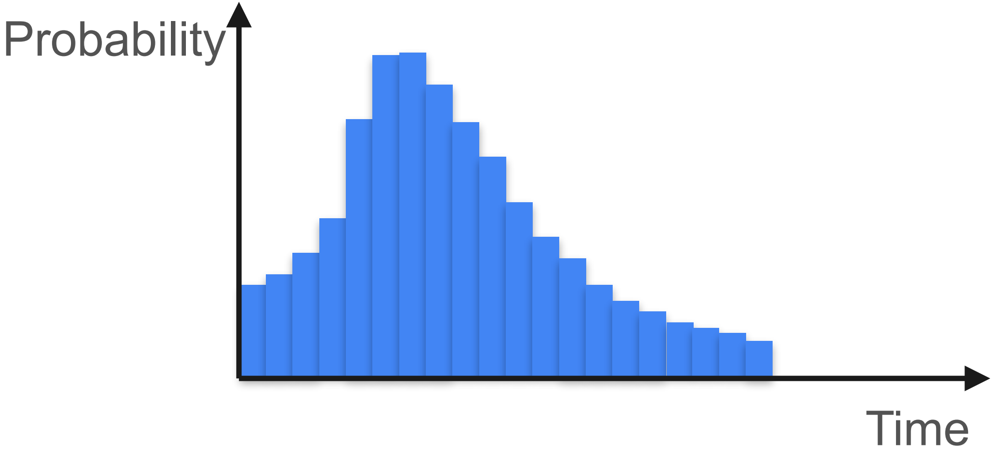
$\mathbf{P}$(between 0 and 0.5 mins)
$\mathbf{P}$(between 0.5 and 1 mins)
$\mathbf{P}$(between 1 and 1.5 mins)
⋮
$\mathbf{P}$(between 3.5 and 4 mins)
$\mathbf{P}$(between 4 and 4.5 mins)
$\mathbf{P}$(between 4.5 and 5 mins)

From Discrete to Continuous Distributions



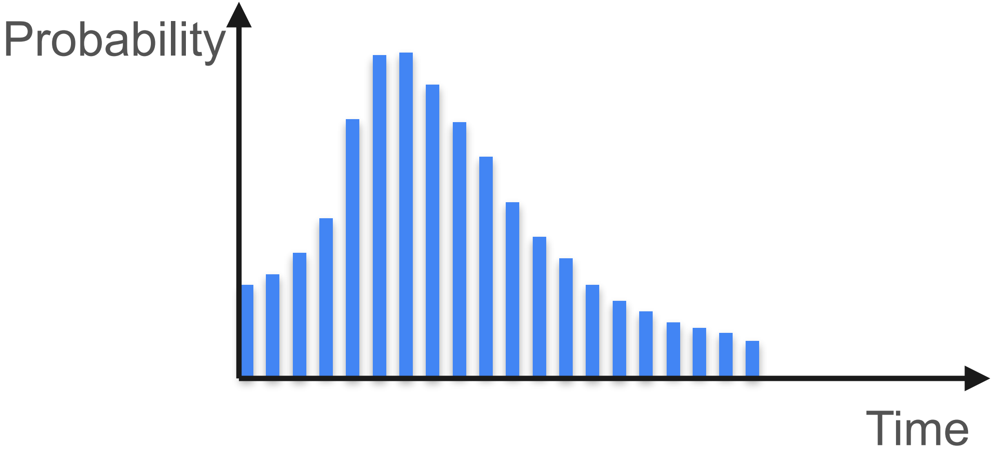
$\mathbf{P}$(between 0 and 0.5 mins)
$\mathbf{P}$(between 0.5 and 1 mins)
$\mathbf{P}$(between 1 and 1.5 mins)
⋮
$\mathbf{P}$(between 3.5 and 4 mins)
$\mathbf{P}$(between 4 and 4.5 mins)
$\mathbf{P}$(between 4.5 and 5 mins)

From Discrete to Continuous Distributions



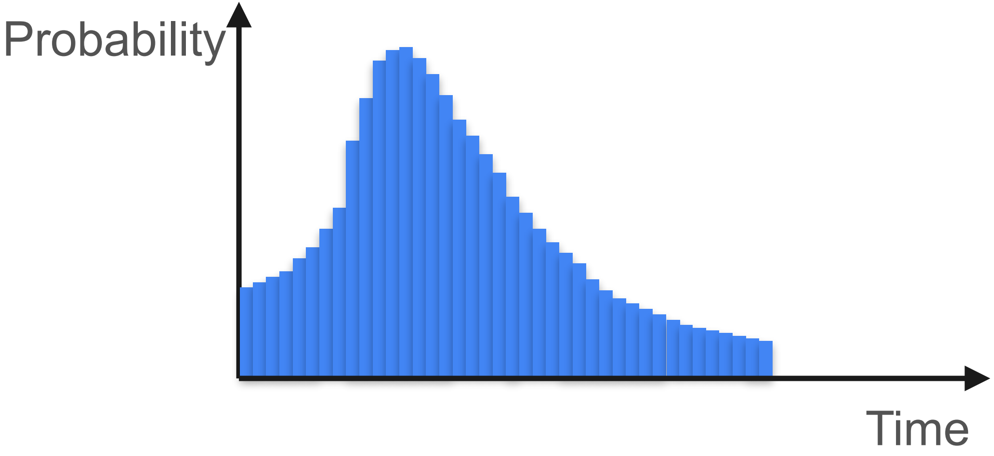
P(between 0 and 0.25 mins)
P(between 0.25 and 0.5 mins)
P(between 0.5 and 0.75 mins)
⋮
P(between 4.25 and 4.5 mins)
P(between 4.5 and 4.75 mins)
P(between 4.75 and 5 mins)

From Discrete to Continuous Distributions



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

From Discrete to Continuous Distributions

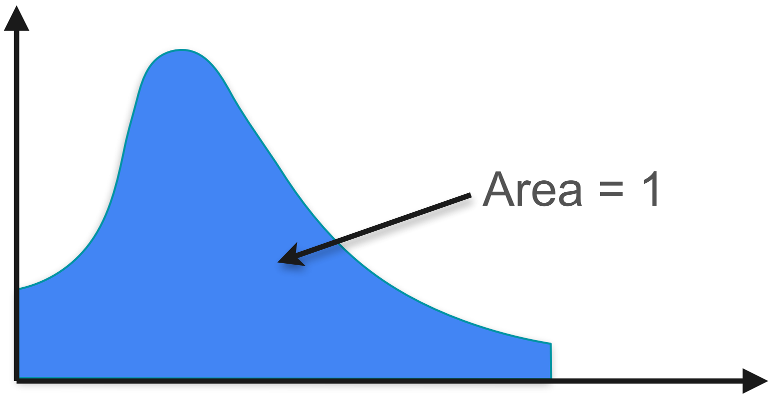


$\mathbf{P}$(between 0 and 0.125 mins)

⋮

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

From Discrete to Continuous Distributions



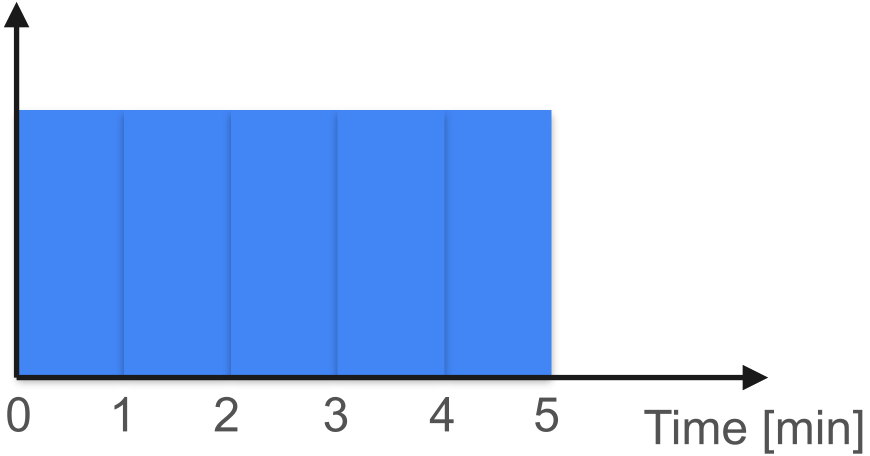
- Discrete:
  - Sum of heights equals 1
- Continuous:
  - Area under the curve equals 1

DeepLearning.AI

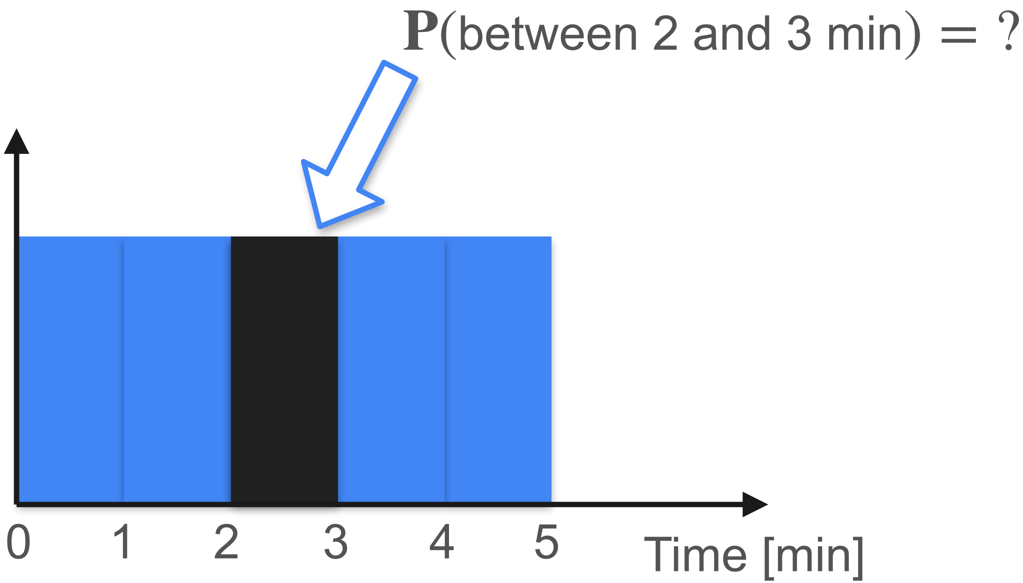
Probability Distributions

Probability density function

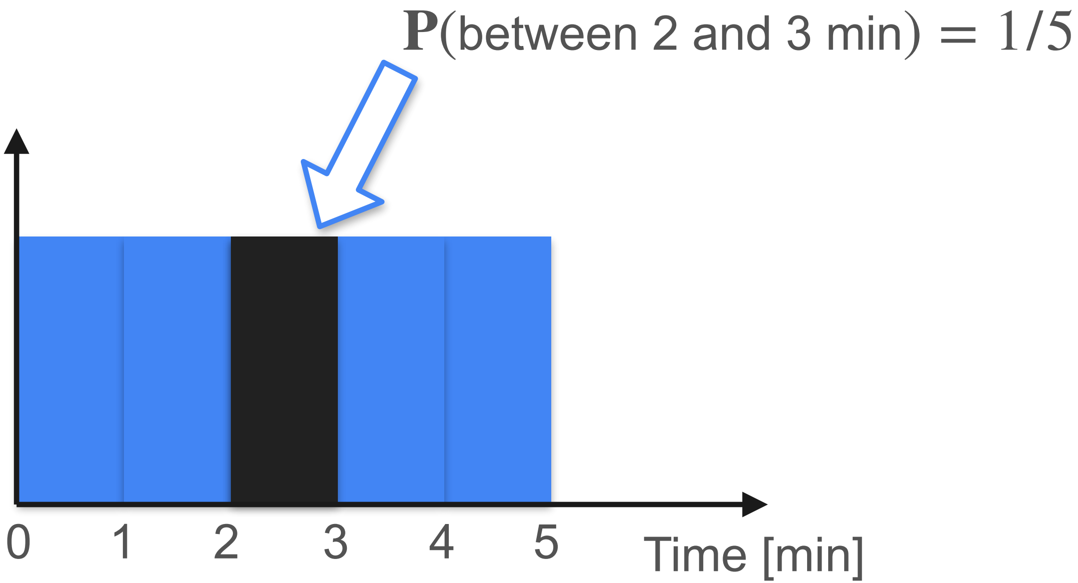
Probability Density Function



Probability Density Function

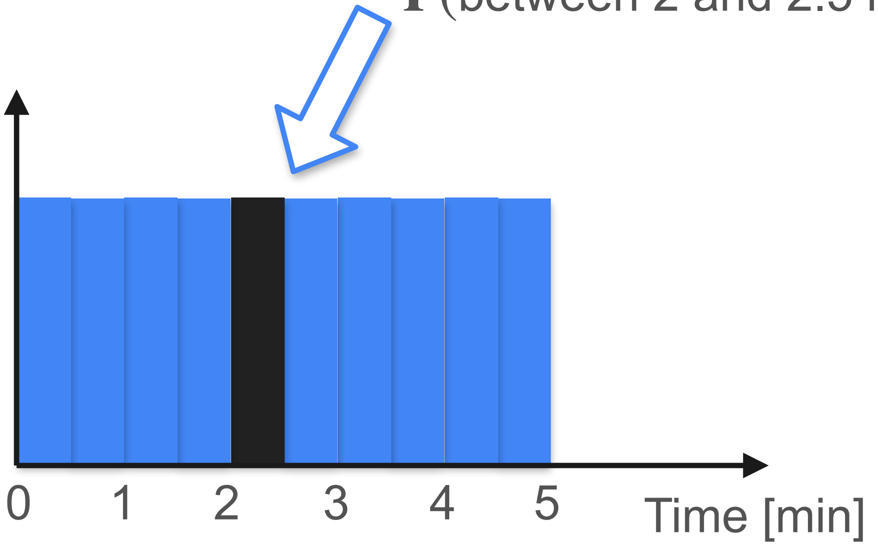


Probability Density Function



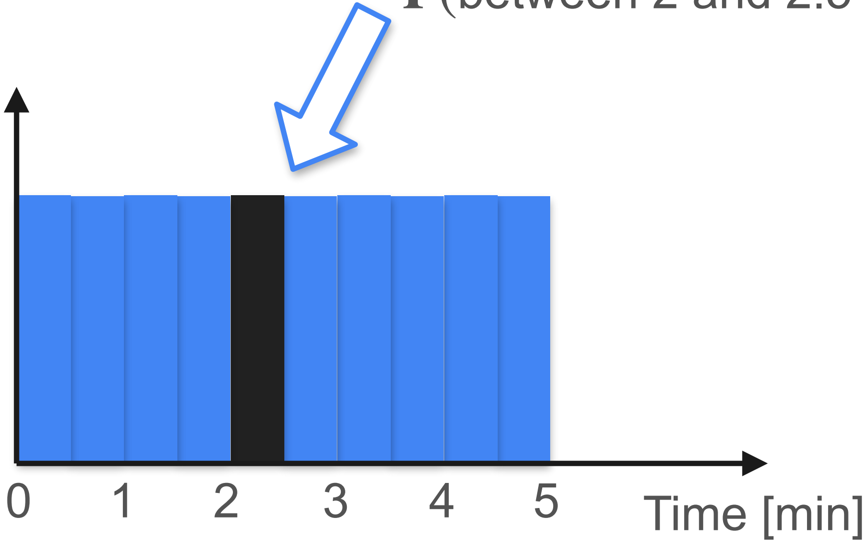
Probability Density Function

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



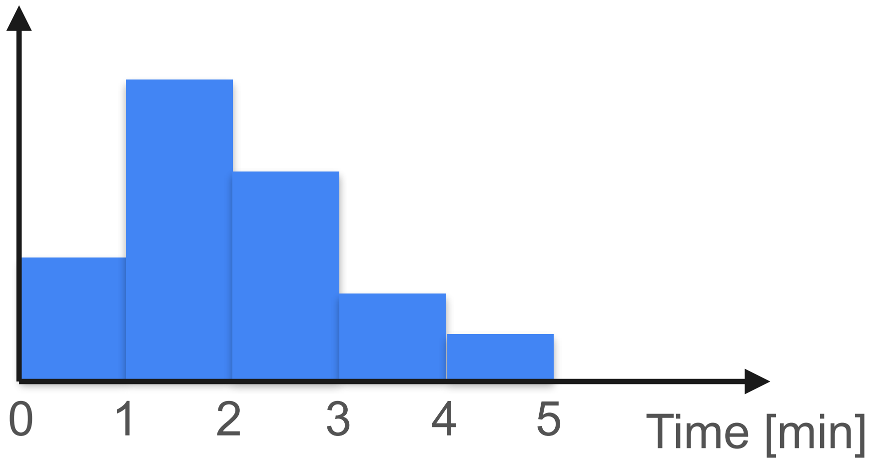
Probability Density Function

$\mathbf{P}(\text{between 2 and 2.5 min}) = 0.1$



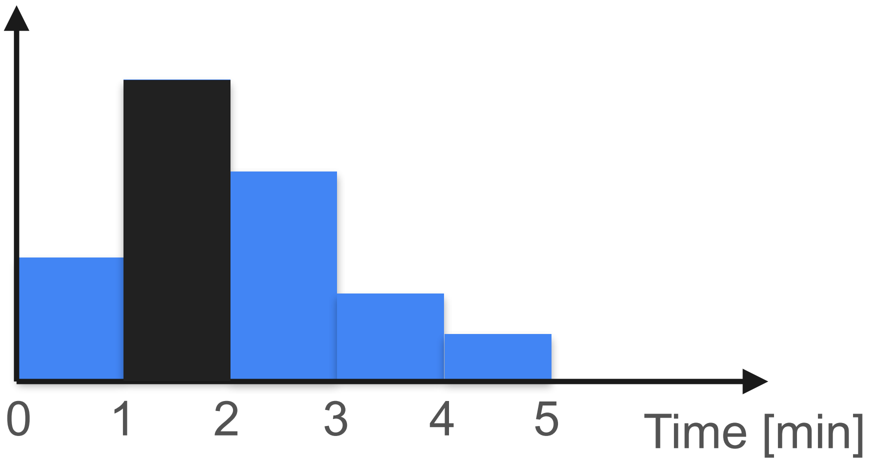
Probability Density Function

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



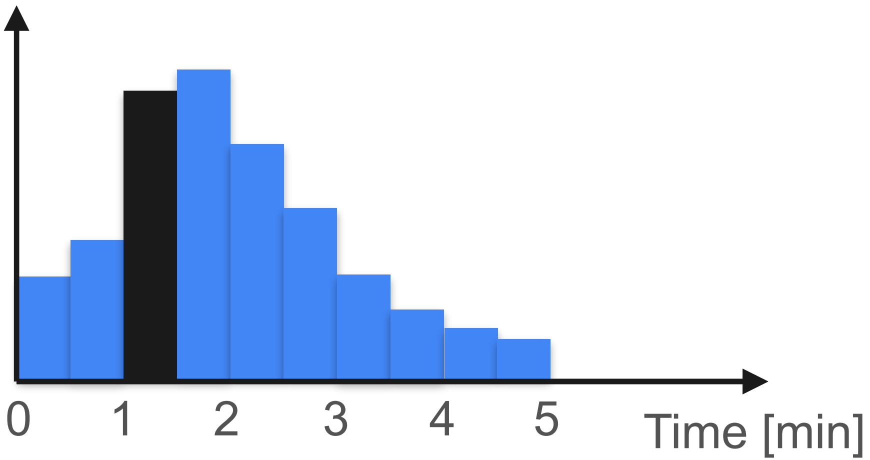
Probability Density Function

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



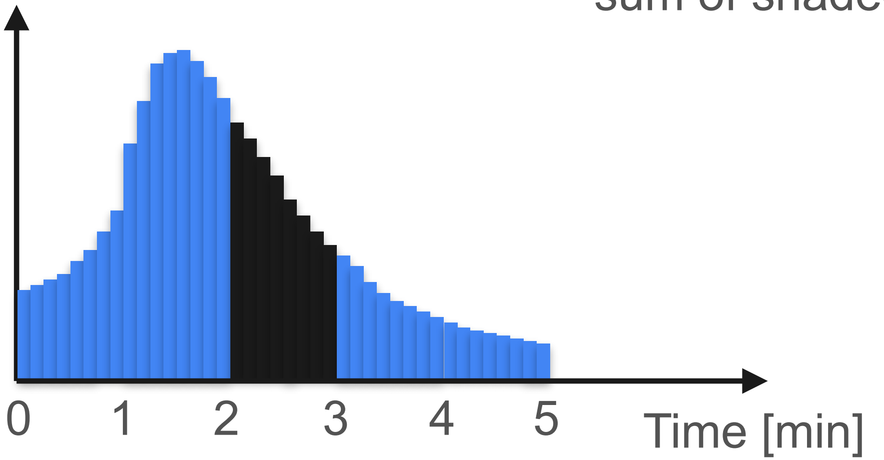
Probability Density Function

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

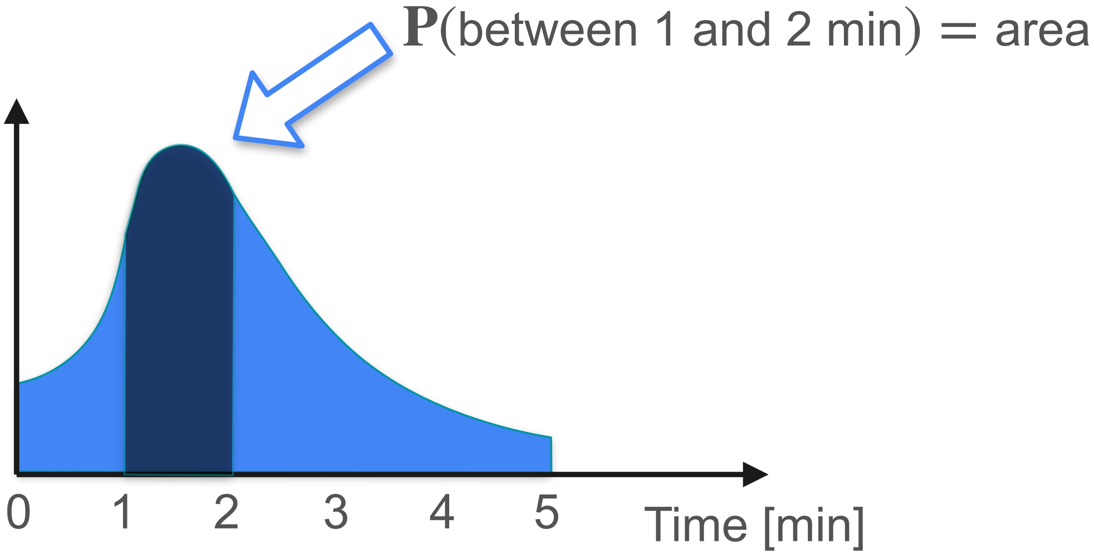


Probability Density Function

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

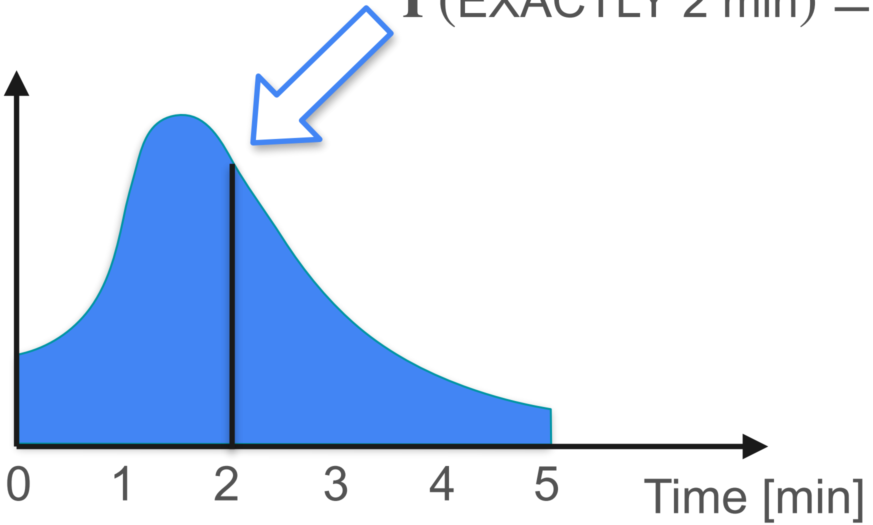


Probability Density Function

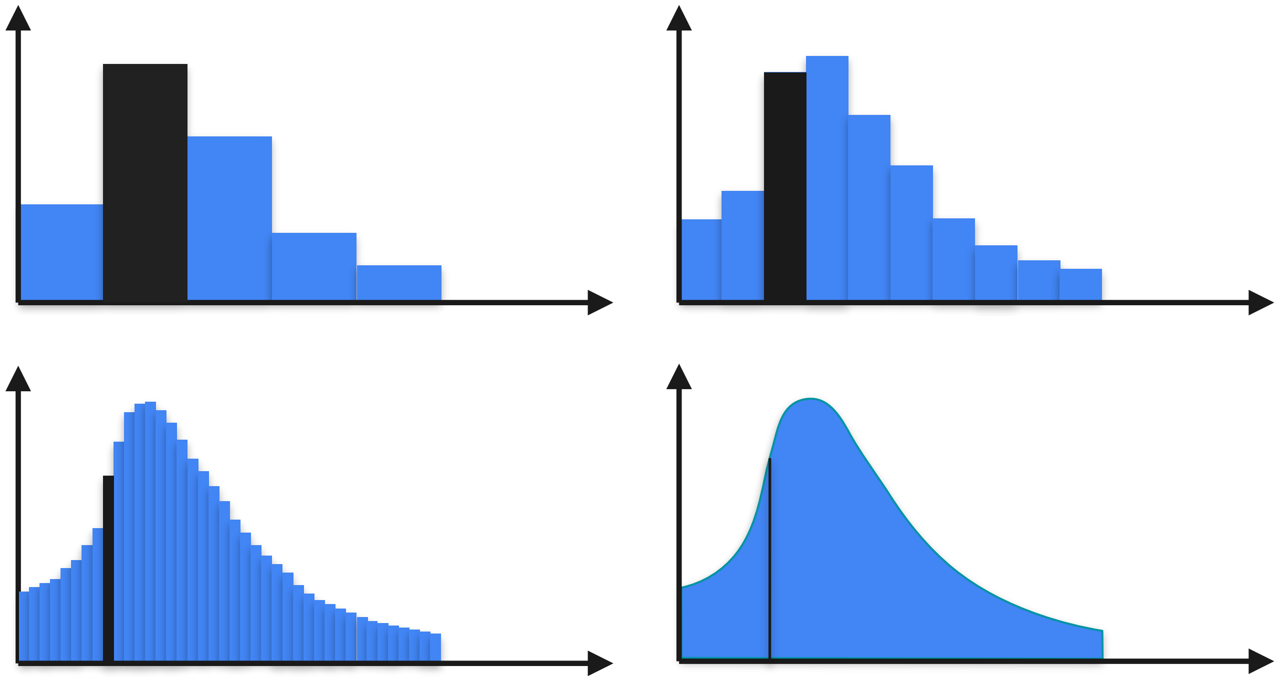


Probability Density Function

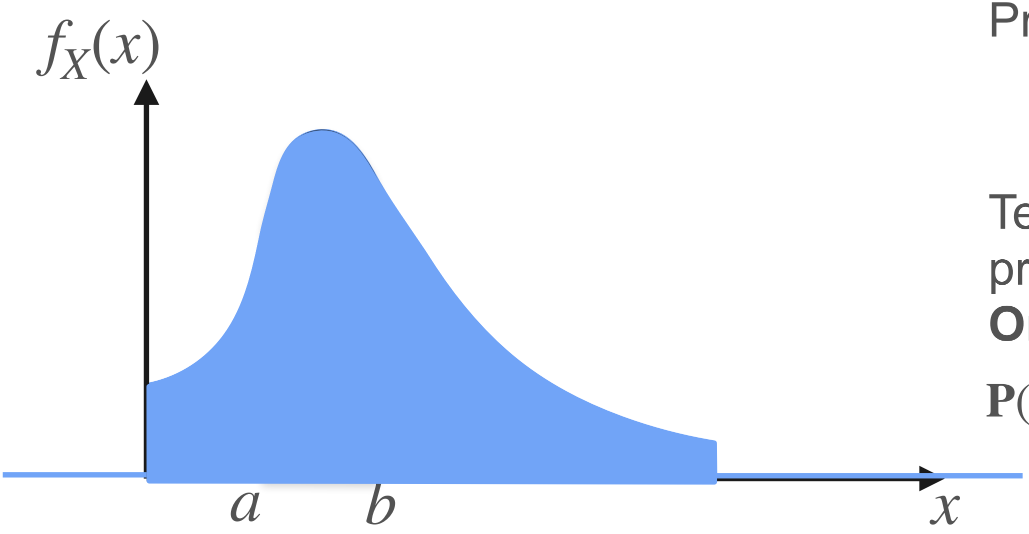
$\mathbf{P}(\text{EXACTLY 2 min}) = \text{area} \; = 0$



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!

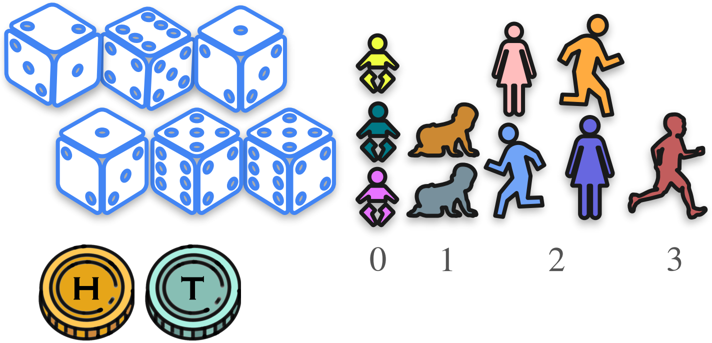
$\mathbf{P}(a < X < b) =$ area under $f_X(x)$

$f_X(x)$ 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

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

Continuous random variables

Takes values on an interval (infinite possibilities!)

PDF: $f_X(x)$

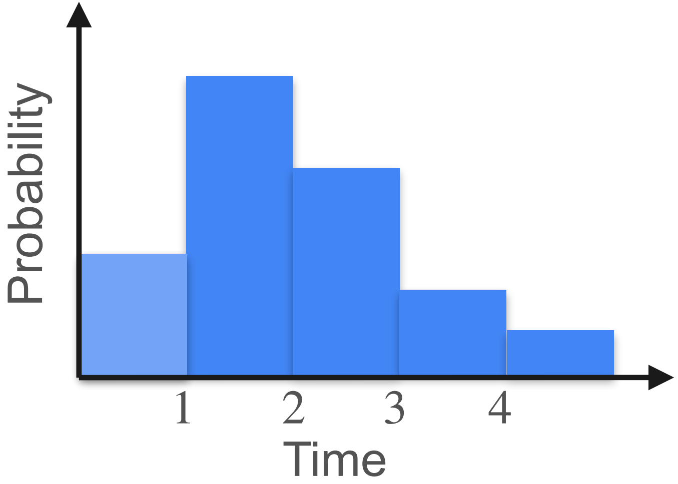
$\mathbf{P}(X = x) = 0 \quad \forall x$

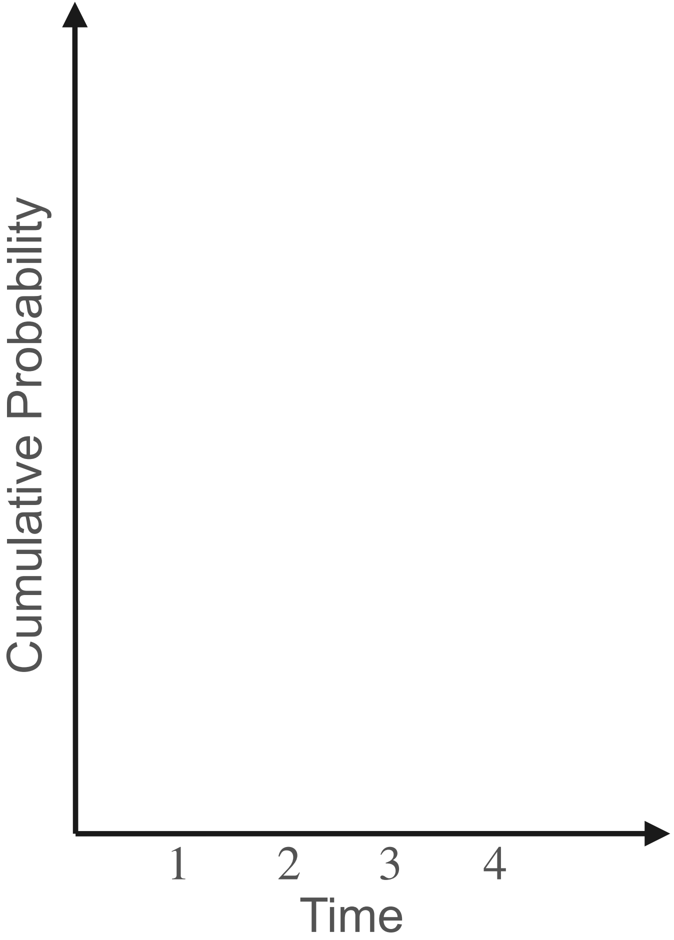
DeepLearning.AI

Probability Distributions

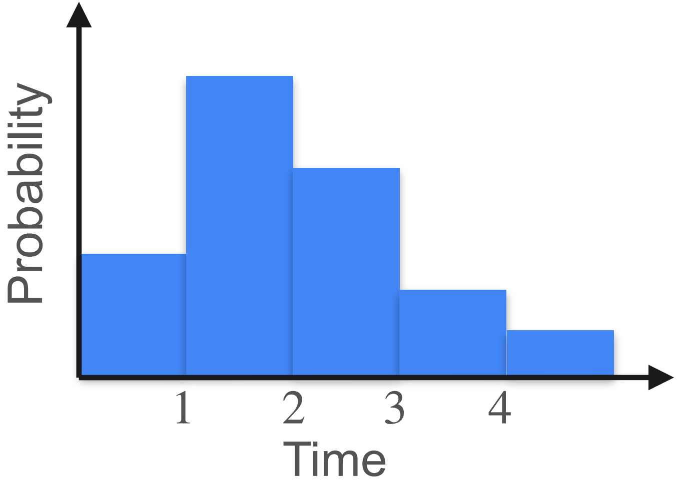
Cumulative Distribution Function

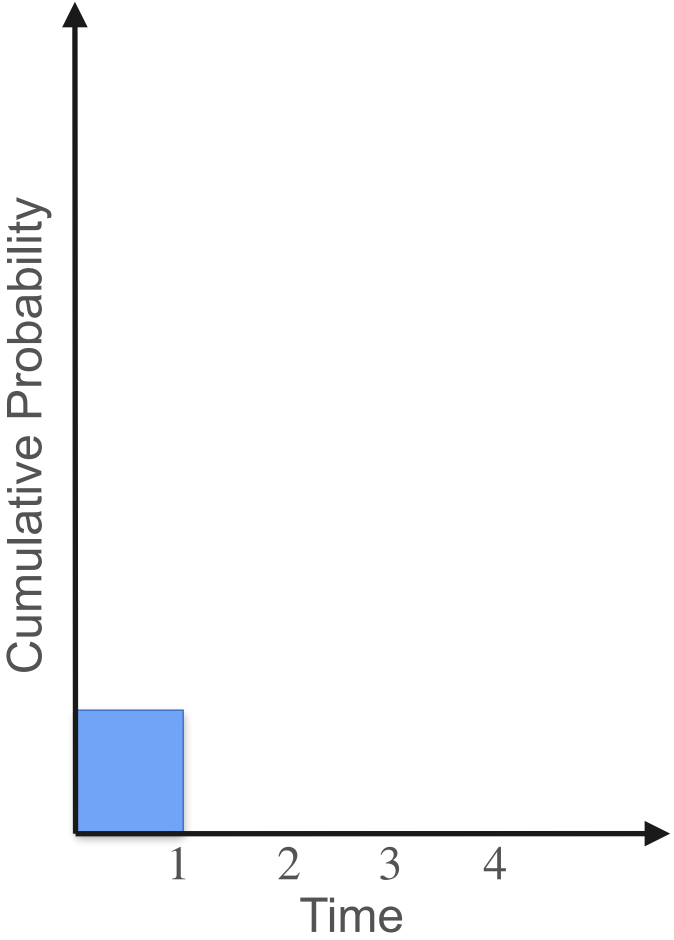
Cumulative Distribution





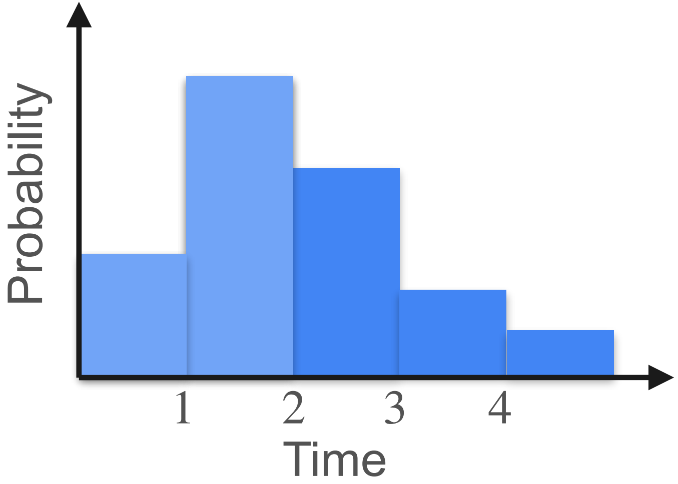
Cumulative Distribution

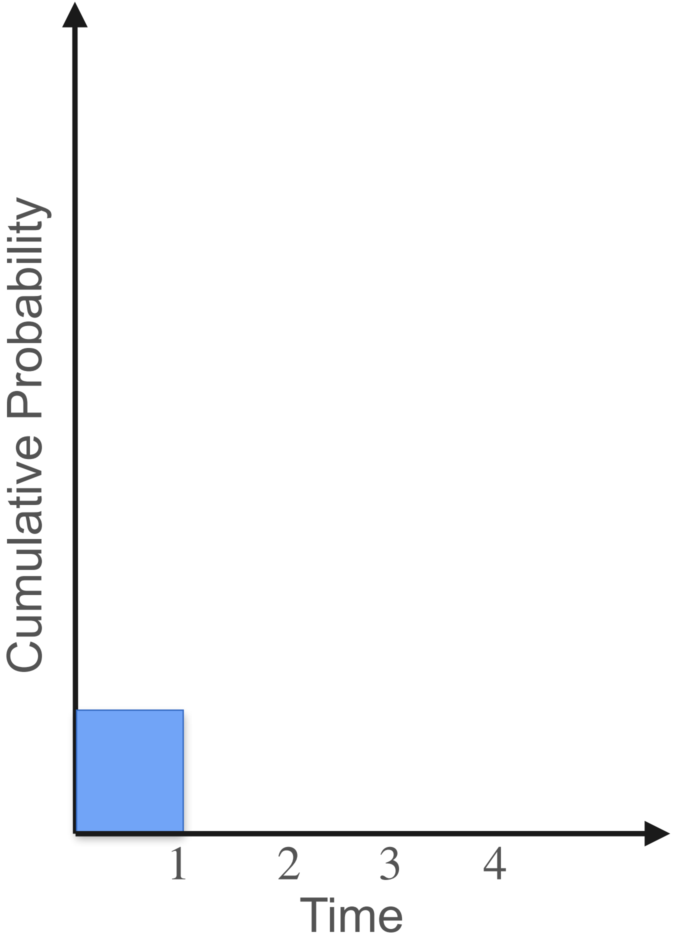




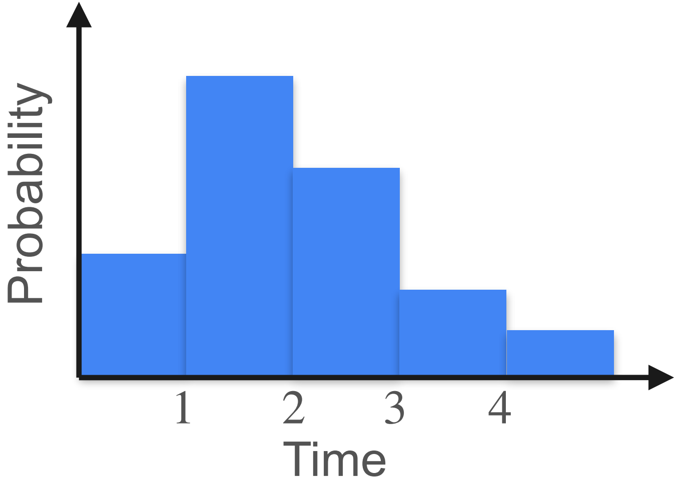
DeepLearning.AI

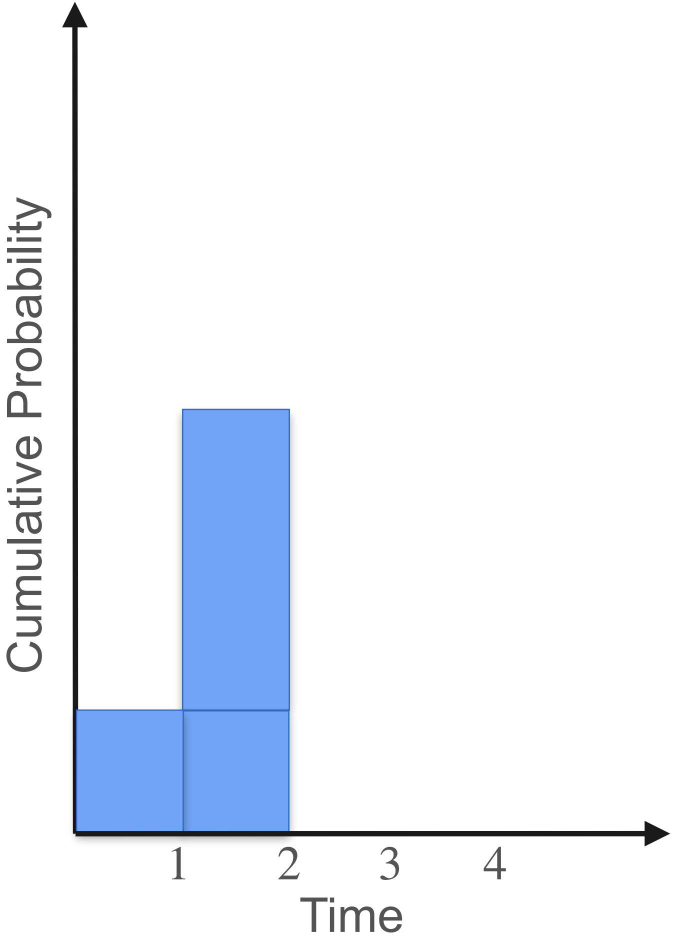
Cumulative Distribution



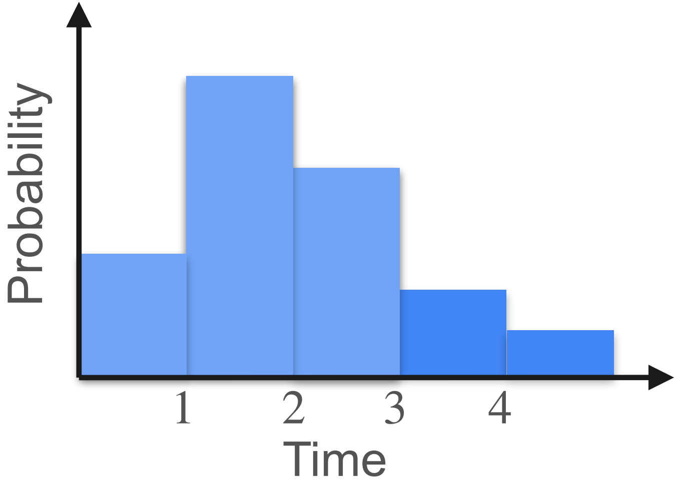


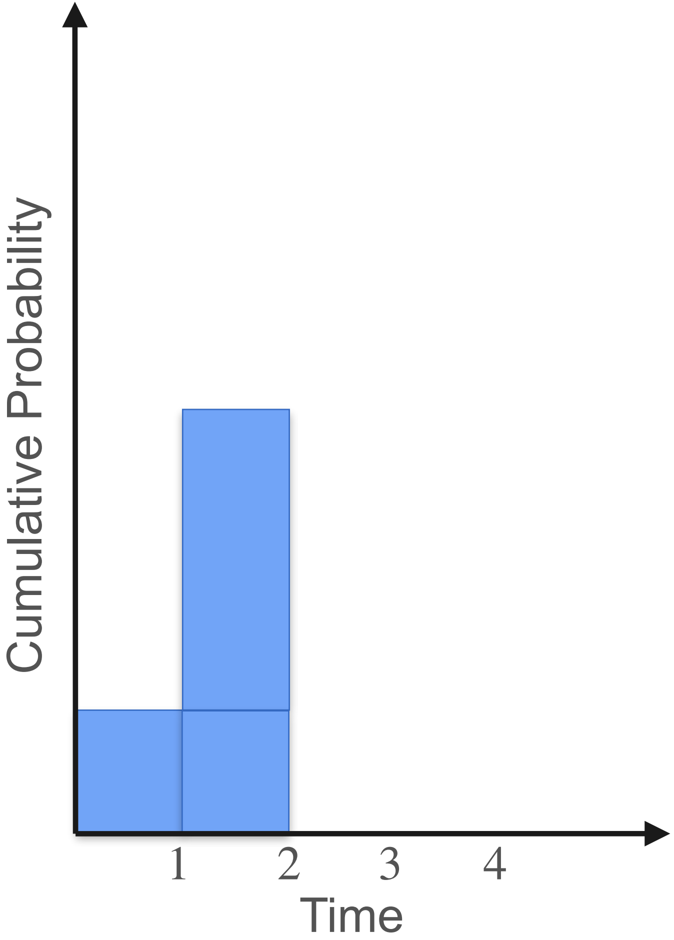
Cumulative Distribution



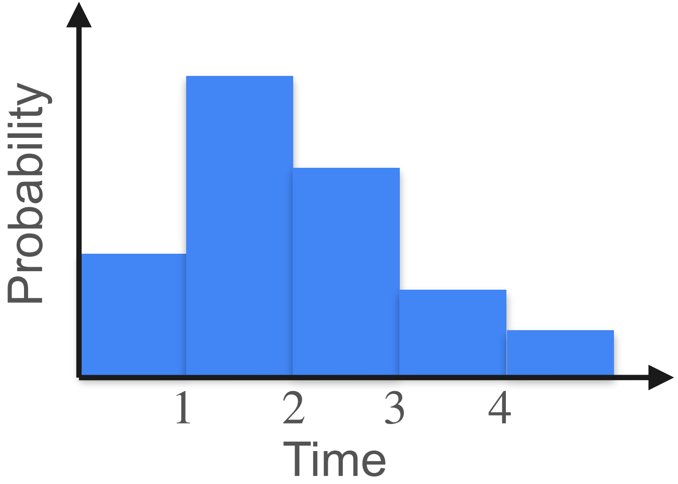


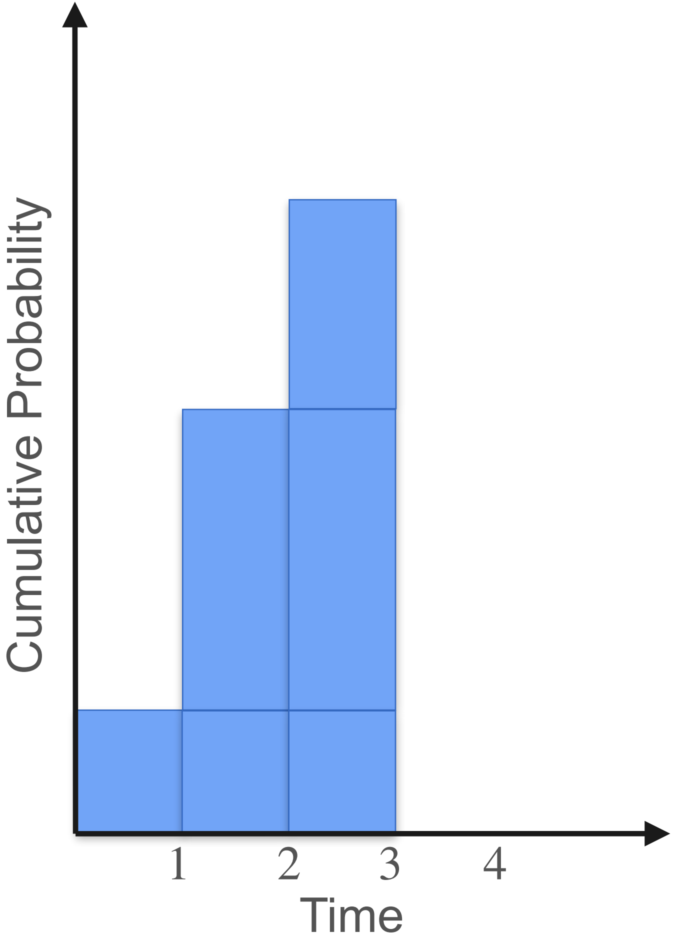
Cumulative Distribution



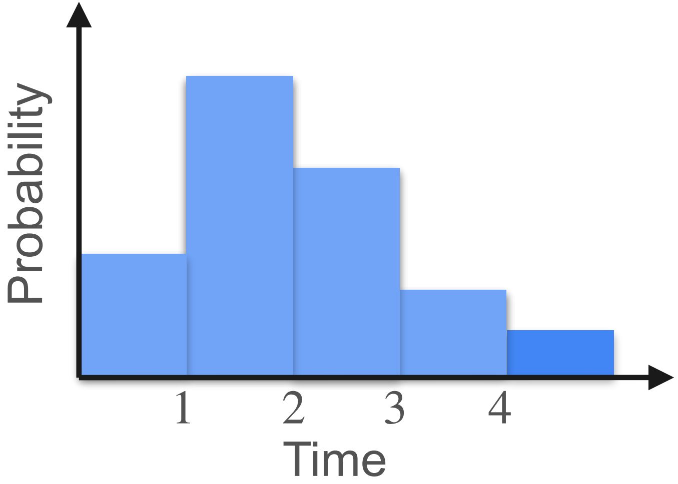


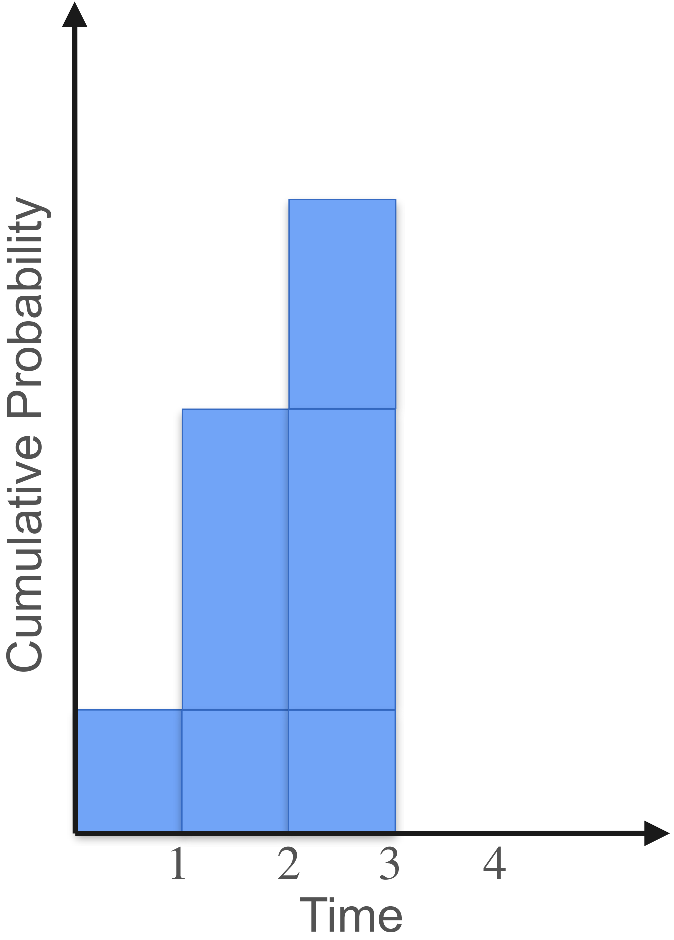
Cumulative Distribution



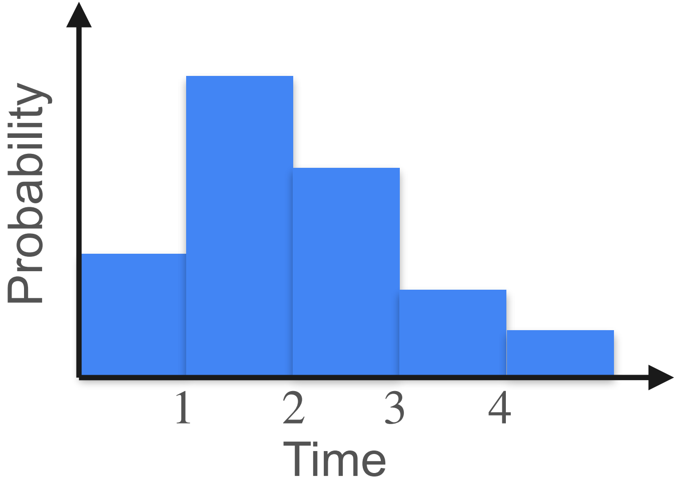


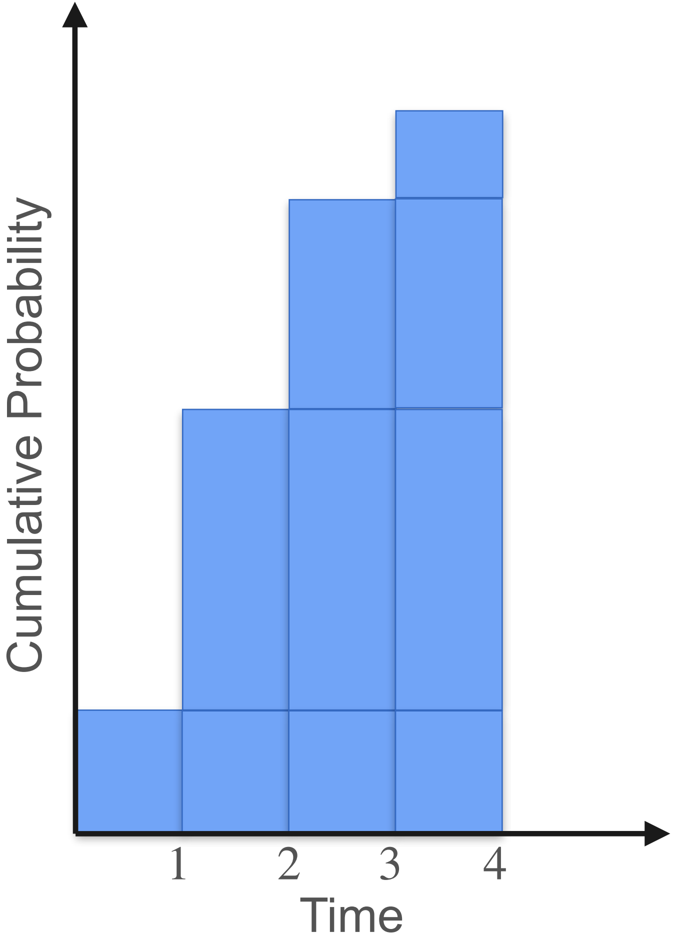
Cumulative Distribution



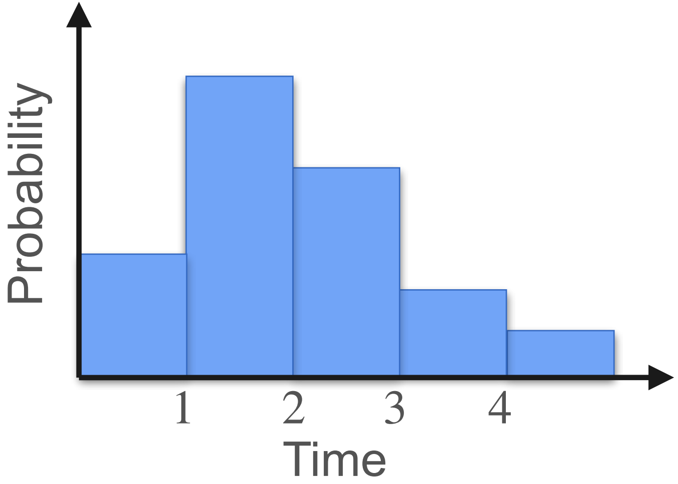


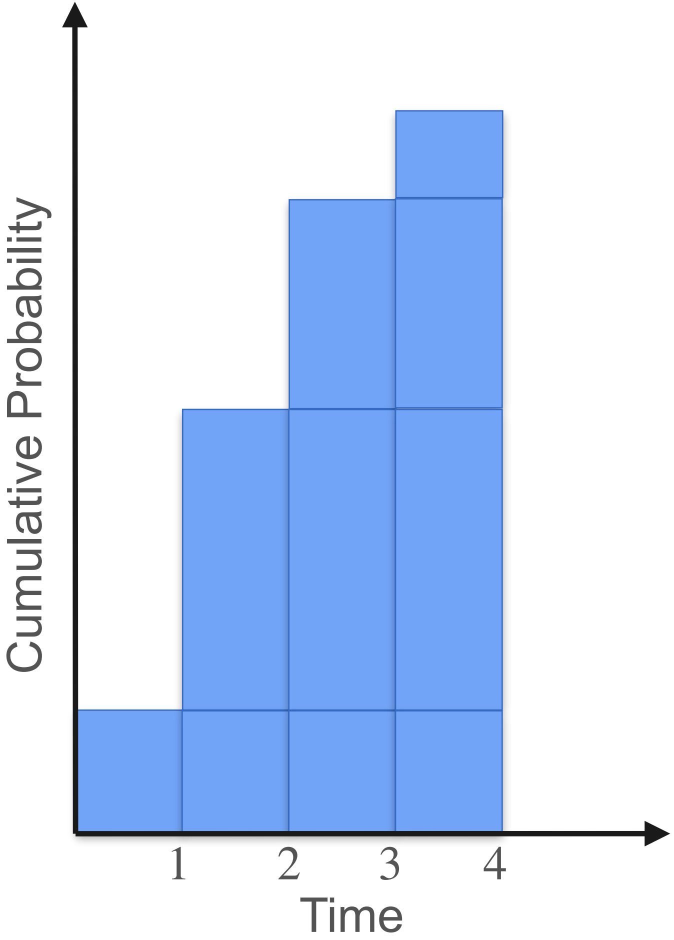
Cumulative Distribution



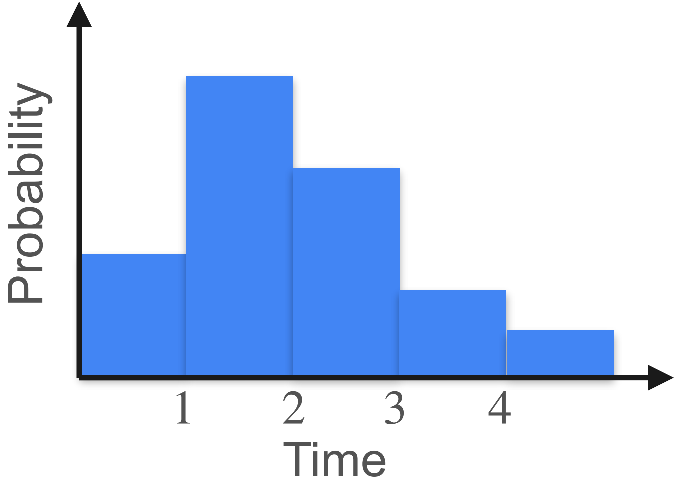


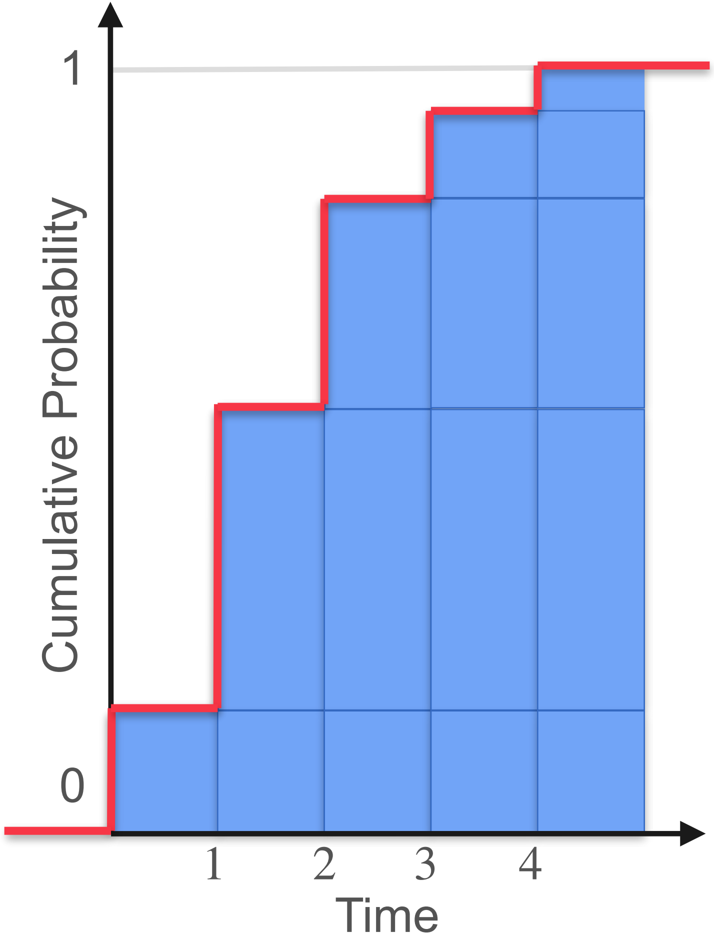
Cumulative Distribution





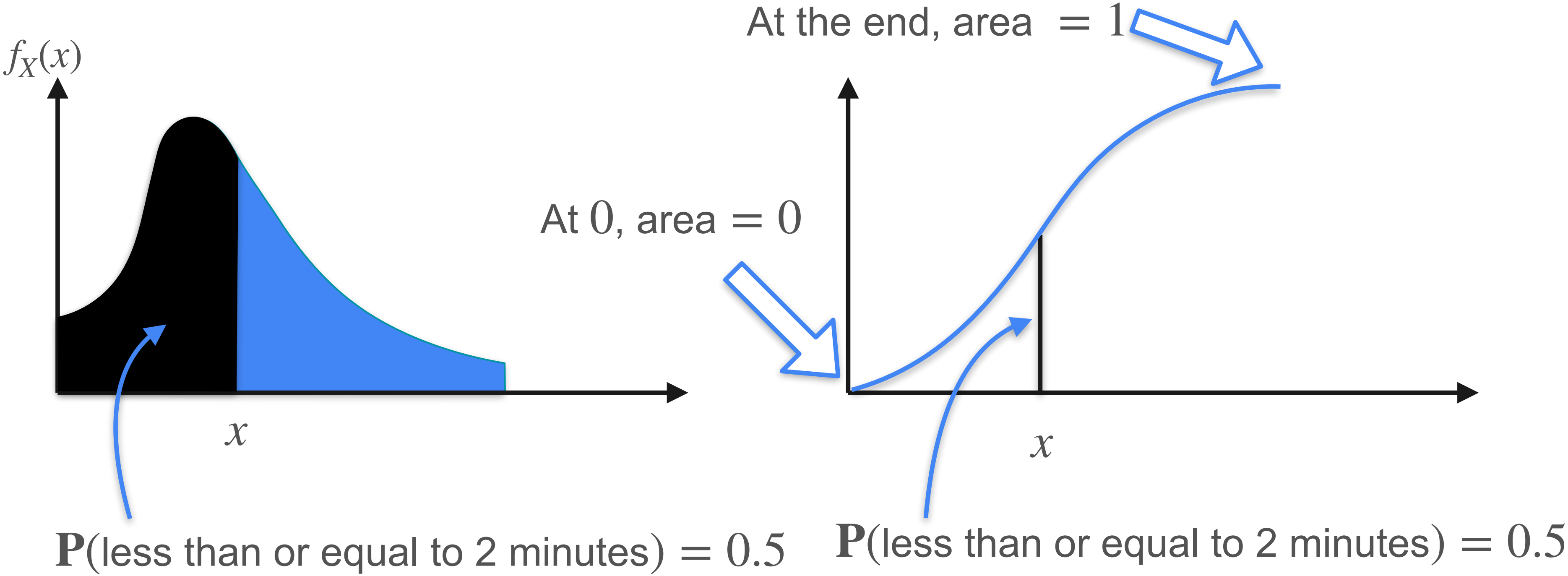
Cumulative Distribution





Cumulative Distribution

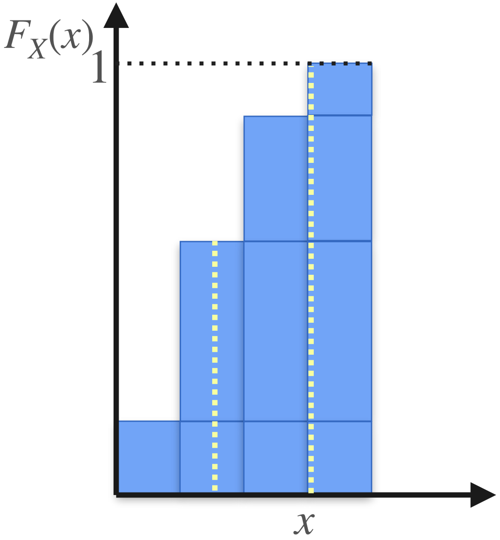
CDF: Cumulative distribution function

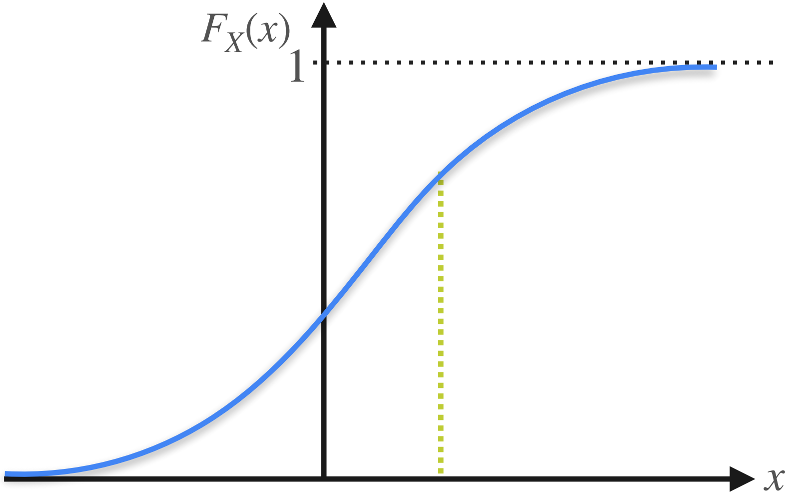


Cumulative Distribution Function: Formal Definition

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

That means that $F_X(x) = \mathbf{P}(X \leq x)$ ← It is defined for every real number

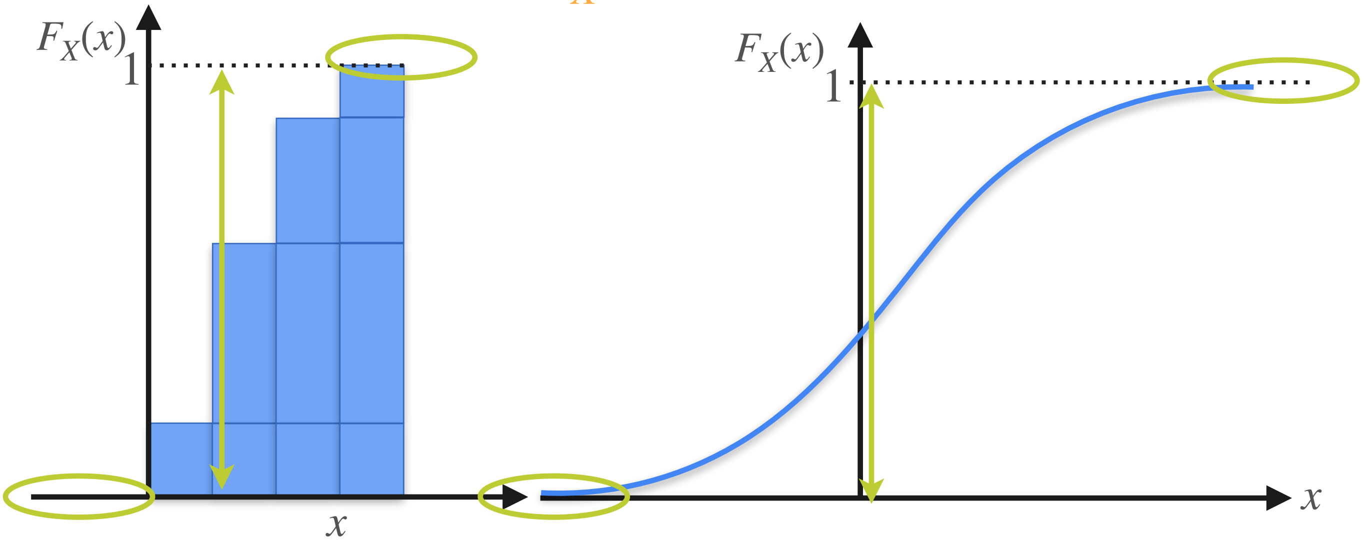




Cumulative Distribution Function: Formal Definition

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

That means that $F_X(x) = \mathbf{P}(X \leq x)$ ← It is defined for every real number

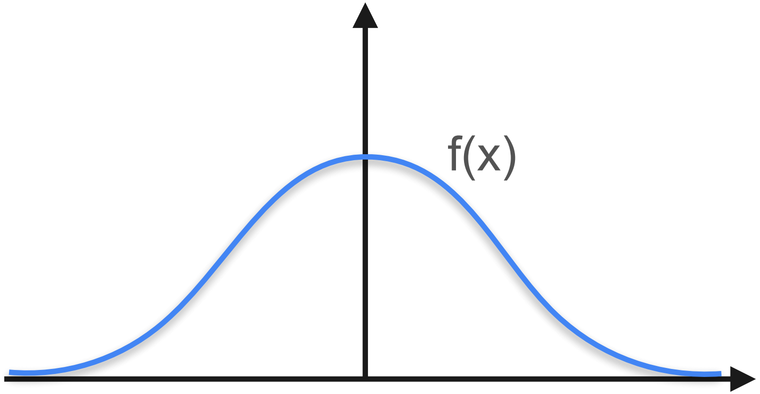


Properties

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

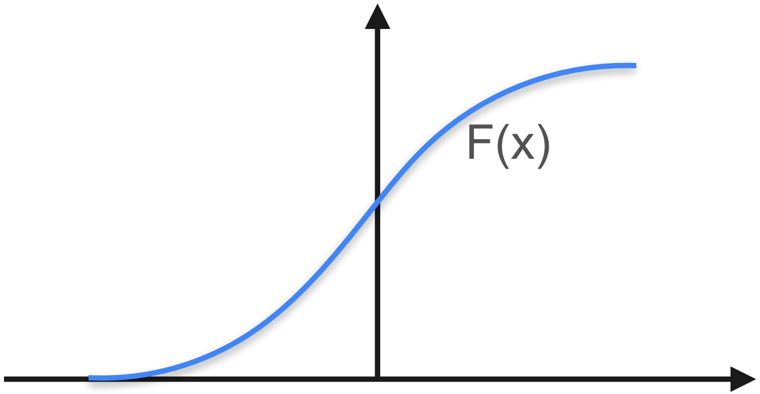
PDF and CDF Summary

PDF



- area = 1
- Always positive

CDF



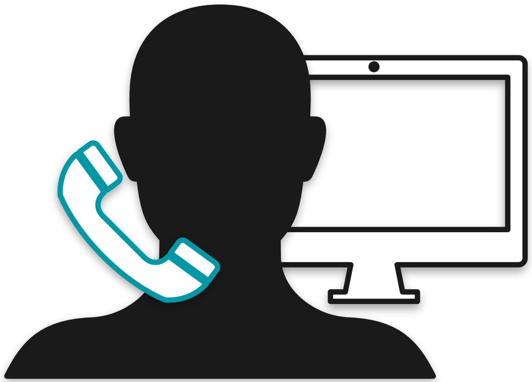
- left “endpoint” is 0
- right “endpoint" is 1
- (endpoints can be at infinity)
- Always positive and increasing

DeepLearning.AI

Probability Distributions

Uniform Distribution

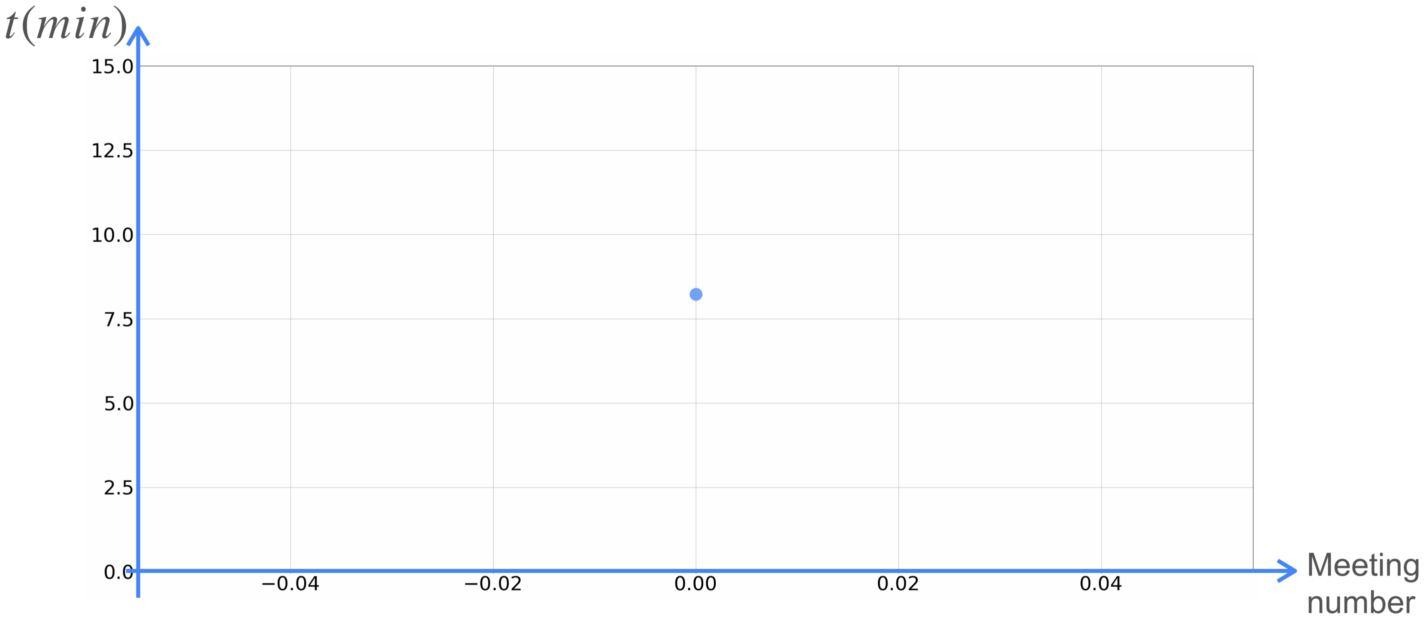
Uniform Distribution: Motivation



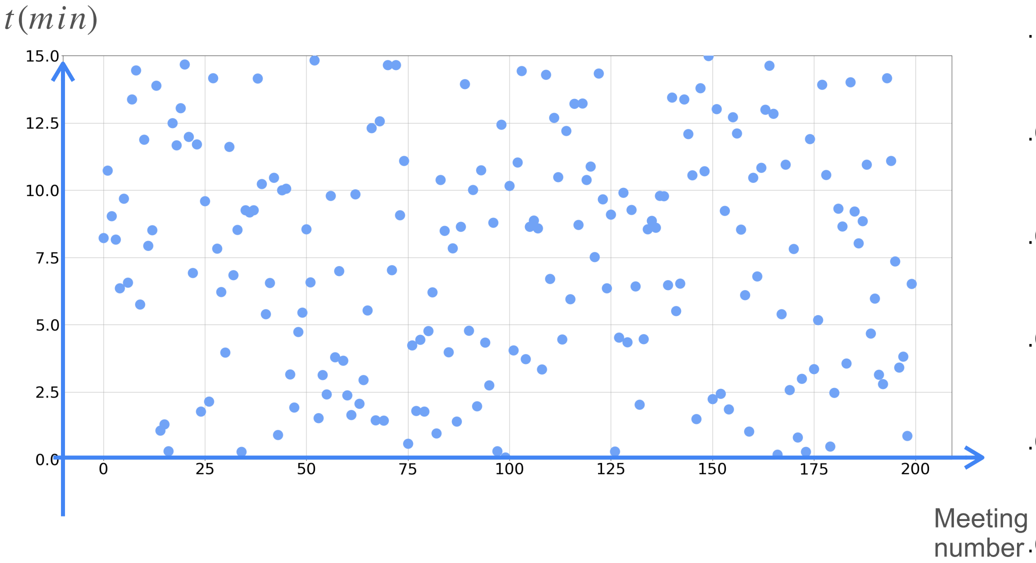
You’re calling a tech support line.
They can answer any time between zero and 15 minutes and if they don't answer in this time, the line is disconnected.

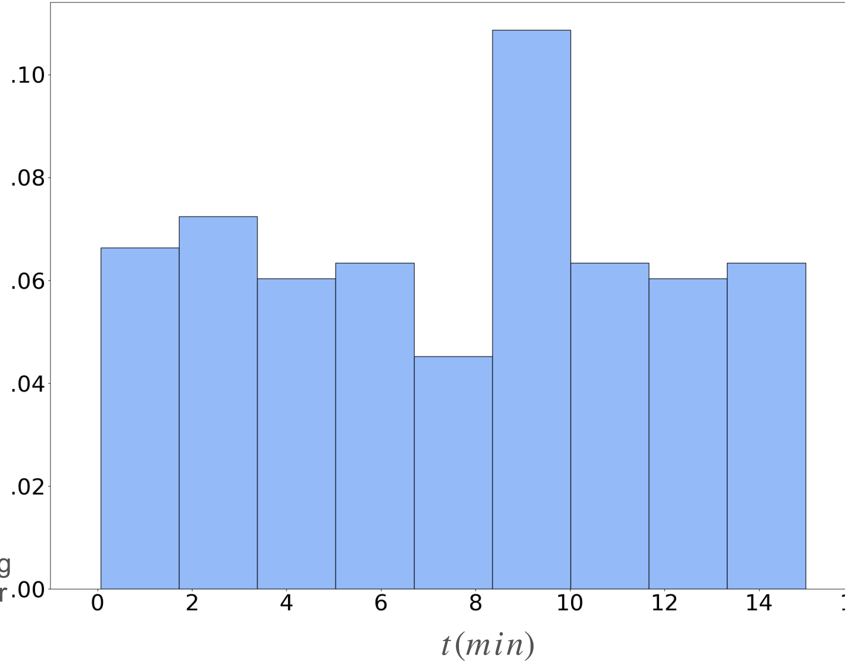
Uniform Distribution: Motivation

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



Uniform Distribution: Motivation





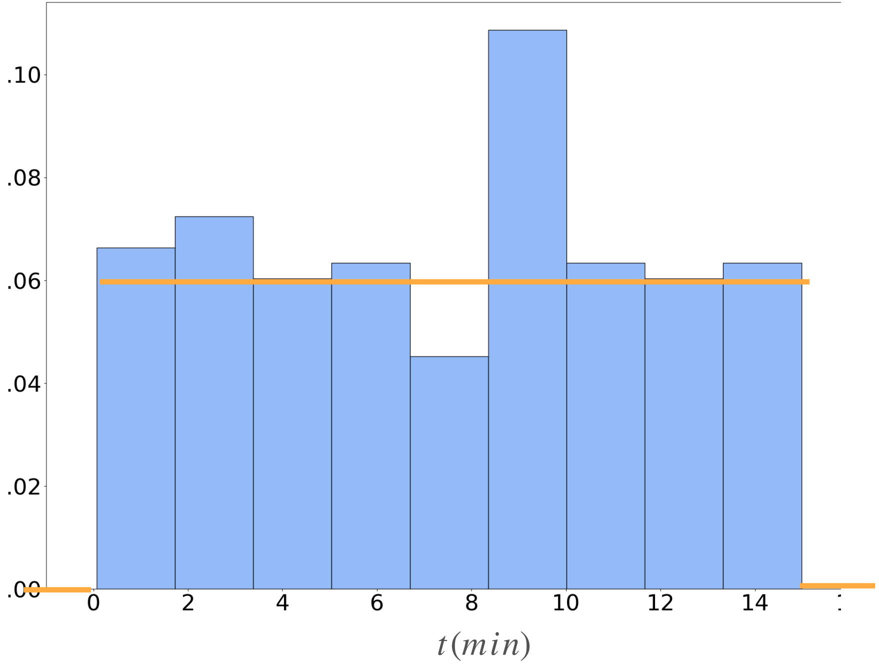
Uniform Distribution: Motivation

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



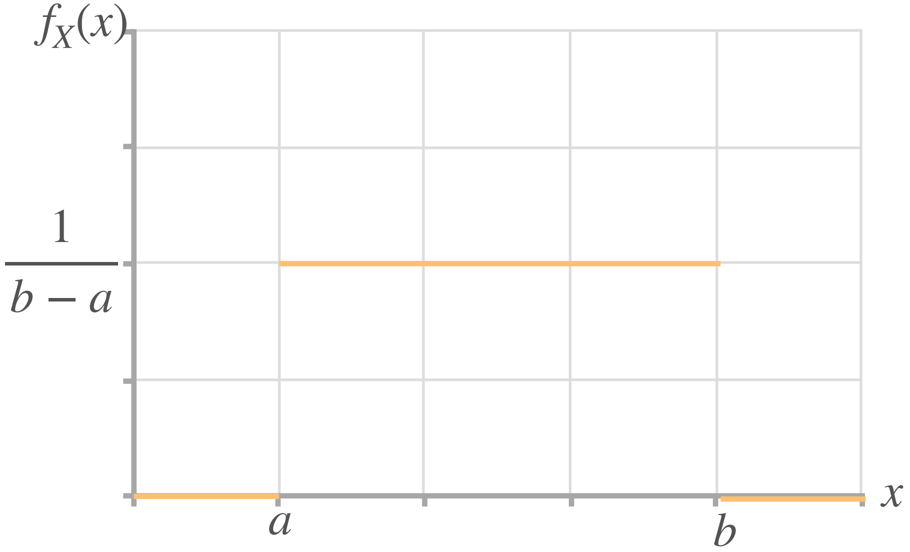
Uniform Distribution: Model

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

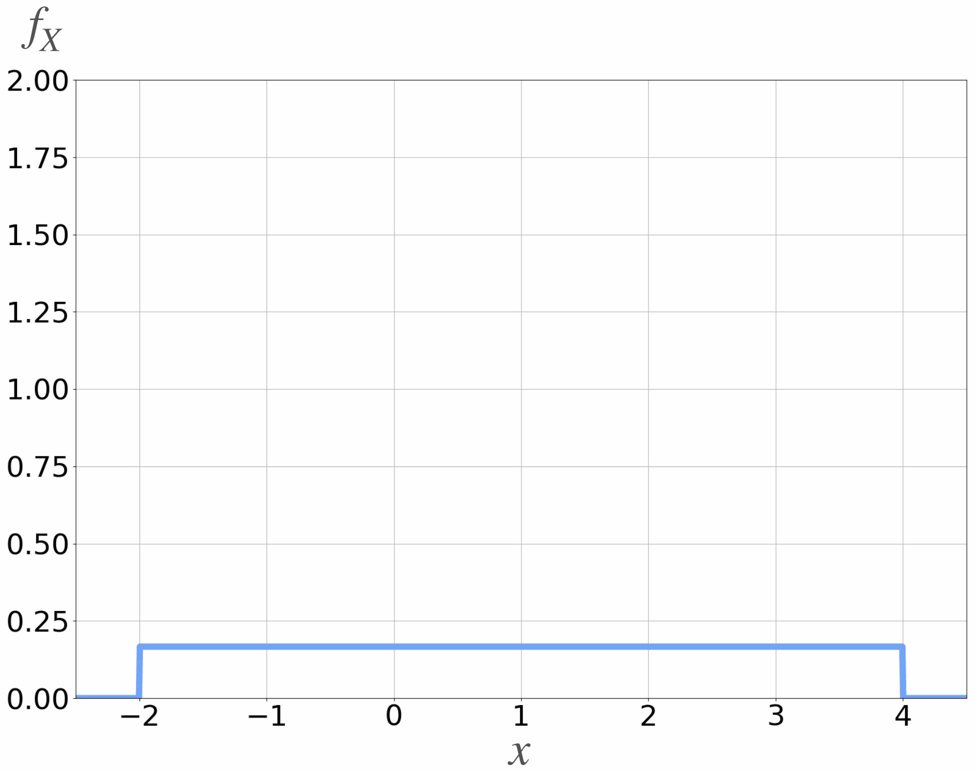
Parameters:

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

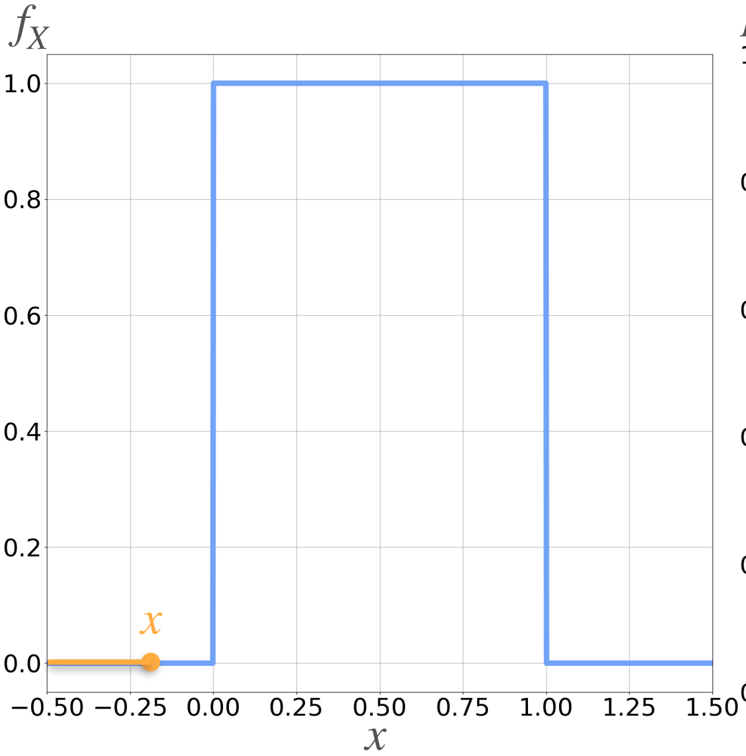
$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x \notin (a, b) \end{cases}$$

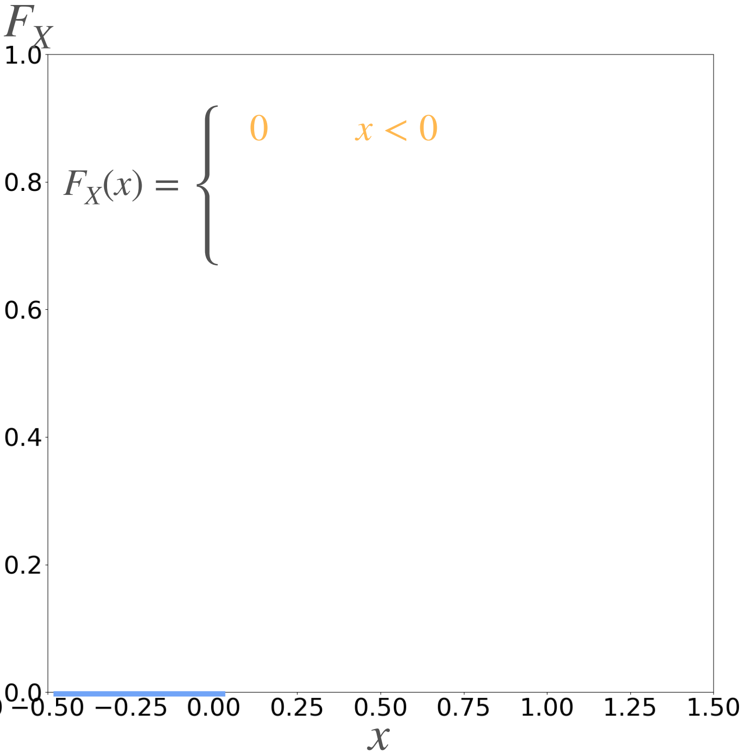


Uniform Distribution: PDF



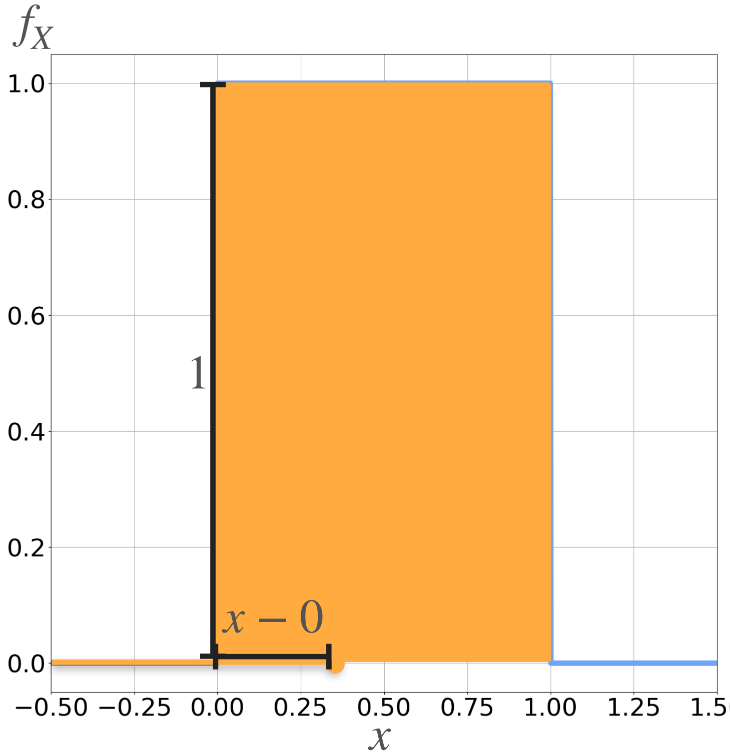
Uniform Distribution: CDF

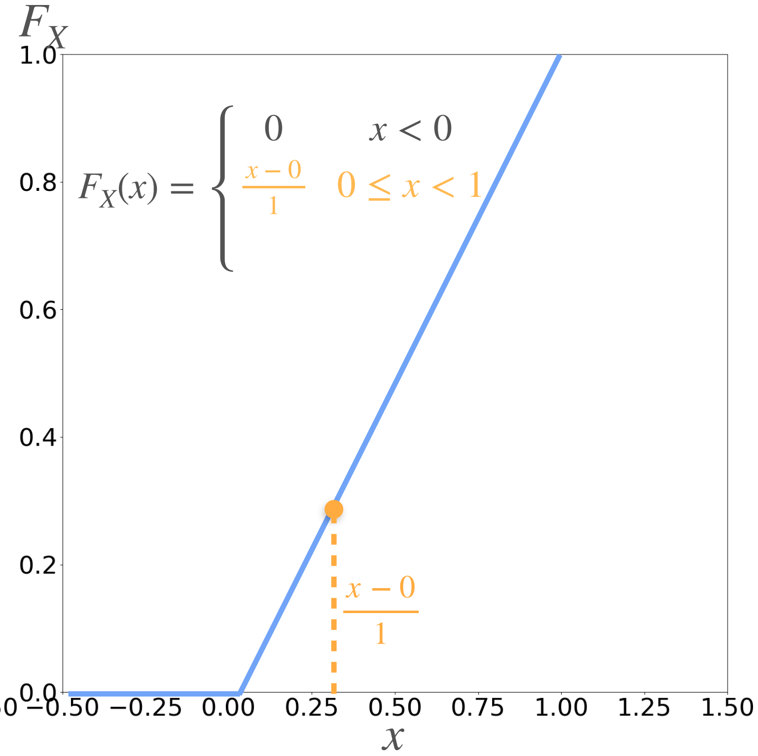




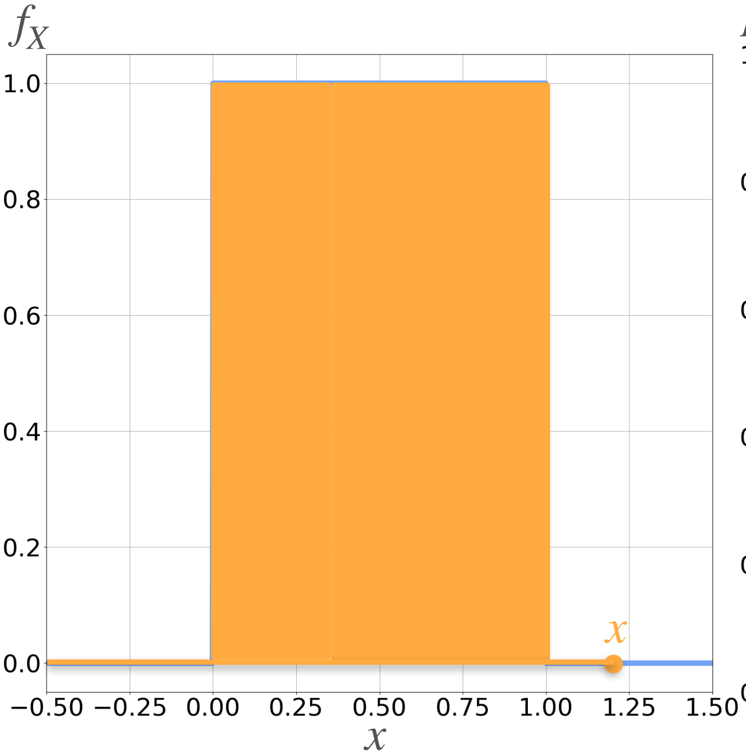
$$F_X(x) = \begin{cases} 0 & x < 0 \\ \end{cases}$$

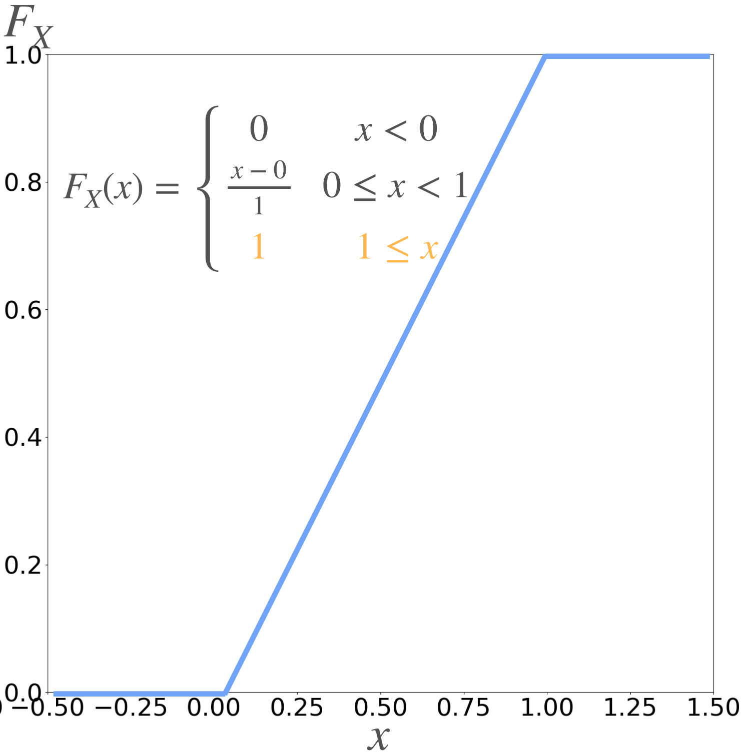
Uniform Distribution: CDF





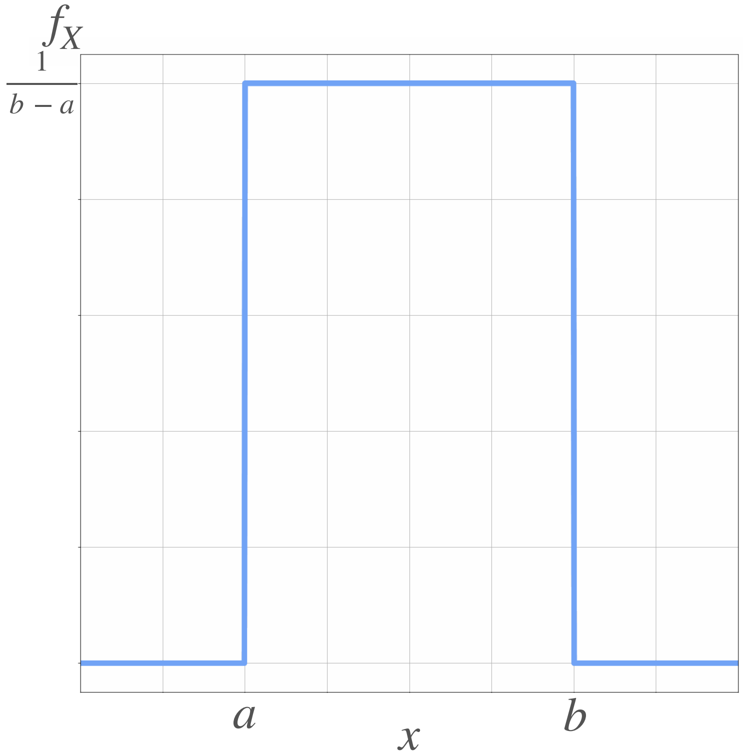
Uniform Distribution: CDF

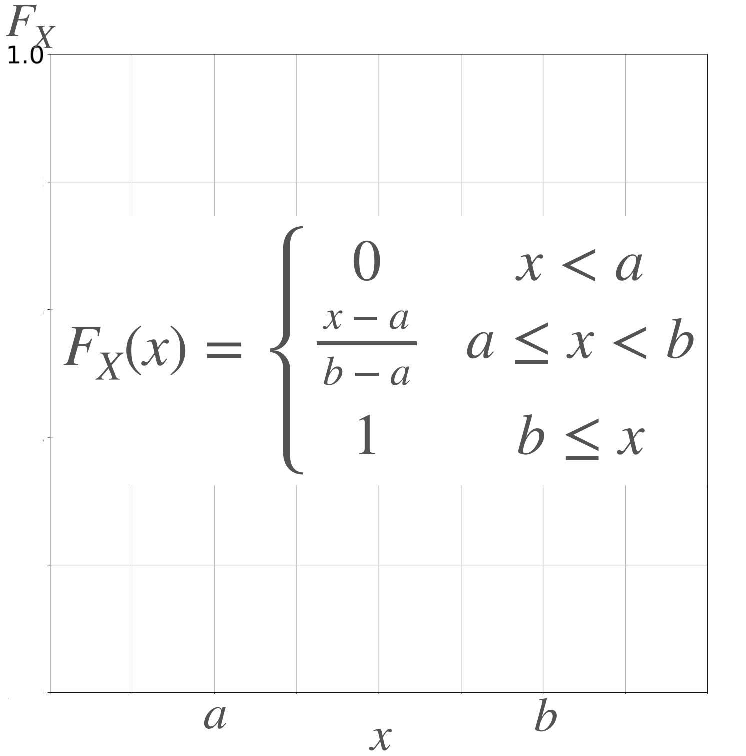




$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{x-0}{1} & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

Uniform Distribution: CDF





$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & b \le x \end{cases}$$

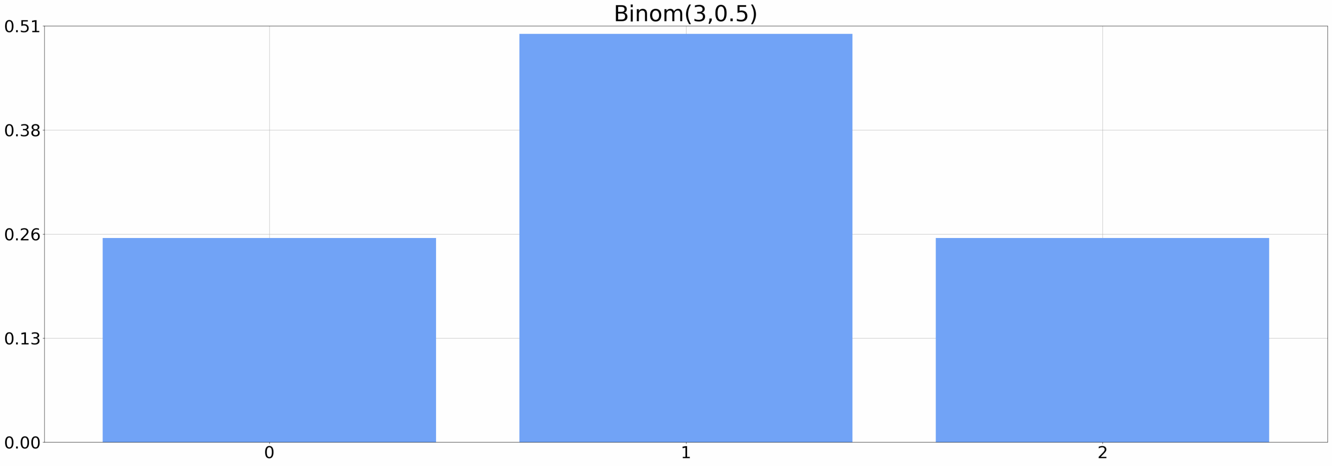
DeepLearning.AI

DeepLearning.AI

Probability Distributions

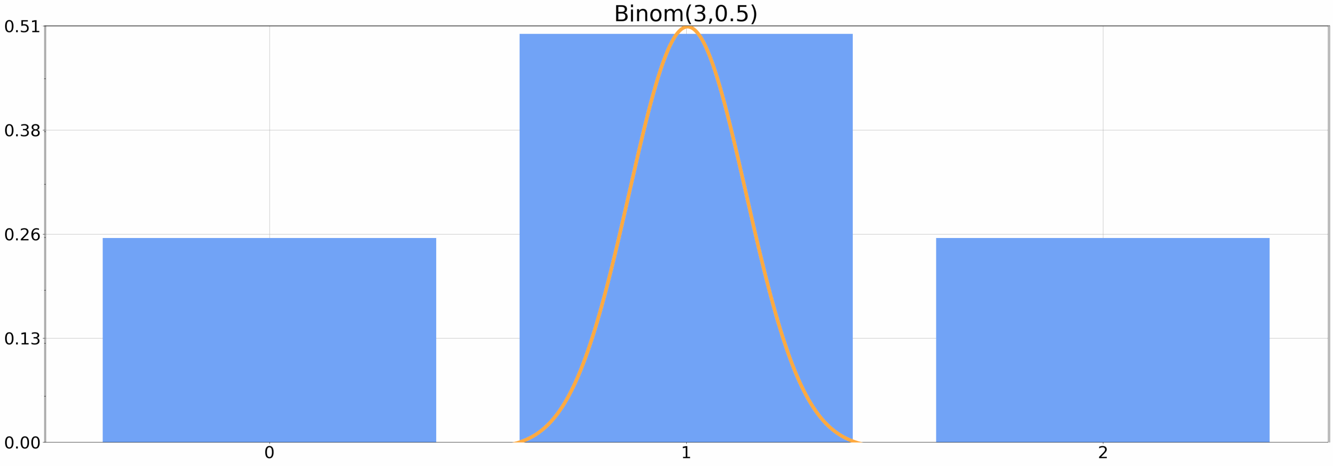
Normal distribution

Binomial Distribution With Very Large n



Number of throws = 2

Binomial Distribution With Very Large n

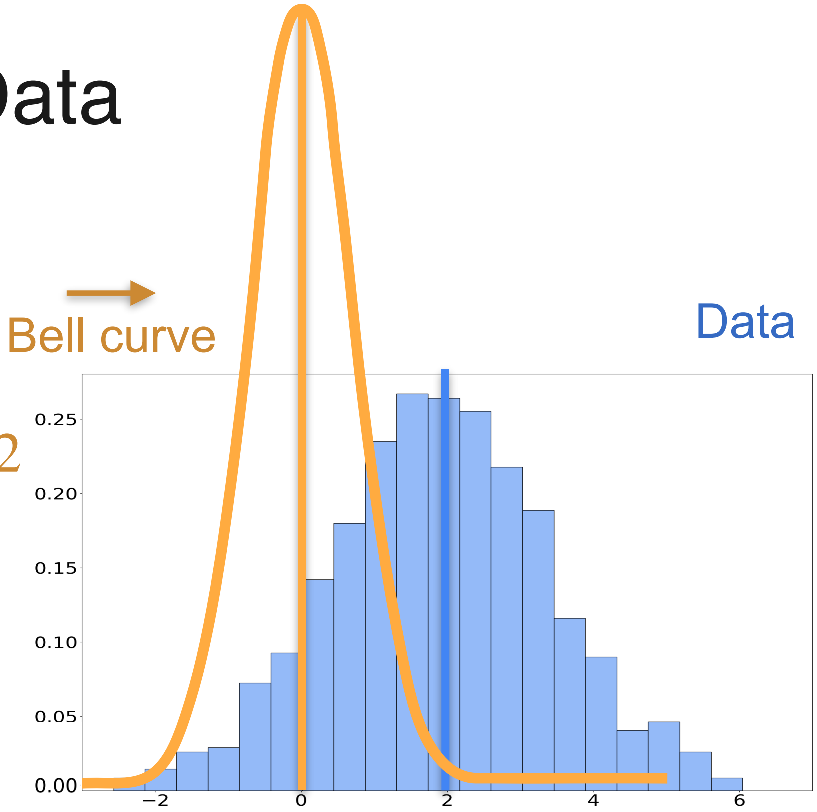


Number of throws $= n$

Bell Shaped Data

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

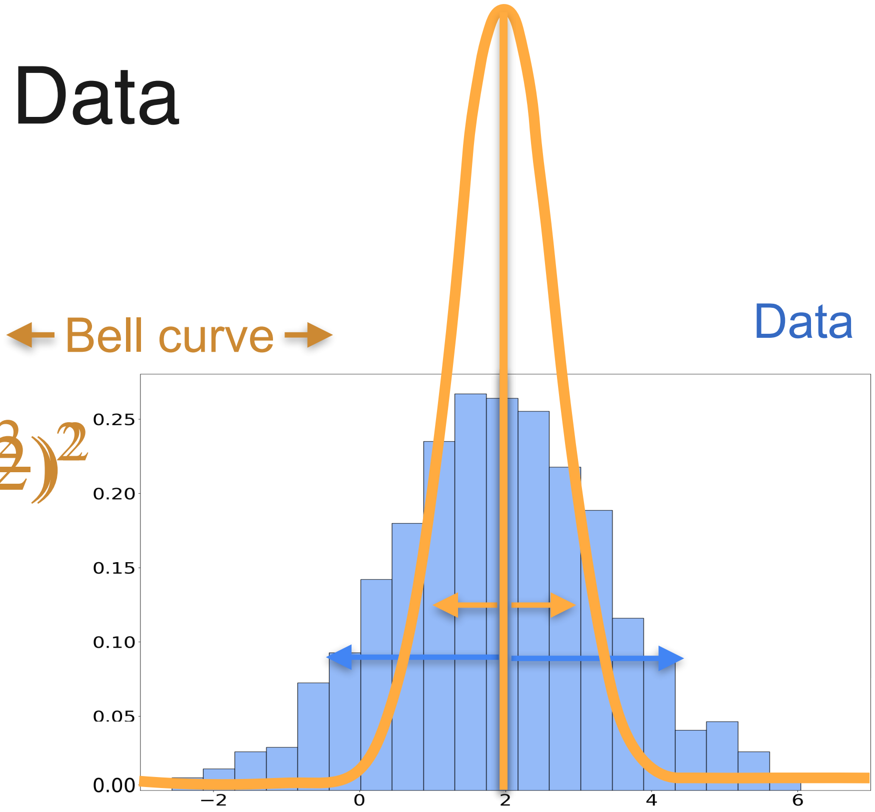
$$e^{-\frac{1}{2}(x-2)^2}$$



$\mu = 0$ $\quad \mu = 2$

Bell Shaped Data

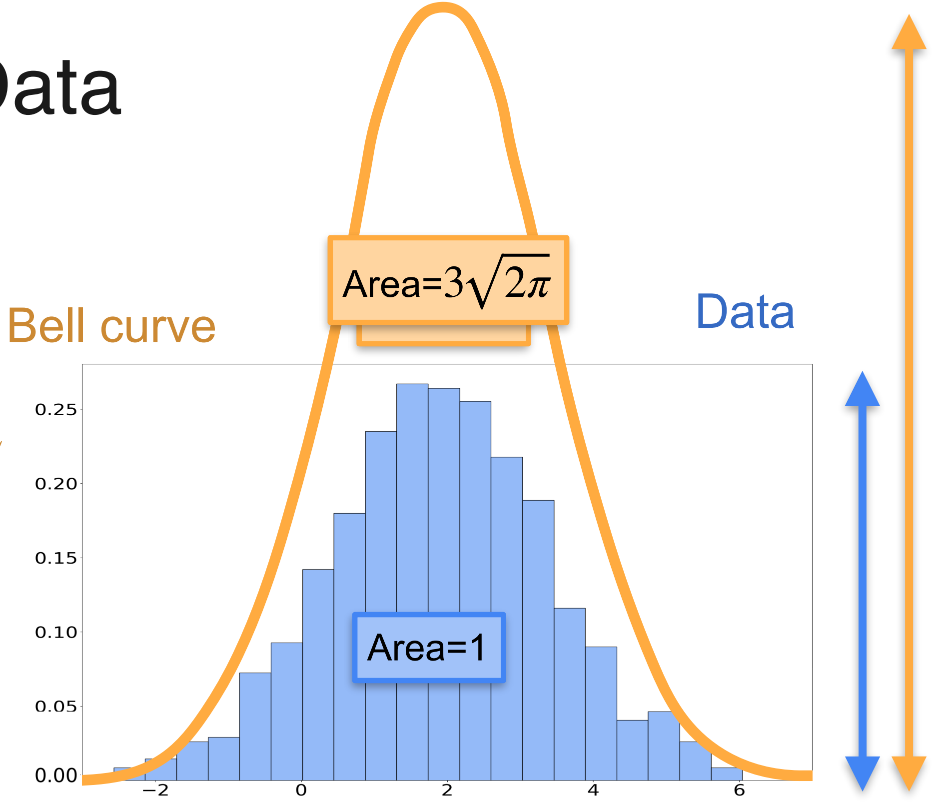
$$e^{-\frac{1}{2}\left(\frac{x-2}{3}\right)^2}$$



$\sigma = 1$ $\sigma = 3$

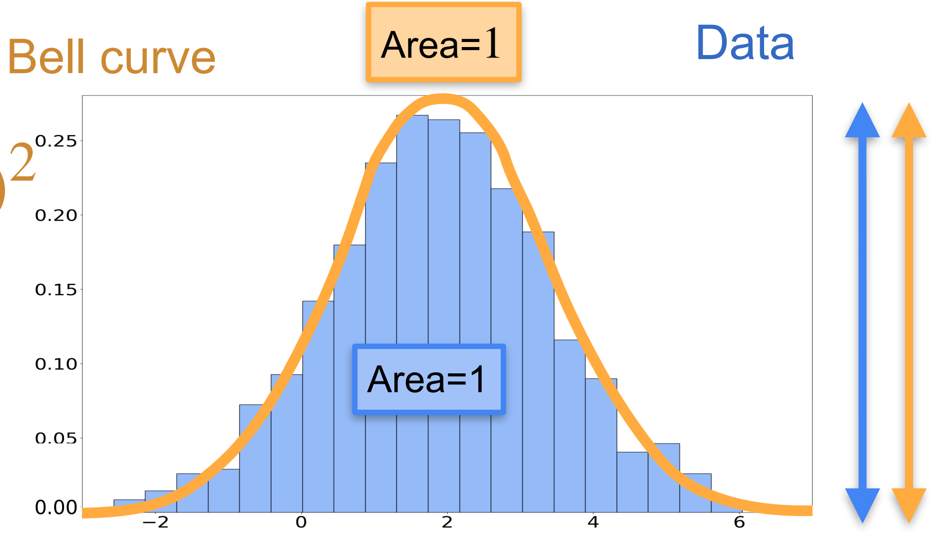
Bell Shaped Data

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



Bell Shaped Data

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

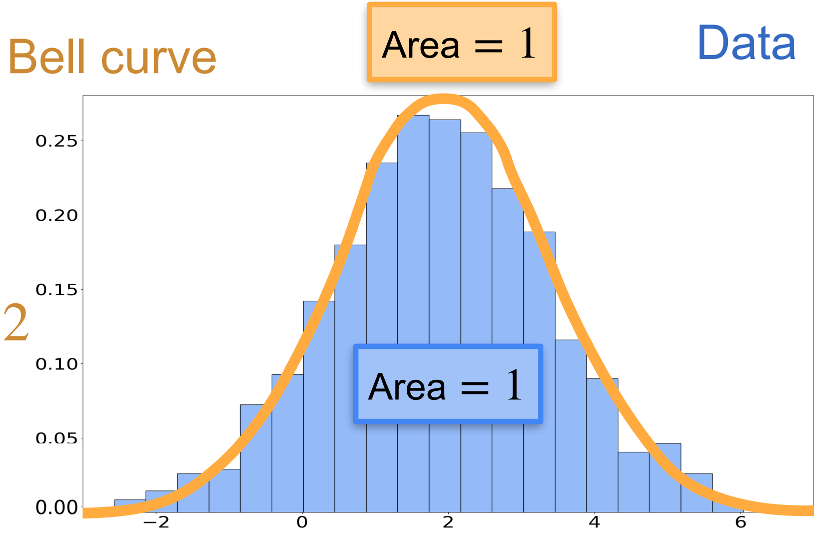


Bell Shaped Data

Mean = μ

Standard deviation = σ

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

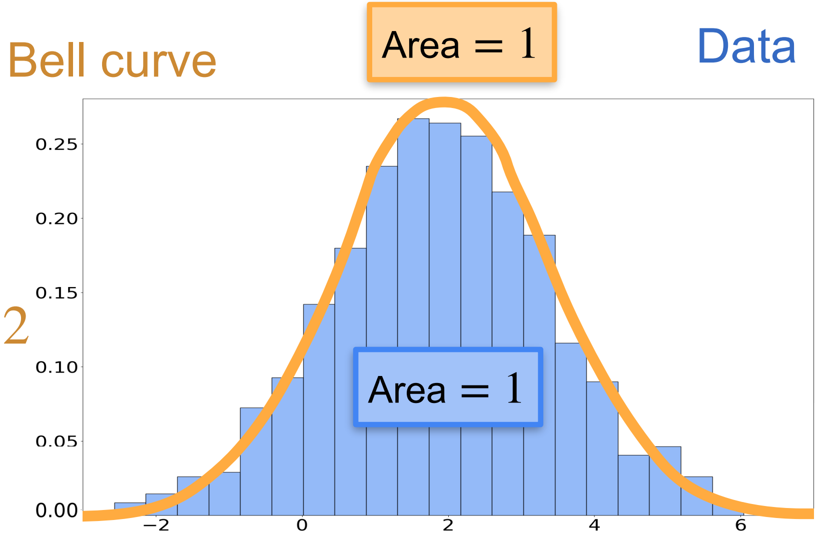


Bell Shaped Data

Mean = μ

Standard deviation = σ

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



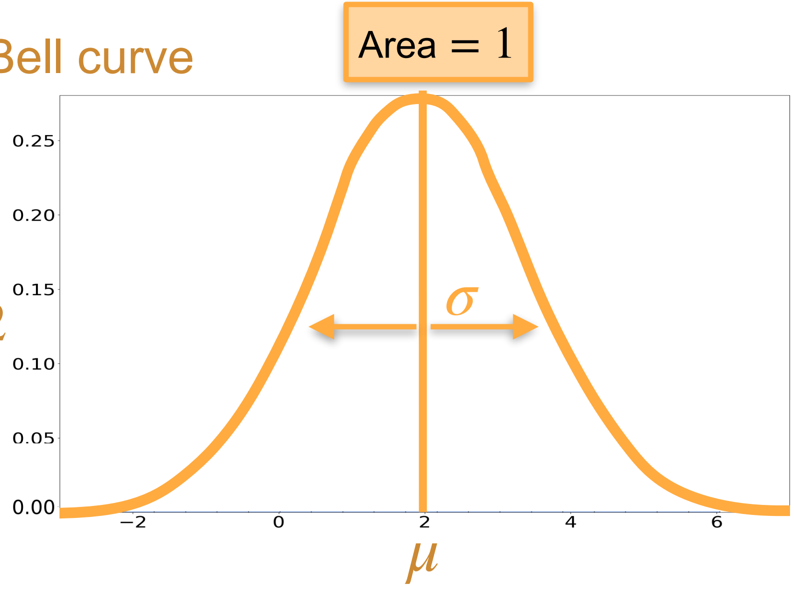
Normal Distribution

Mean = μ

Standard deviation = σ

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

Bell curve



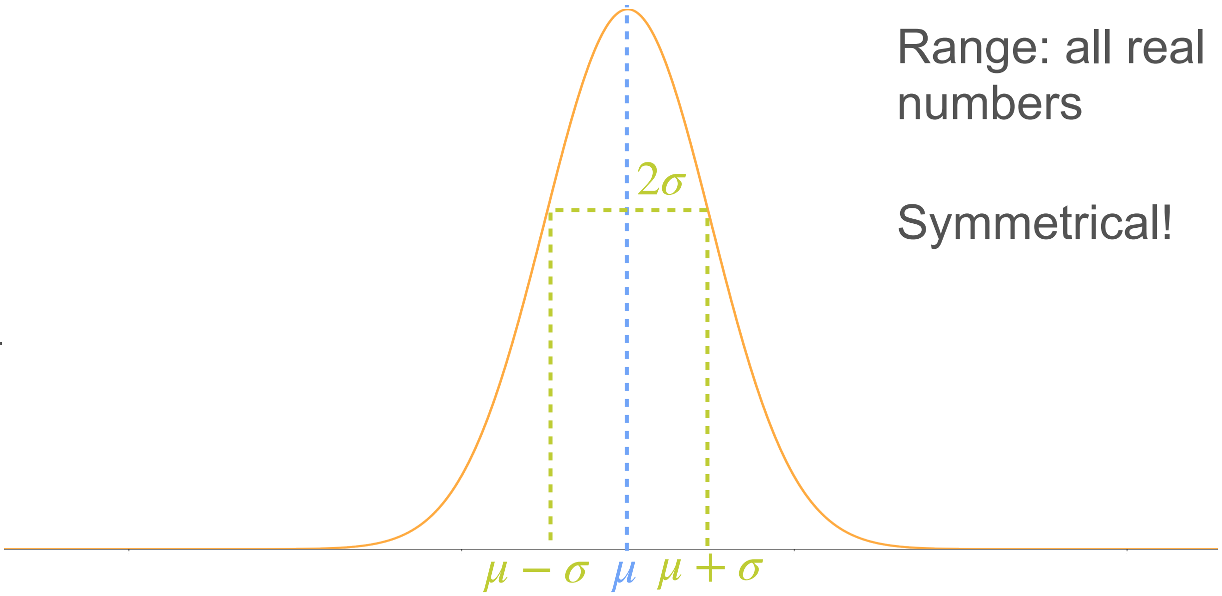
Normal Distribution

Parameters:

- μ: center of the bell
- σ: spread of the bell

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

Scaling constant



Range: all real numbers

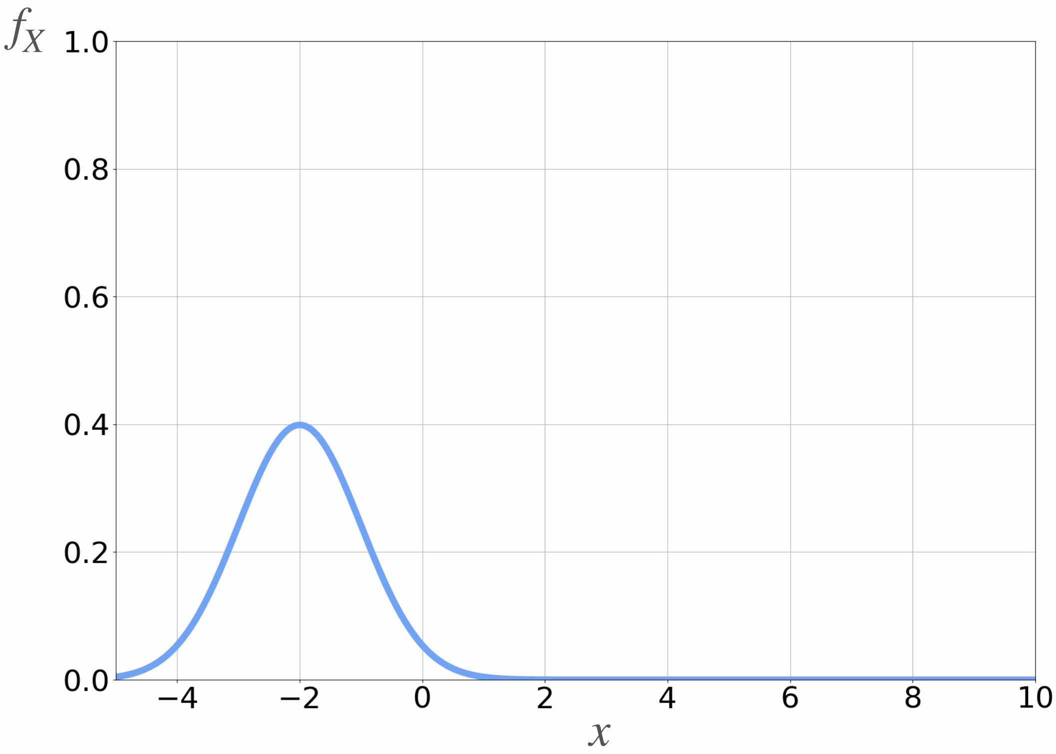
Symmetrical!

Normal Distribution

Parameters:

- μ: center of the bell
- σ: 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:

- μ: center of the bell
- σ: spread of the bell

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

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

Standard Normal Distribution

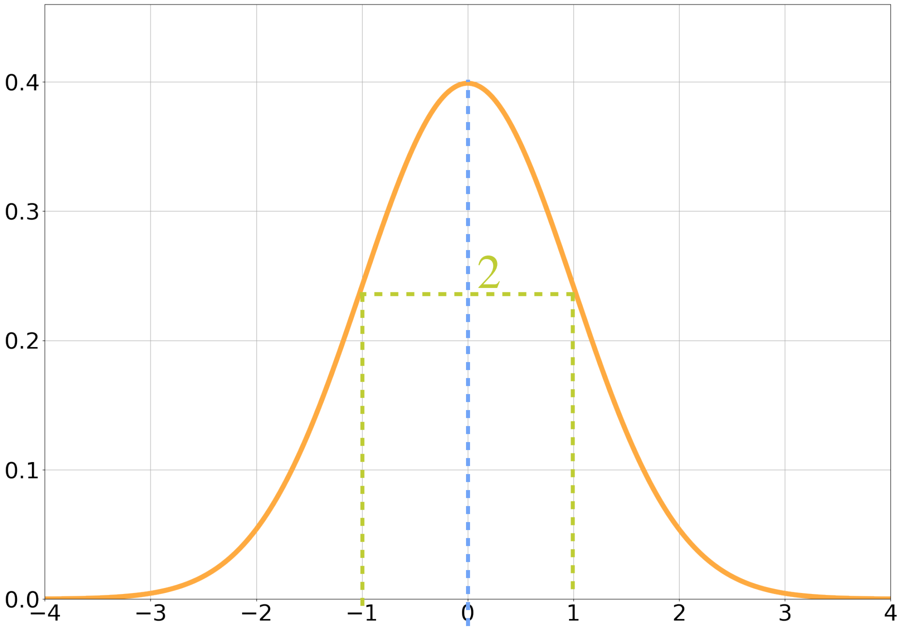
Parameters:

- μ: 0
- σ: 1

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

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

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$



Standardization

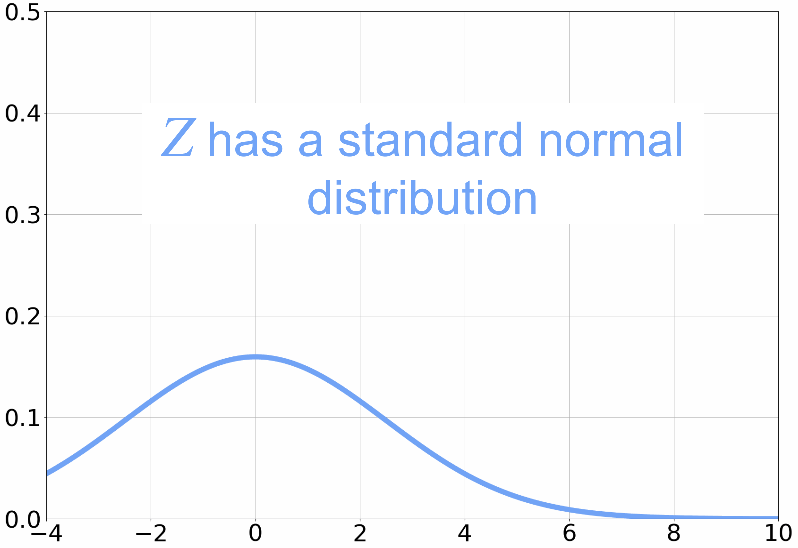
There's a really easy way to convert any normal distribution to the standard one!

X distributes normally with

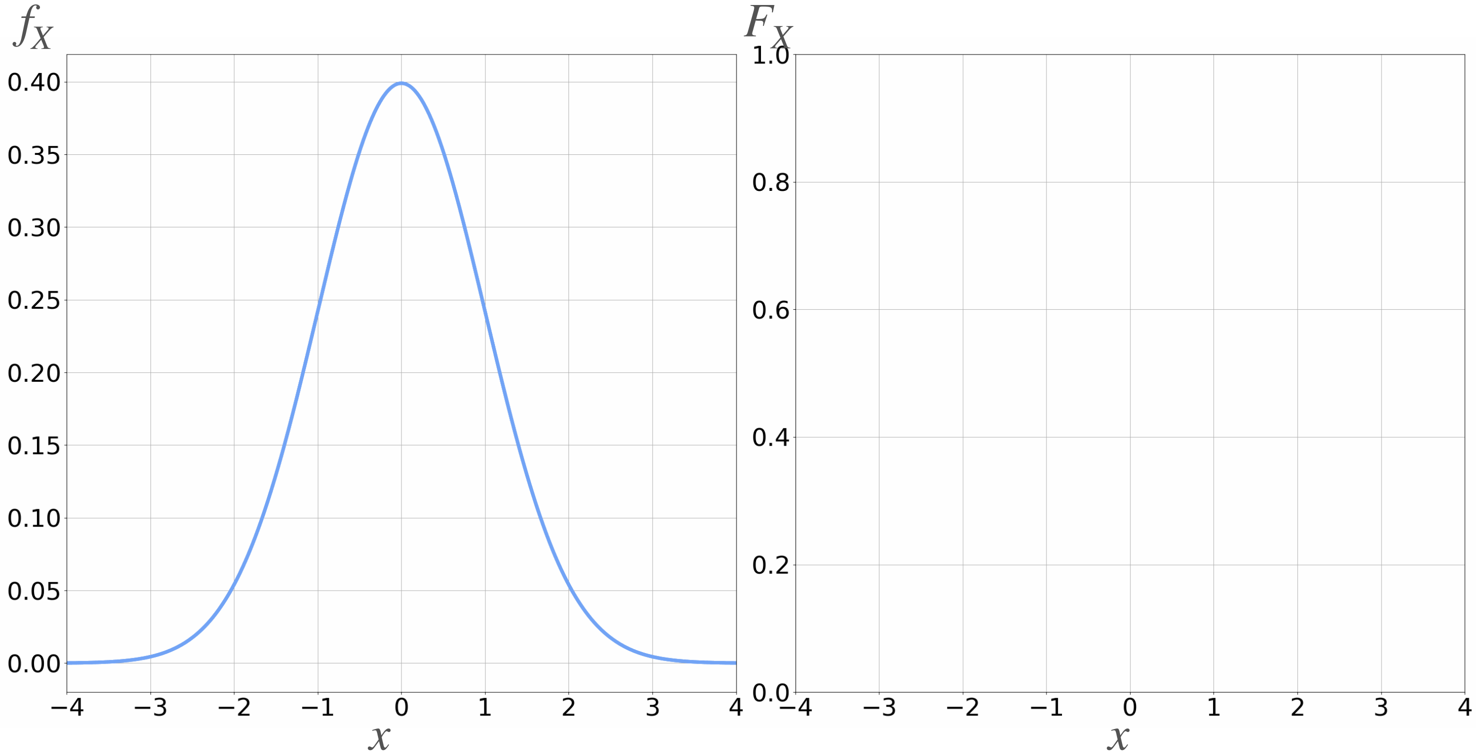
$$\mu = 2, \sigma = 2.5$$

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

Standardization is crucial to compare variables of different magnitudes!

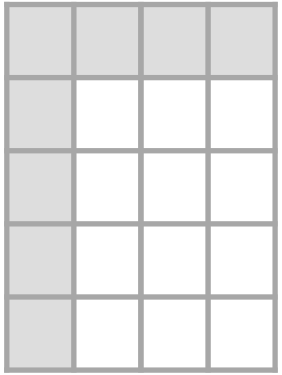


What Does the CDF Look Like?



Computing Probabilities From the PDF

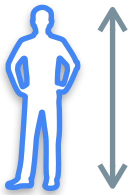
This math can't be done by hand



In the old days, people used tables of data

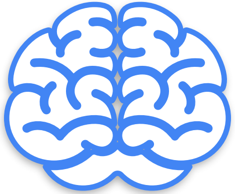
Now, you can use the help of some software to do the approximate area under the curve for you!

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

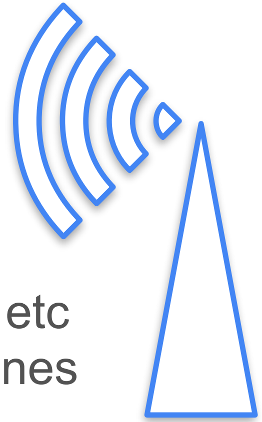
Probability Distributions

(Optional)
Chi-squared distribution

Chi-Square Distribution: Motivation

Communication channel

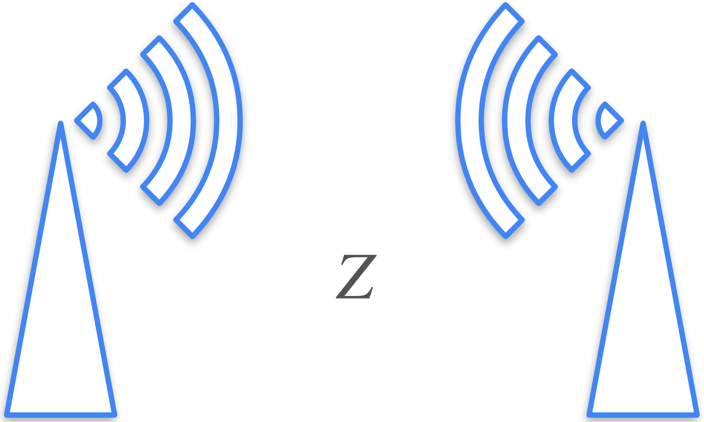
Noise
Interference from other devices
Obstructions like walls, trees, etc.
Atmospheric conditions: rain, humidity, etc
Electrical interference, i.e. from power lines
Others



Message sent: 10010

Message received: 10010 $+Z$

Chi-Square Distribution: Motivation



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

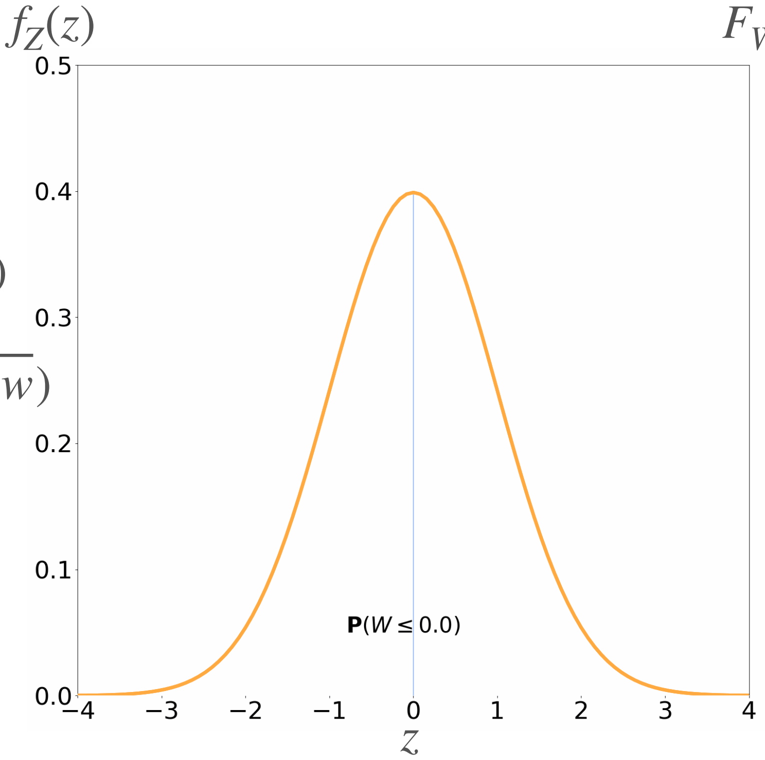
$$W = Z^2$$

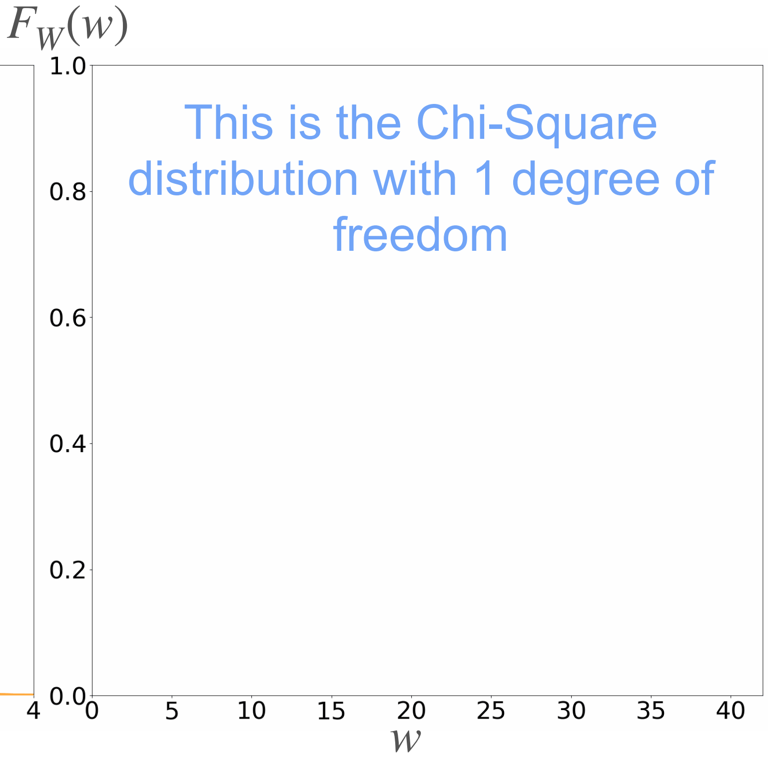
The communication channel has noise with a standard normal distribution

What is the distribution of W?

Chi Square Distribution

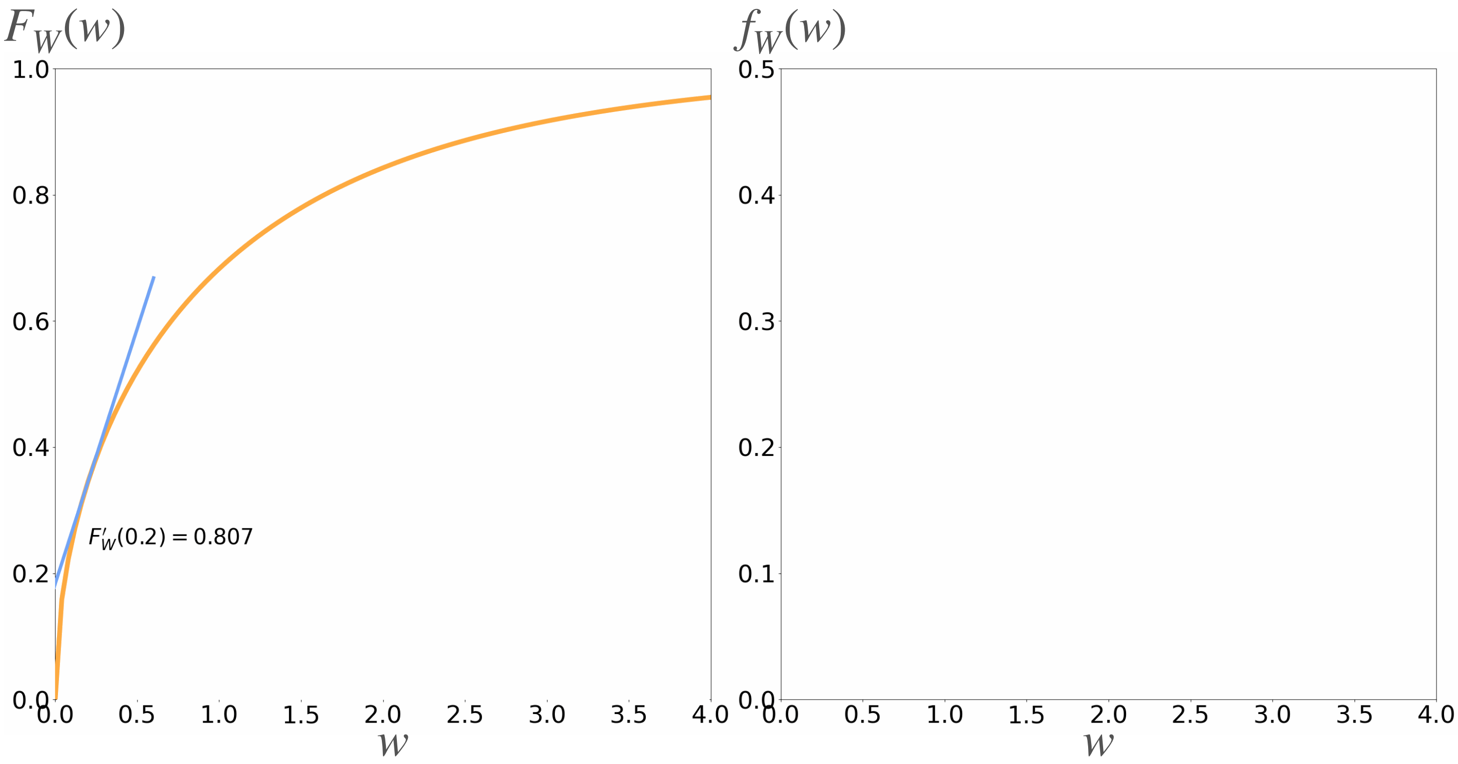
$$F_W(w) = \mathbf{P}(W \leq w)$$
$$= \mathbf{P}(Z^2 \leq w)$$
$$= \mathbf{P}(|Z| \leq \sqrt{w})$$
$$= \mathbf{P}(-\sqrt{w} \leq Z \leq \sqrt{w})$$





Chi Square Distribution

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



Chi-Square Distribution

Accumulated power over 2 transmissions?

$$W_2 = Z_1^2 + Z_2^2$$

Chi-Square with 2 df

Accumulated power over 5 transmissions?

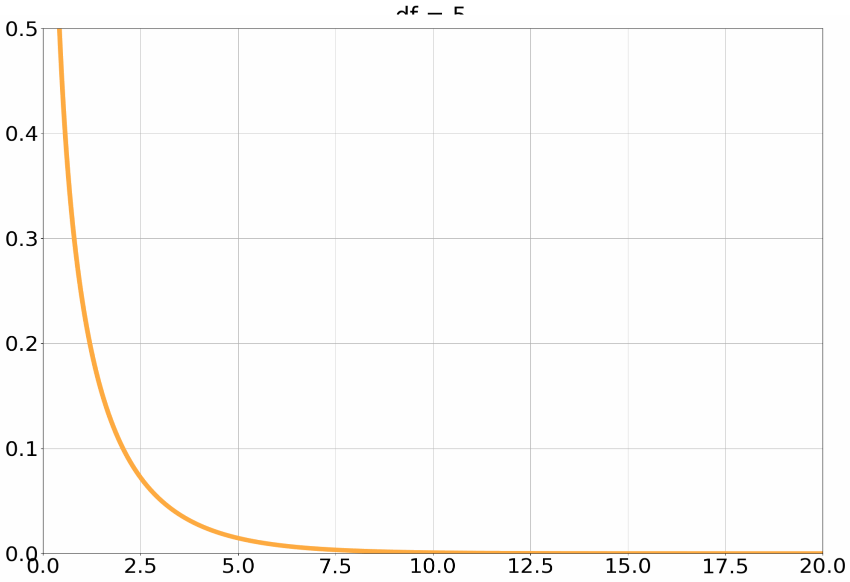
$$W_5 = Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2 + Z_5^2$$

Chi-Square with 5 df

Accumulated power over k transmissions?

$$W_k = \sum_{i=1}^{k} Z_i^2$$

Chi-Square with k df

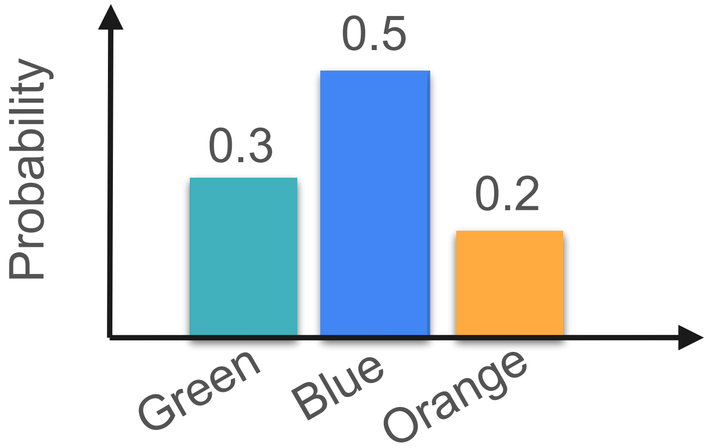


DeepLearning.AI

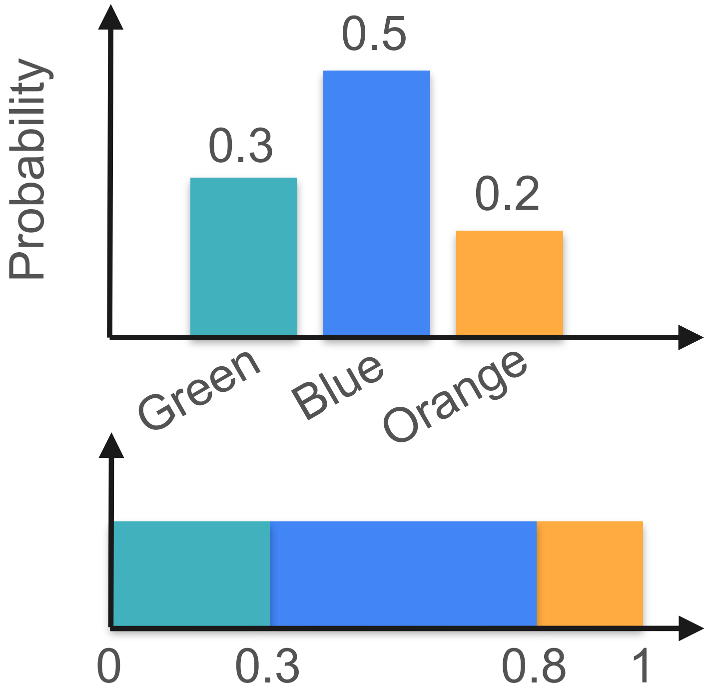
Probability Distributions

Sampling from a Distribution

Sampling From a Distribution

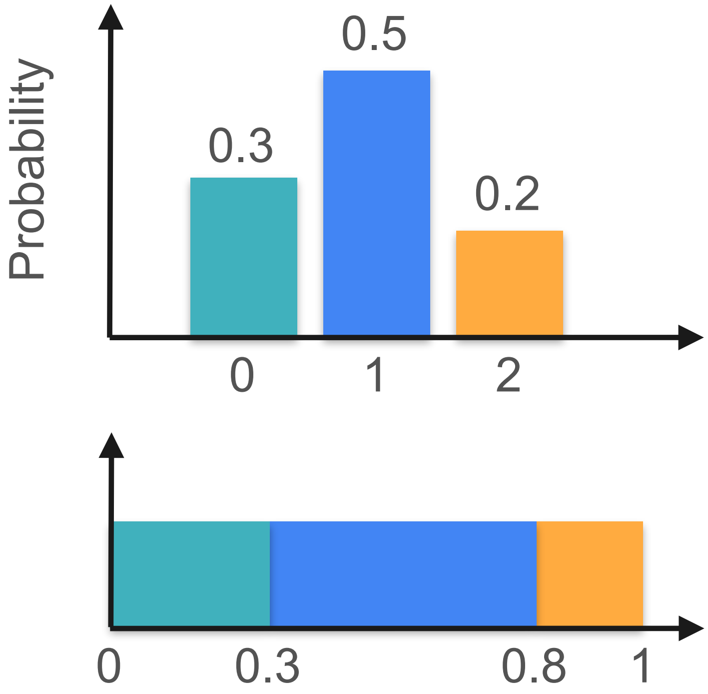


Sampling From a Distribution



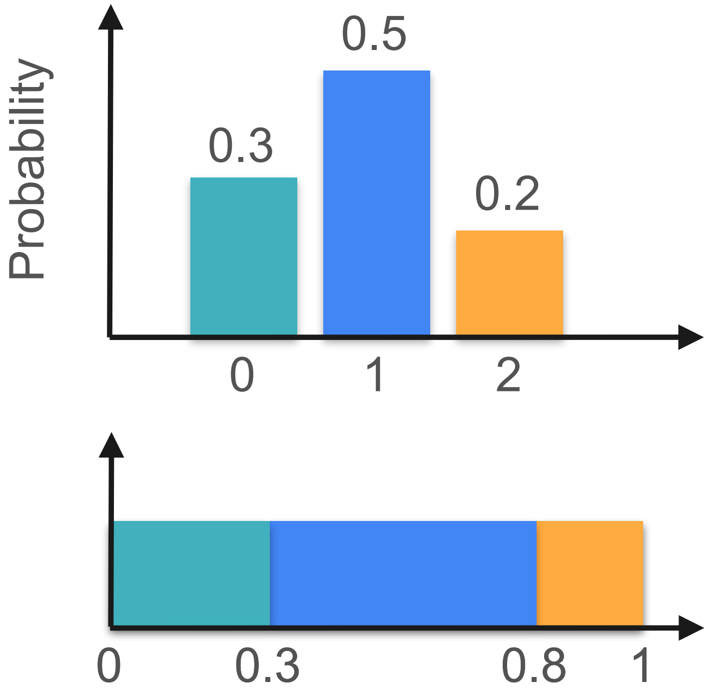
- **Step 1**: generate a random number between 0 and 1
- **Step 2**: find out which interval the number belongs to
  - [0, 0.3)
  - [0.3, 0.8)
  - [0.8, 1]
- **Step 3**: Assign an outcome based on the interval

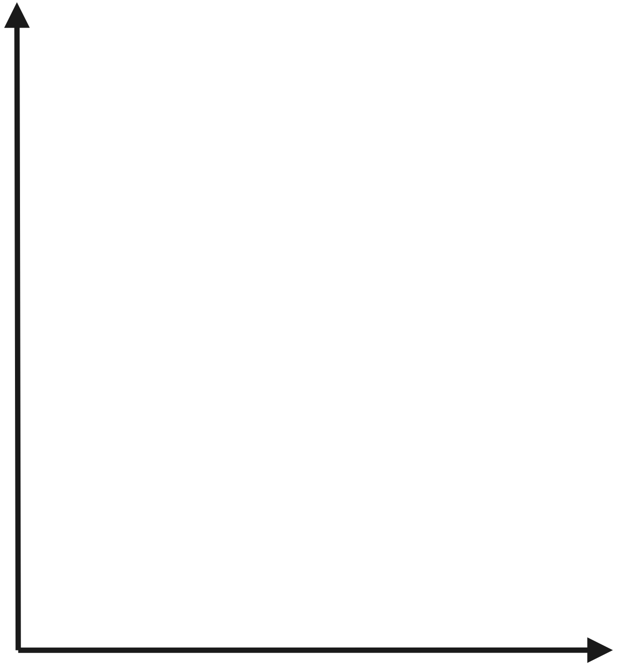
Sampling From a Distribution



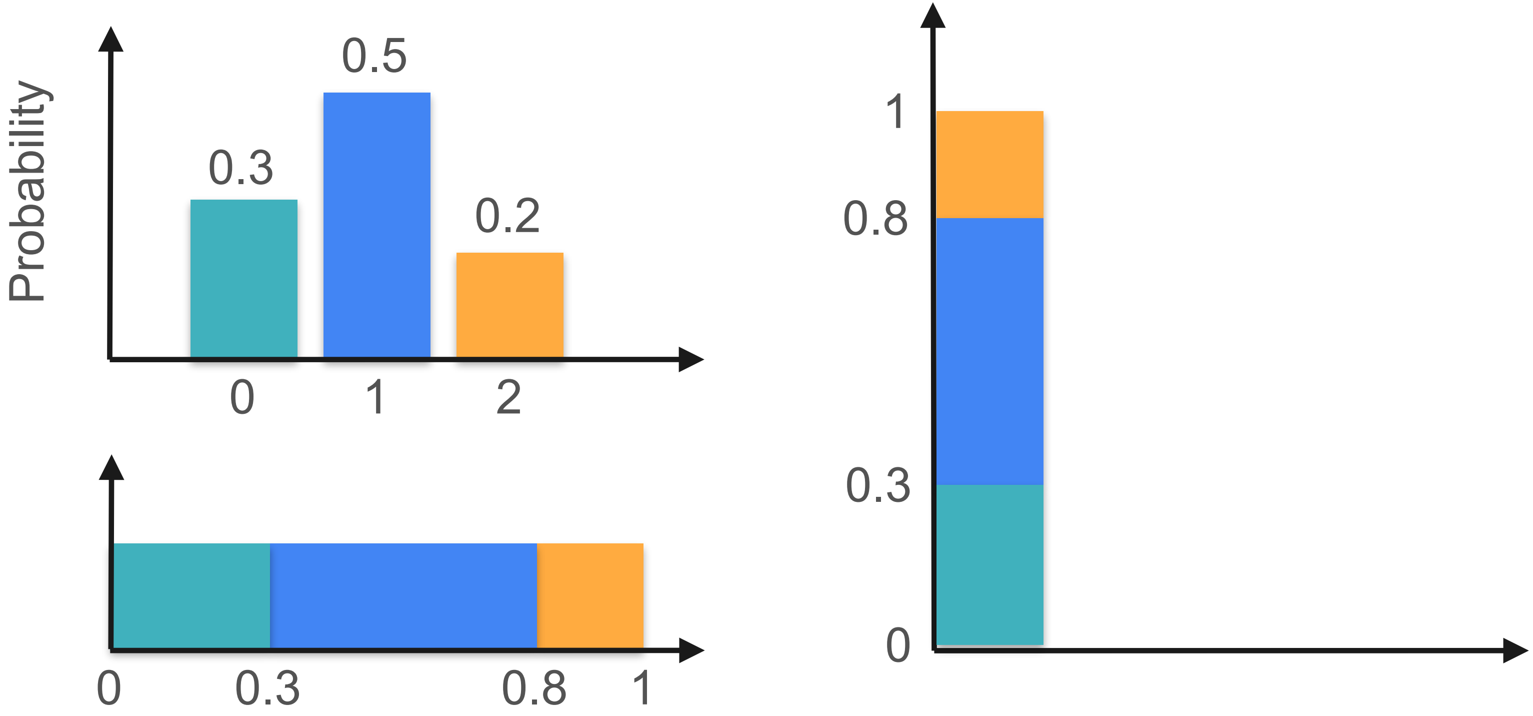
- **Step 1**: generate a random number between 0 and 1
- **Step 2**: find out which interval the number belongs to
  - [0, 0.3)
  - [0.3, 0.8)
  - [0.8, 1]
- **Step 3**: Assign an outcome based on the interval

Sampling From a Distribution



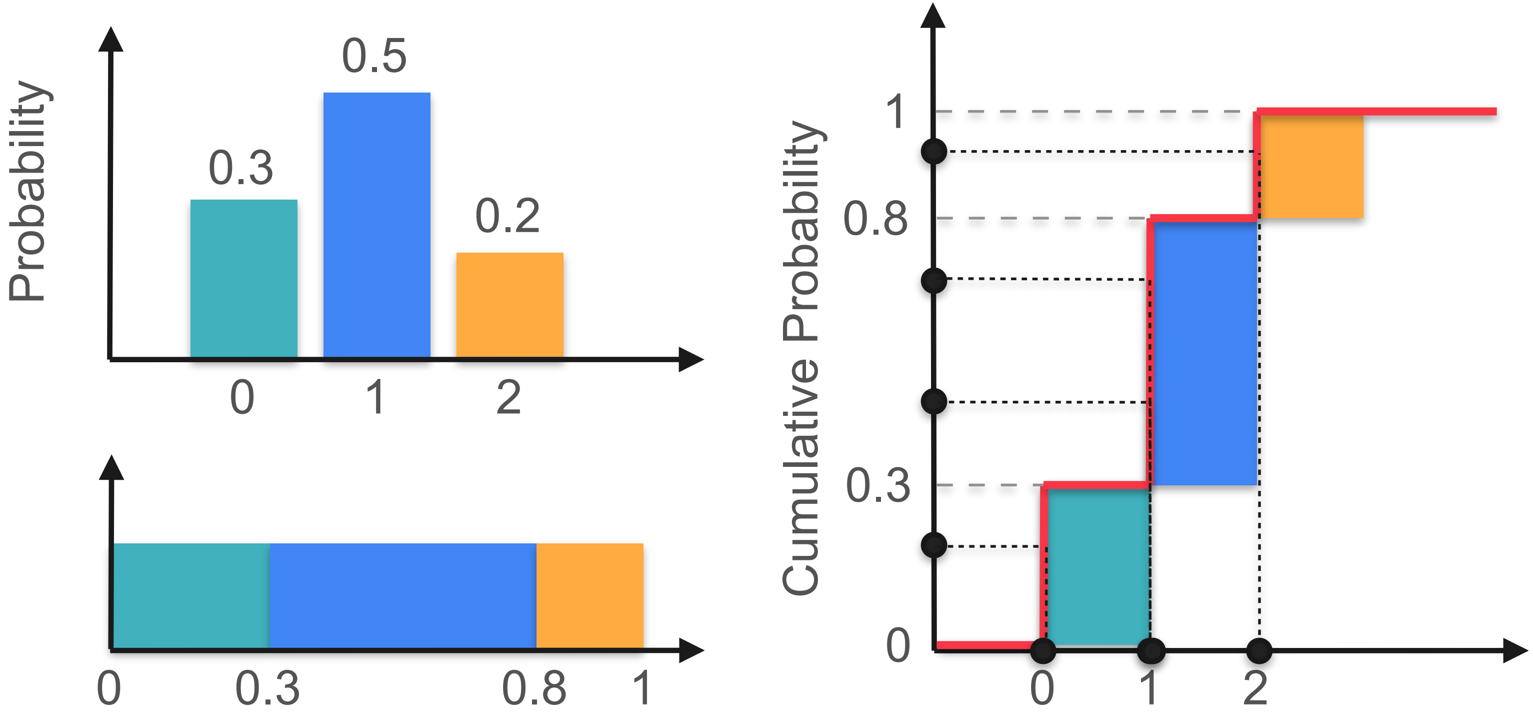


Sampling From a Distribution



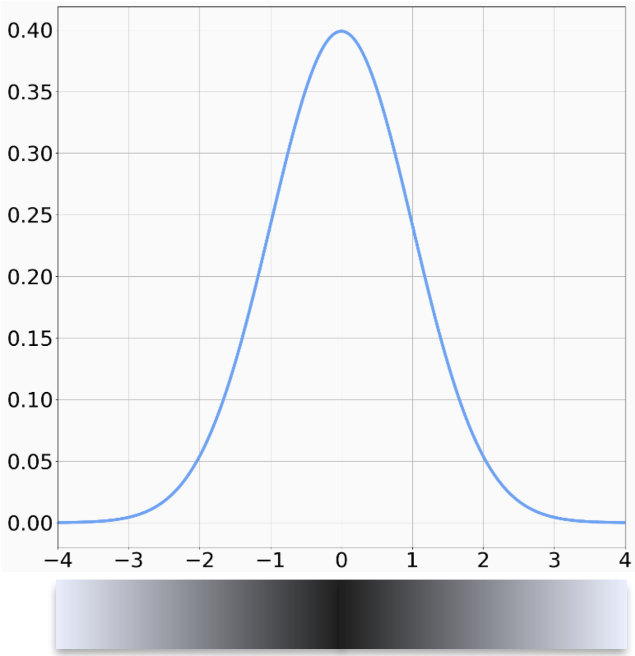
DeepLearning.AI

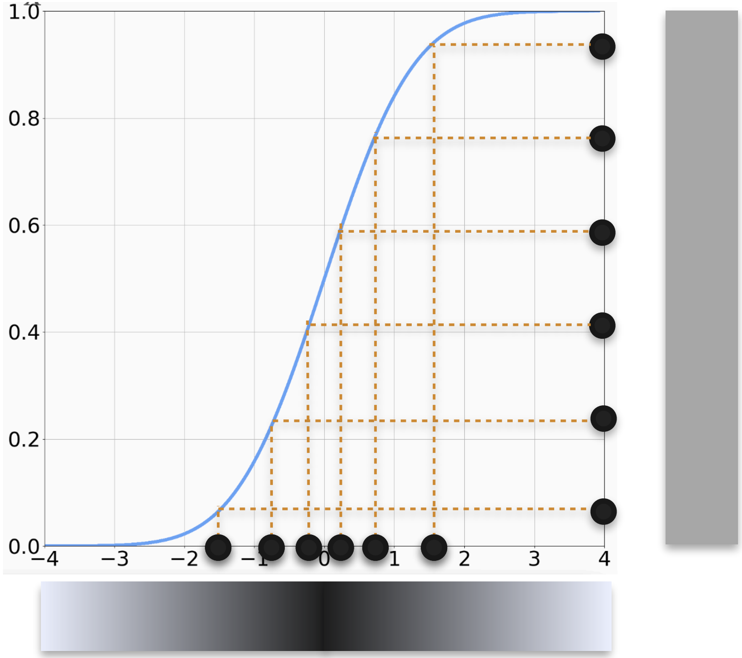
Sampling From a Distribution



DeepLearning.AI

Sampling From a Normal Distribution





DeepLearning.AI

Probability Distributions

Conclusion

Week 1 - Conclusions

Talking head