



DeepLearning.AI

Math for Machine Learning

Probability and Statistics



DeepLearning.AI

Math for Machine Learning

Probability and Statistics - Week 2

W2 Lesson 1



DeepLearning.AI

Describing Distributions

**Measures of
Central Tendency**

Mean: Example

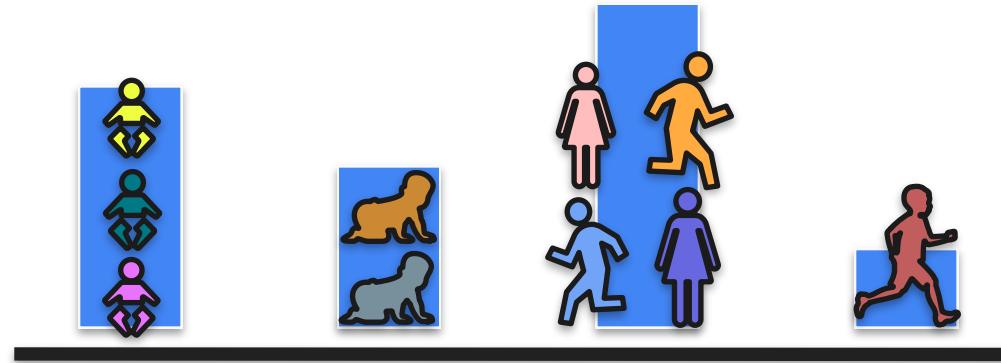
Age:

0

1

2

3



Mean: Example

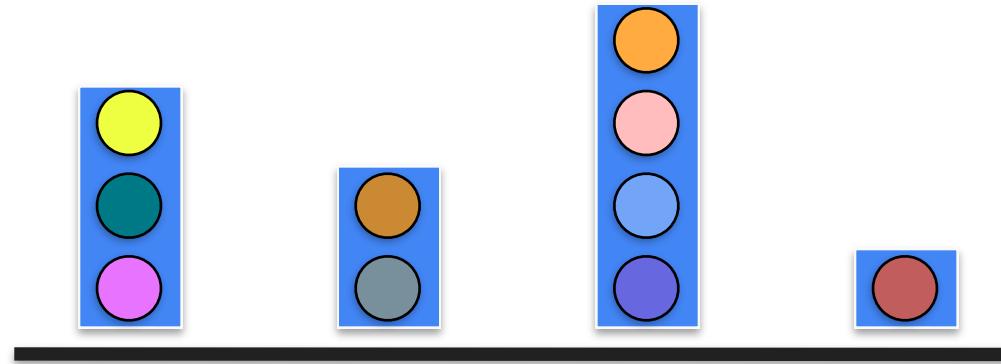
Age:

0

1

2

3



Mean: Example

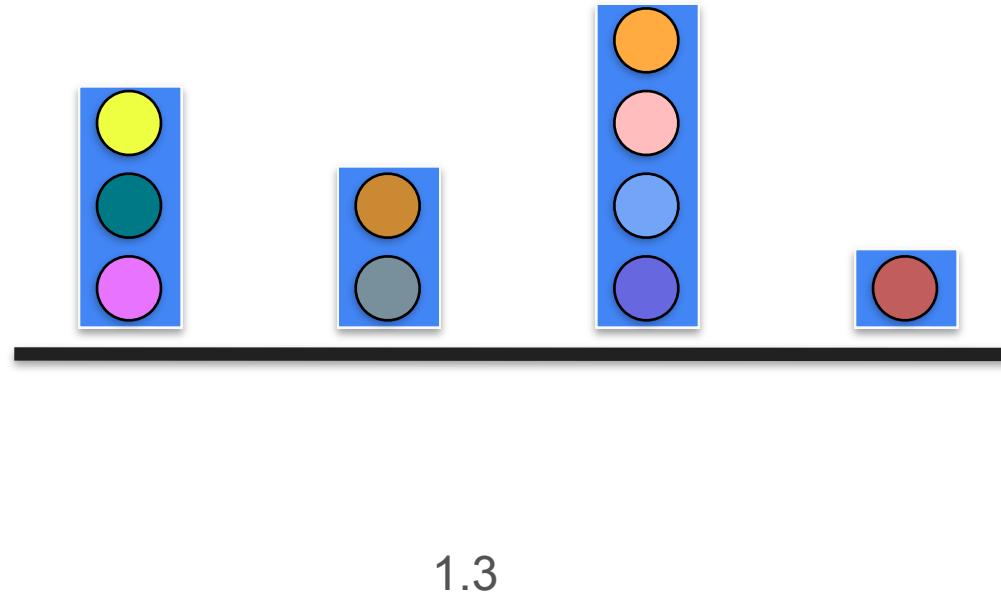
Age:

0

1

2

3



Mean: Example

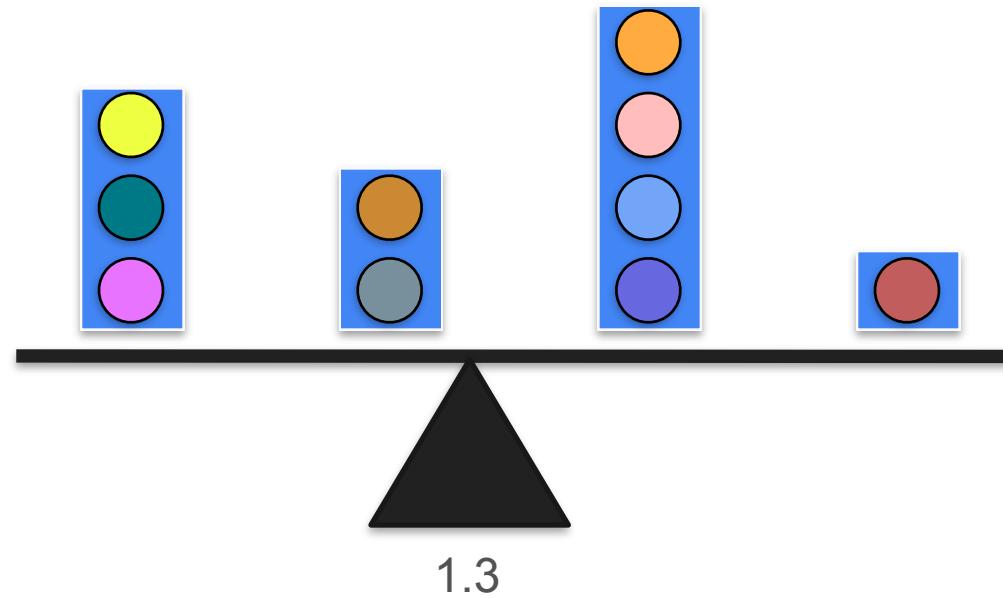
Age:

0

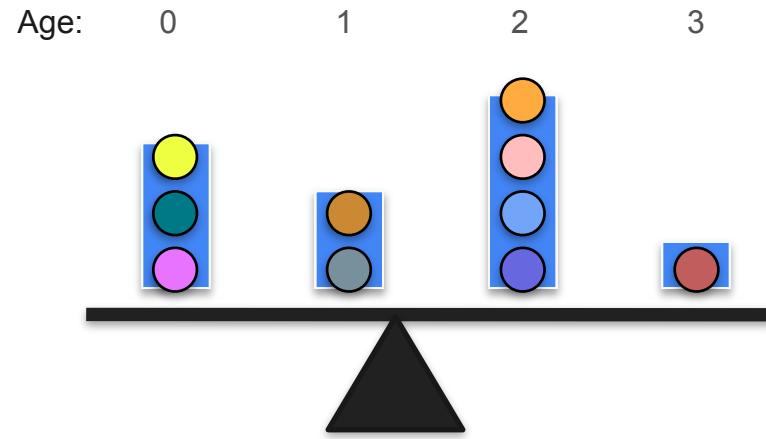
1

2

3

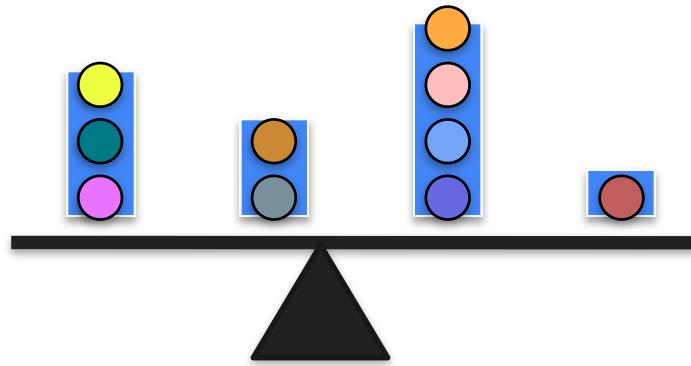


Mean: Example



Mean: Example

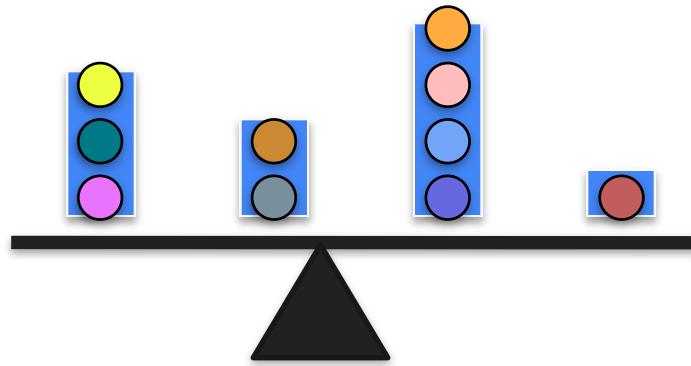
Age: 0 1 2 3 $0 + 0 + 0$



Mean: Example

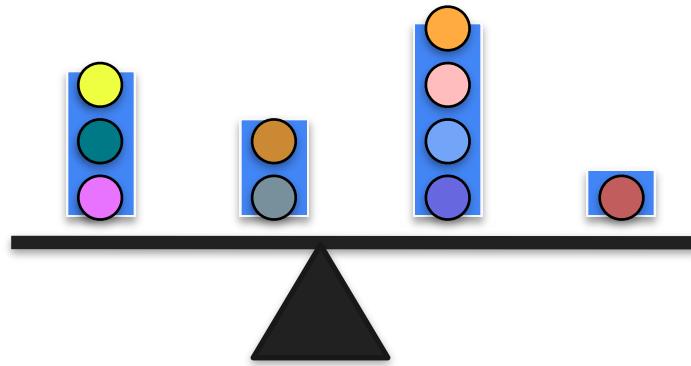
Age: 0 1 2 3

$0 + 0 + 0 + 1 + 1$



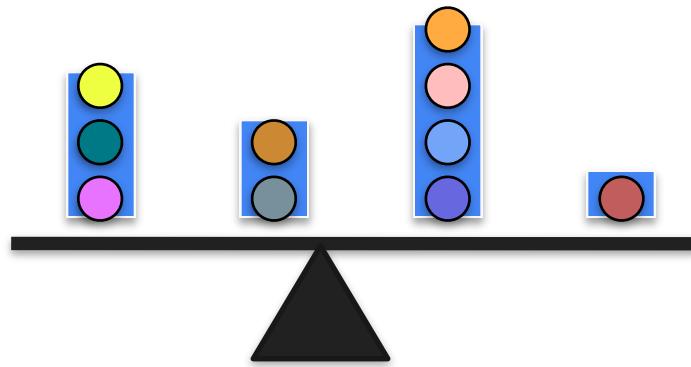
Mean: Example

Age: 0 1 2 3 $0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 2$



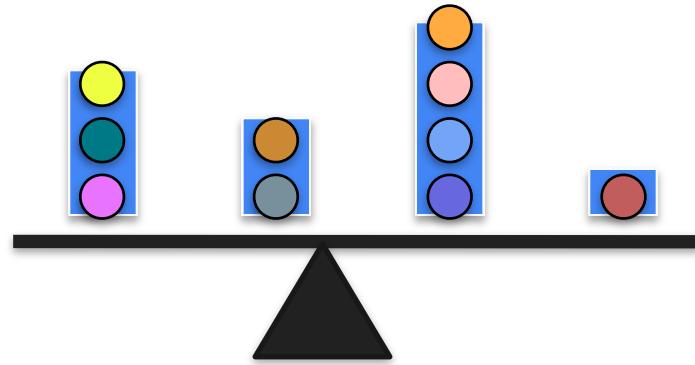
Mean: Example

Age: 0 1 2 3 $0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 2 + 3$



Mean: Example

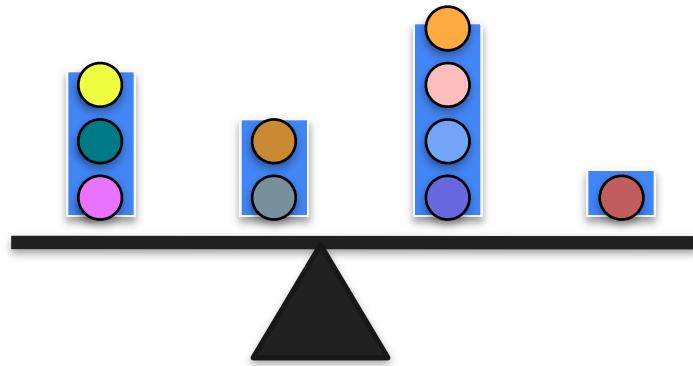
Age: 0 1 2 3



$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

Mean: Example

Age: 0 1 2 3

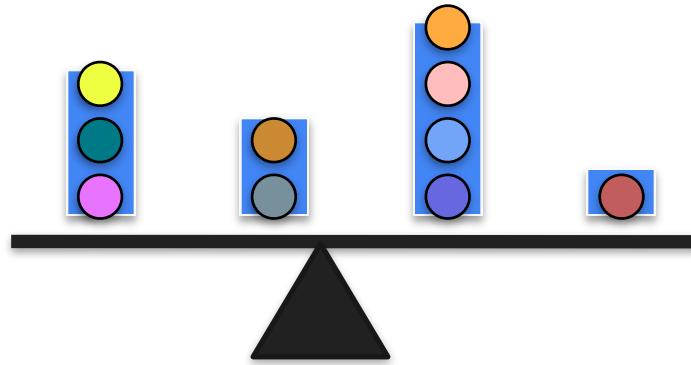


$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

$$= \frac{13}{10}$$

Mean: Example

Age: 0 1 2 3



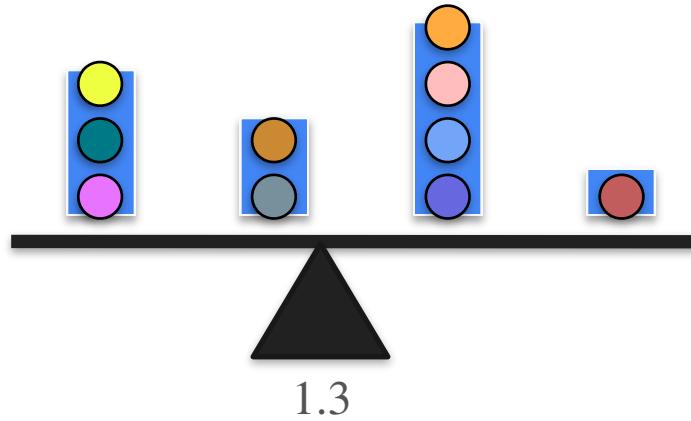
$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

$$= \frac{13}{10}$$

$$= 1.3$$

Mean: Example

Age: 0 1 2 3

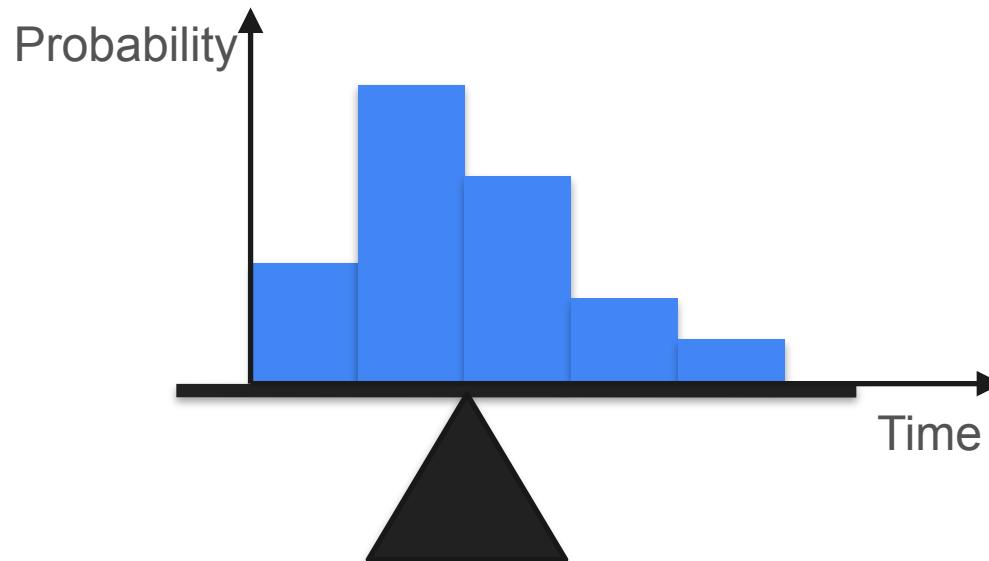
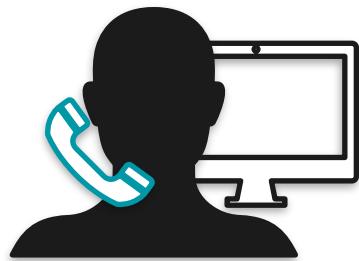


$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

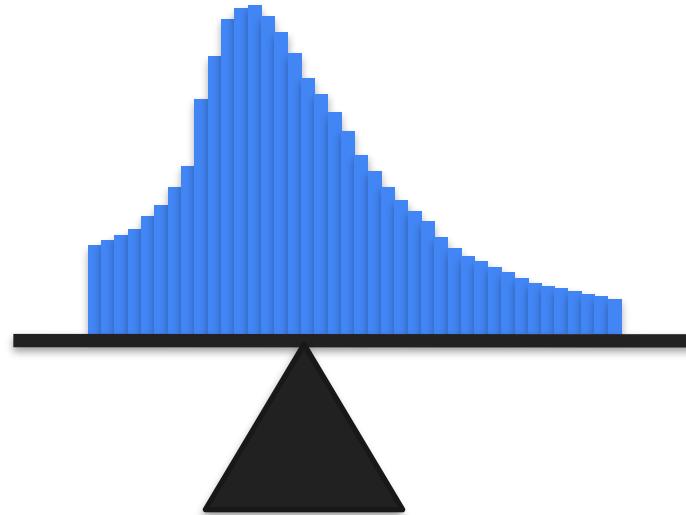
$$= \frac{13}{10}$$

$$= 1.3$$

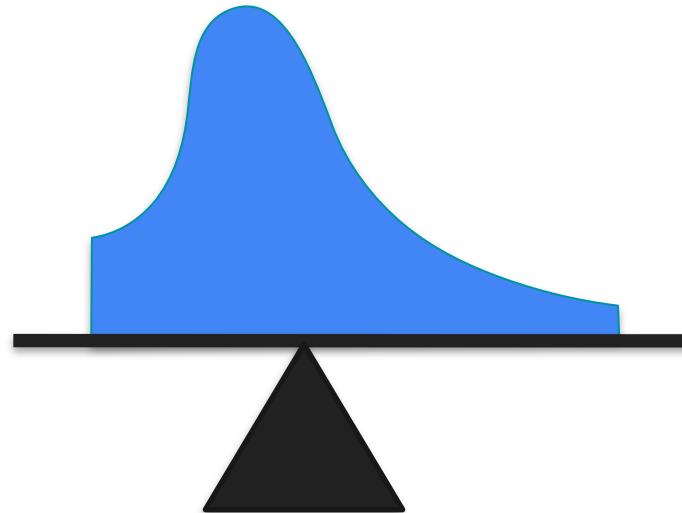
Mean



Mean



Mean



Median: Motivation

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money



Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
 - For the rest of the country, the starting salary for a geography graduate was \$22,000.
 - Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money
- 



Michael Jordan

Outliers

Outliers

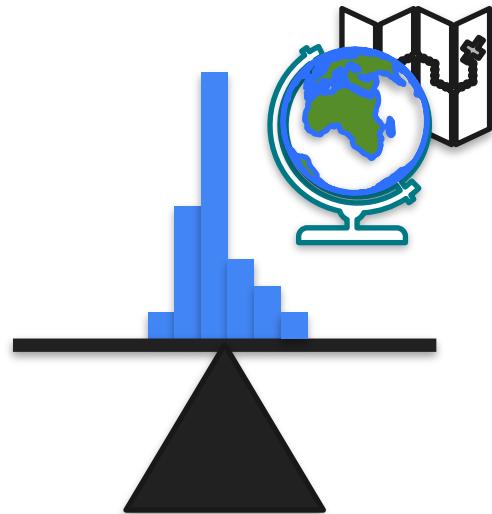


Outliers



Outliers

Graduates



Salary

Outliers

Graduates



Outliers

Graduates



Outliers

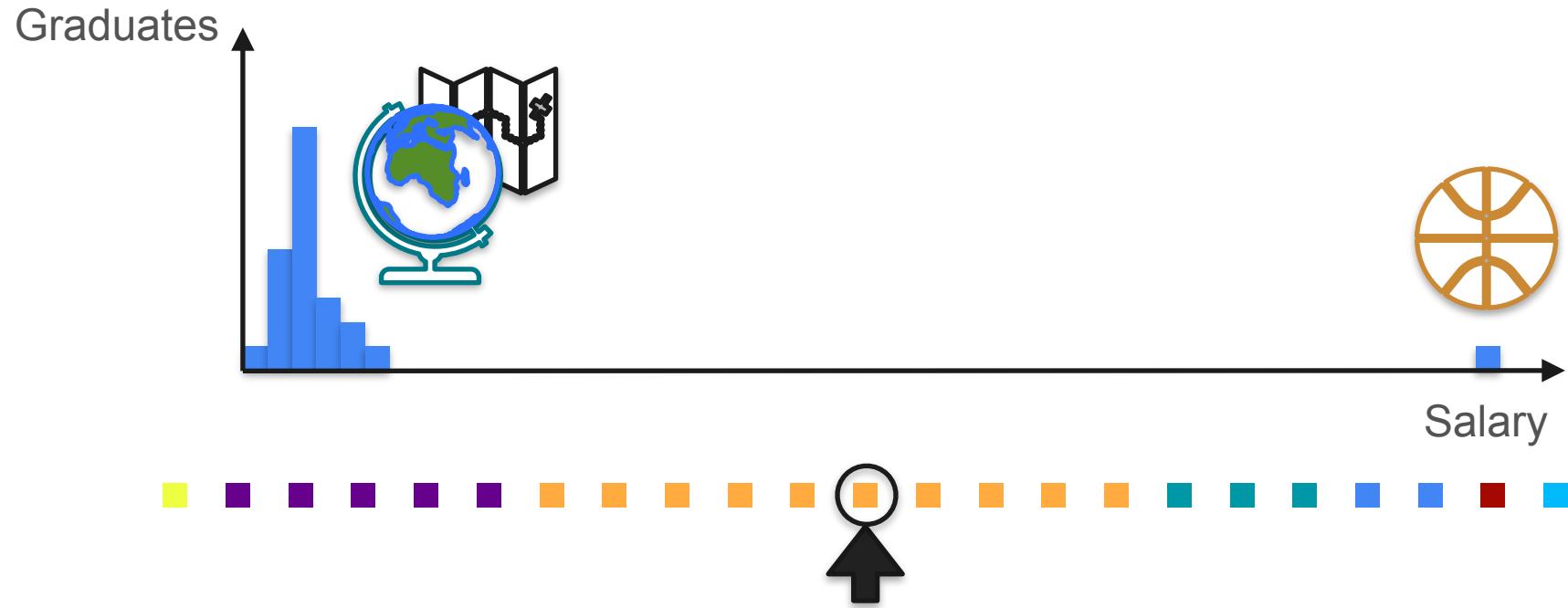


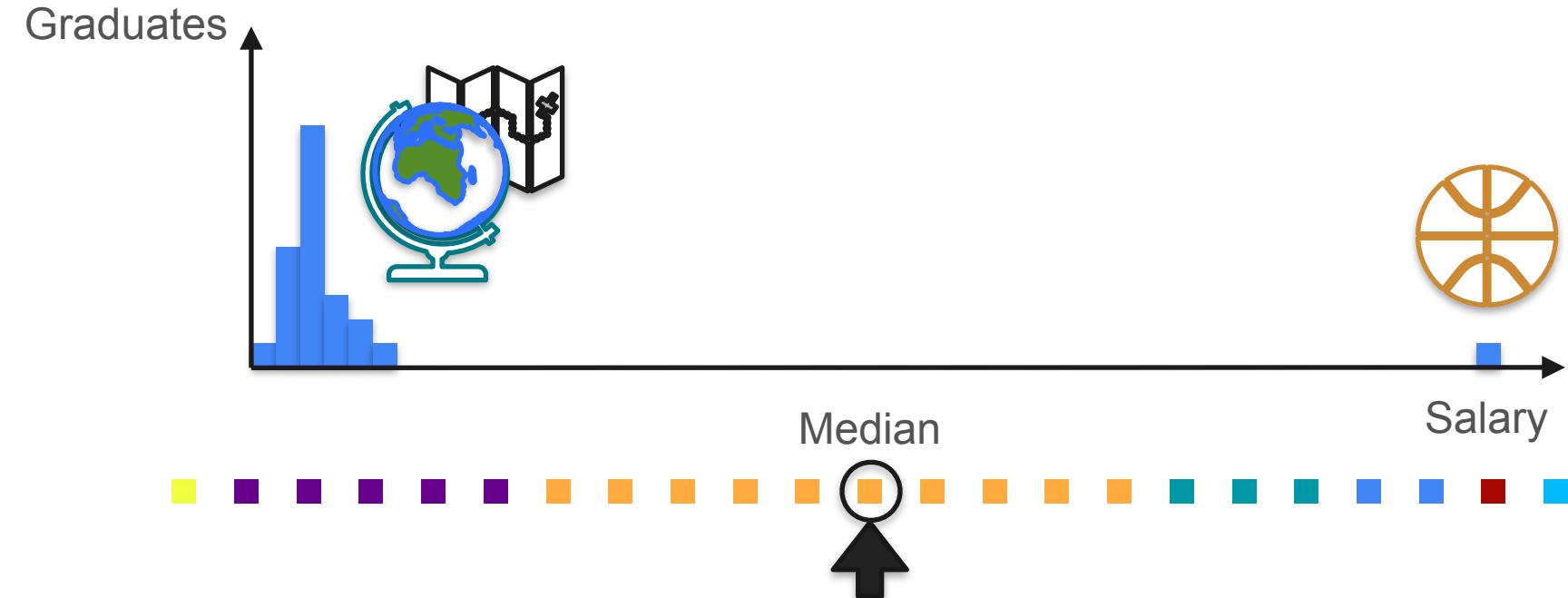
Median

Graduates



Median





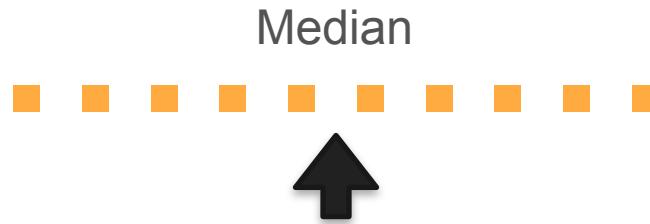
Median



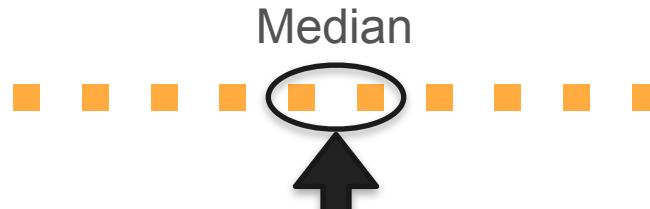
Median



Median

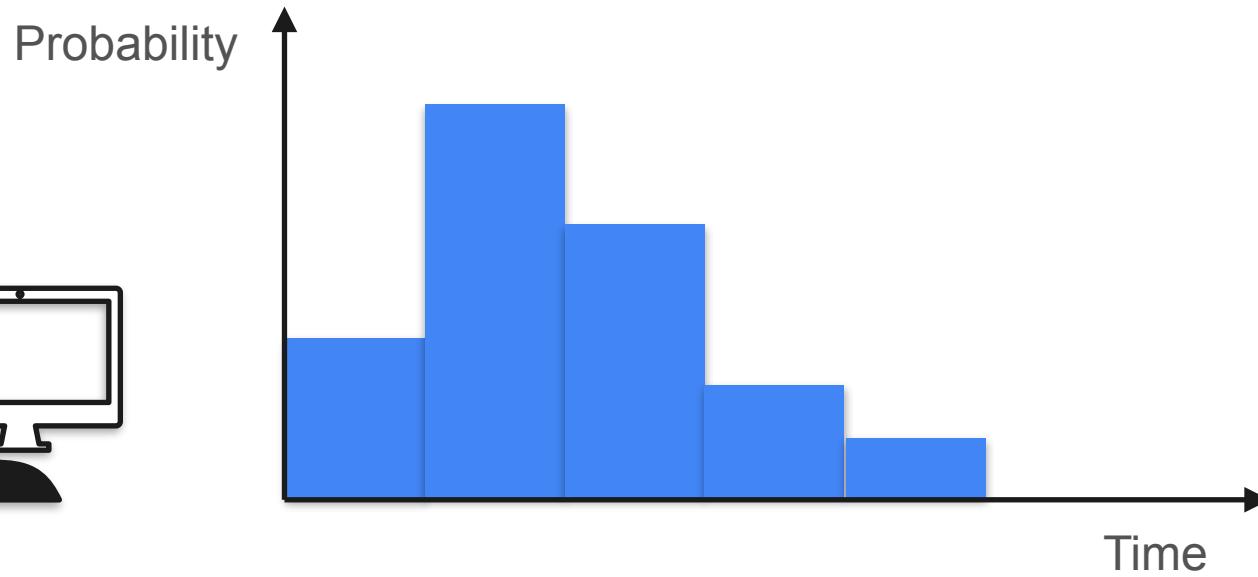


Median

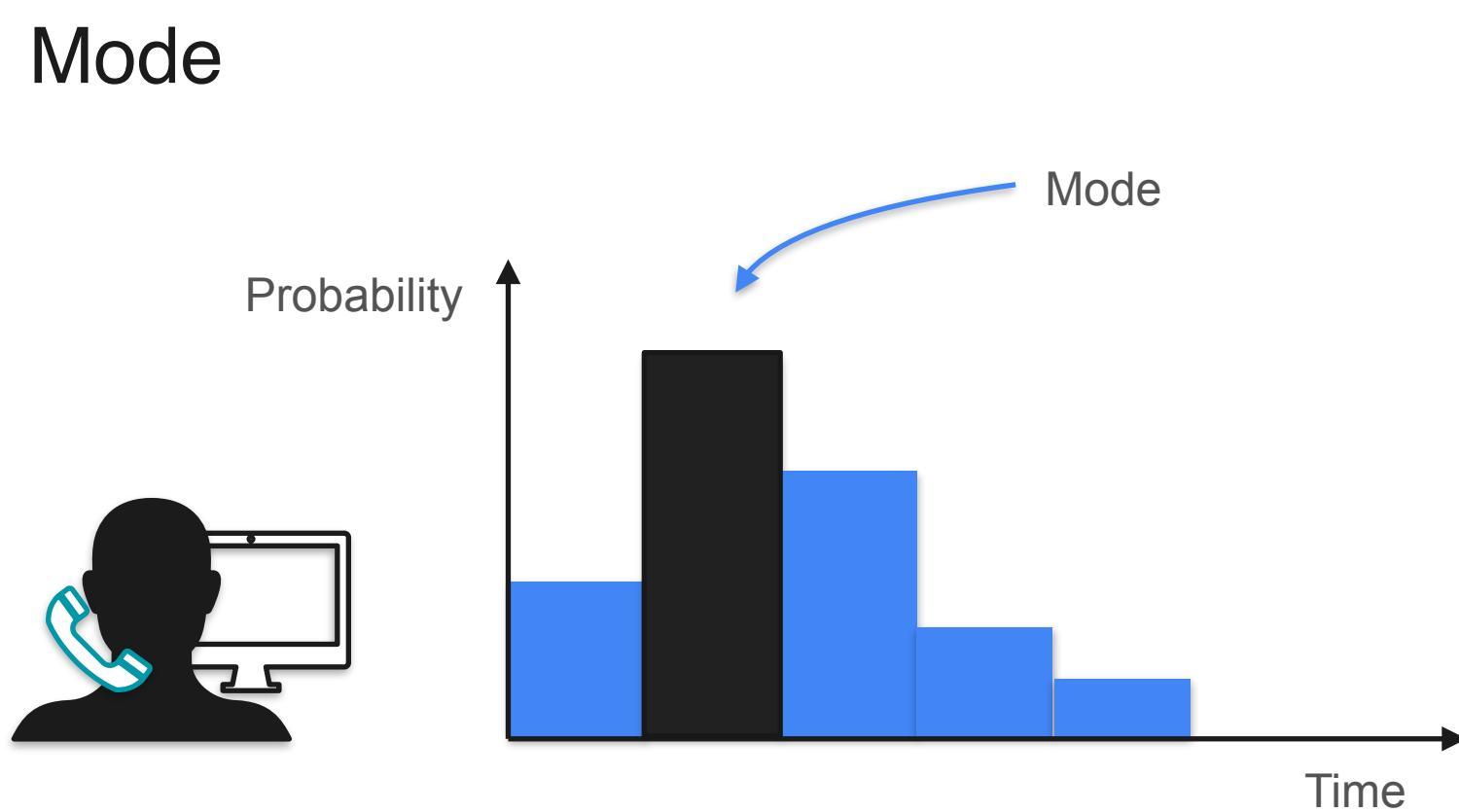


Median
Average of the
two middle ones

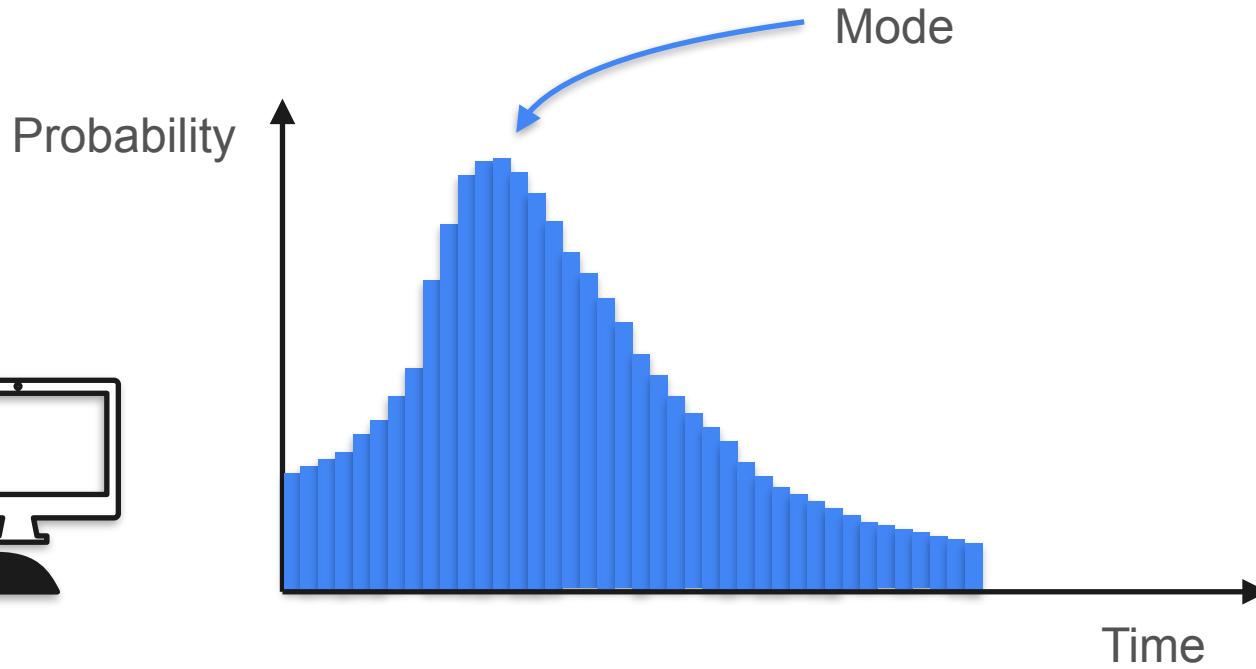
Mode



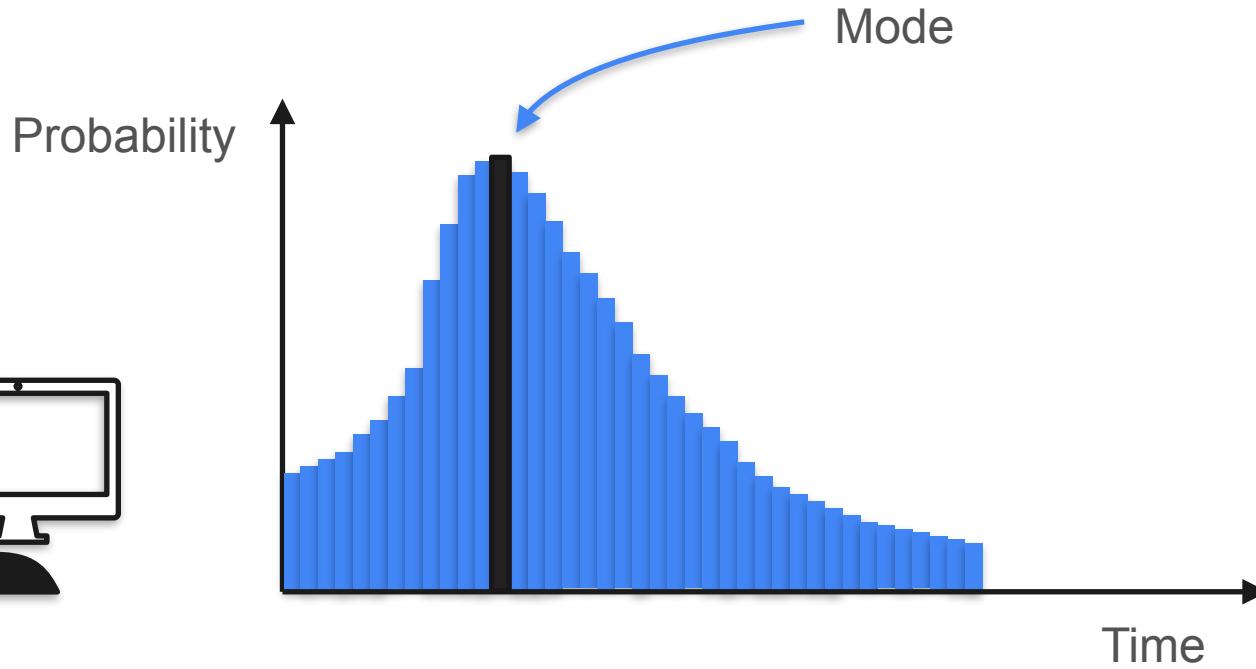
Mode



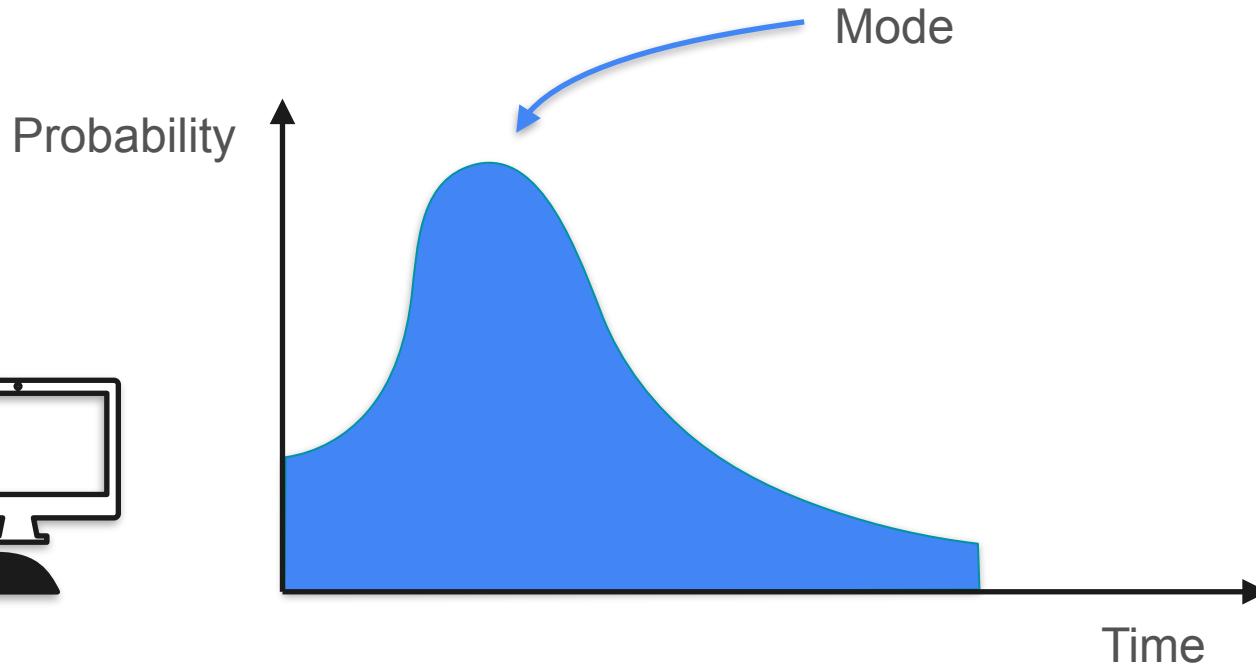
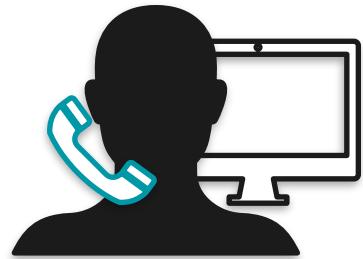
Mode



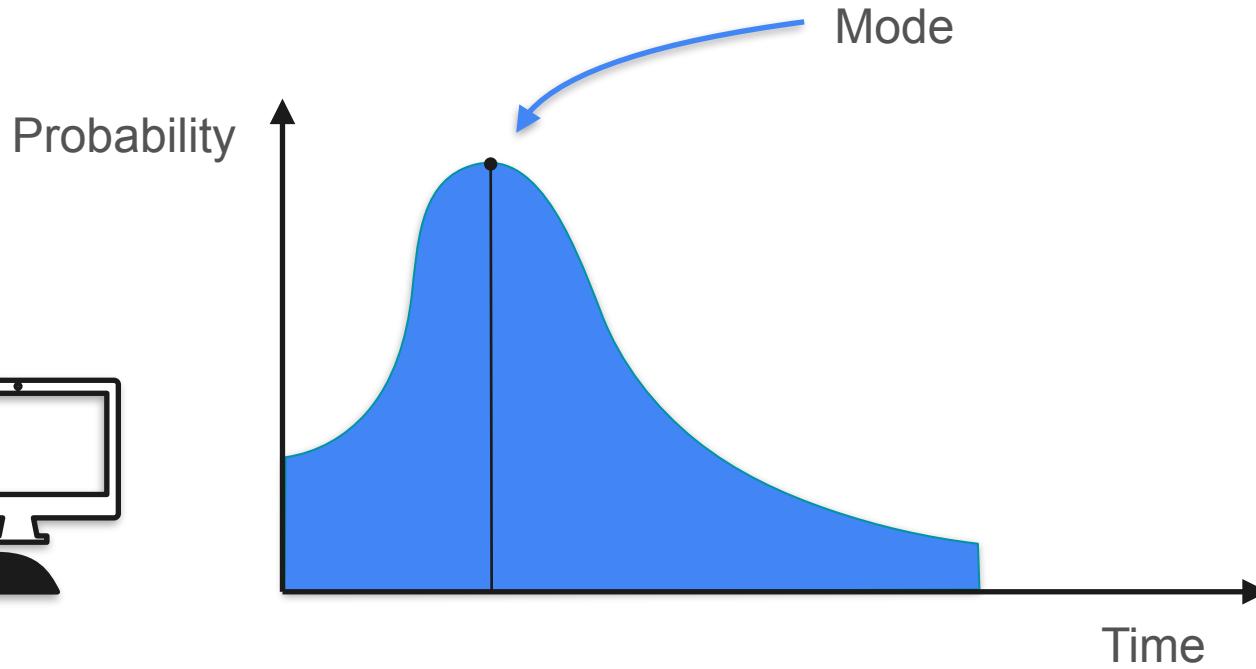
Mode



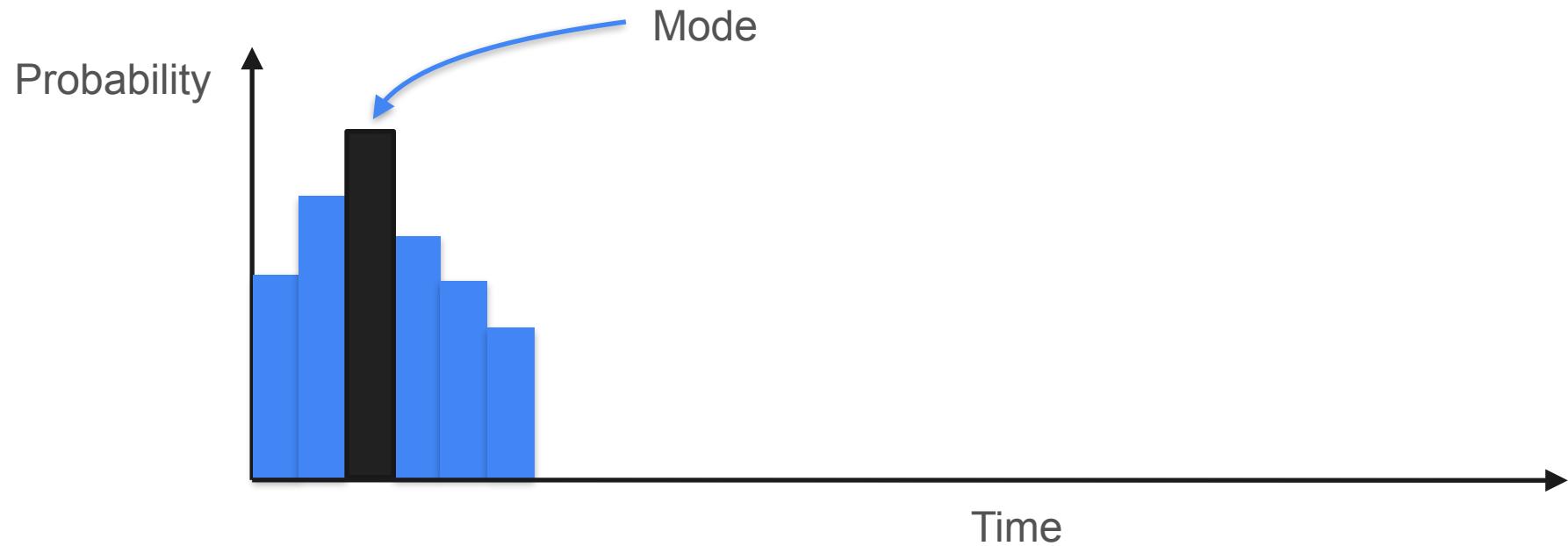
Mode



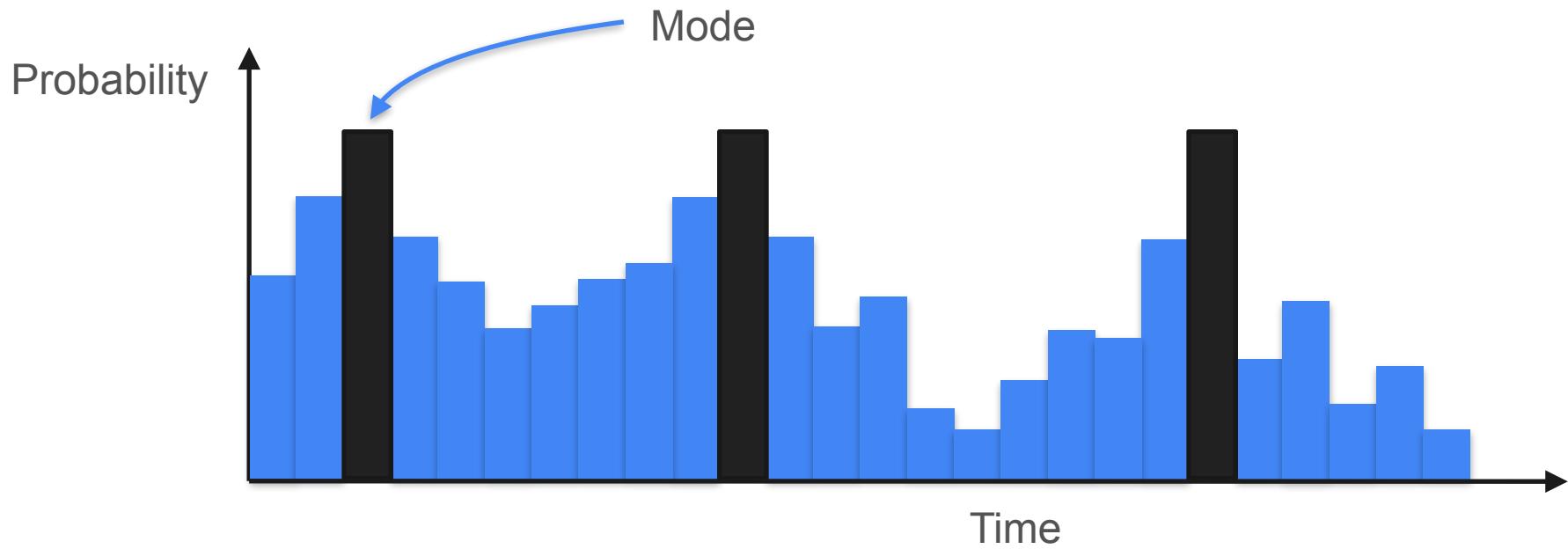
Mode



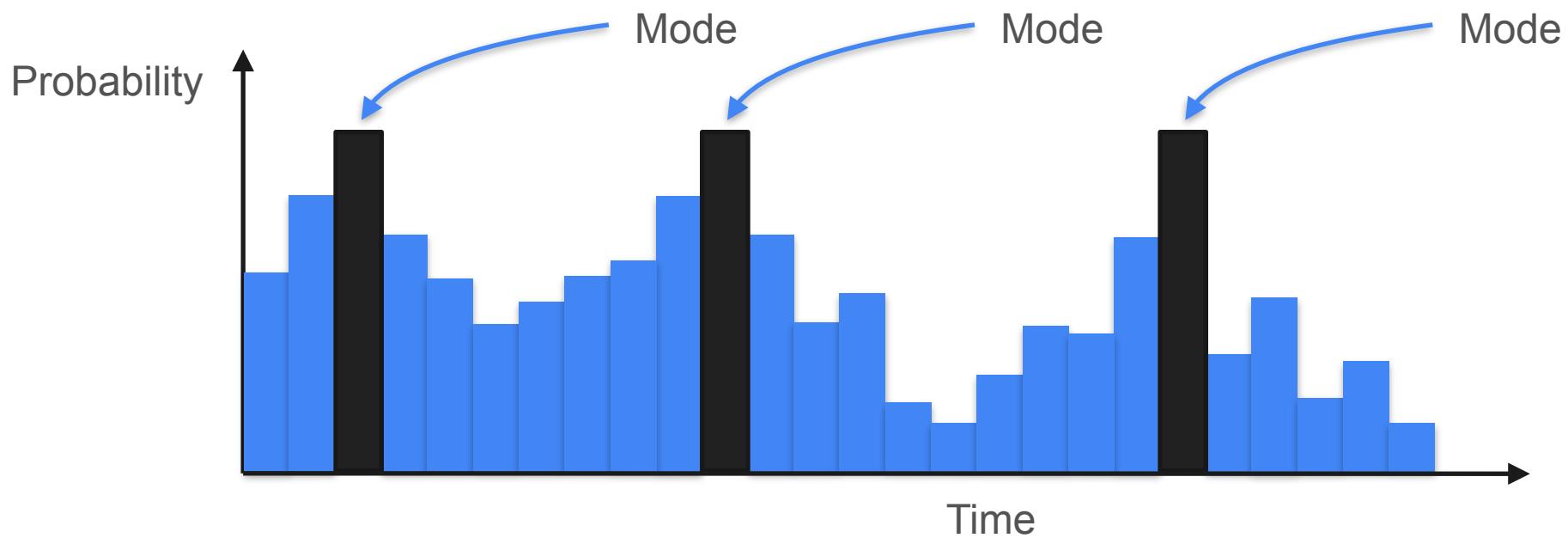
Mode: Multimodal Distribution



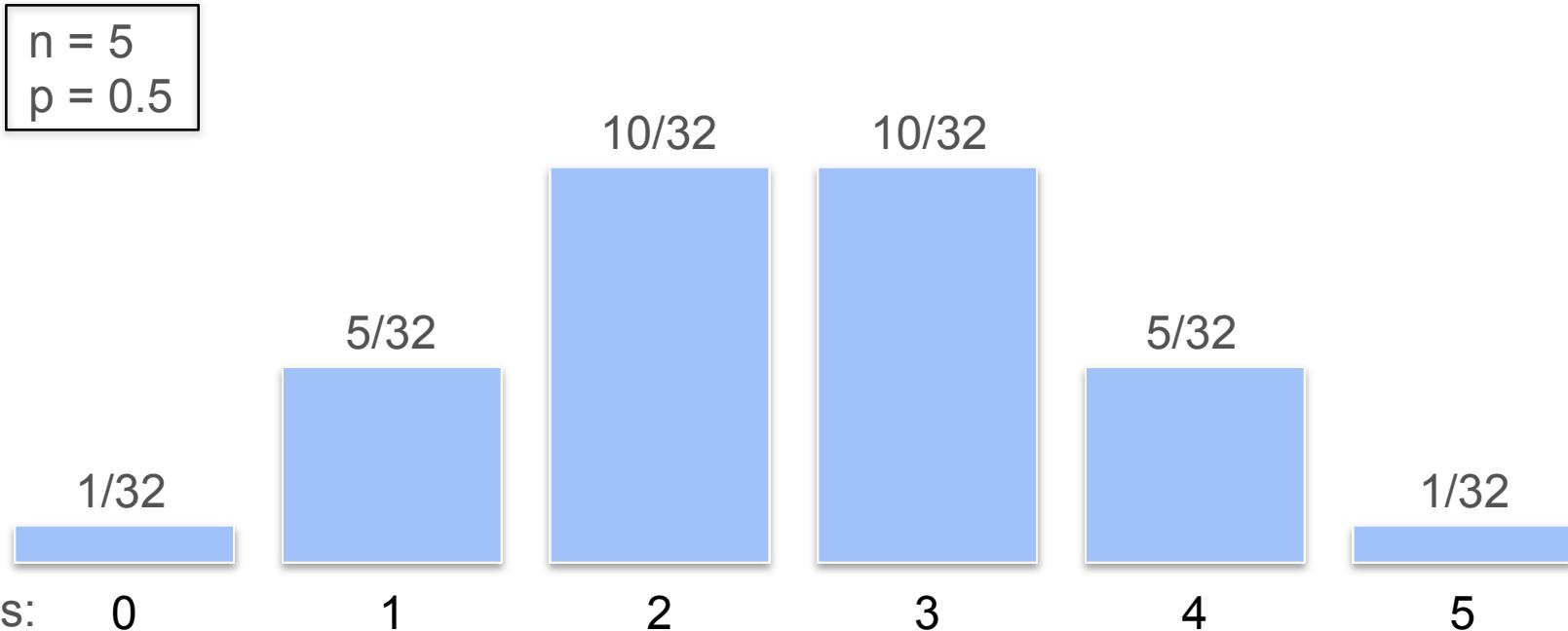
Mode: Multimodal Distribution



Mode: Multimodal Distribution

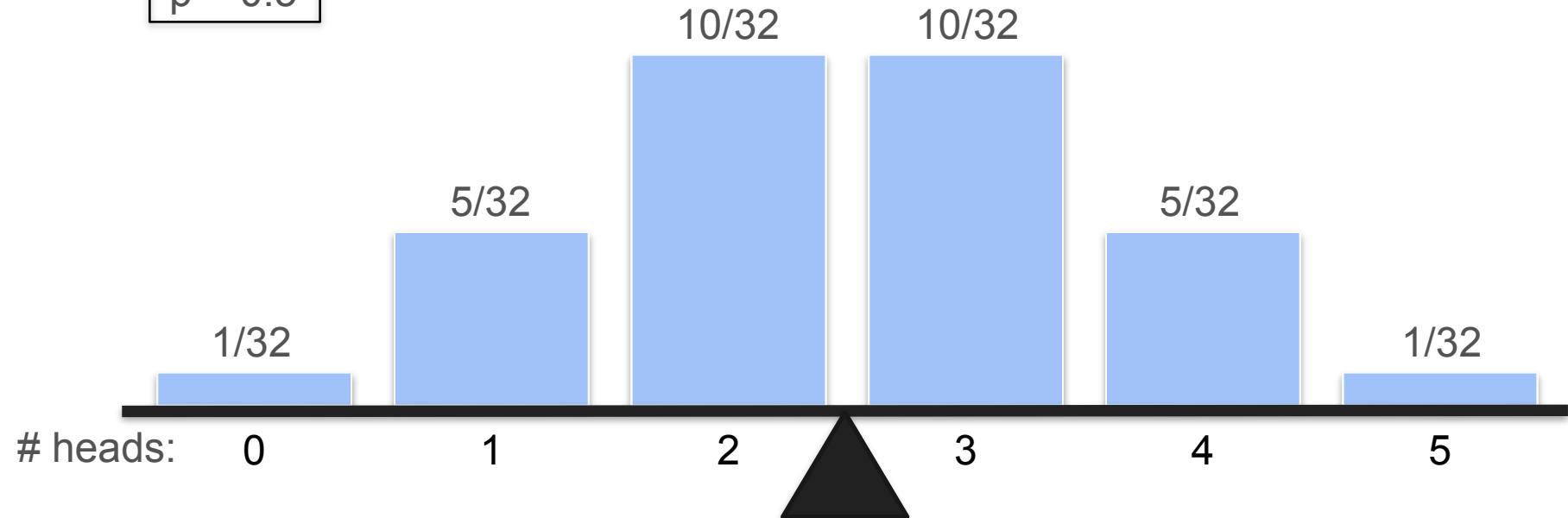


Mean, Median and Mode in Binomial Distribution

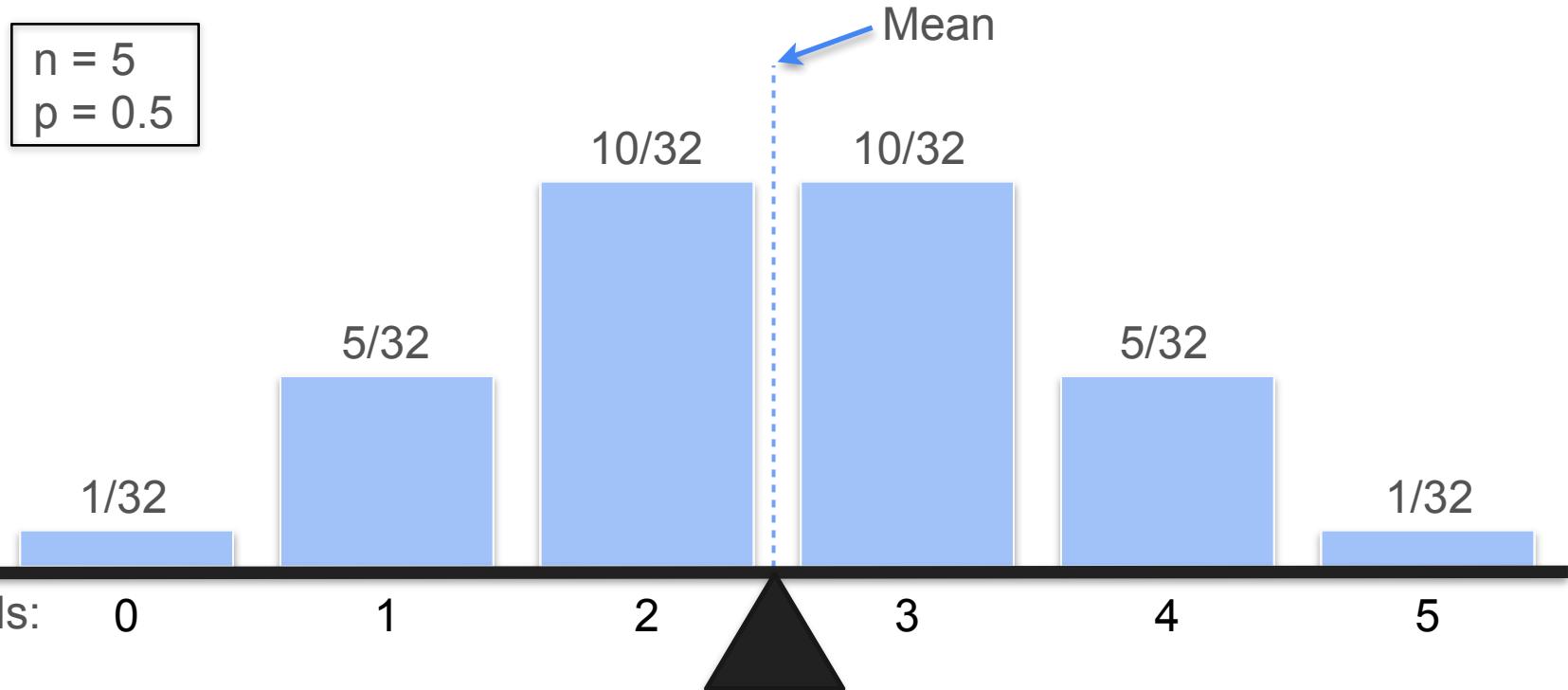


Mean, Median and Mode in Binomial Distribution

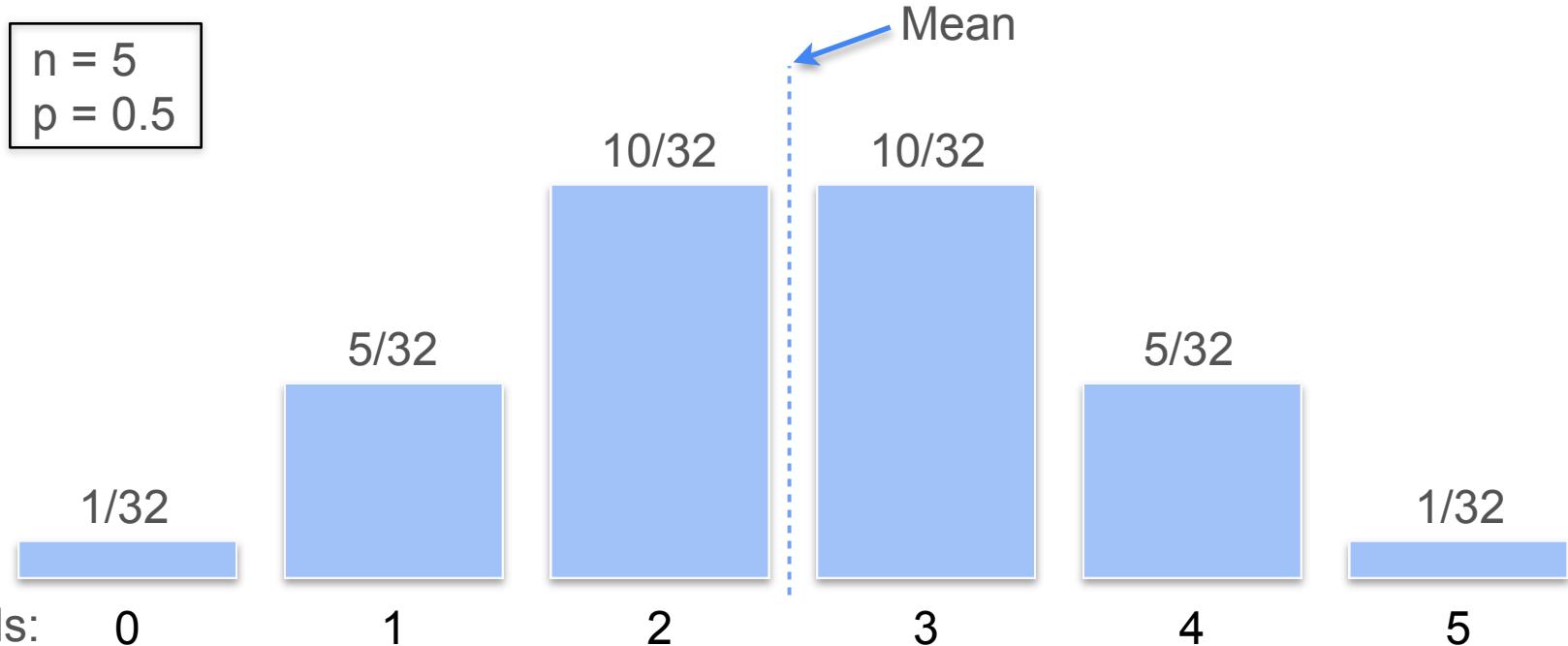
$n = 5$
 $p = 0.5$



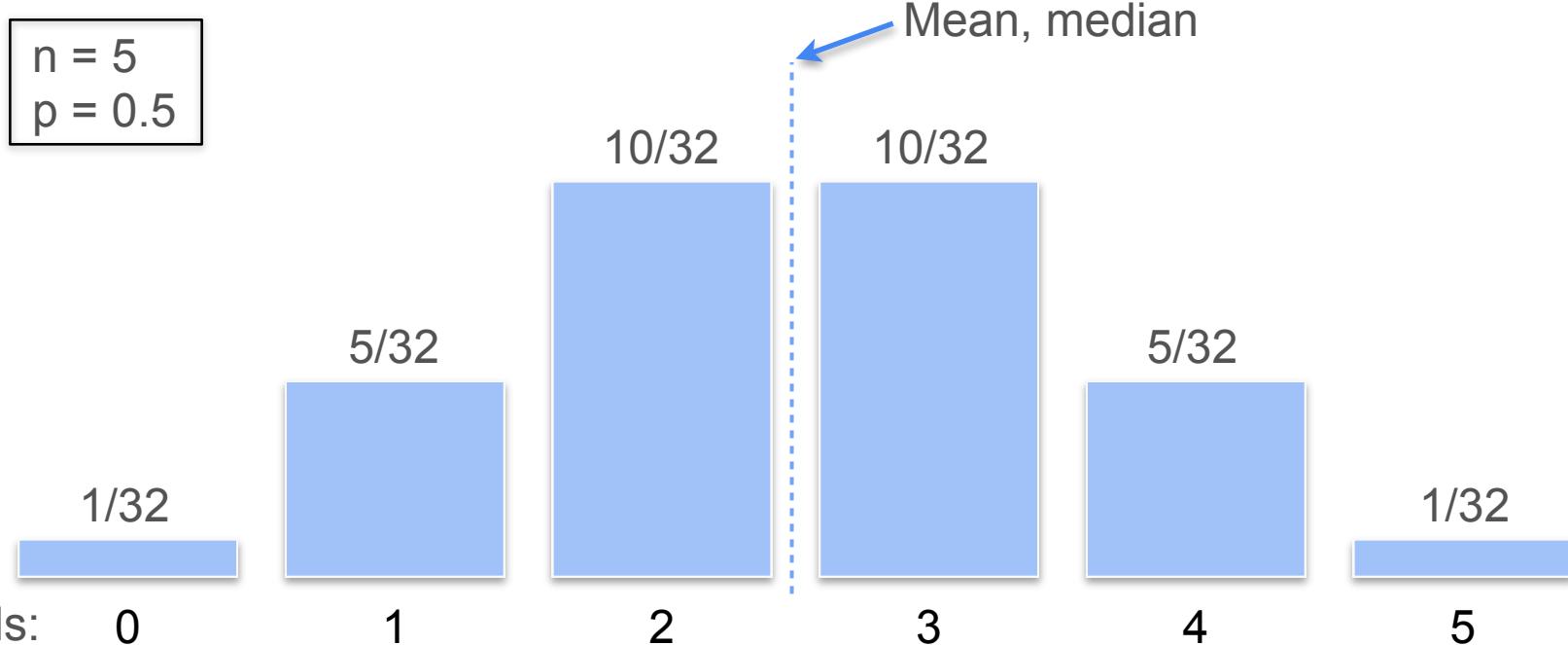
Mean, Median and Mode in Binomial Distribution



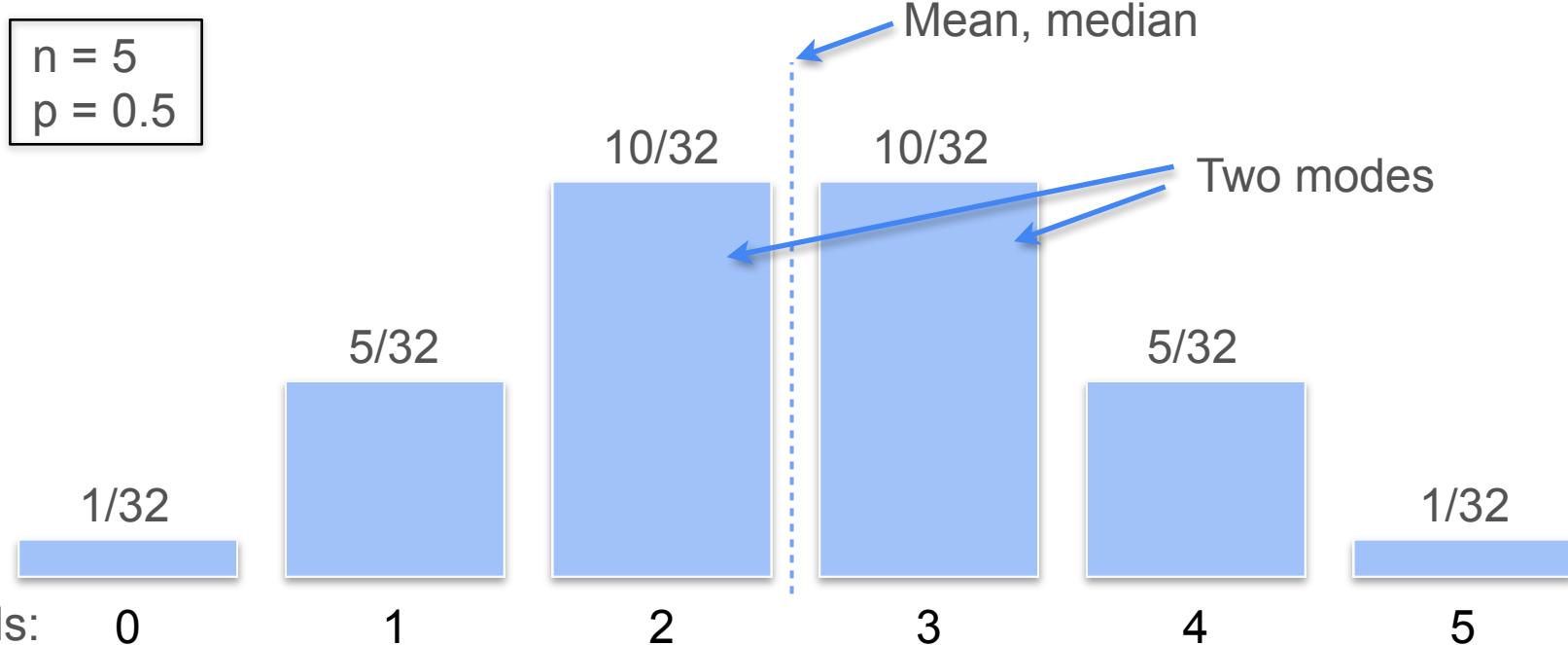
Mean, Median and Mode in Binomial Distribution



Mean, Median and Mode in Binomial Distribution

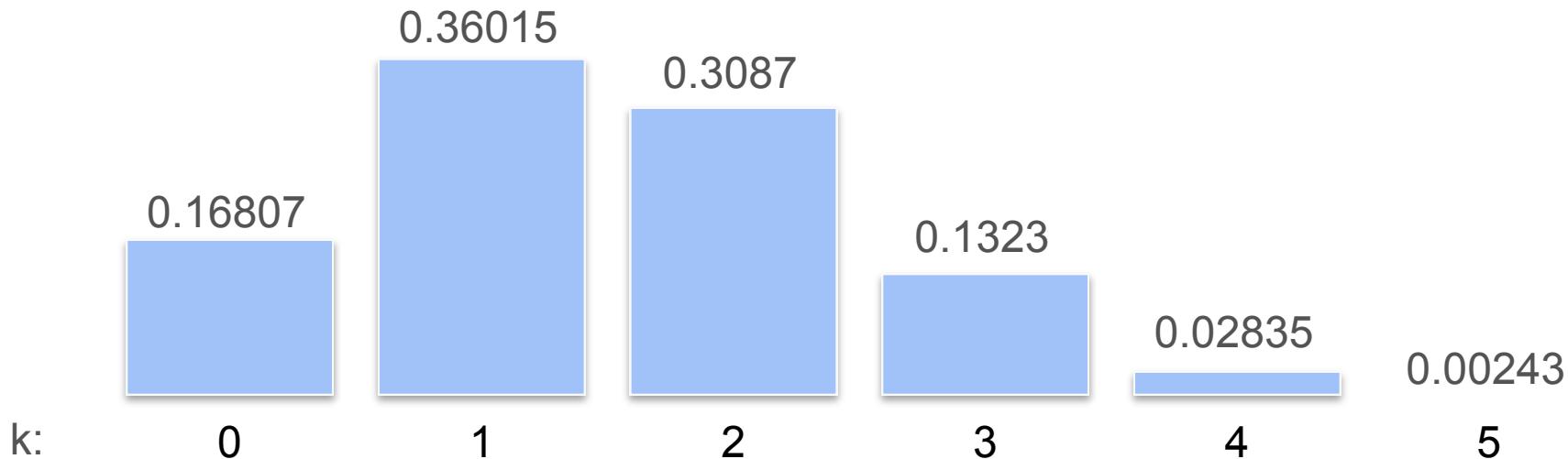


Mean, Median and Mode in Binomial Distribution



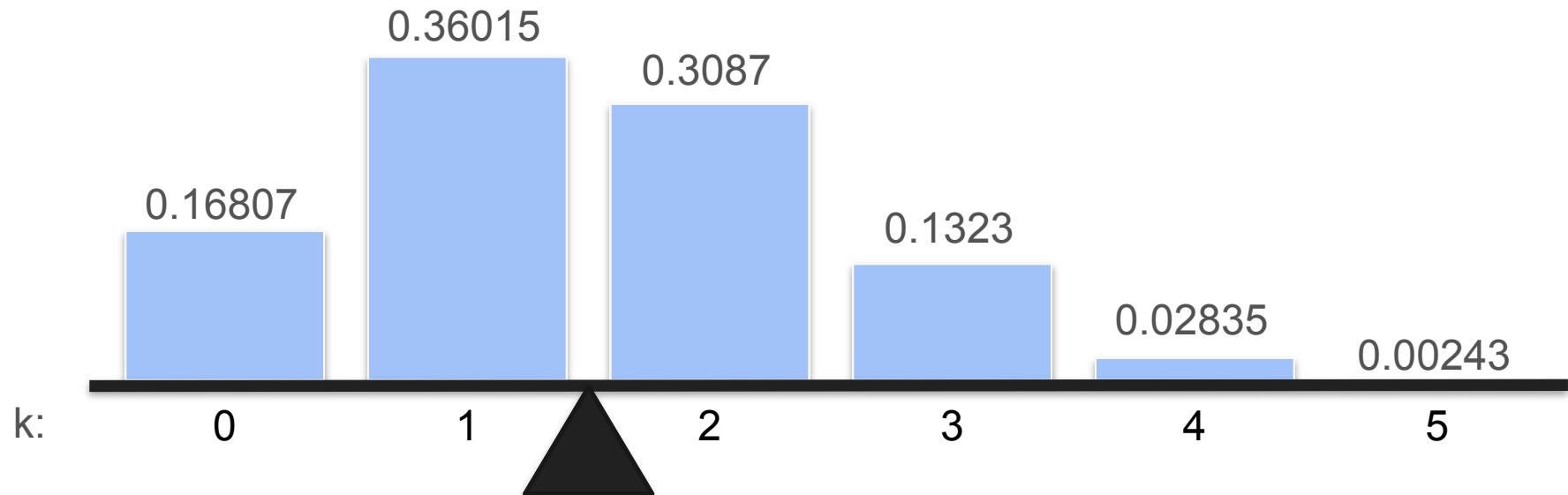
Mean, Median and Mode in Binomial Distribution

$n = 5$
 $p = 0.3$

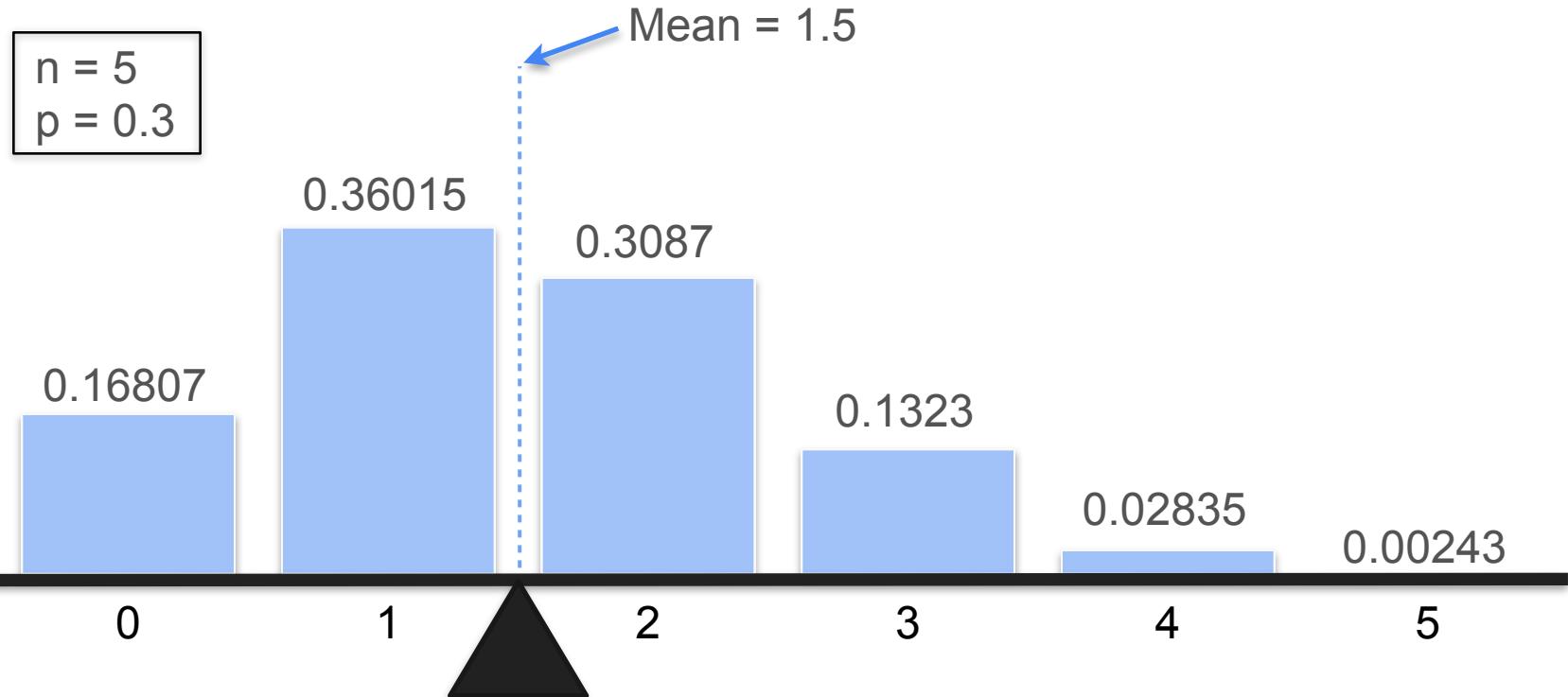


Mean, Median and Mode in Binomial Distribution

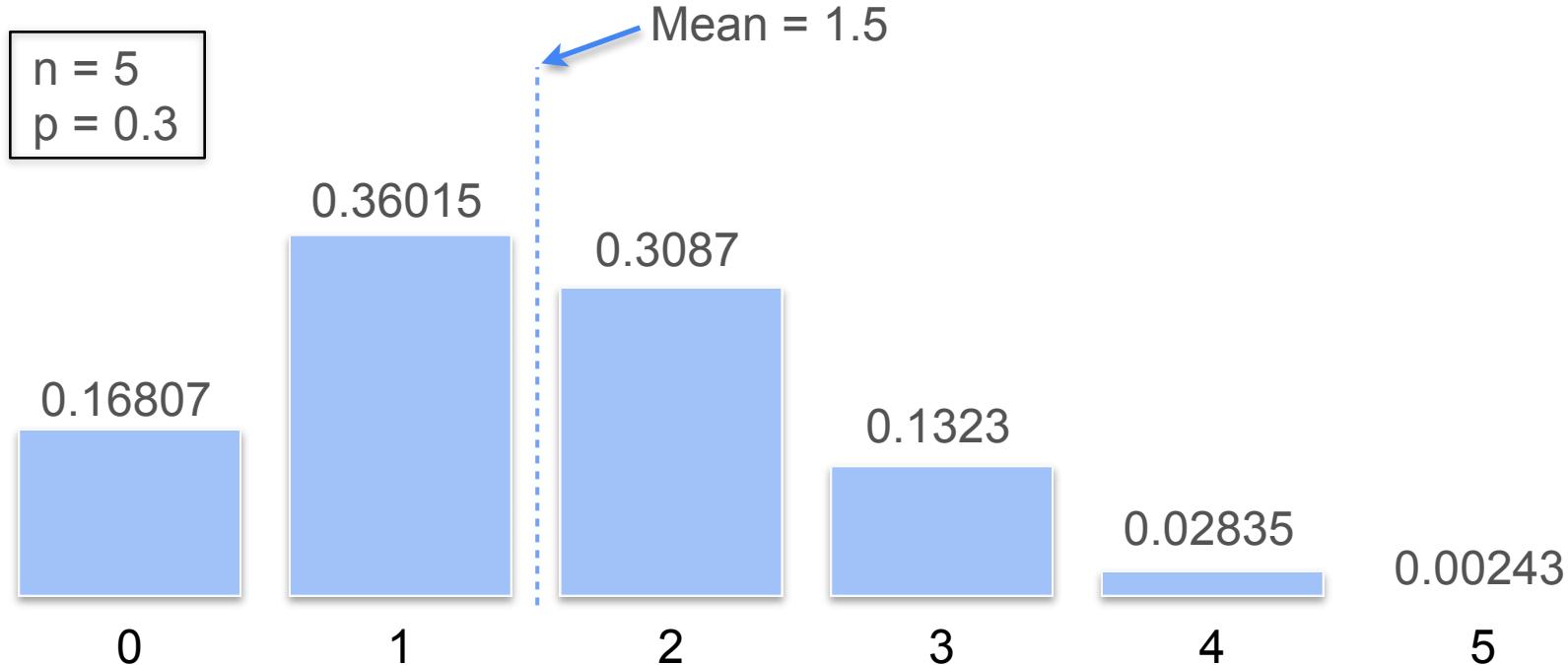
$n = 5$
 $p = 0.3$



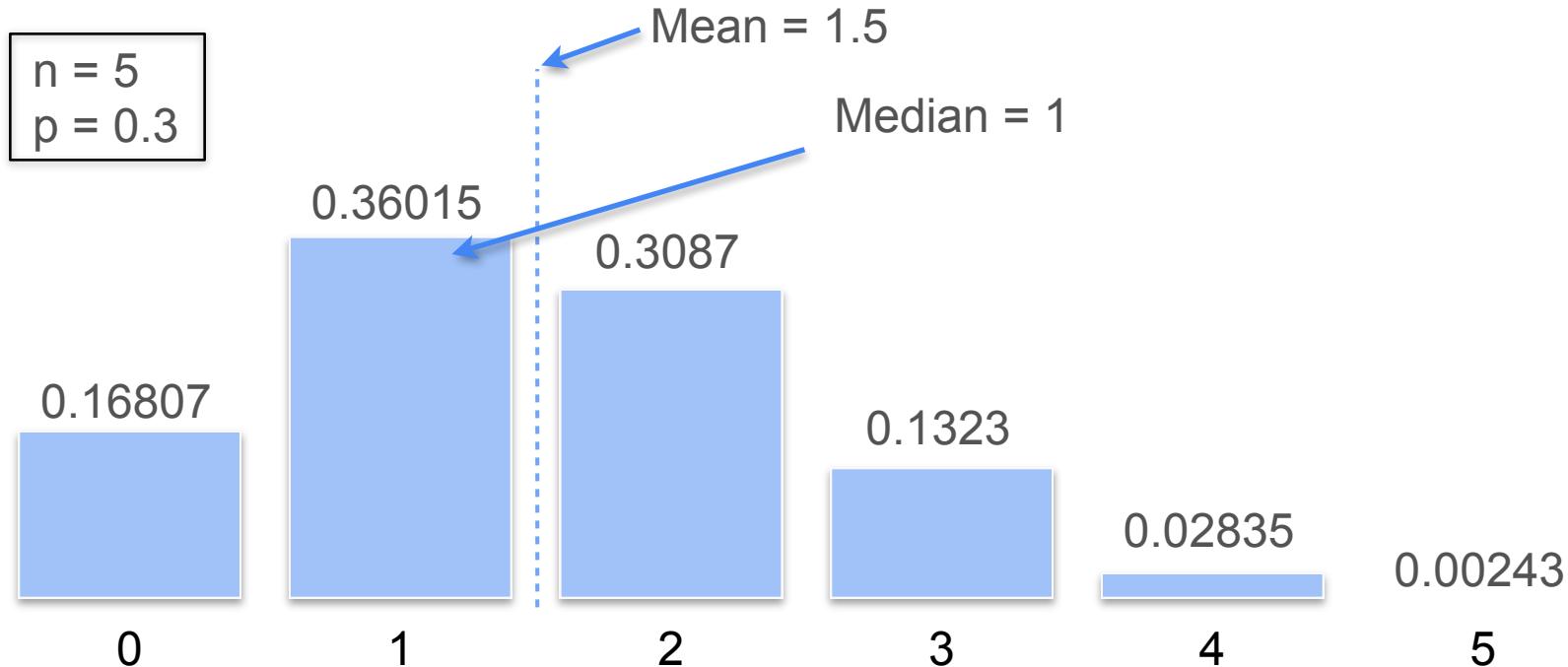
Mean, Median and Mode in Binomial Distribution



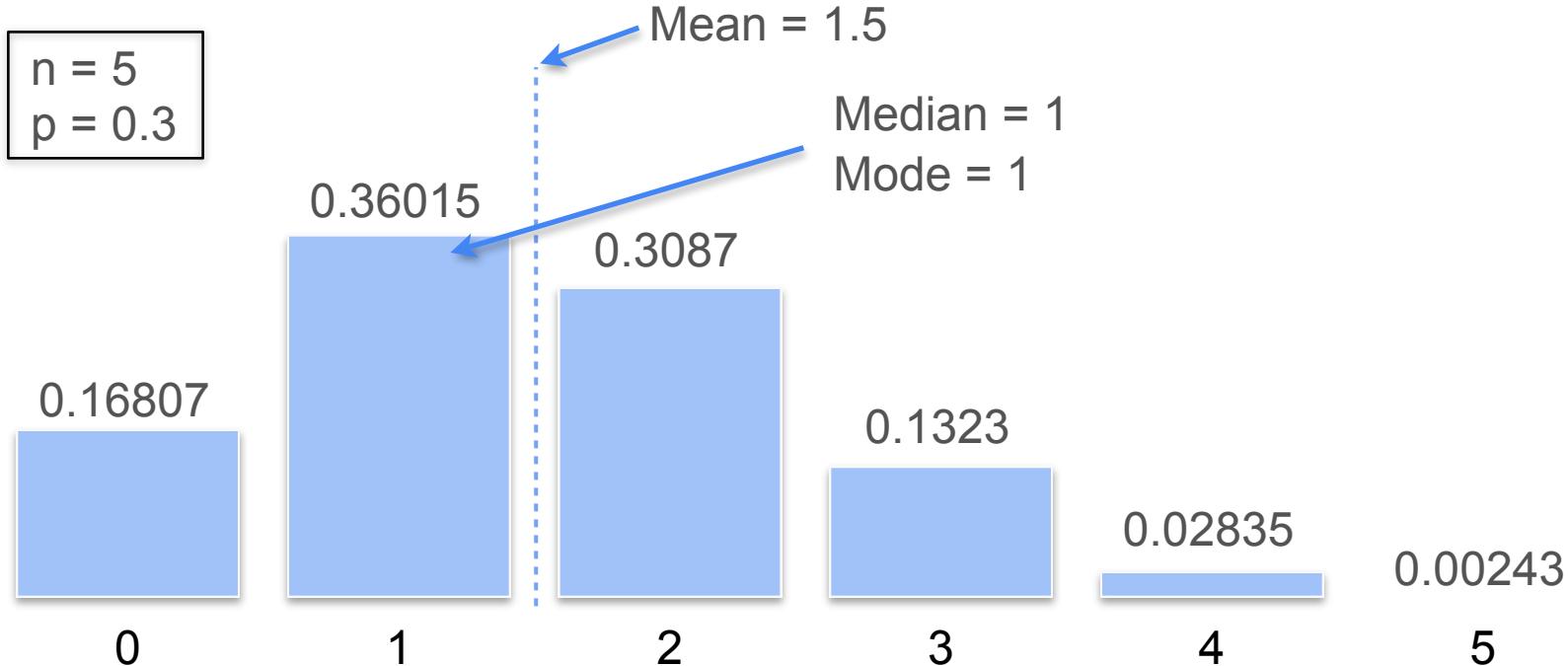
Mean, Median and Mode in Binomial Distribution



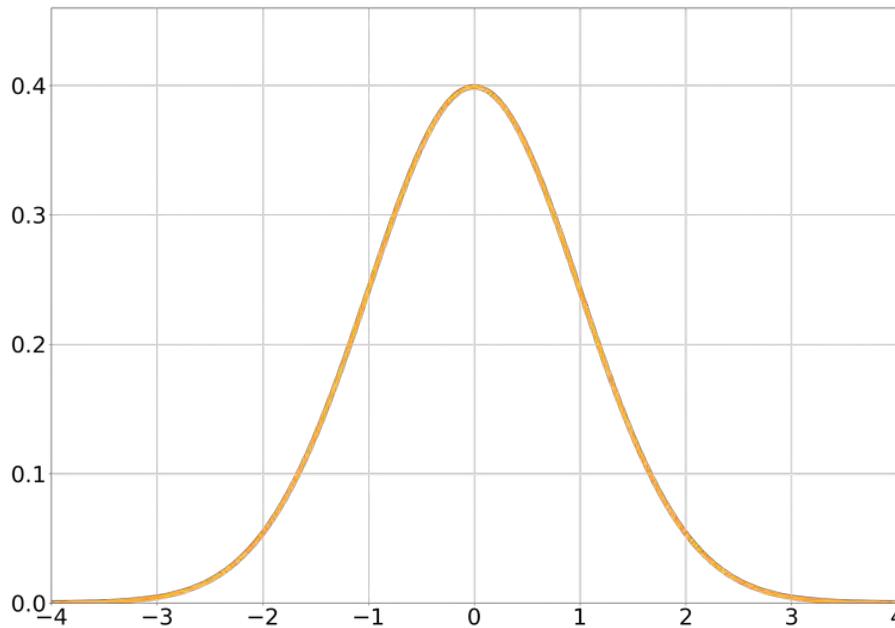
Mean, Median and Mode in Binomial Distribution



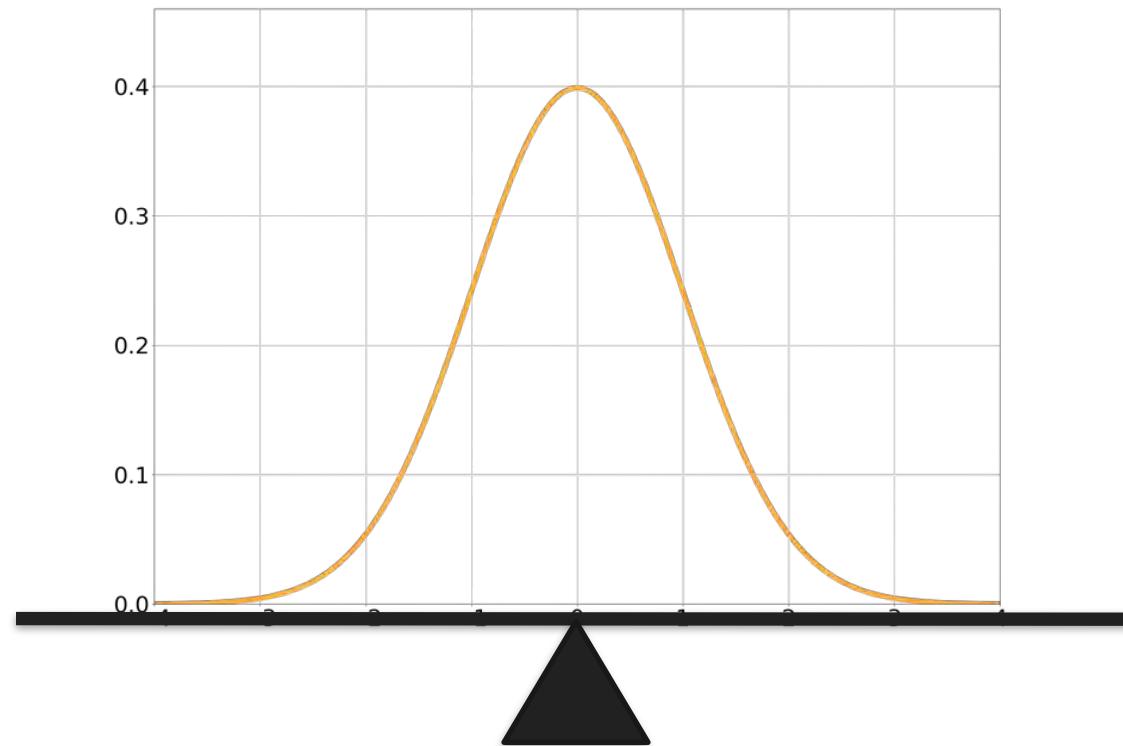
Mean, Median and Mode in Binomial Distribution



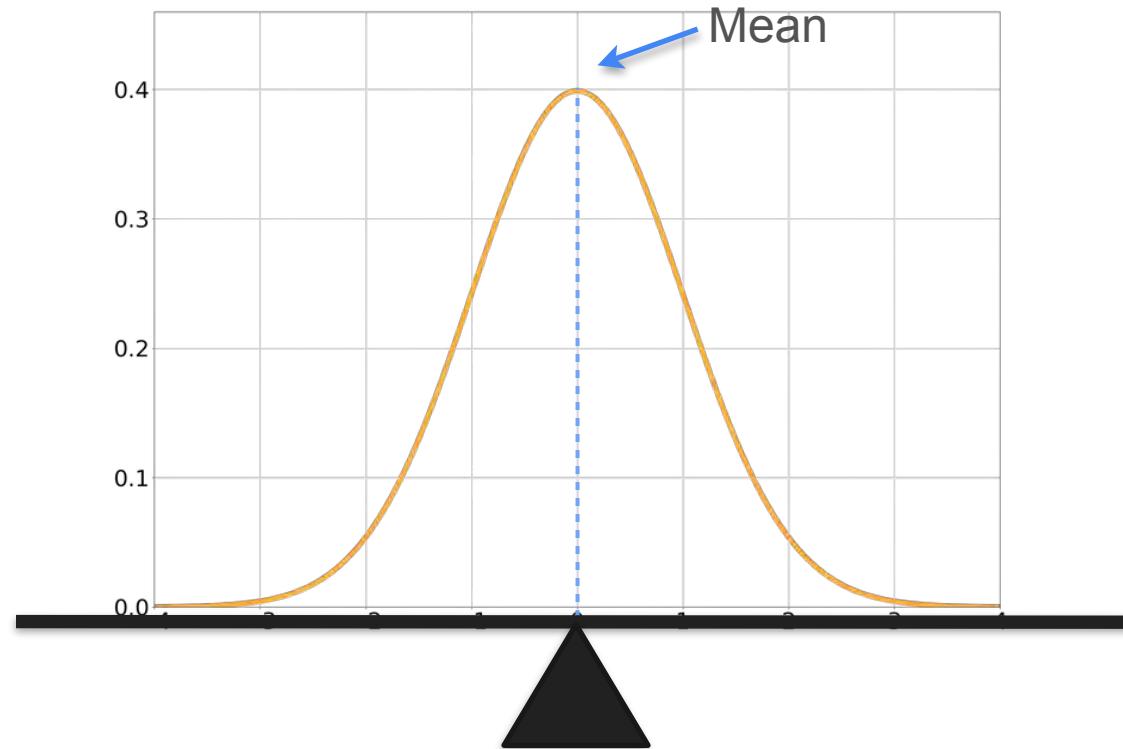
Mean, Median and Mode in Normal Distribution



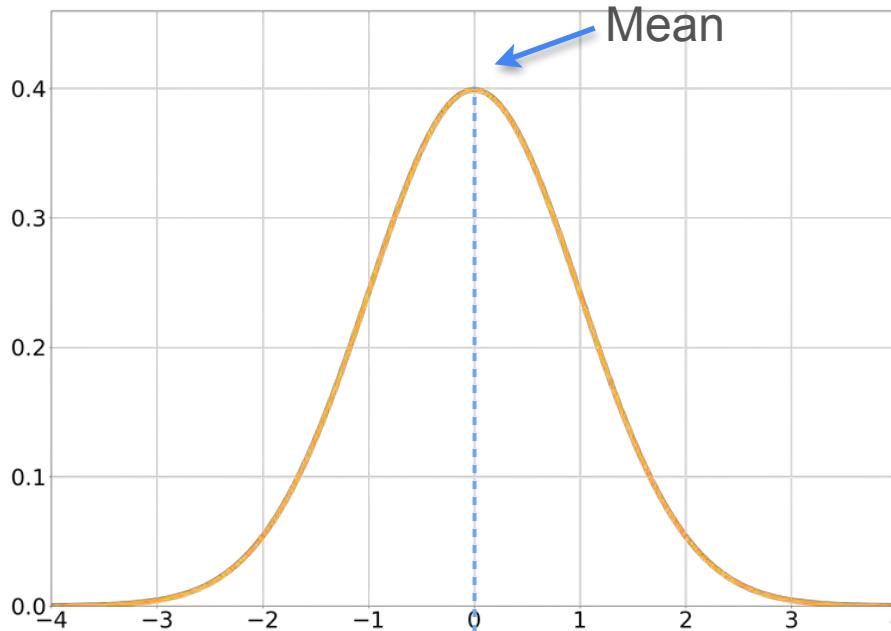
Mean, Median and Mode in Normal Distribution



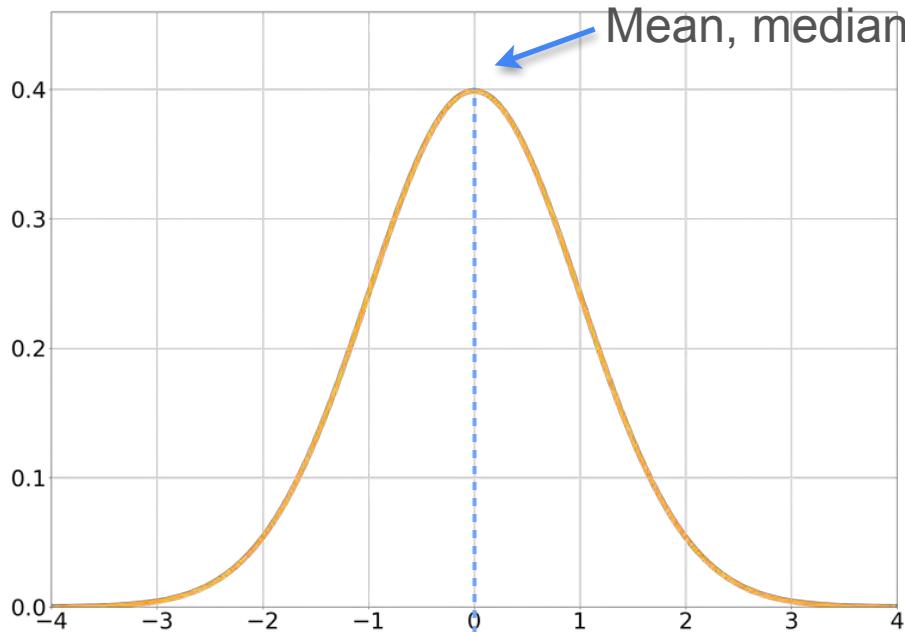
Mean, Median and Mode in Normal Distribution



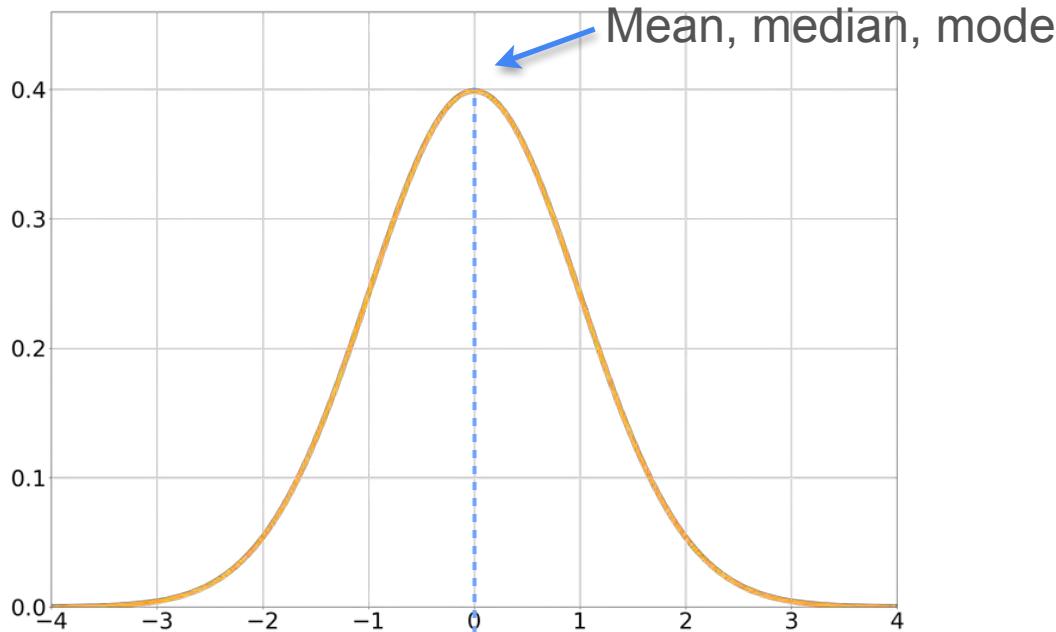
Mean, Median and Mode in Normal Distribution



Mean, Median and Mode in Normal Distribution



Mean, Median and Mode in Normal Distribution





DeepLearning.AI

Describing Distributions

Expected Value

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$6

Do you play the game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$4

Do you play the game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$4

Do you play the game?

What is the maximum amount of money you would pay to play this game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term:

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \0

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5 \rightarrow$ You expect to win \$5 on average
 $\mathbb{E}[X] = 5$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$5

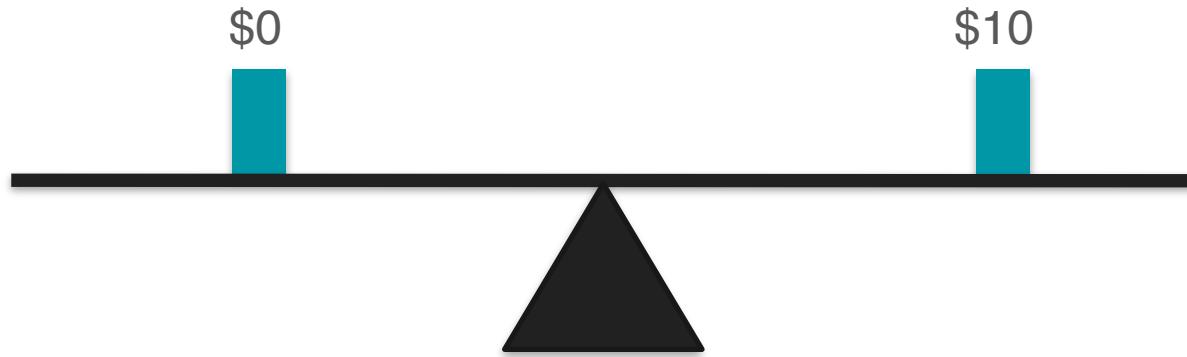
Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5$ 

You expect to win \$5 on average
 $E[X] = 5$

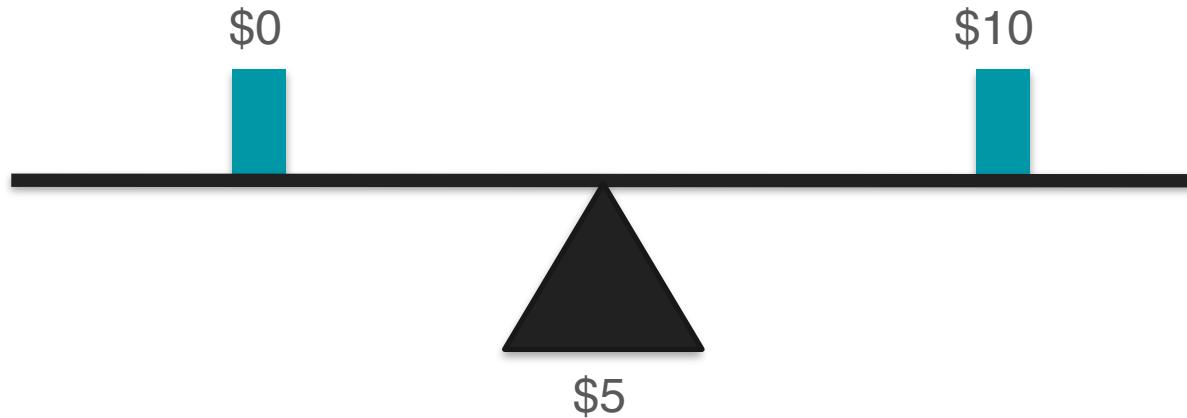
Expected Value: Motivation Example 1



Expected Value: Motivation Example 1



Expected Value: Motivation Example 1



Expected Value: Motivation Example 2

Play another game



Flip three coins. For each heads you win \$1

What is the maximum amount of money you would pay to play this game?

Expected Value: Motivation Example 2

Number of heads:

0



1



2



3



Expected Value: Motivation Example 2

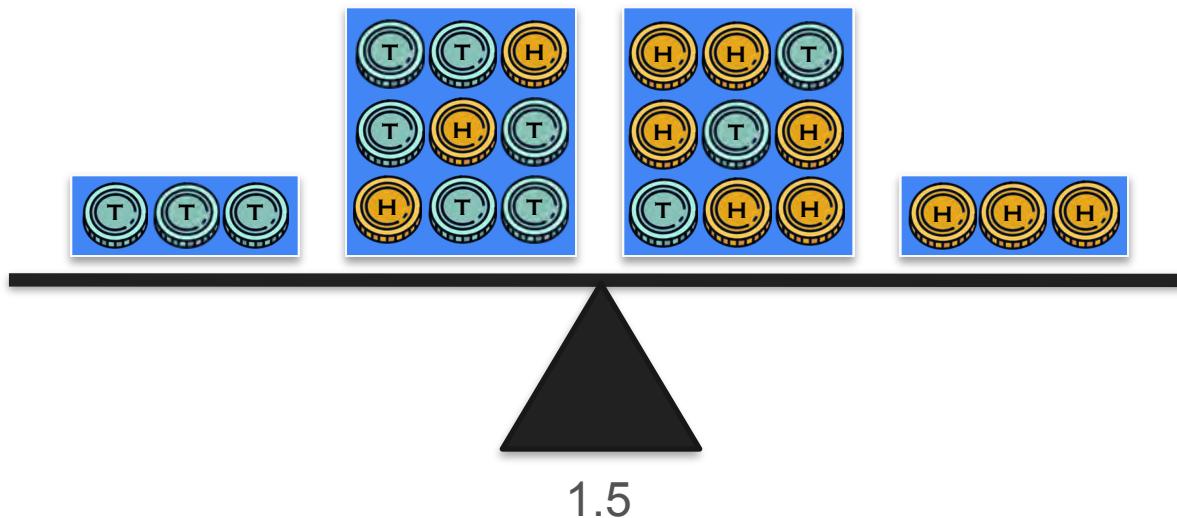
Number of heads:

0

1

2

3



Expected Value: Motivation Example 2

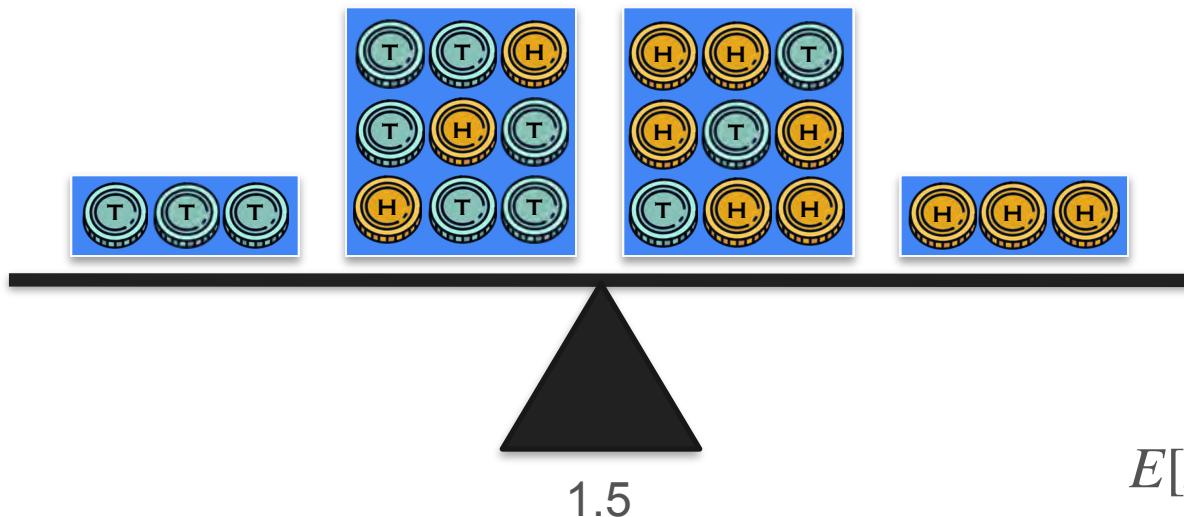
Number of heads:

0

1

2

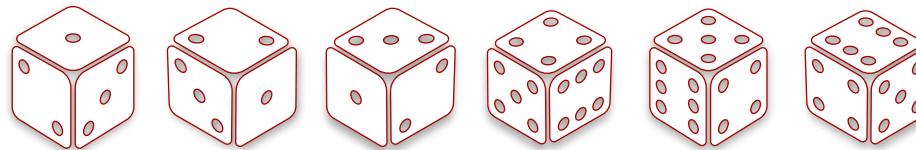
3



Expected Value: Motivation Example 3

Probability: 1/6 1/6 1/6 1/6 1/6 1/6

Roll: 1 2 3 4 5 6

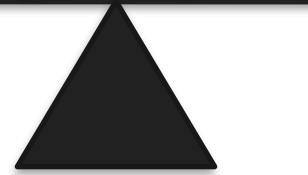
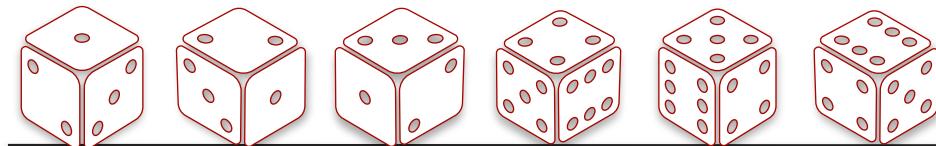


$$\frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

Expected Value: Motivation Example 3

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6

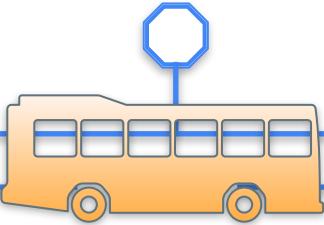


3.5

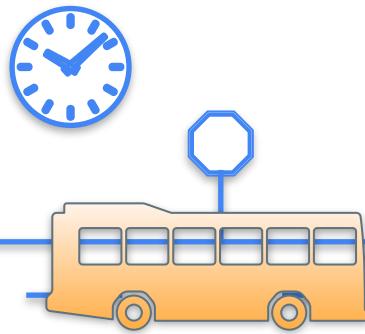
Expected Value



Expected Value



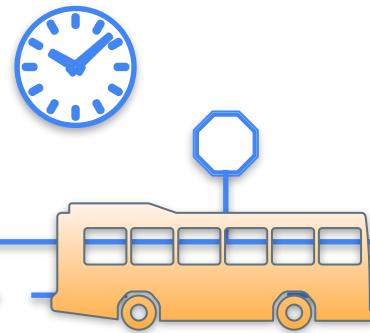
Expected Value



Expected Value

Waiting Time

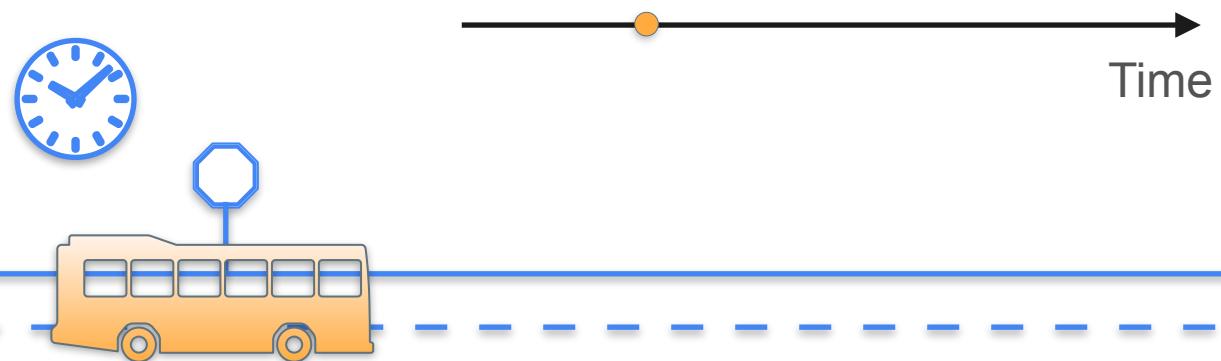
15 min



Expected Value

Waiting Time

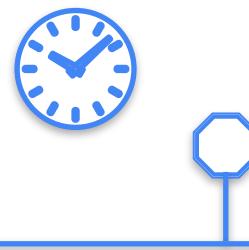
15 min



Expected Value

Waiting Time

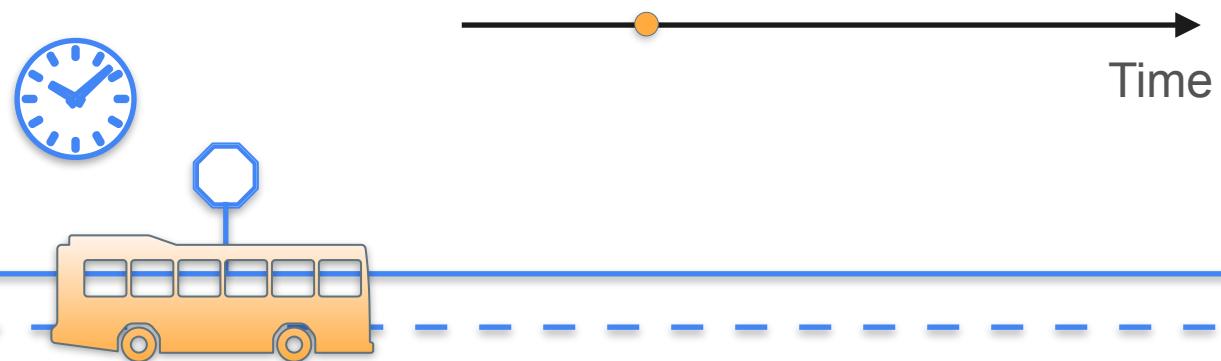
15 min



Expected Value

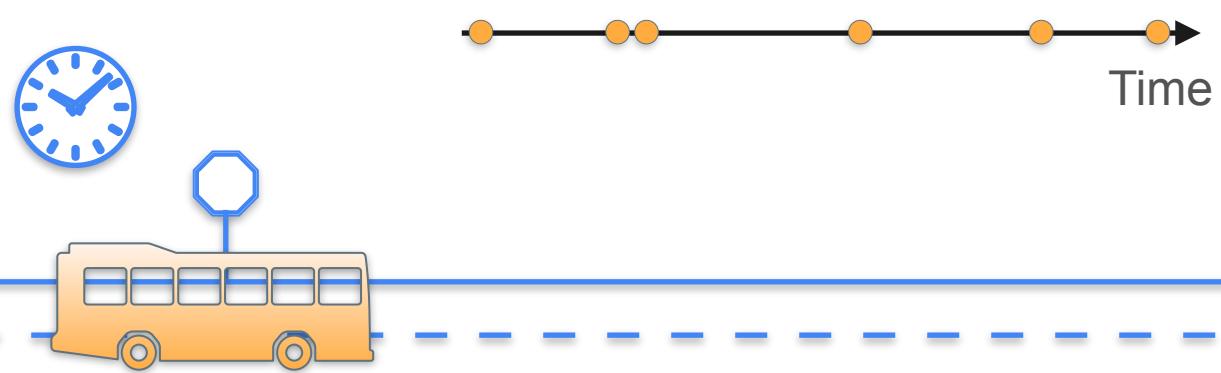
Waiting Time

15 min



Expected Value

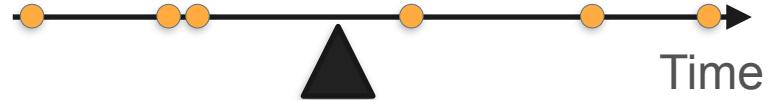
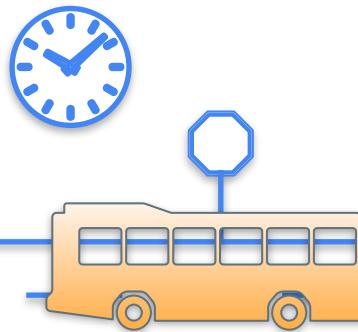
Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min

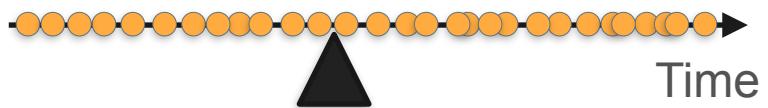
Average = 27.833



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

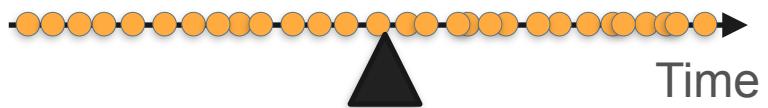
Average = 27.833



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

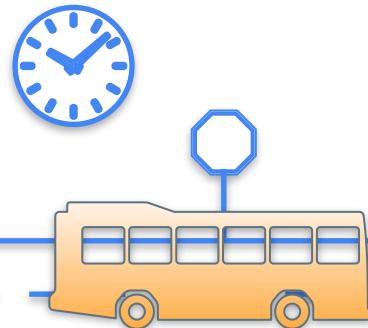
Average = 30



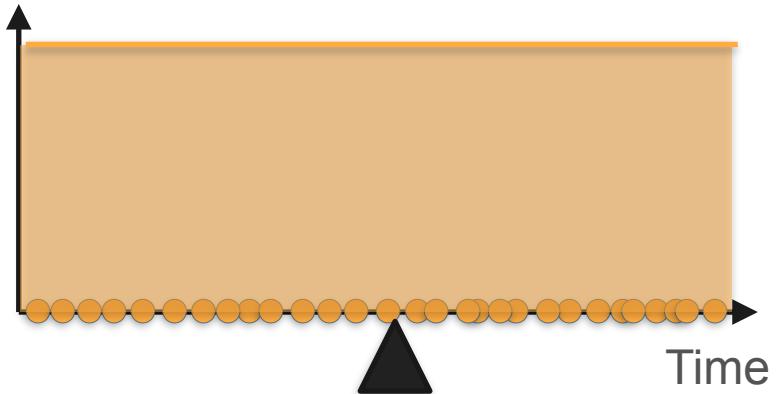
Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

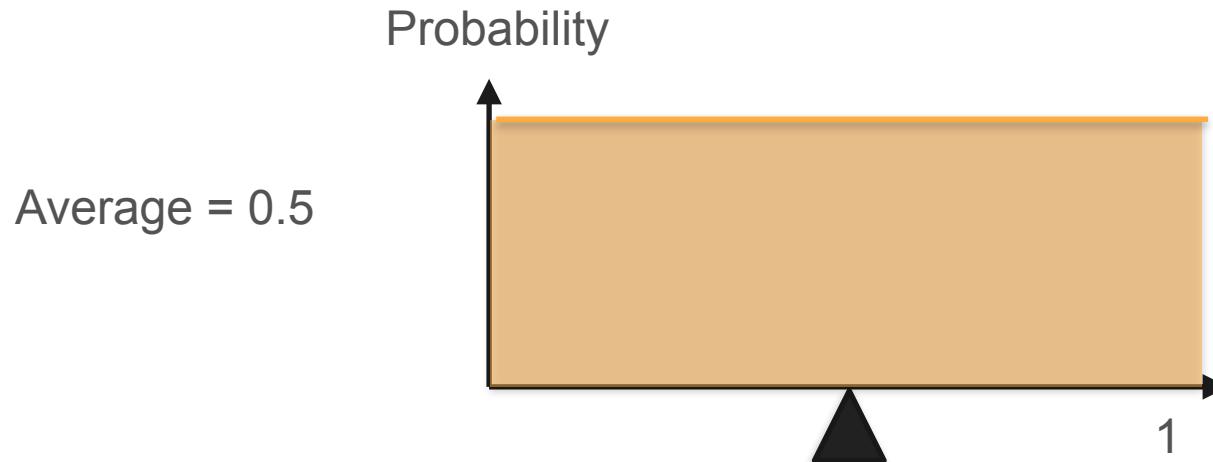
Average = 30



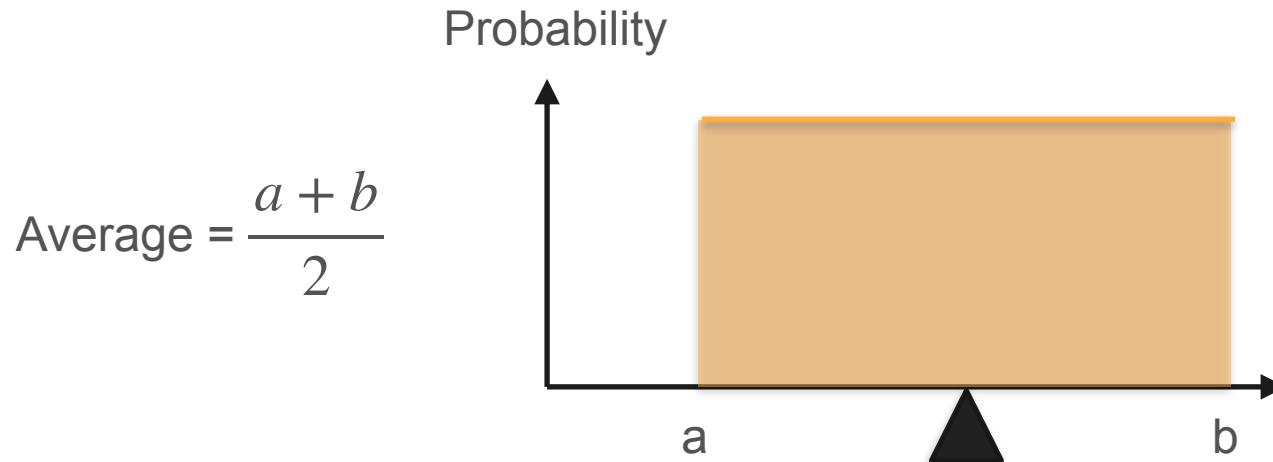
Probability



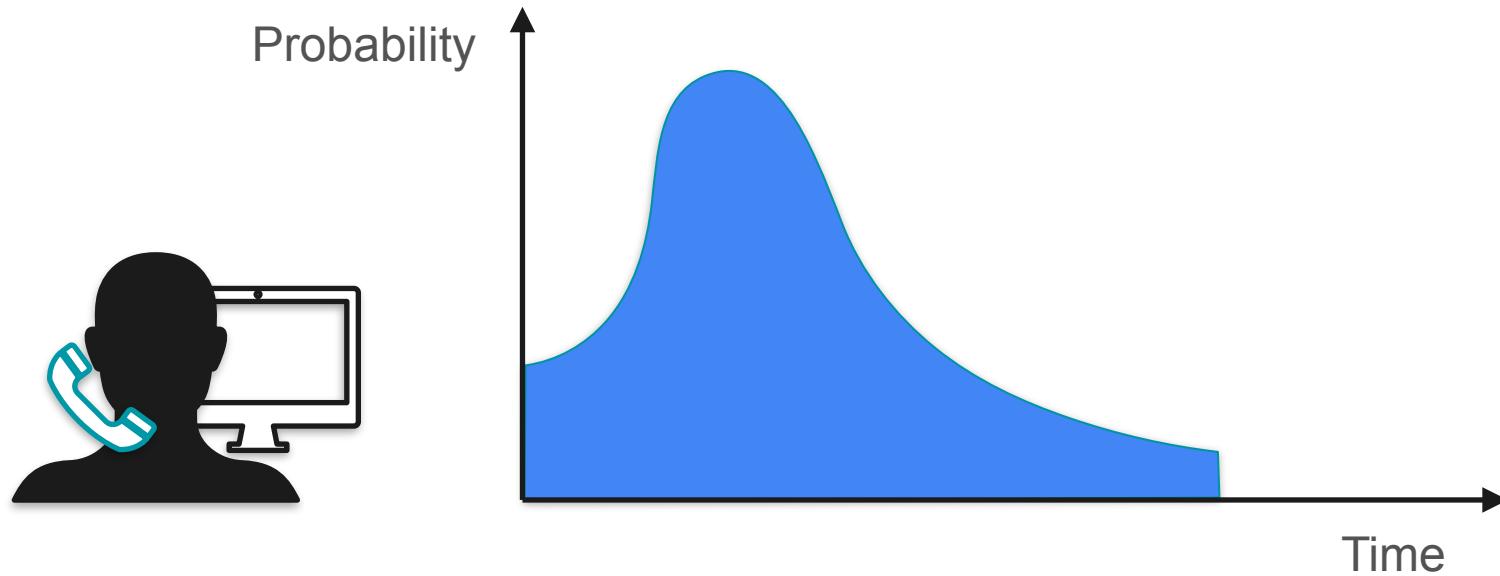
Expected Value: Uniform Distribution



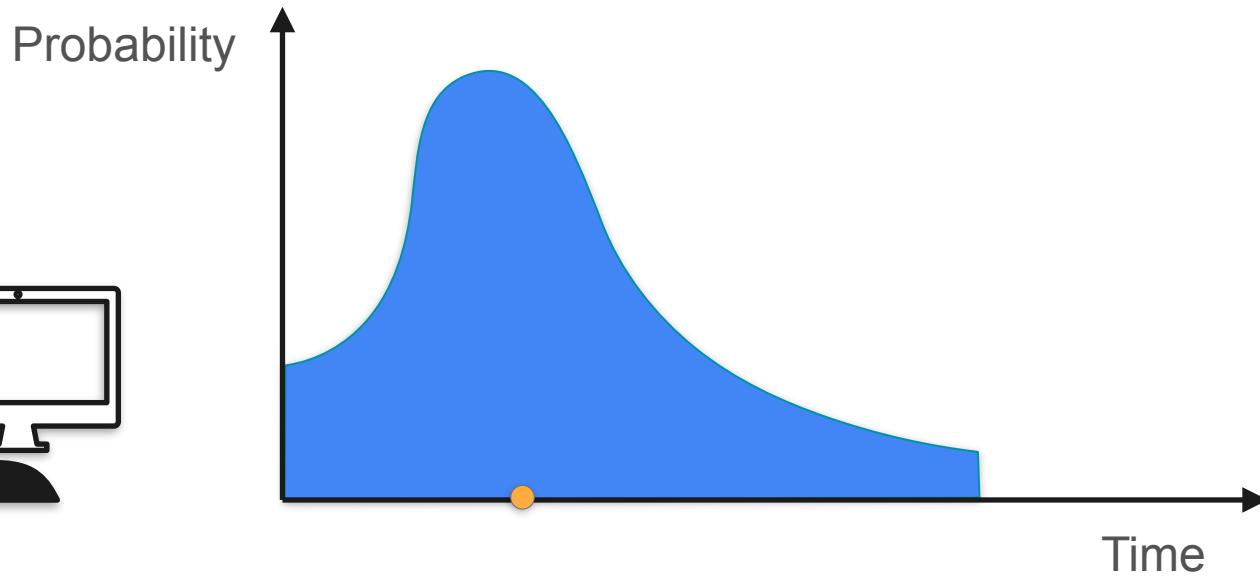
Expected Value: Uniform Distribution



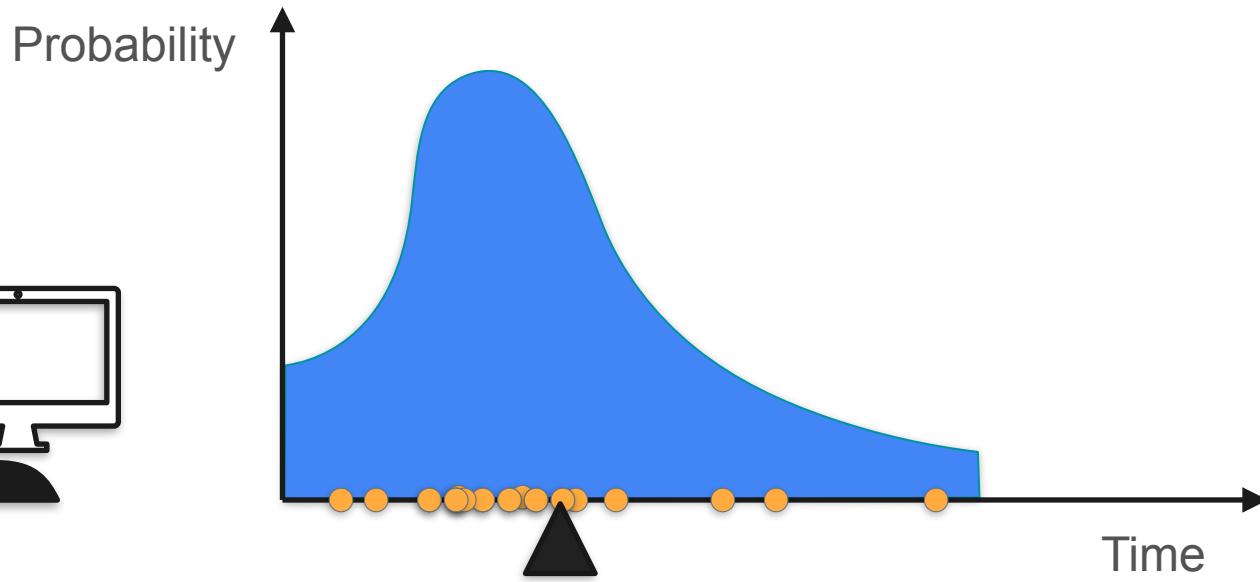
Expected Value



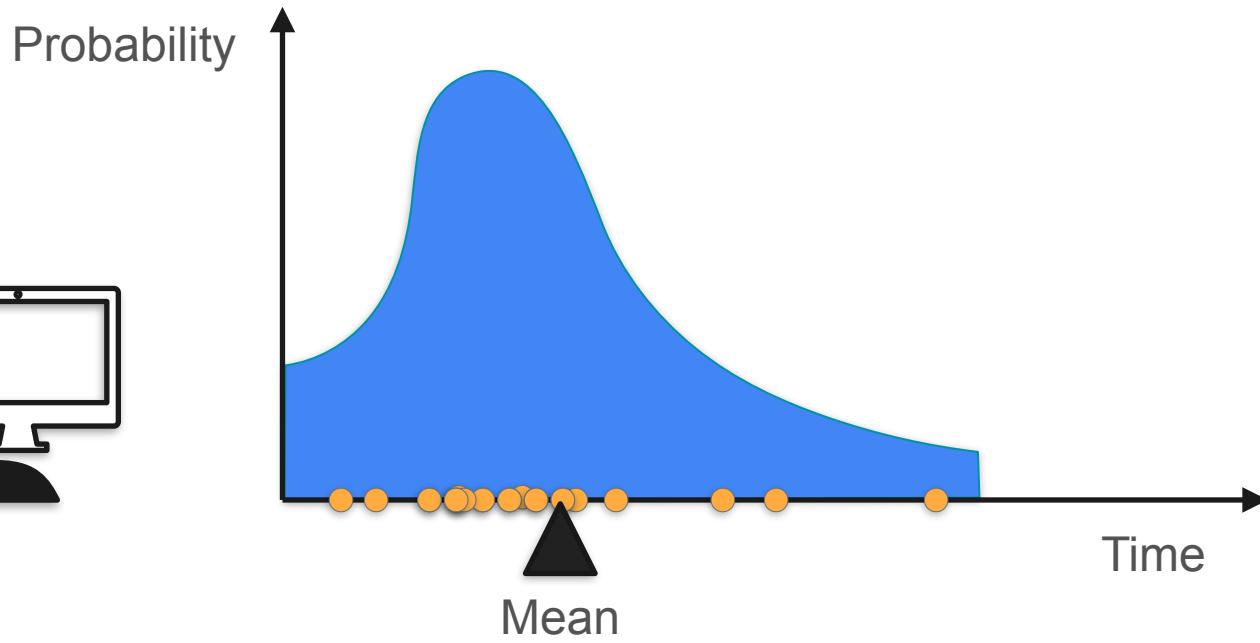
Expected Value



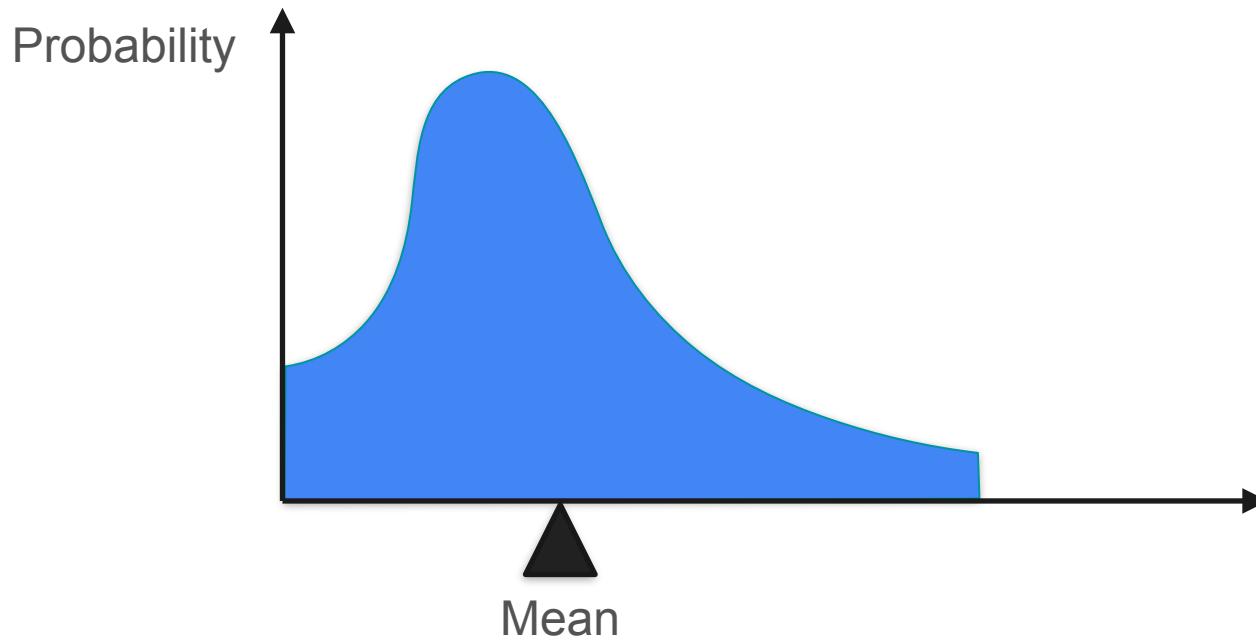
Expected Value



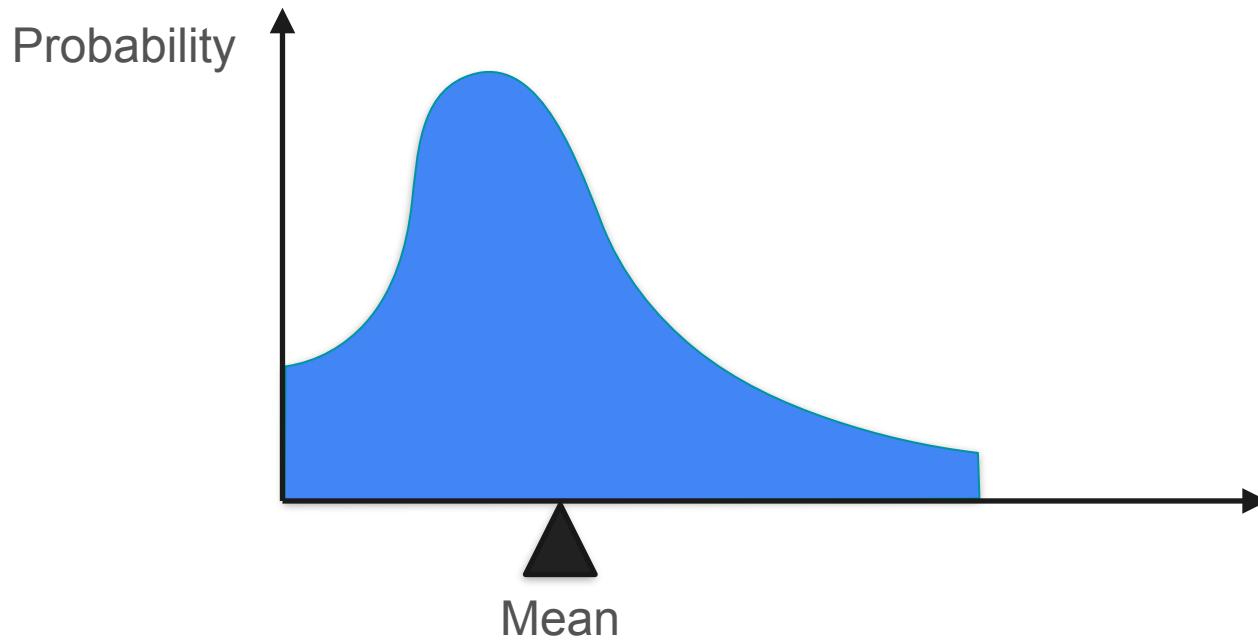
Expected Value



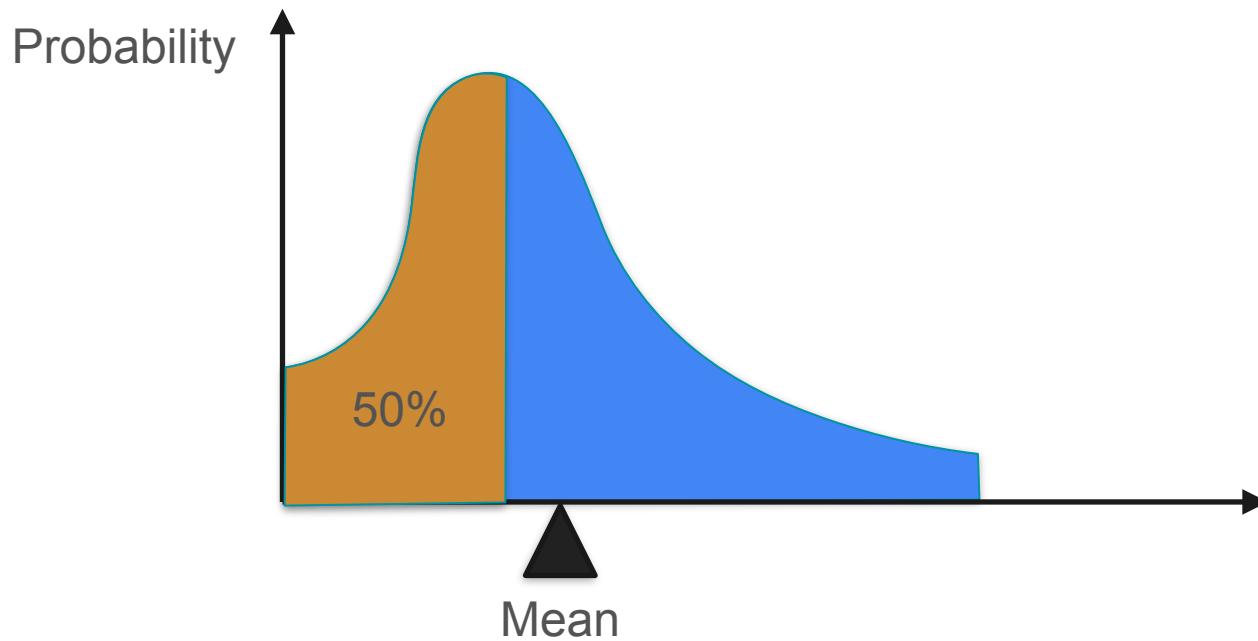
Expected Value: General Case



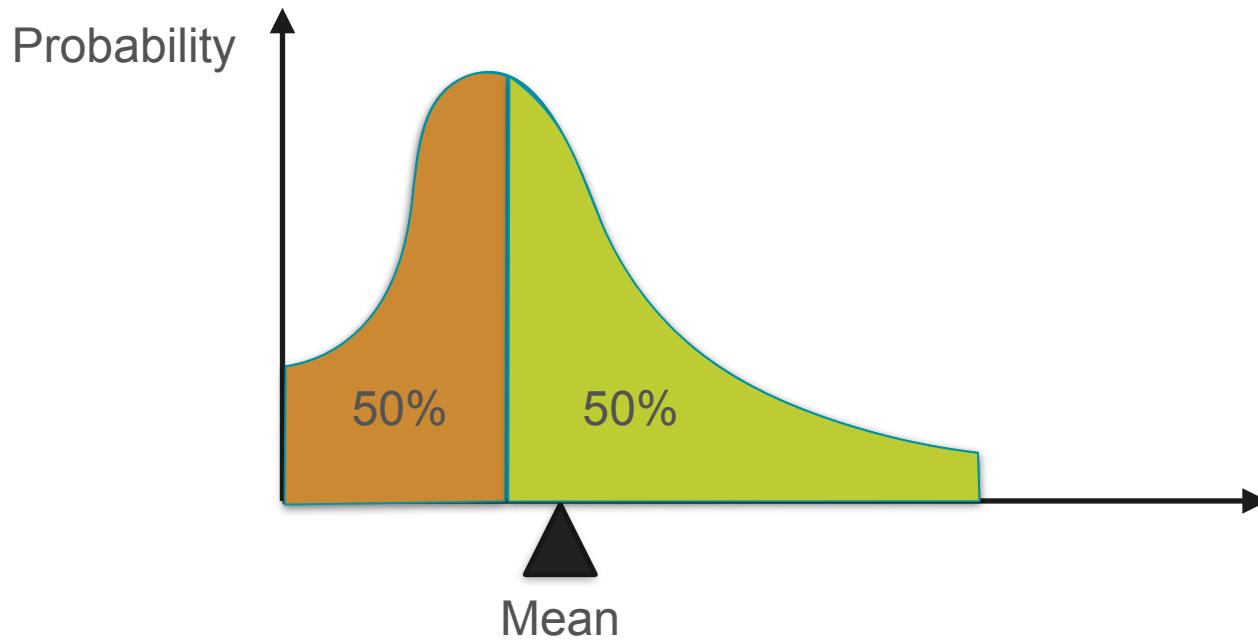
Expected Value: Common Misconception



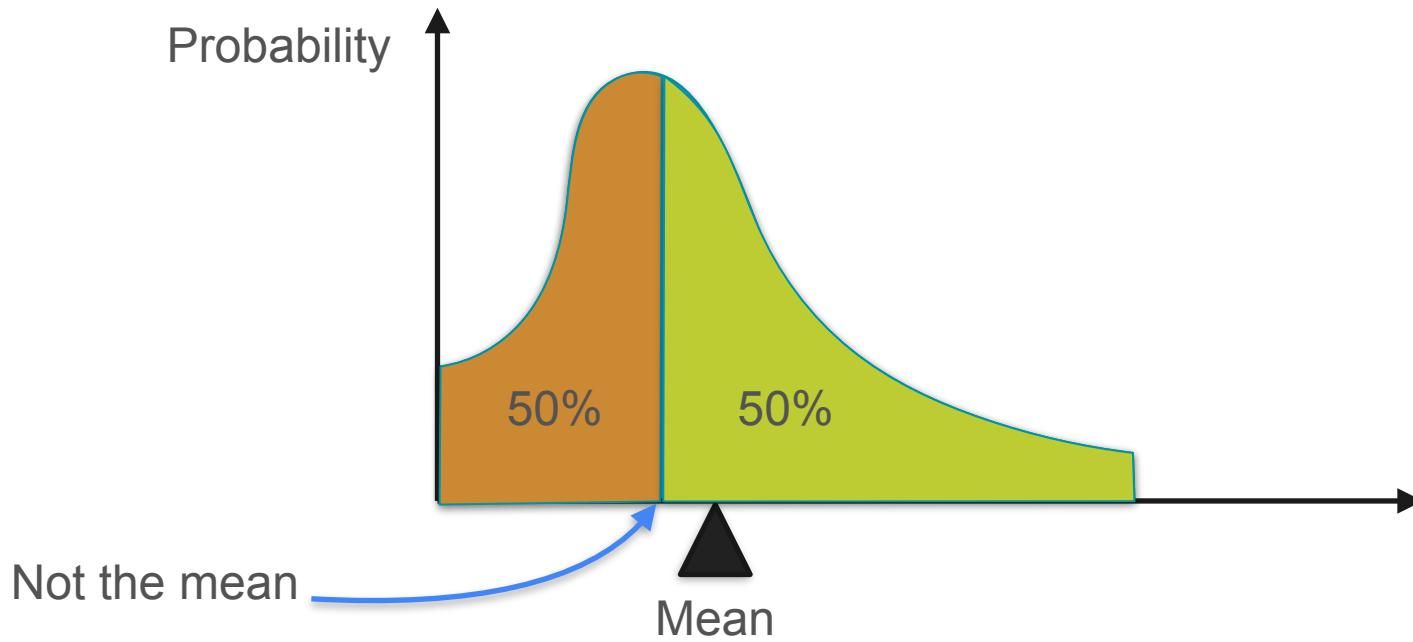
Expected Value: Common Misconception



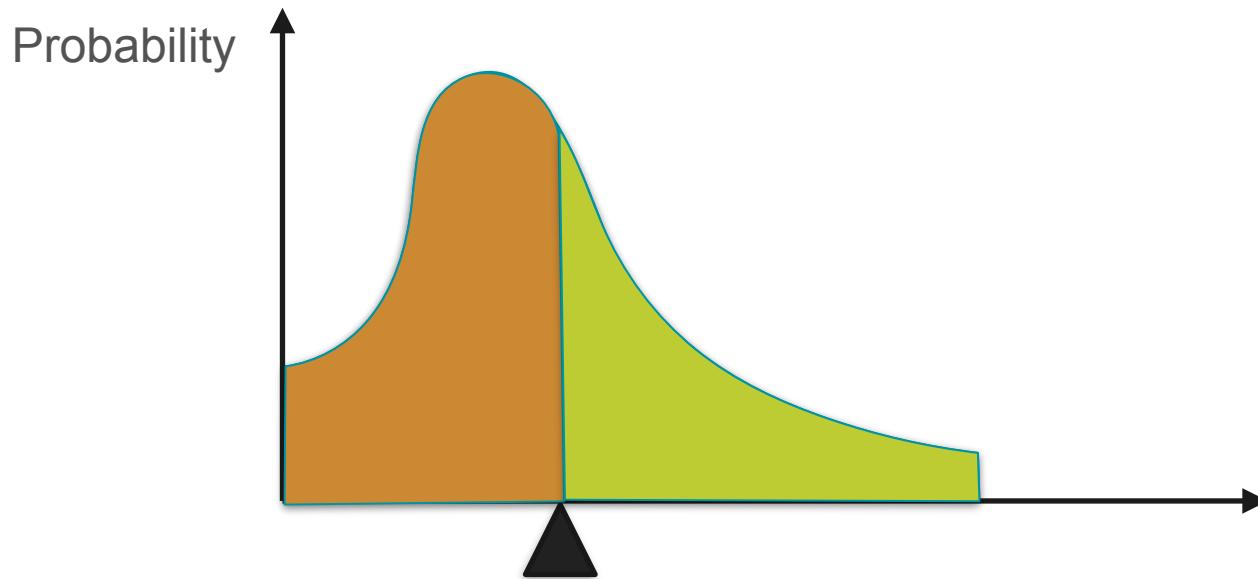
Expected Value: Common Misconception



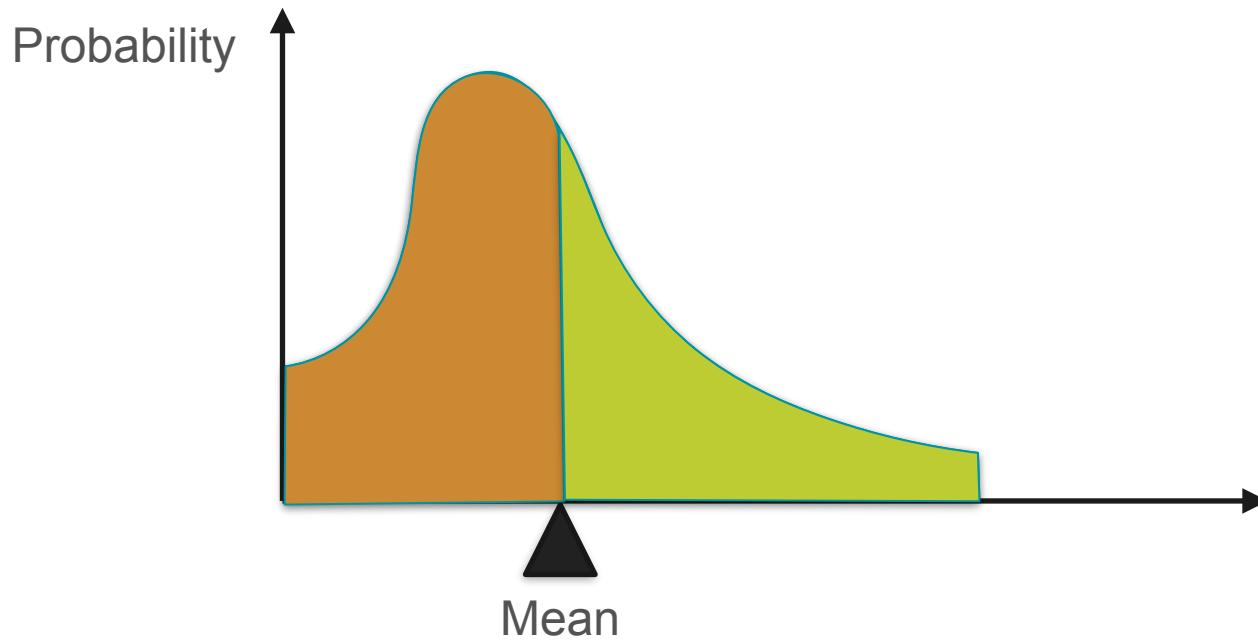
Expected Value: Common Misconception



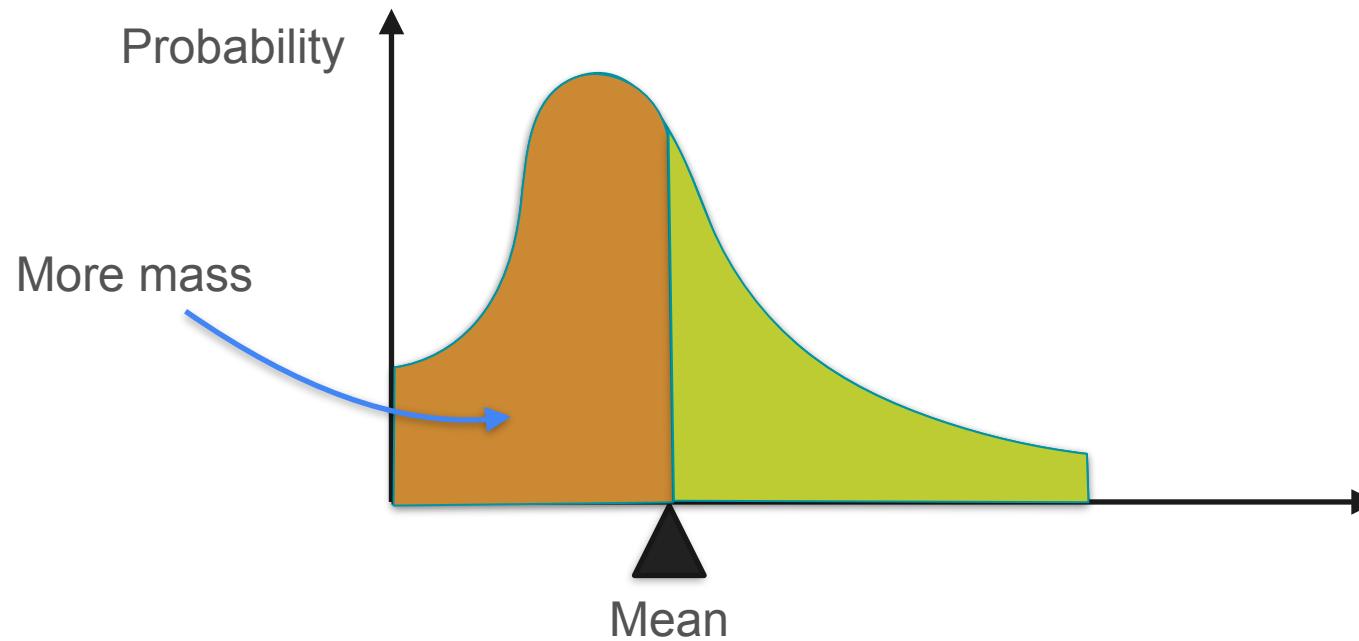
Expected Value: Common Misconception



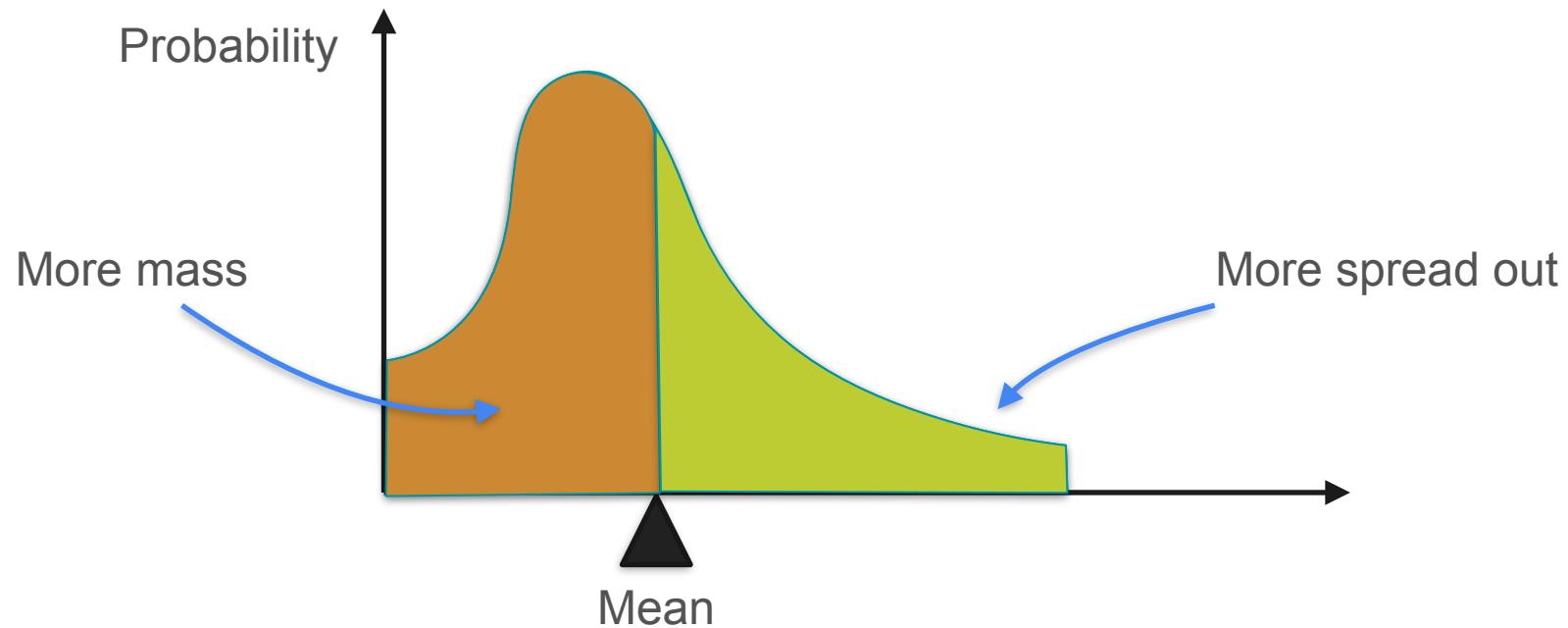
Expected Value: Common Misconception



Expected Value: Common Misconception



Expected Value: Common Misconception

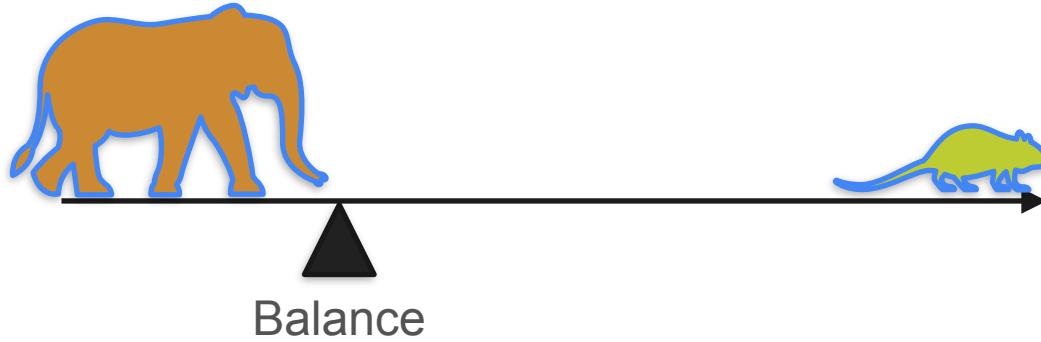


Expected Value: Common Misconception

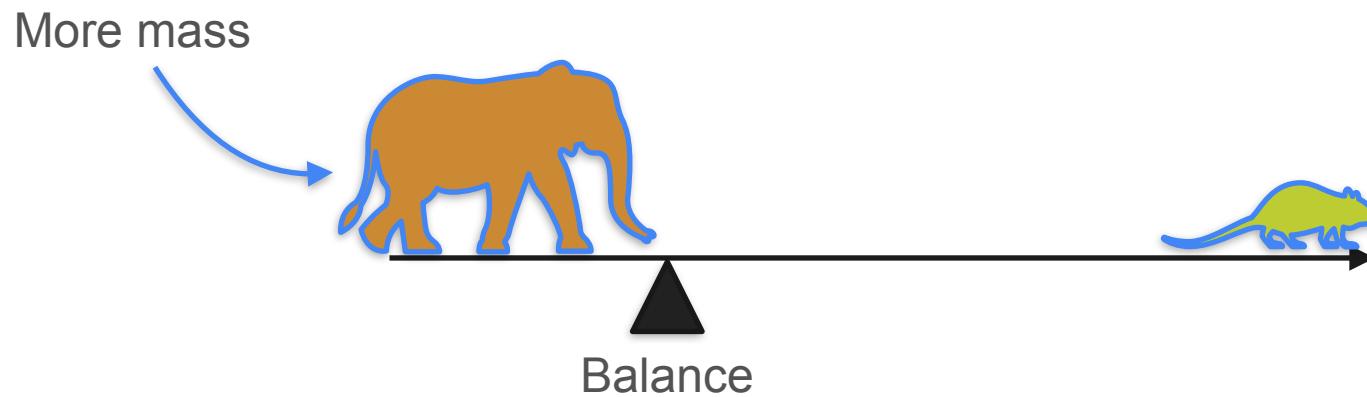


Balance

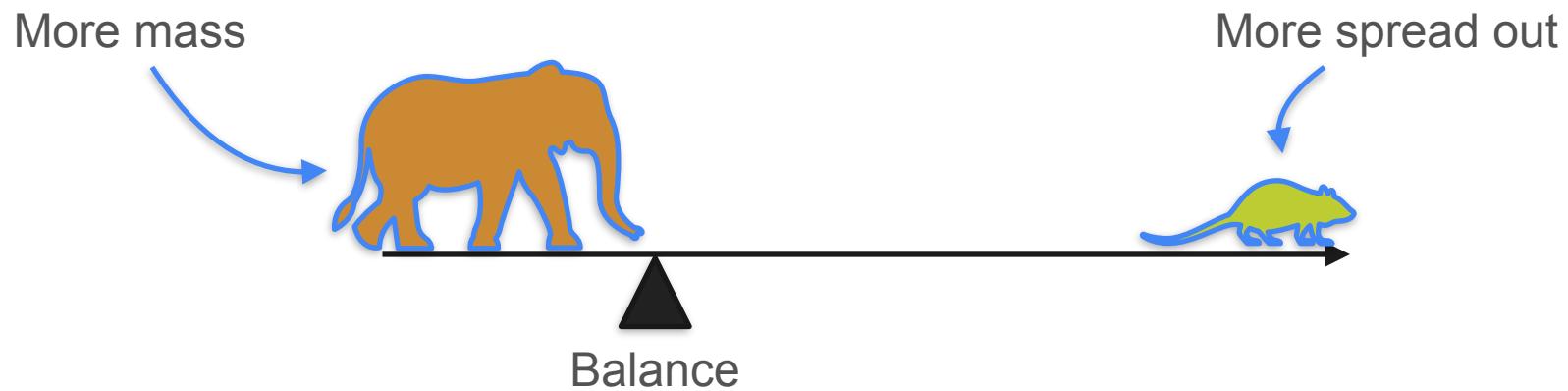
Expected Value: Common Misconception



Expected Value: Common Misconception



Expected Value: Common Misconception



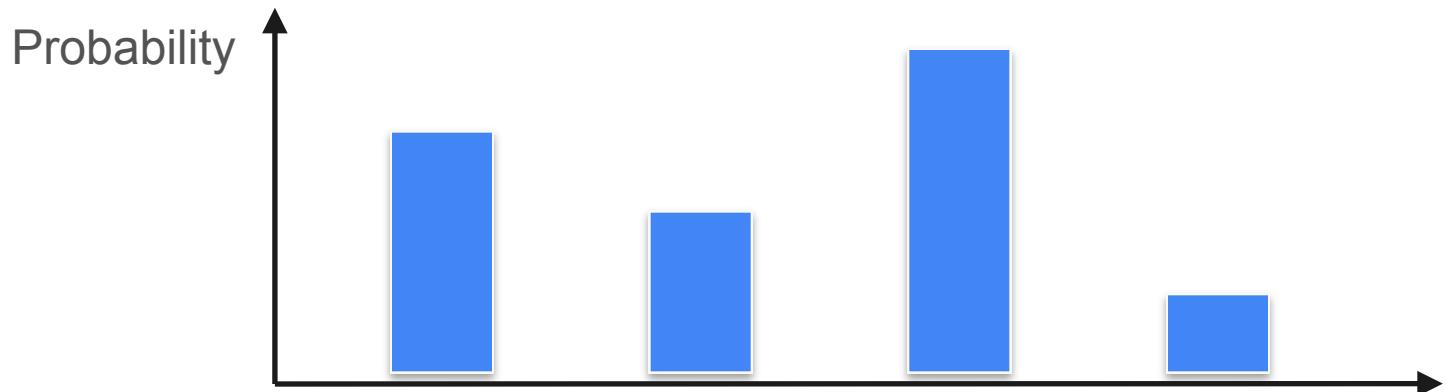


DeepLearning.AI

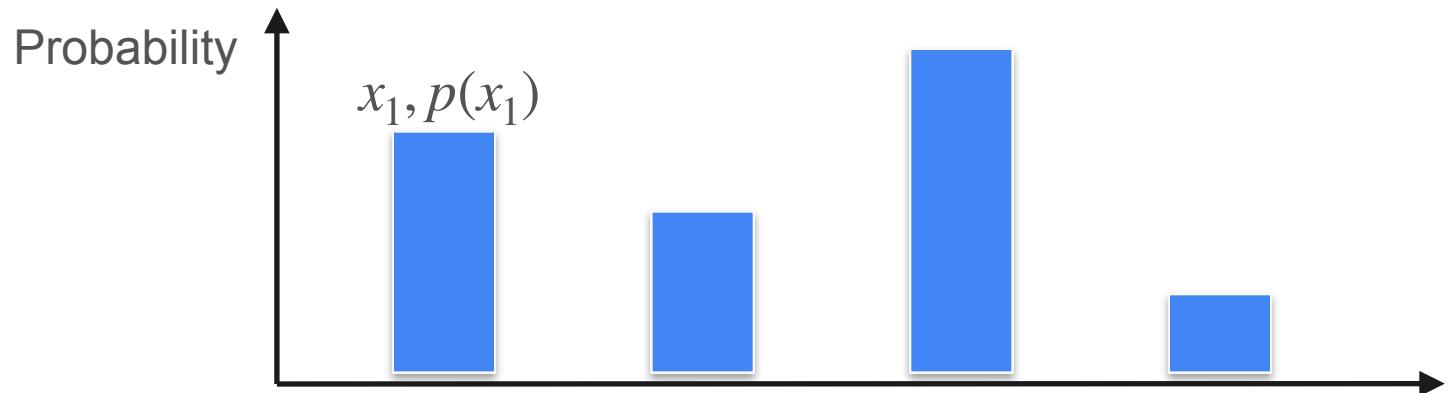
Describing Distributions

Expected value of a function

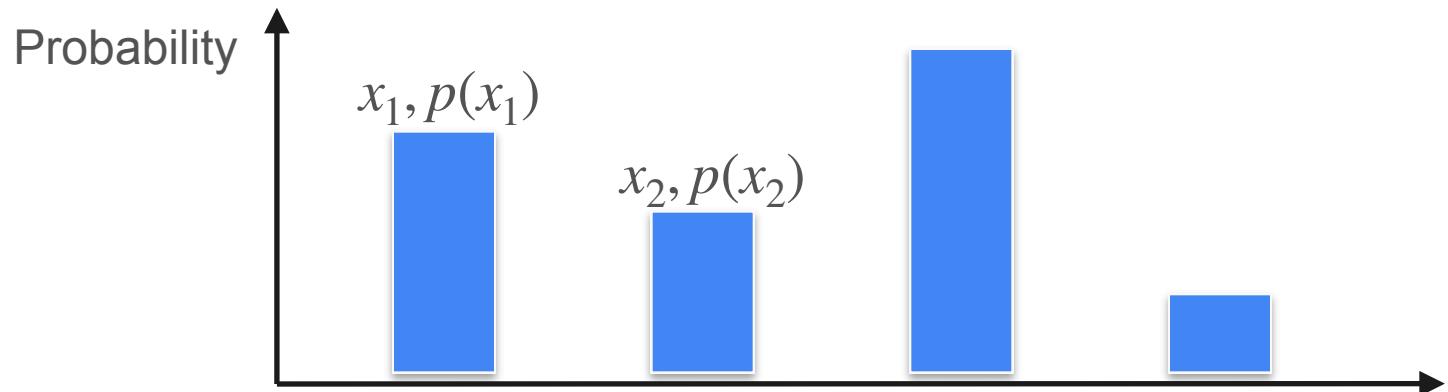
Expected Value of a Function



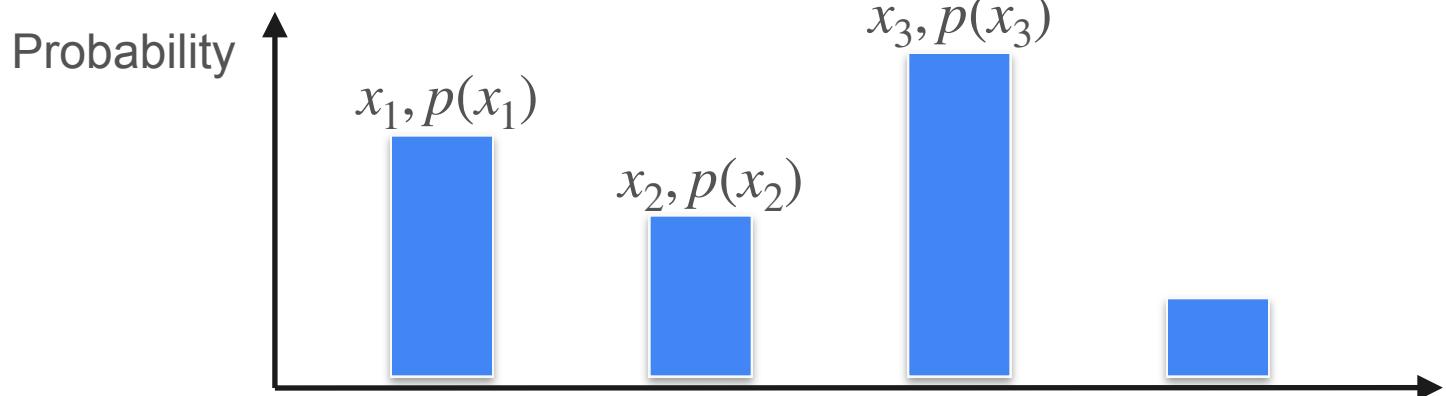
Expected Value of a Function



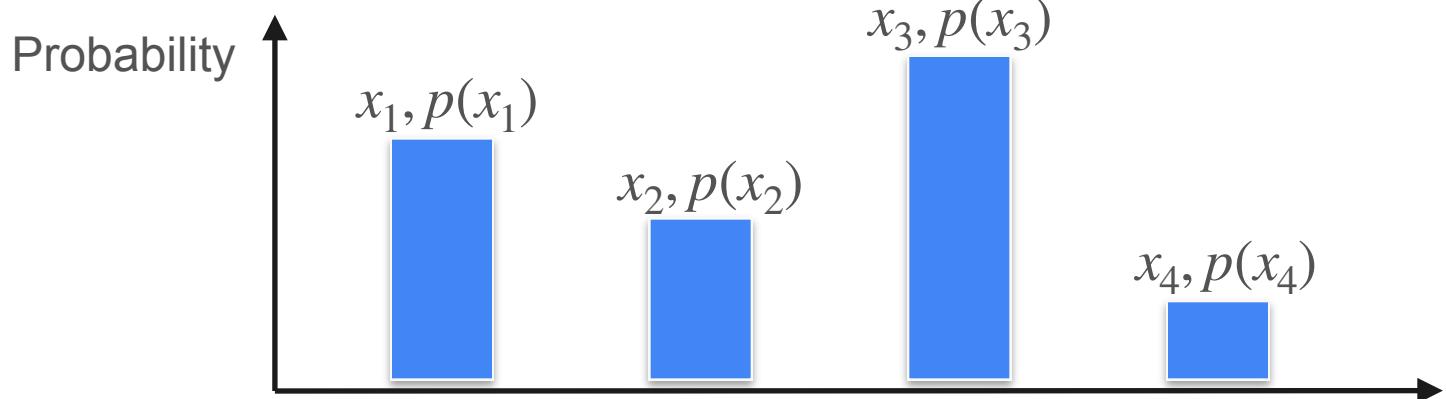
Expected Value of a Function



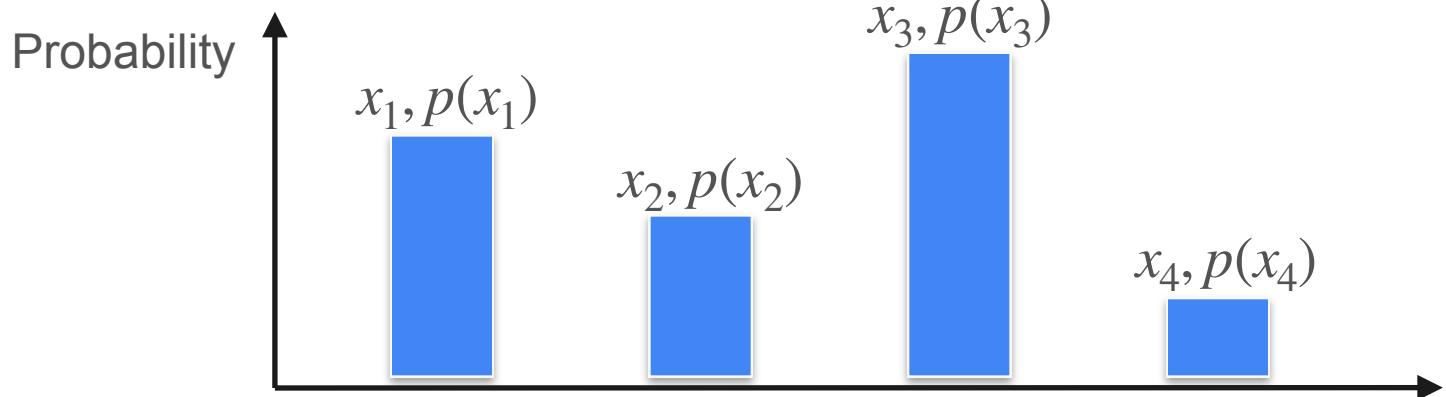
Expected Value of a Function



Expected Value of a Function

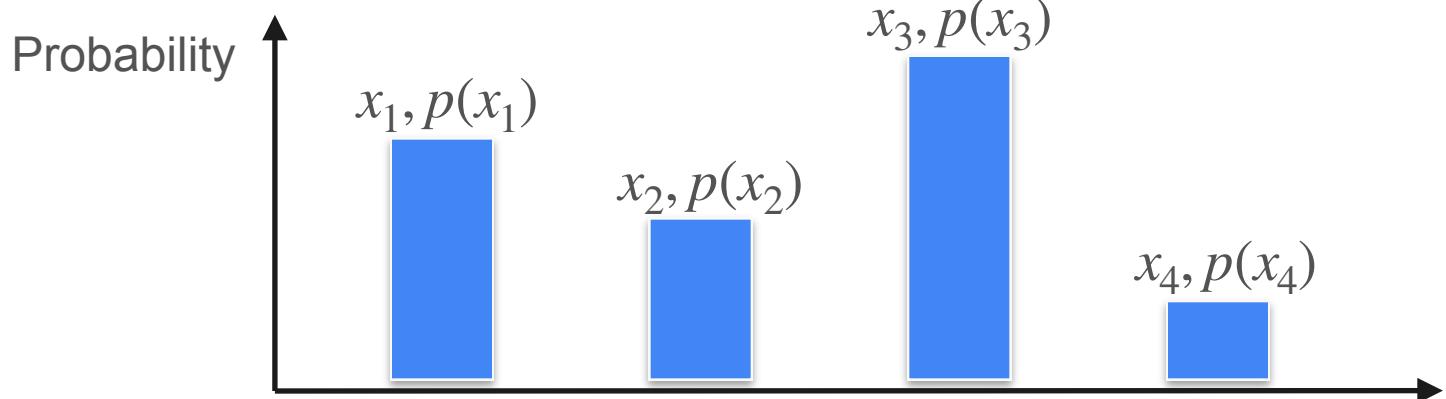


Expected Value of a Function



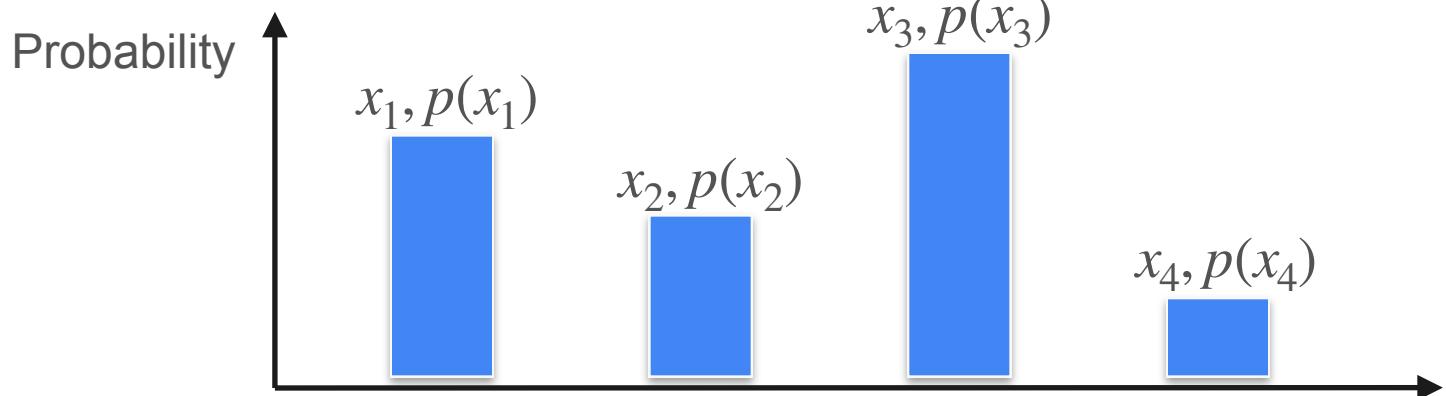
$$\mathbb{E}[X] =$$

Expected Value of a Function



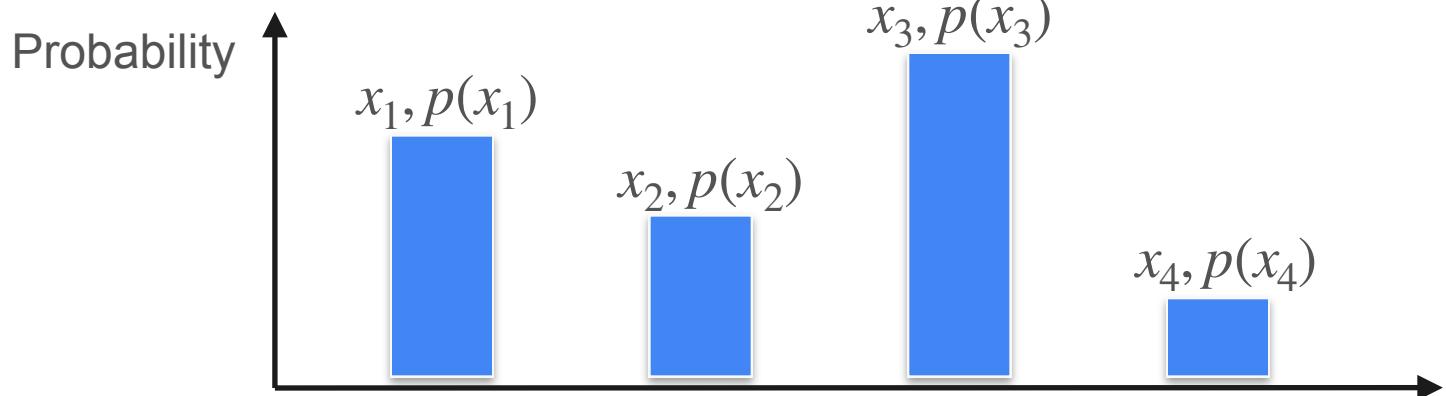
$$\mathbb{E}[X] = x_1 p(x_1)$$

Expected Value of a Function



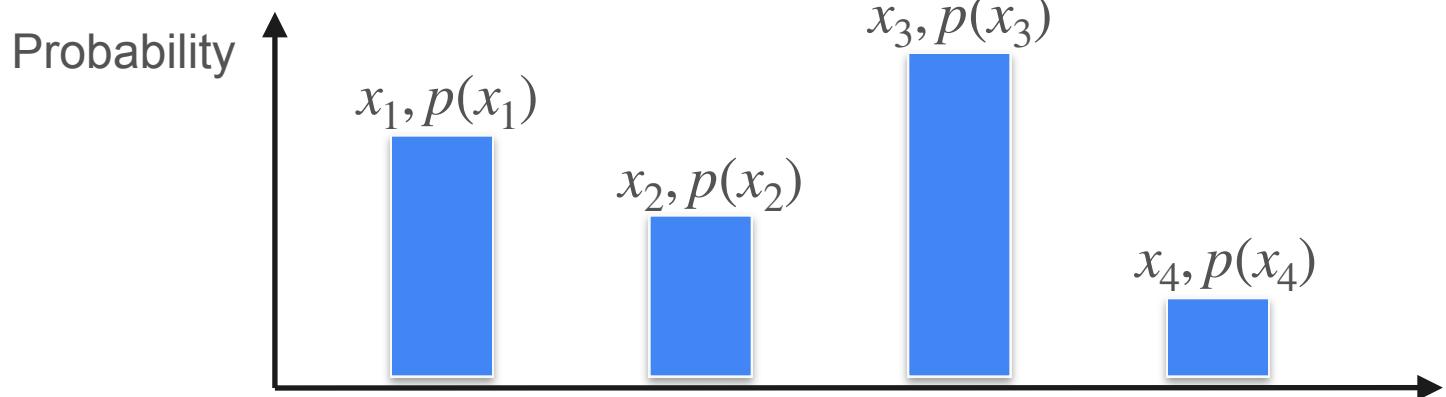
$$\mathbb{E}[X] = x_1 p(x_1) + x_2 p(x_2)$$

Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

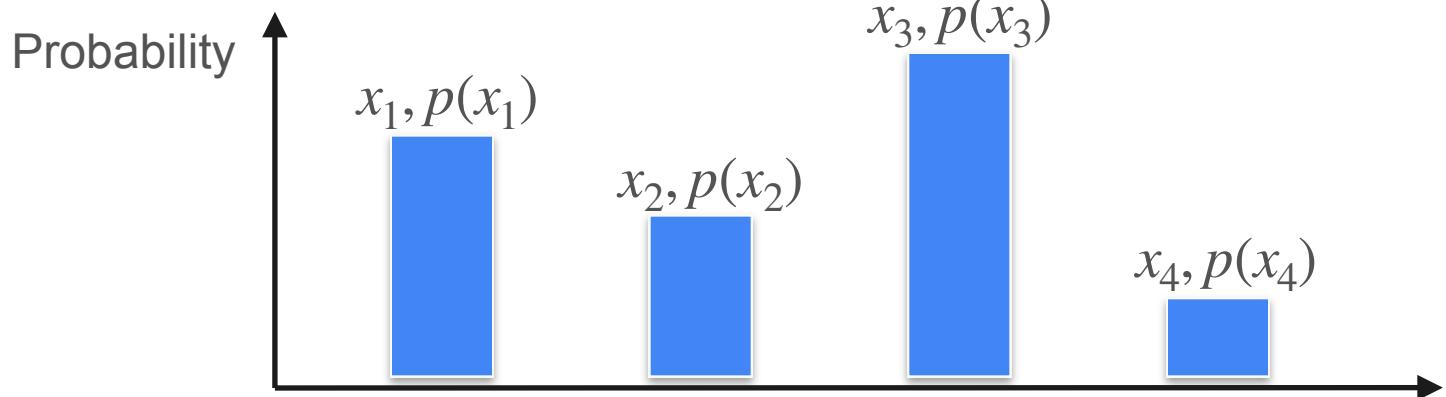
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] =$$

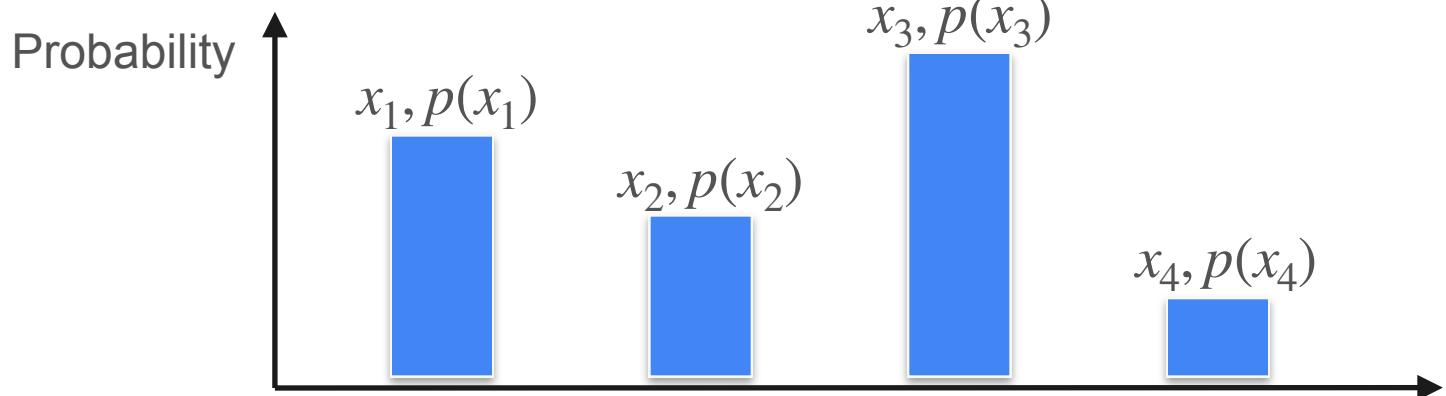
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1)$$

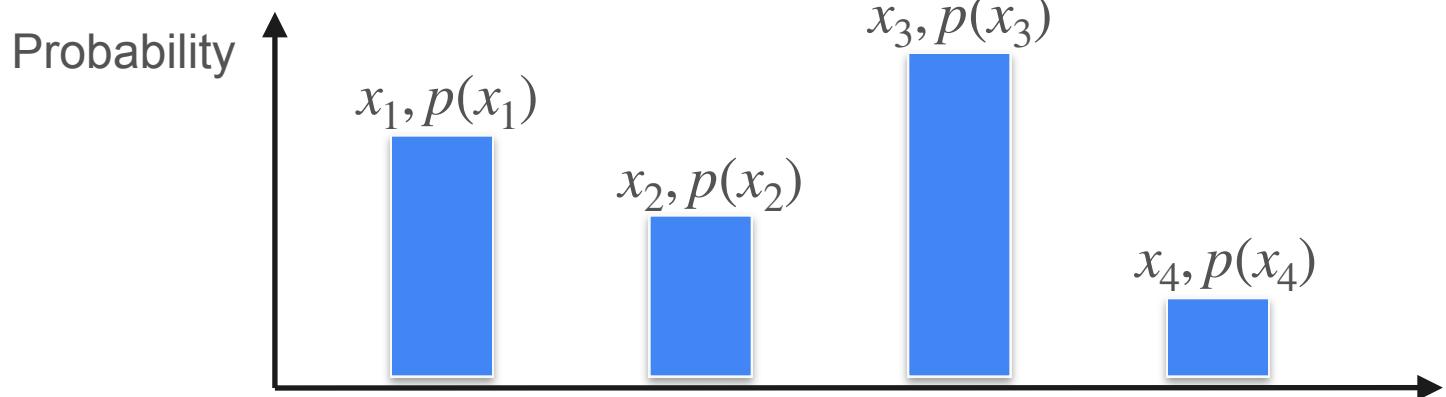
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1) + f(x_2)p(x_2)$$

Expected Value of a Function



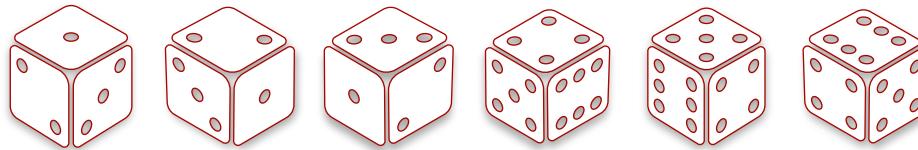
$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1) + f(x_2)p(x_2) + f(x_3)p(x_3) + f(x_4)p(x_4)$$

Expected Value of a Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

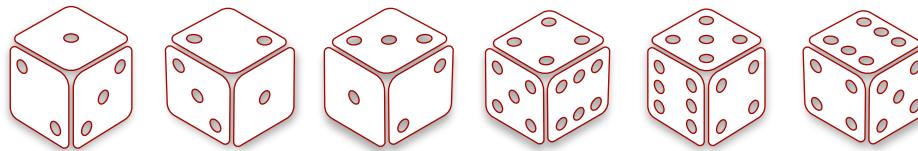
Roll: 1 2 3 4 5 6



Expected Value of a Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

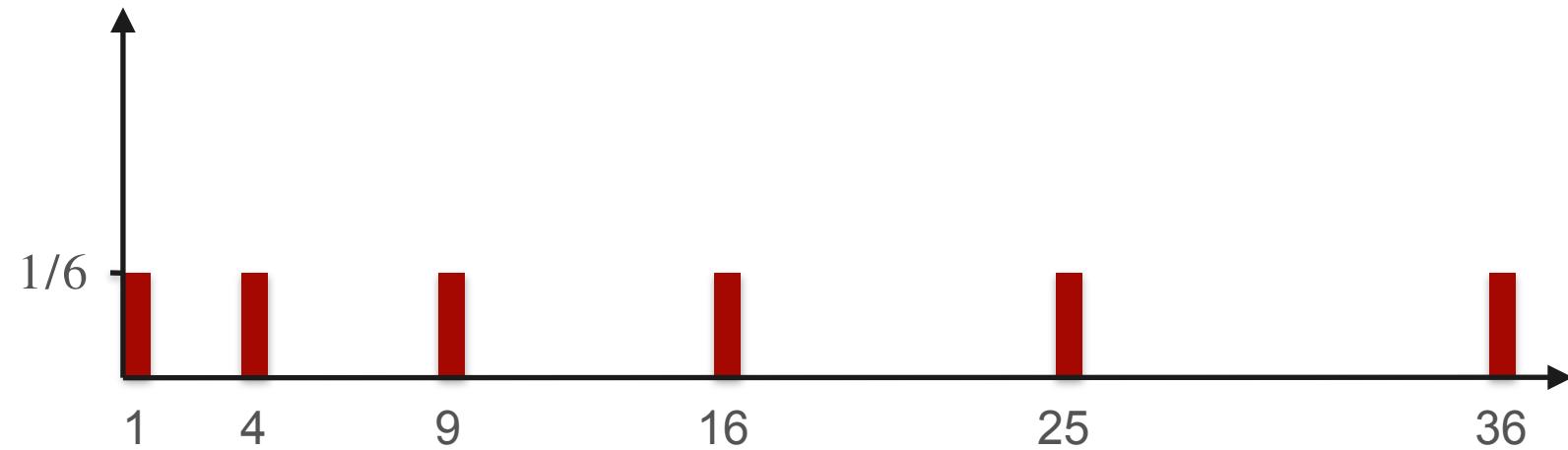
Roll: 1 2 3 4 5 6



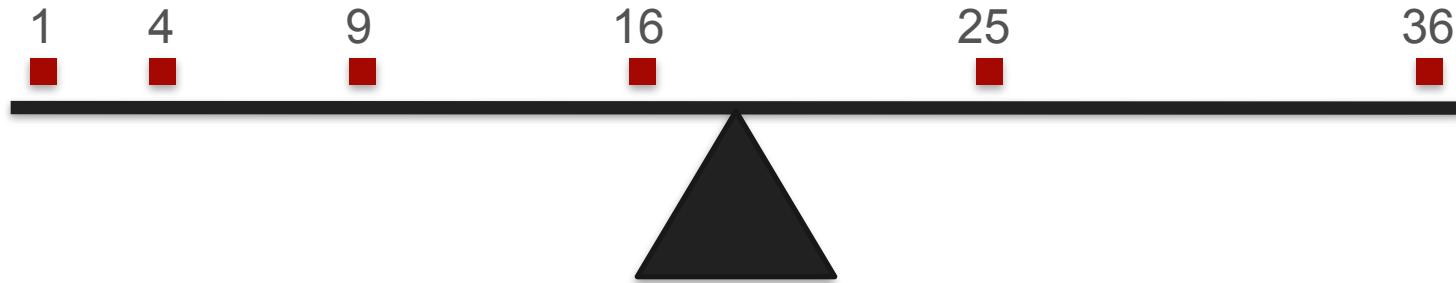
Square: 1 4 9 16 25 36

Expected Value of a Function

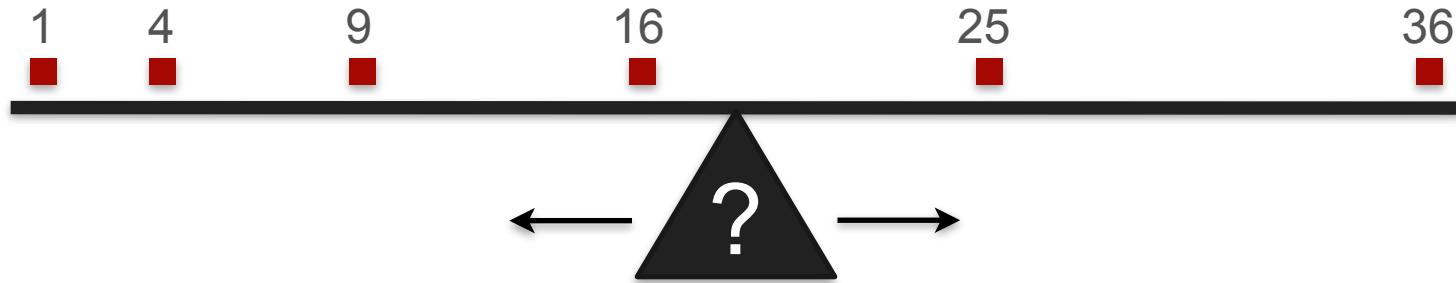
Probability



Expected Value of a Function

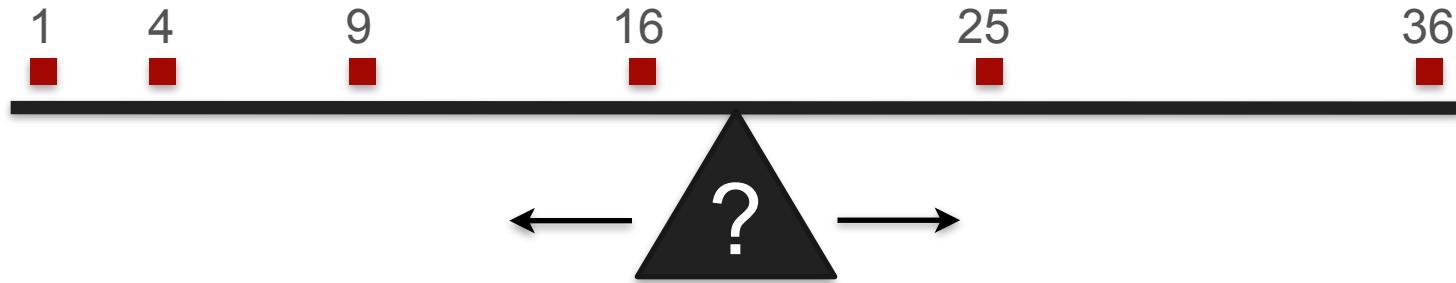


Expected Value of a Function



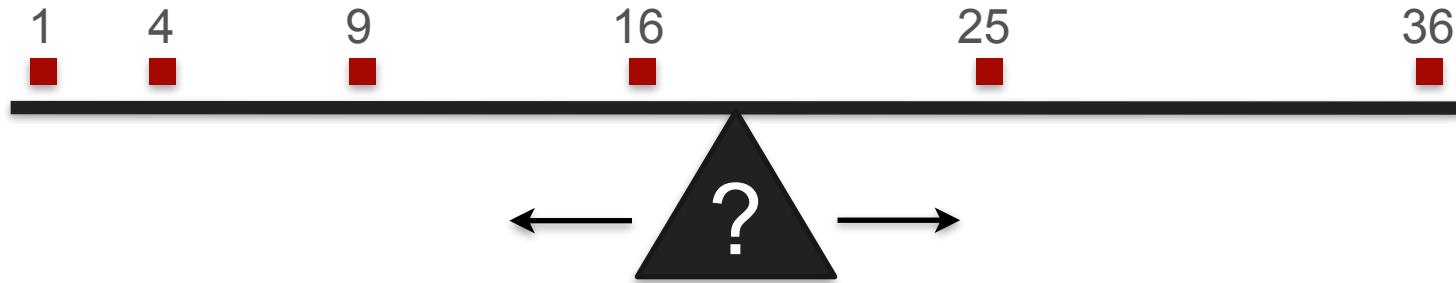
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6}$$



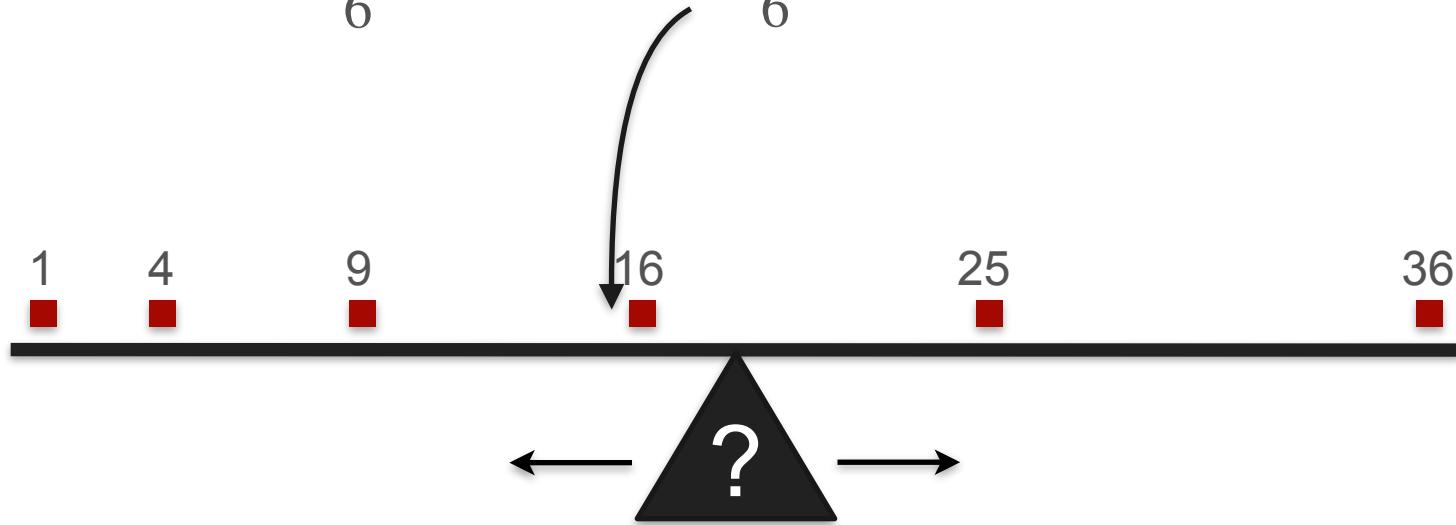
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



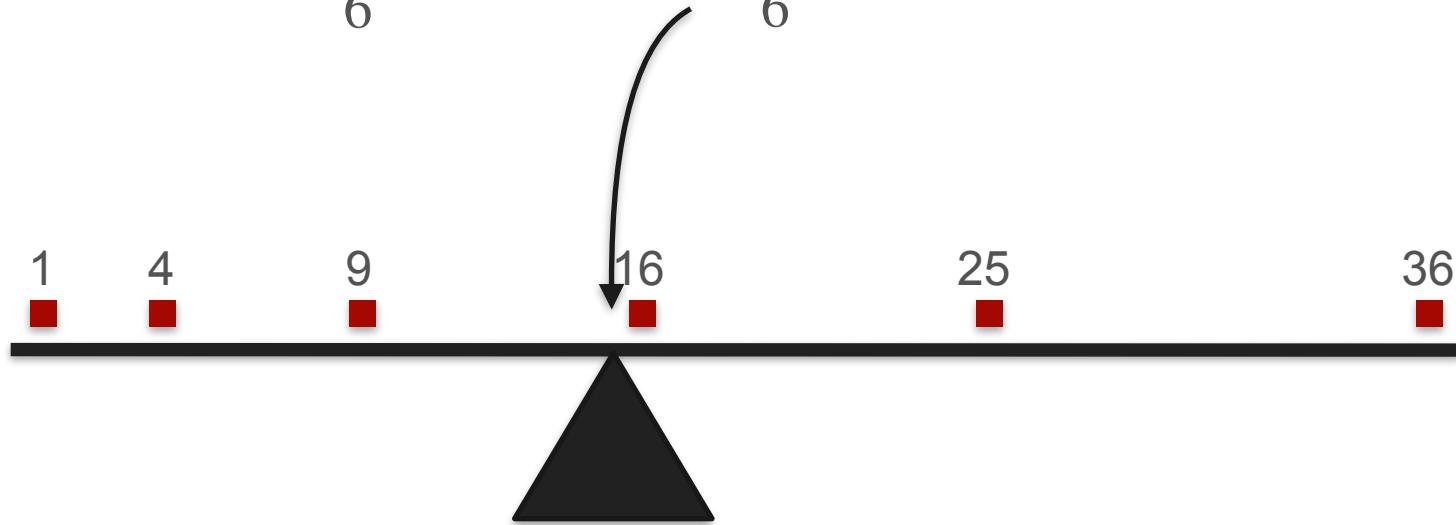
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



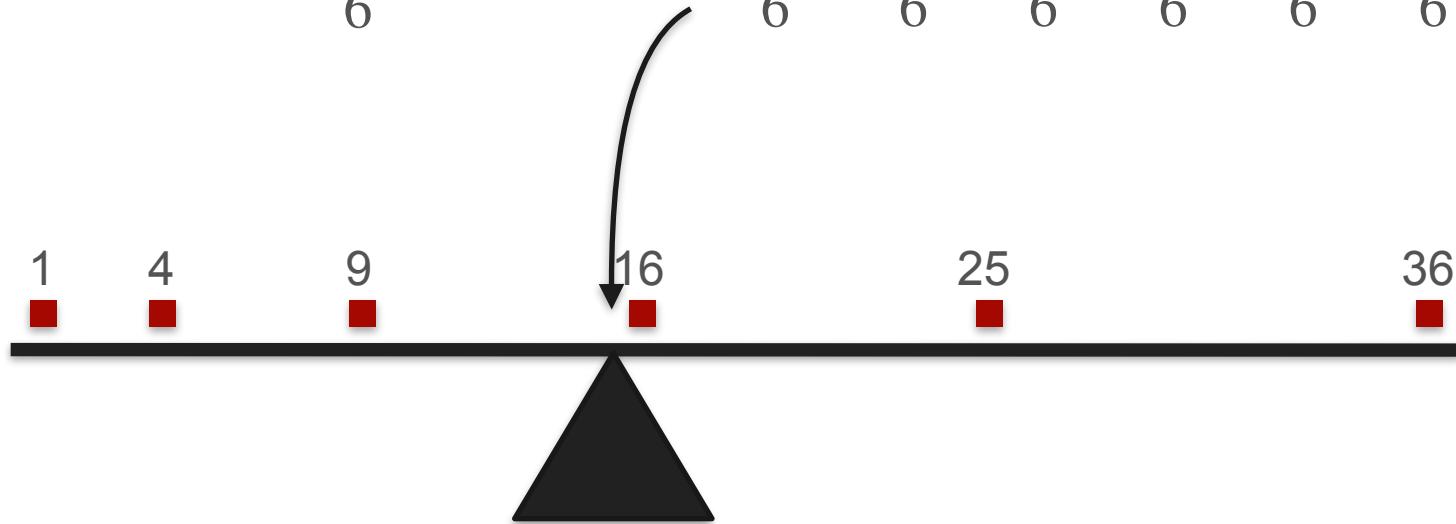
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



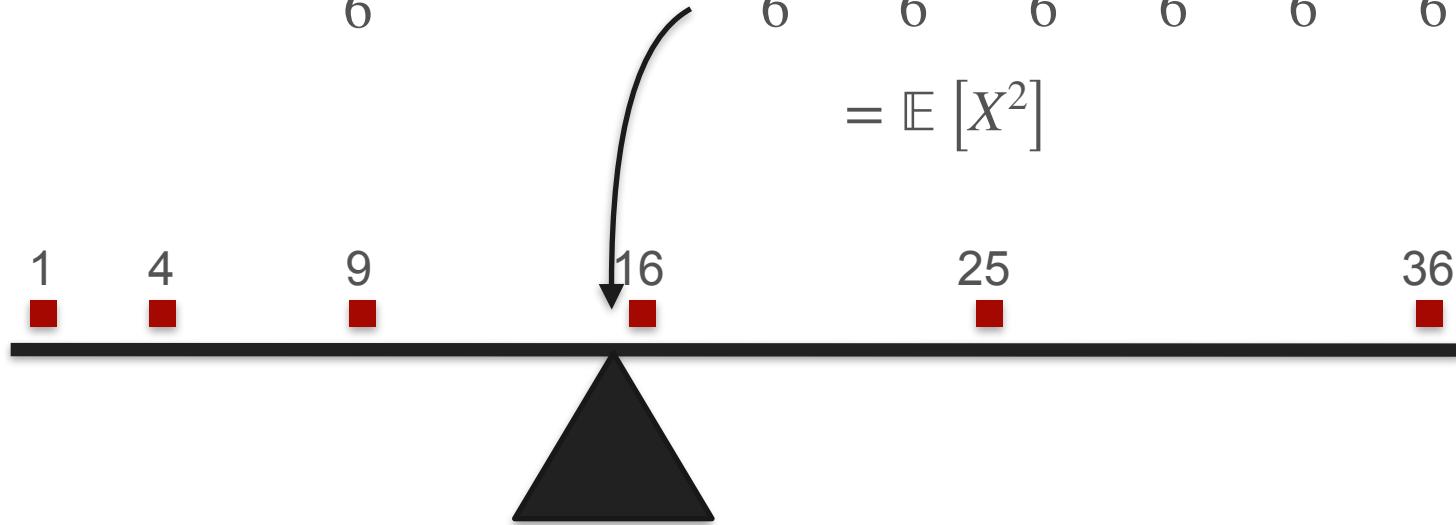
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6} = \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6}$$



Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6} = \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6}$$
$$= \mathbb{E}[X^2]$$





DeepLearning.AI

Describing Distributions

Sum of expectations

Sum of Expectations

You play a game:

Sum of Expectations

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Sum of Expectations

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.



Win \$1



Win nothing

Sum of Expectations

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then roll a die. You win the amount you roll.



Win \$1



Win nothing

Sum of Expectations

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then roll a die. You win the amount you roll.



Win \$1



Win nothing

Win

\$1

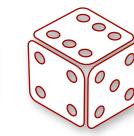
\$2

\$3

\$4

\$5

\$6



Sum of Expectations

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then roll a die. You win the amount you roll.



Win \$1



Win nothing

Win

\$1

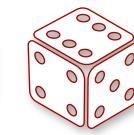
\$2

\$3

\$4

\$5

\$6



What are your expected winnings for the game?

Sum of Expectations



Win \$1



Win nothing

Win

\$1

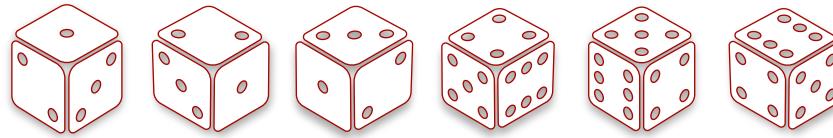
\$2

\$3

\$4

\$5

\$6



Sum of Expectations

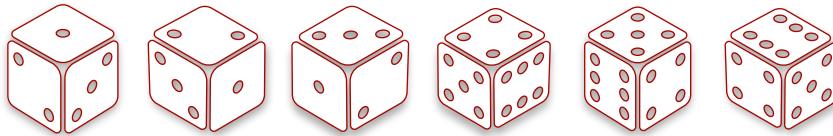


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E} [X_{coin}] = \$0.5$$

Sum of Expectations



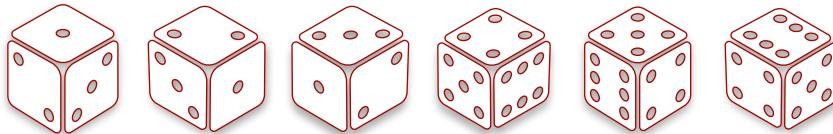
Win \$1



Win nothing

$$\mathbb{E} [X_{coin}] = \$0.5$$

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E} [X_{dice}] = \$3.5$$

Sum of Expectations

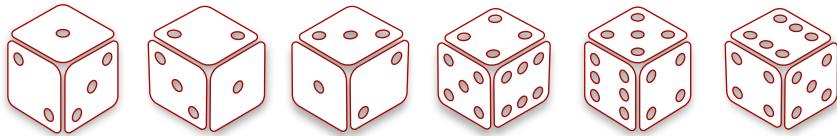


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E}[X_{coin}] = \$0.5$$

$$\mathbb{E}[X_{dice}] = \$3.5$$

$$\mathbb{E}[X] = \mathbb{E}[X_{coin}] + \mathbb{E}[X_{dice}] = \$0.5 + \$3.5 = \$4$$

Sum of Expectations

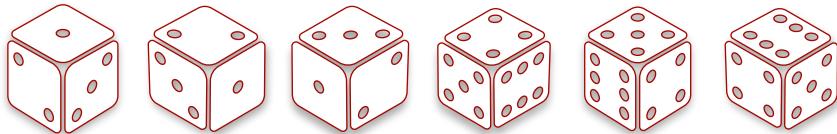


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E}[X_{coin}] = \$0.5$$

$$\mathbb{E}[X_{dice}] = \$3.5$$

$$\mathbb{E}[X] = \mathbb{E}[X_{coin}] + \mathbb{E}[X_{dice}] = \$0.5 + \$3.5 = \$4$$

In general: $\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$

Sum of Expectations



8 billion people

Sum of Expectations



8 billion people

Sum of Expectations



8 billion people

Sum of Expectations



Expected number of
correct assignments?



8 billion people

Sum of Expectations



1

Expected number of
correct assignments?



8 billion people

Sum of Expectations



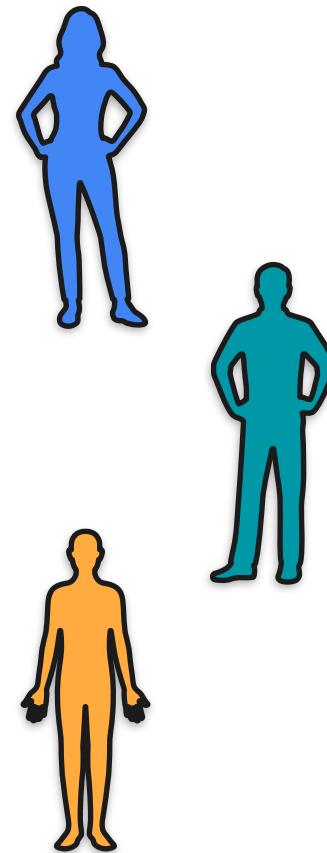
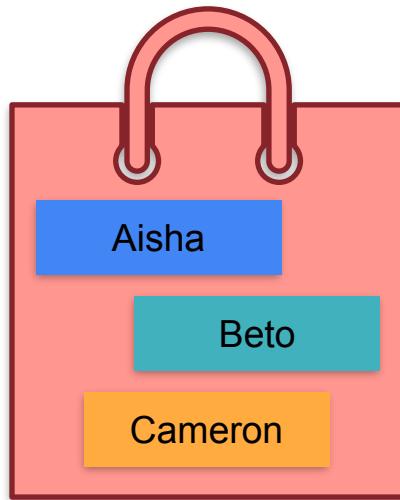
1

Expected number of
correct assignments?

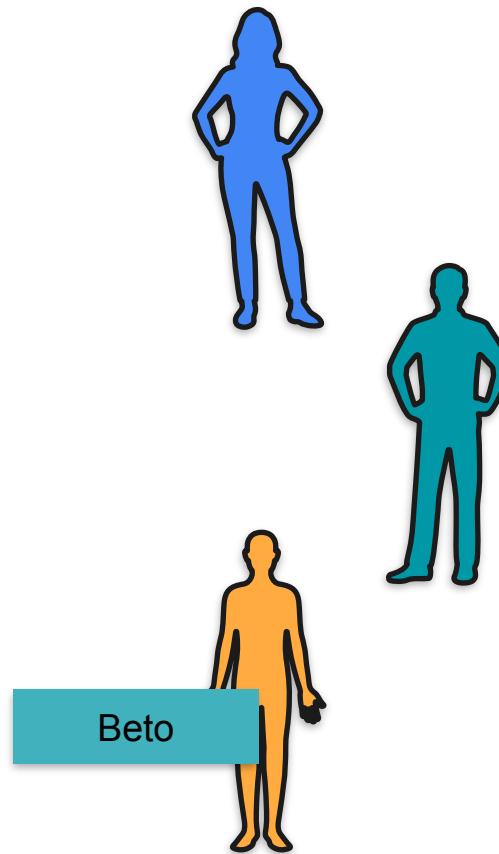
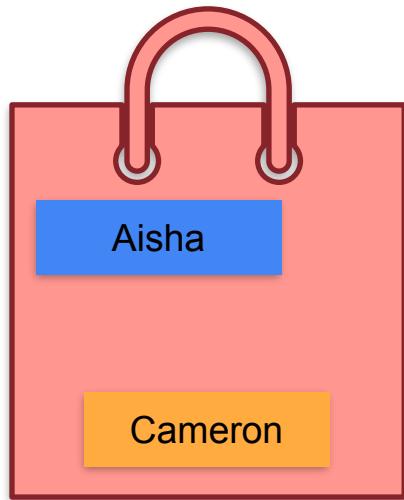


8 billion people

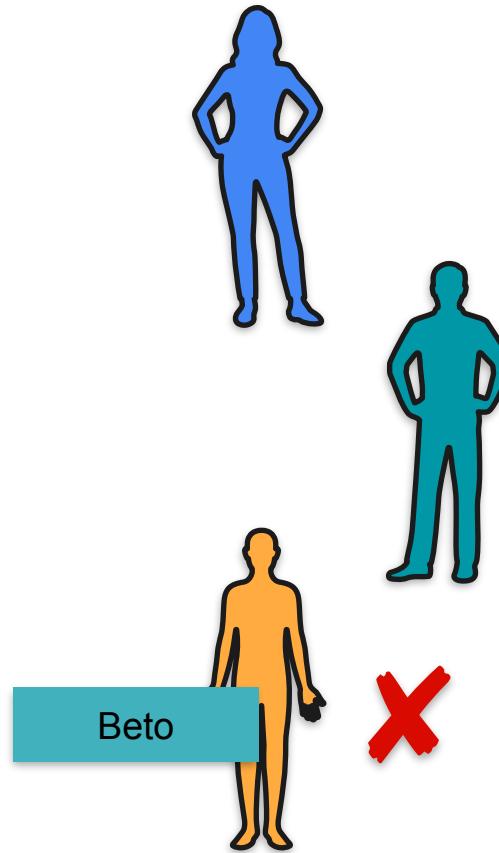
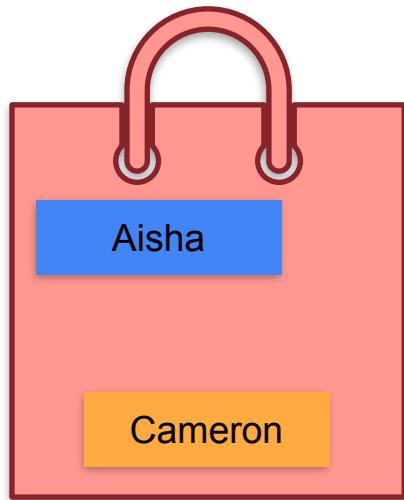
Sum of Expectations



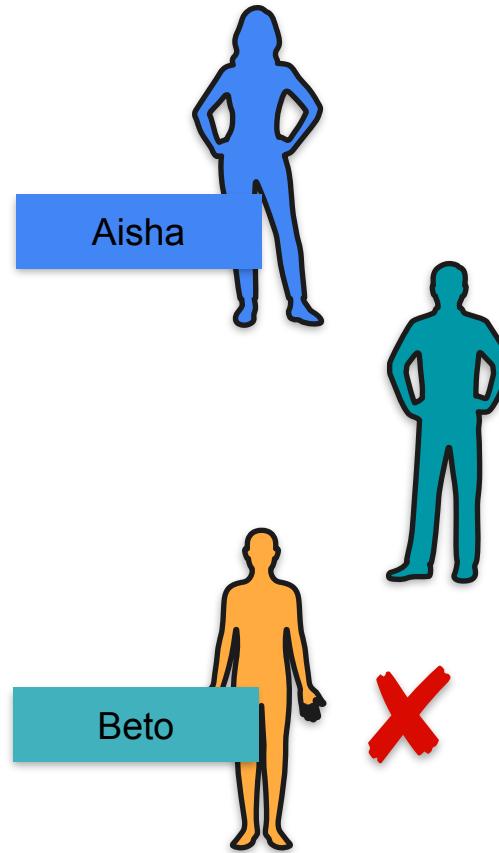
Sum of Expectations



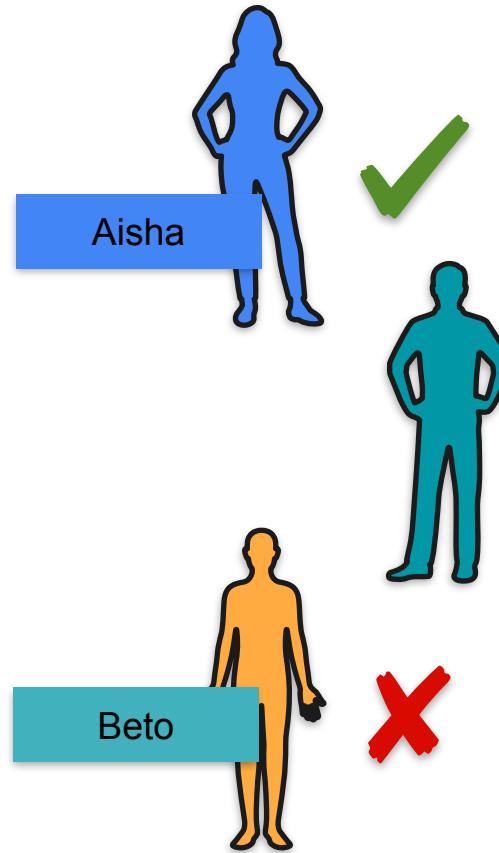
Sum of Expectations



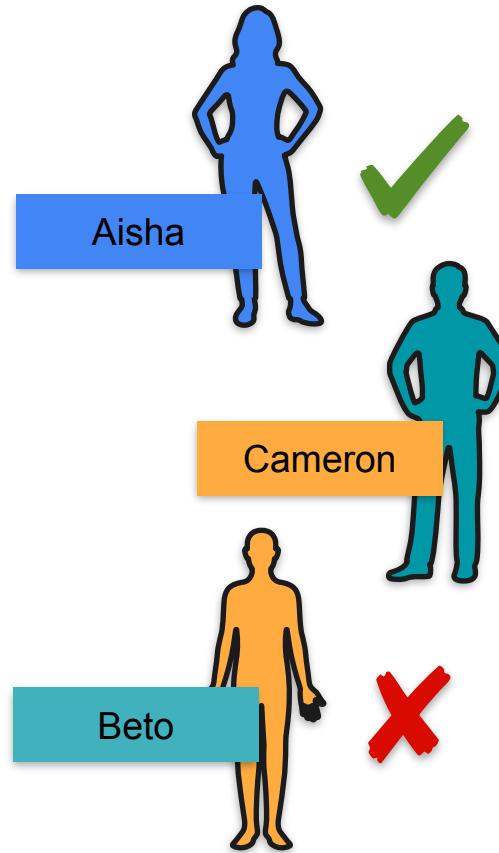
Sum of Expectations



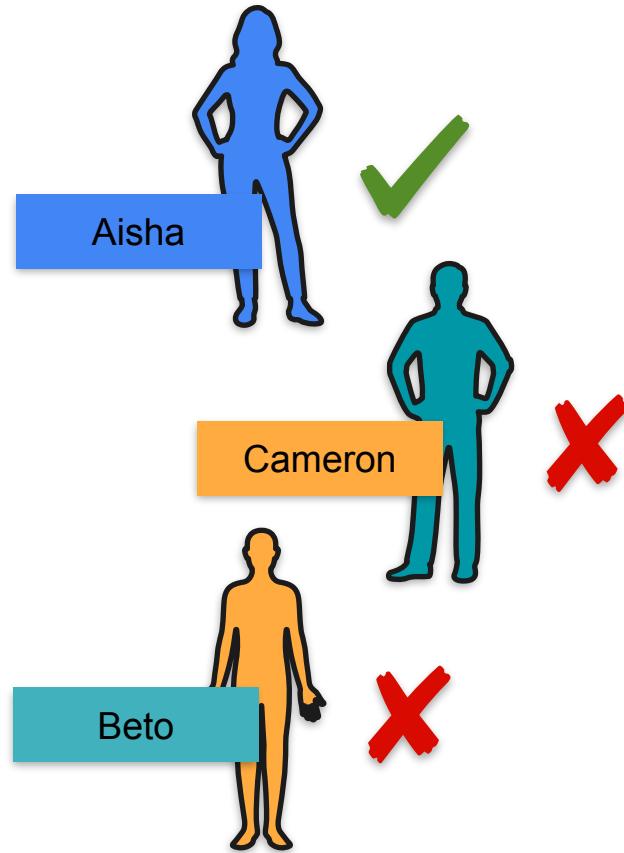
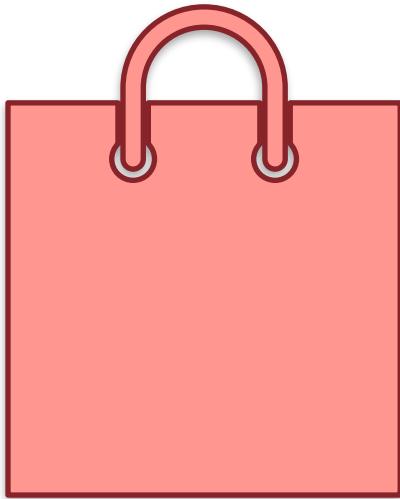
Sum of Expectations



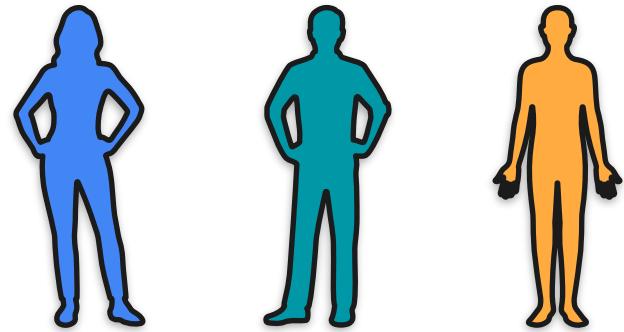
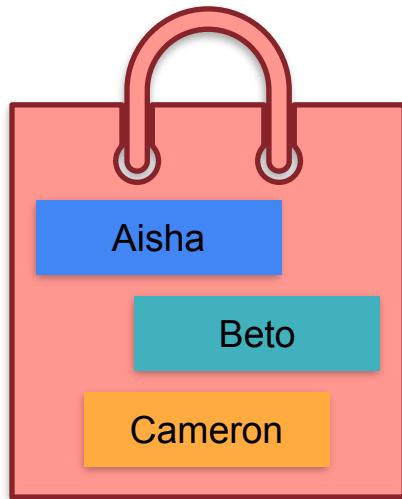
Sum of Expectations



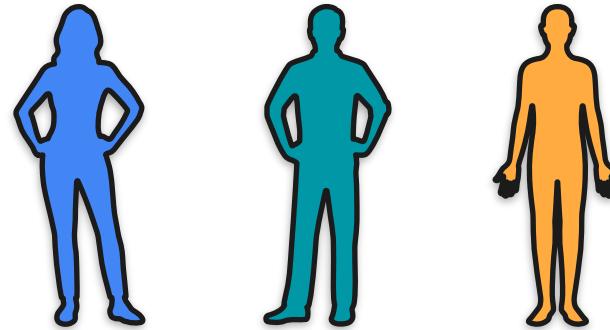
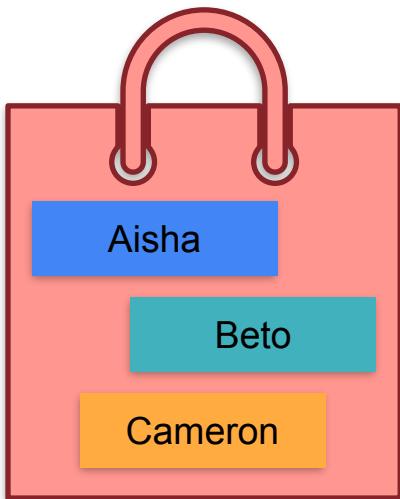
Sum of Expectations



Sum of Expectations

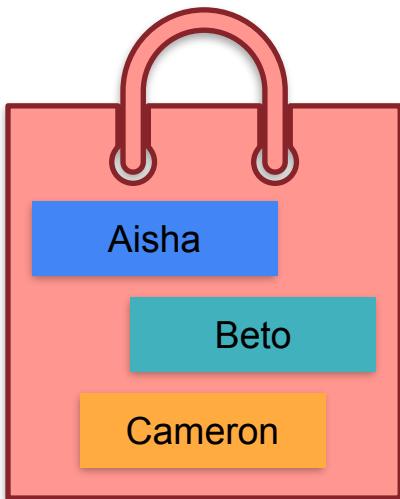


Sum of Expectations

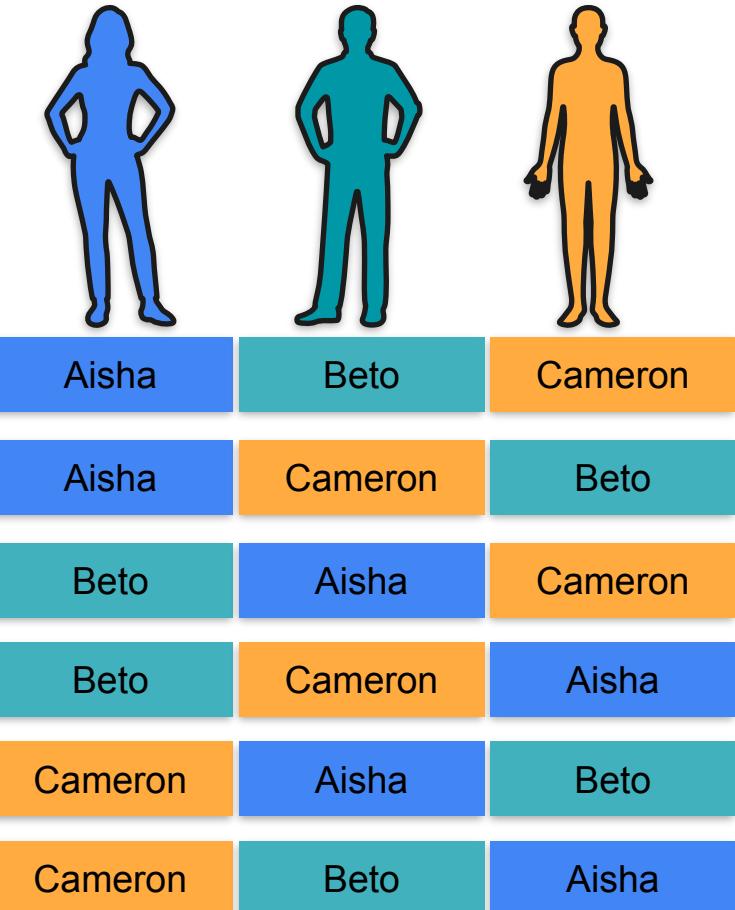


Aisha	Beto	Cameron
Aisha	Cameron	Beto
Beto	Aisha	Cameron
Beto	Cameron	Aisha
Cameron	Aisha	Beto
Cameron	Beto	Aisha

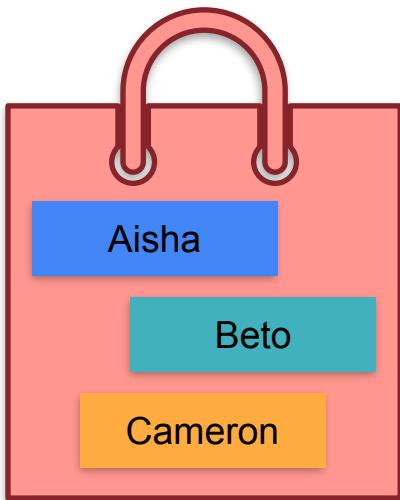
Sum of Expectations



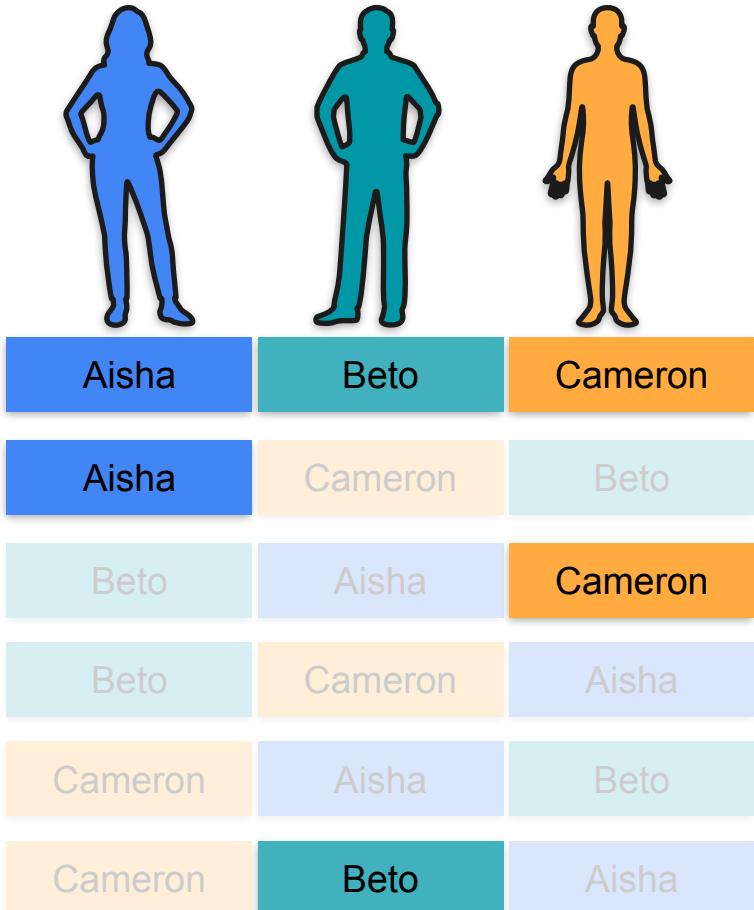
Correct



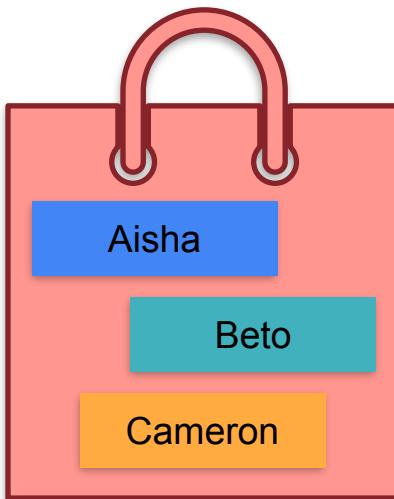
Sum of Expectations



Correct

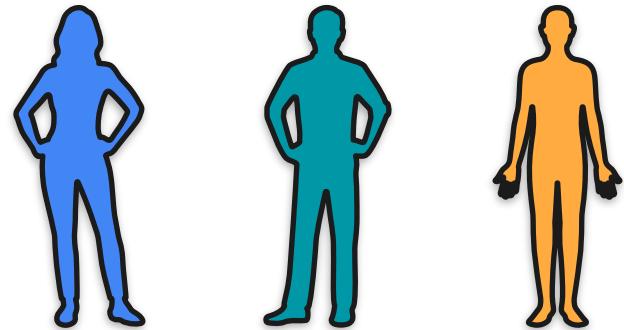


Sum of Expectations

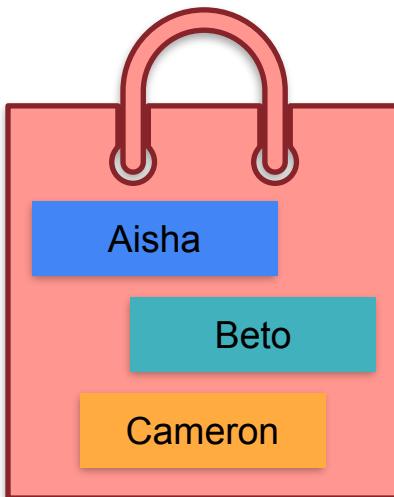


Correct

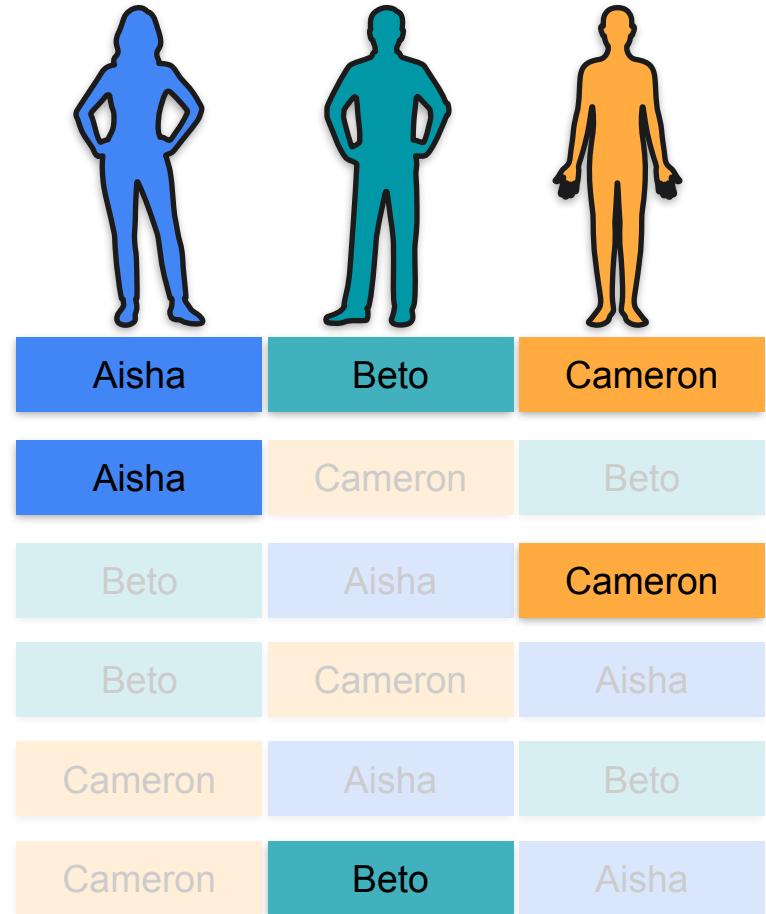
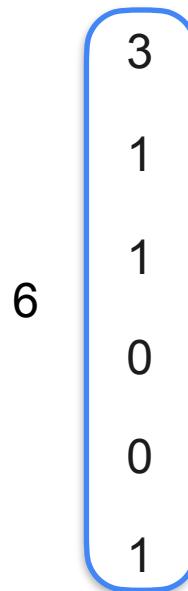
3	Aisha	Beto	Cameron
1	Aisha	Cameron	Beto
1	Beto	Aisha	Cameron
0	Beto	Cameron	Aisha
0	Cameron	Aisha	Beto
1	Cameron	Beto	Aisha



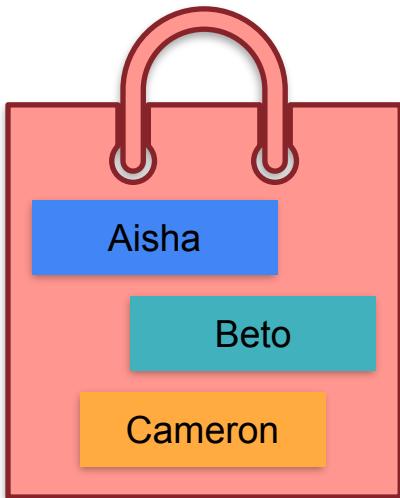
Sum of Expectations



Correct



Sum of Expectations

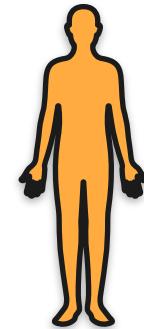
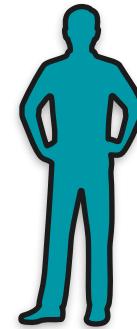
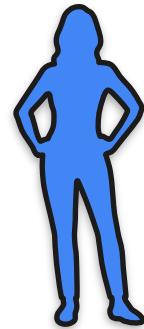
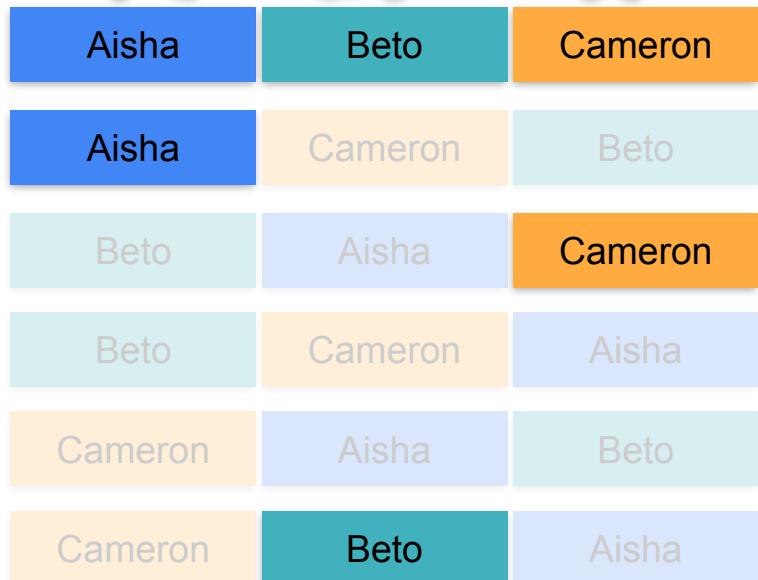


Average
1

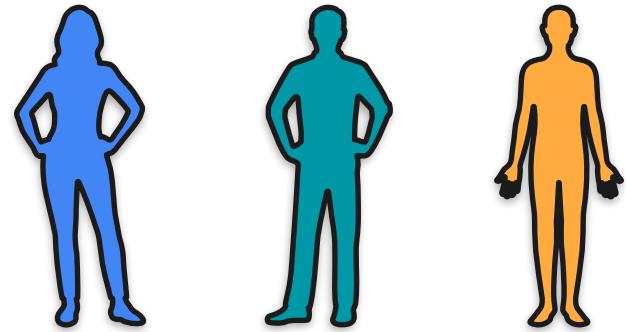
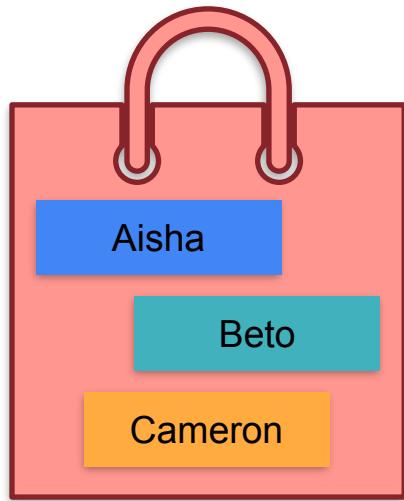
Correct

3
1
1
0
0
1

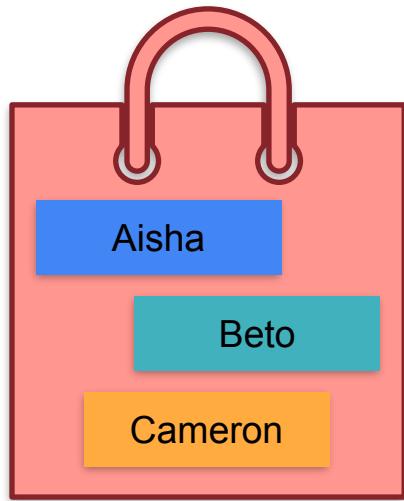
6



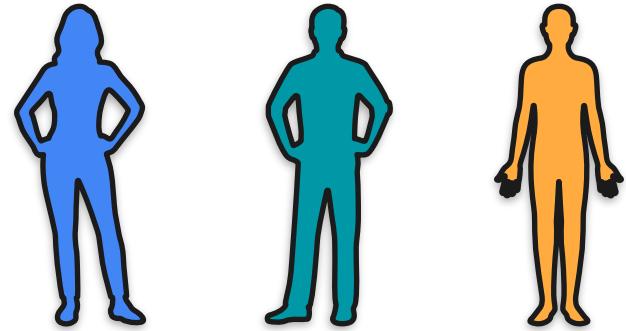
Sum of Expectations



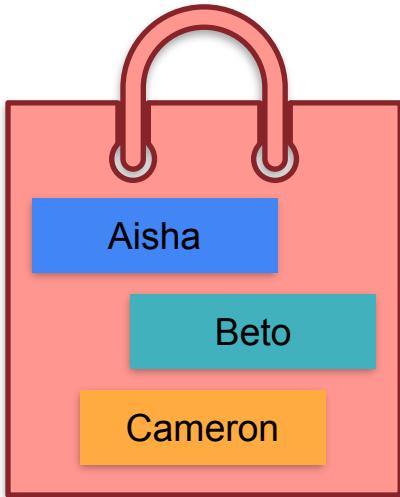
Sum of Expectations



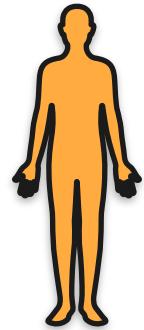
$\mathbb{E}[\text{Matches}]$



Sum of Expectations



$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A]$$



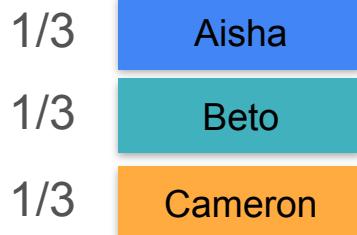
Sum of Expectations



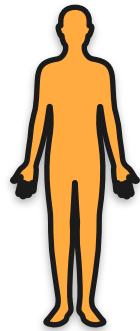
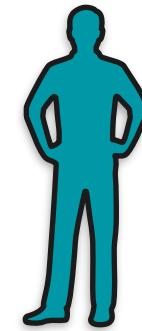
$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A]$$



Sum of Expectations

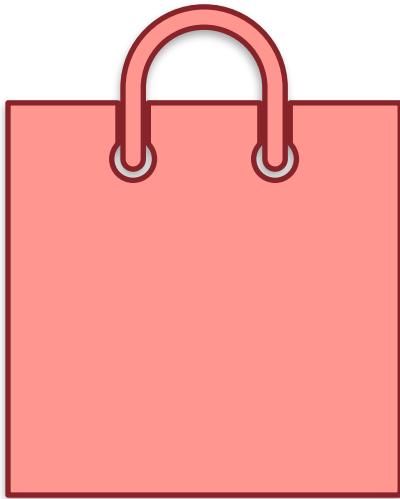


$\mathbb{E}[\text{Matches}]$



$= \mathbb{E}[A]$

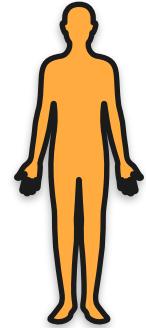
Sum of Expectations



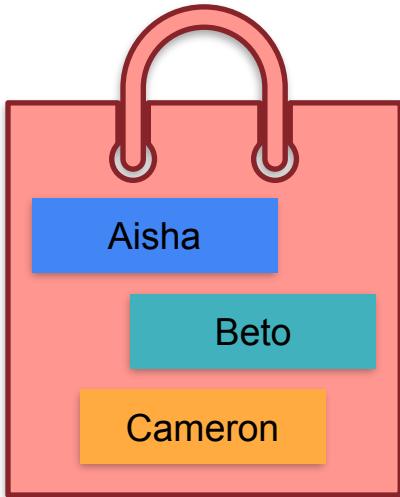
$\mathbb{E}[\text{Matches}]$

$= \mathbb{E}[A]$

$$= 1/3$$



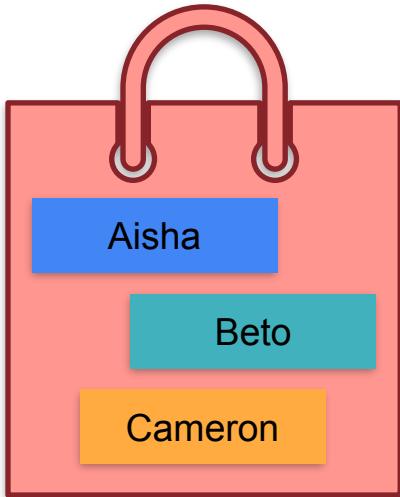
Sum of Expectations



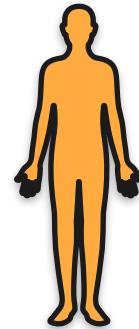
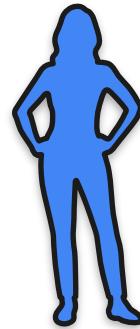
$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A] = 1/3$$



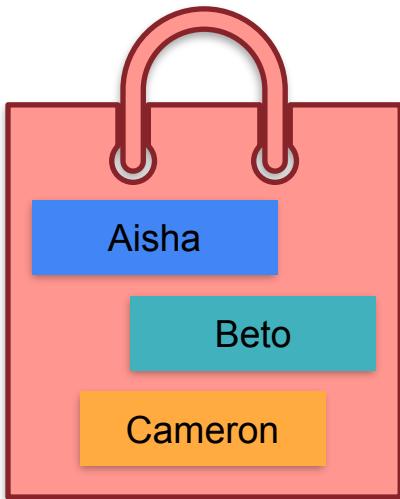
Sum of Expectations



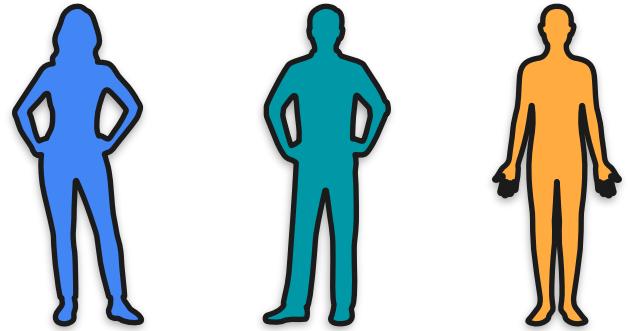
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] \\ &= 1/3 + 1/3\end{aligned}$$



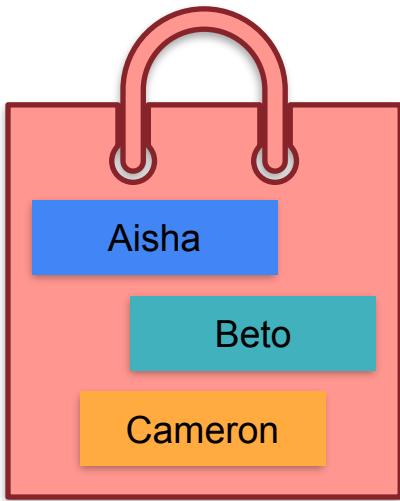
Sum of Expectations



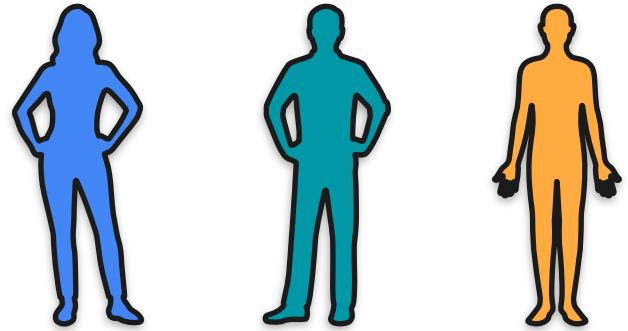
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3\end{aligned}$$



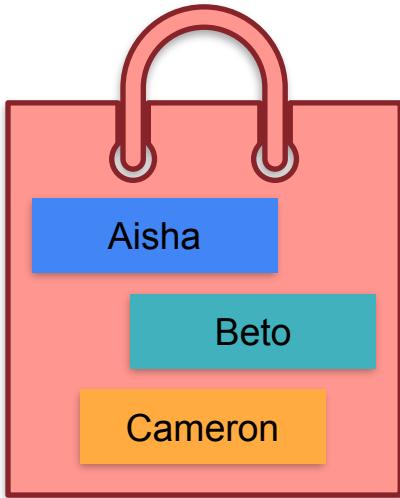
Sum of Expectations



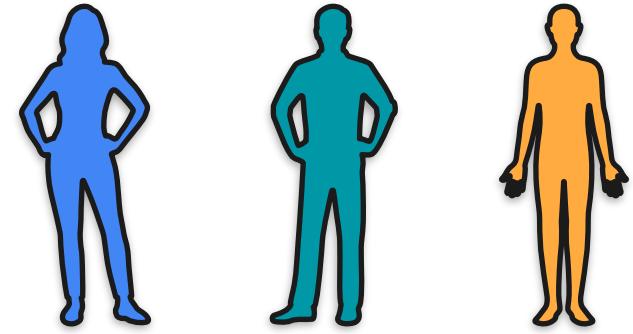
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3 \\ &= 1\end{aligned}$$



Sum of Expectations



$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3 \\ &= 1\end{aligned}$$



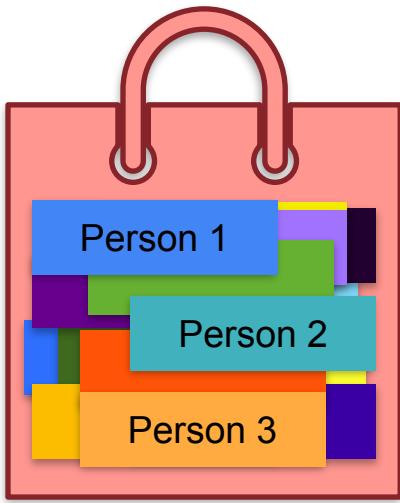
Average
1

Sum of Expectations



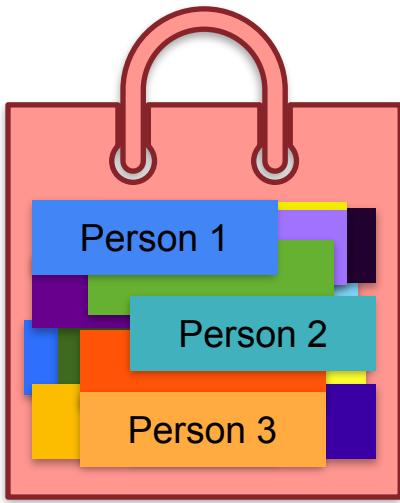
8 billion people

Sum of Expectations



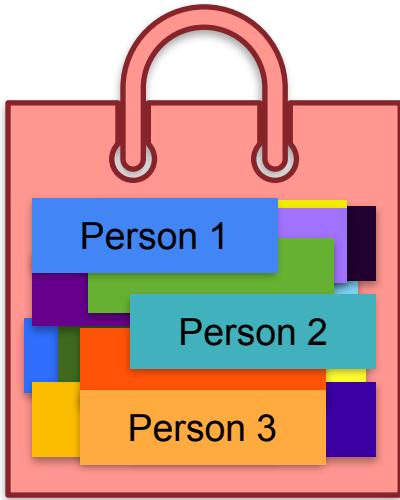
8 billion people

Sum of Expectations



8 billion people

Sum of Expectations



Expected number = ?



8 billion people

Sum of Expectations



Sum of Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}]$$

Sum of Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}]$$

n people ($n = 8$ billion)

Sum of Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \quad \overbrace{\qquad \qquad \qquad}^{\text{n people (n = 8 billion)}} = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}$$

Sum of Expectations



$$\begin{aligned} \mathbb{E} [\text{Matches}] &= \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \\ &\quad \overbrace{\qquad\qquad\qquad}^{\text{n people (n = 8 billion)}} \\ &= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} \\ &= n \cdot \frac{1}{n} \end{aligned}$$

Sum of Expectations



$$\begin{aligned} \mathbb{E} [\text{Matches}] &= \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \\ &\quad \overbrace{\hspace{10em}}^{\text{n people (n = 8 billion)}} \\ &= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} \\ &= n \cdot \frac{1}{n} = 1 \end{aligned}$$

Sum of Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}]$$

n people ($n = 8$ billion)

$$= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}$$

In general:

$$\mathbb{E} [X_1 + X_2 + \dots + X_n] = \mathbb{E} [X_1] + \mathbb{E} [X_2] + \dots + \mathbb{E} [X_n]$$

$$= n \cdot \frac{1}{n} = 1$$



DeepLearning.AI

Describing Distributions

Variance

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

Game cost:

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

Game cost:

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

How do you tell these two games apart?

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 100 dollars



You lose 100 dollars

Variance Motivation: Fair Price To Play the Game

How do you tell these two games apart?

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



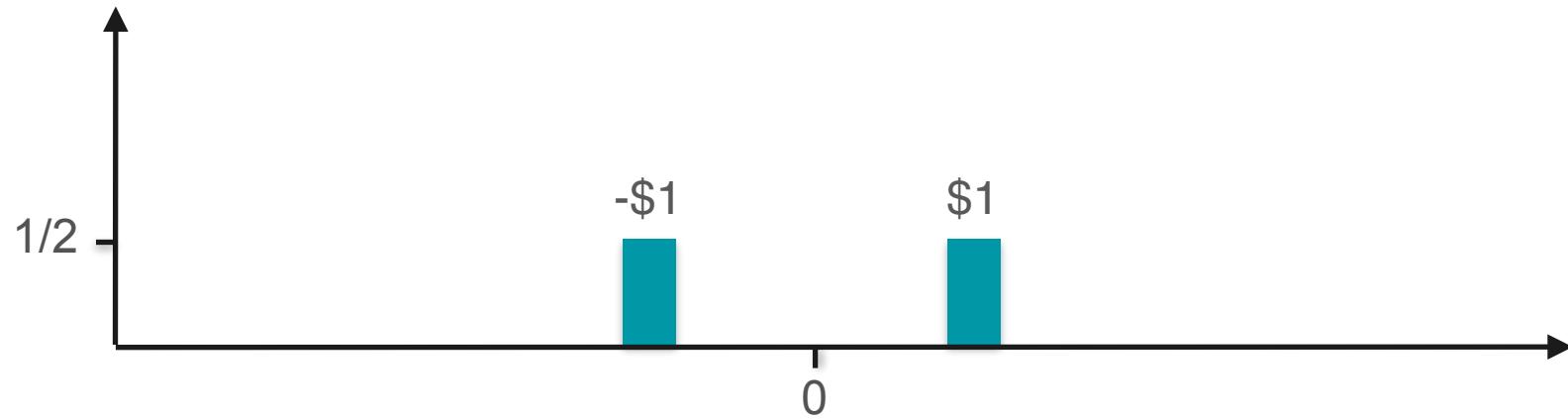
You win 100 dollars

You lose 100 dollars

Variance!

Variance Motivation: Measuring Spread

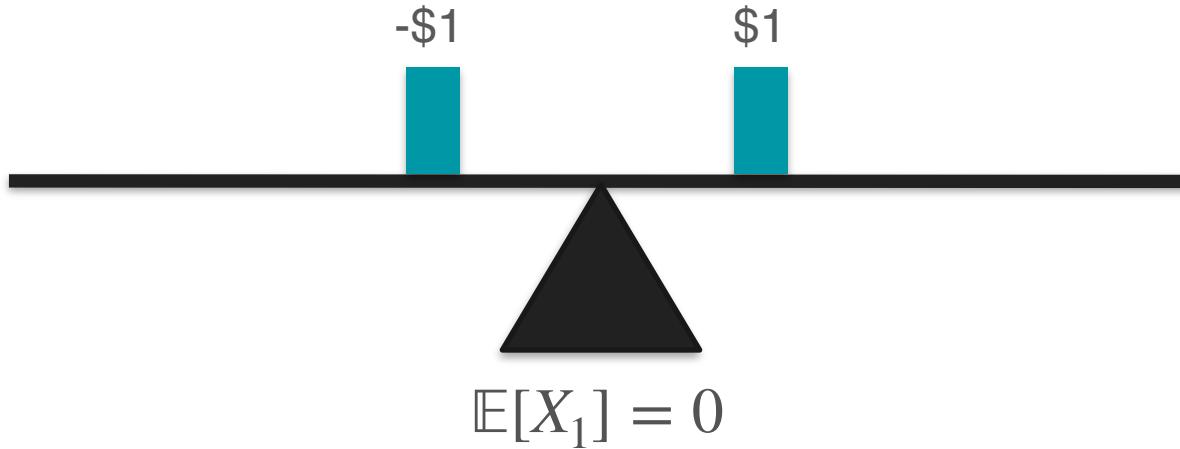
Probability



Variance Motivation: Measuring Spread

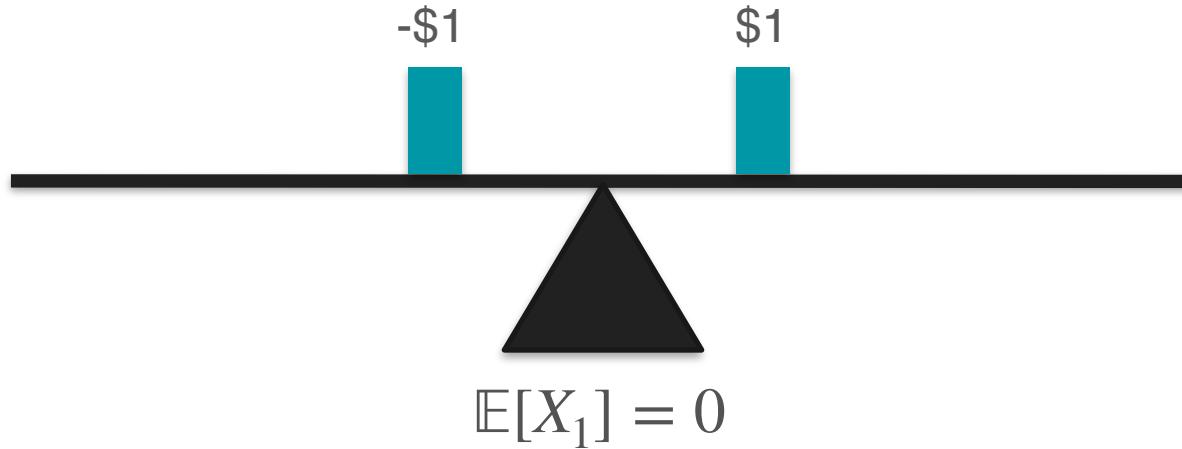


Variance Motivation: Measuring Spread

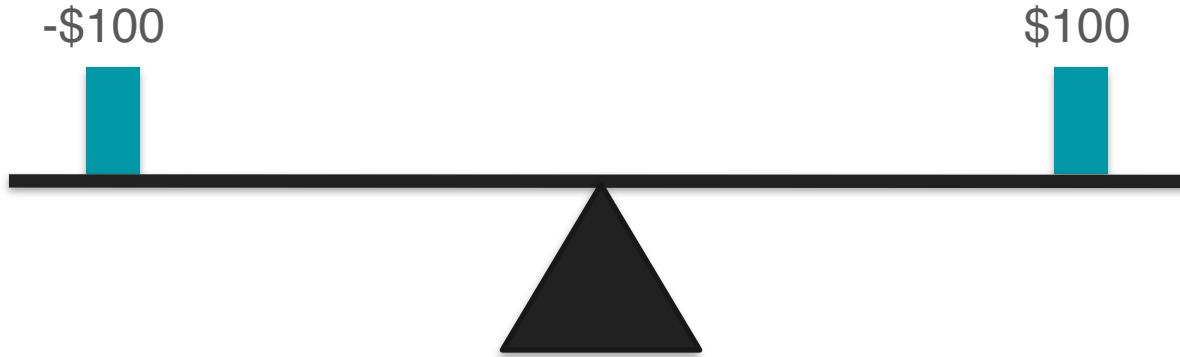


Variance Motivation: Measuring Spread

X_1 = expected amount of money gained in game 1

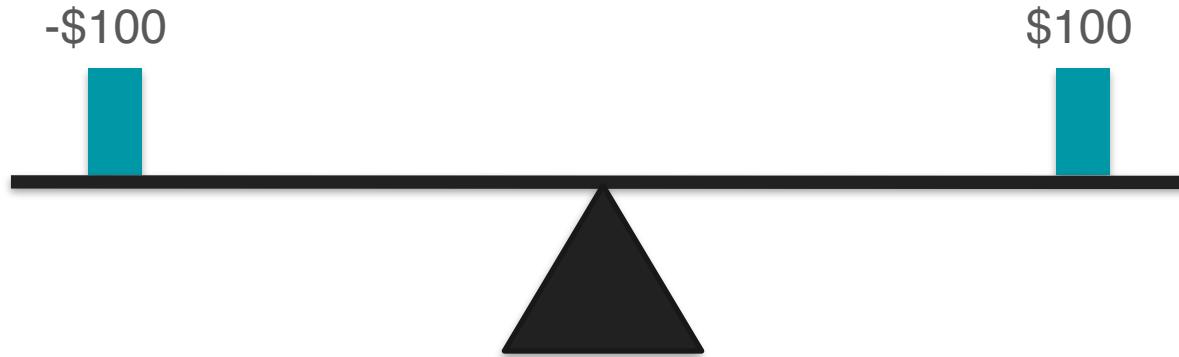


Variance Motivation: Measuring Spread



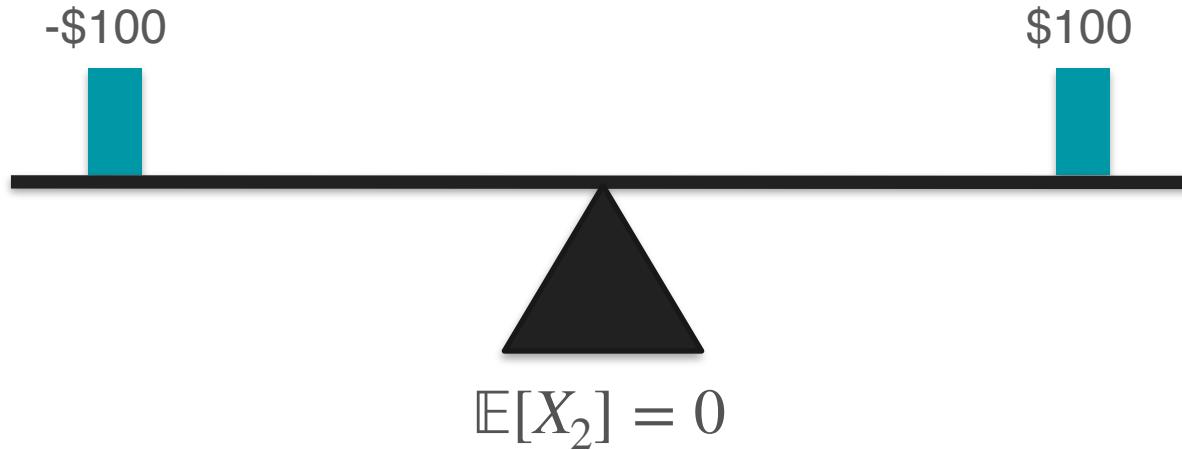
Variance Motivation: Measuring Spread

X_2 = expected amount of money gained in game 2

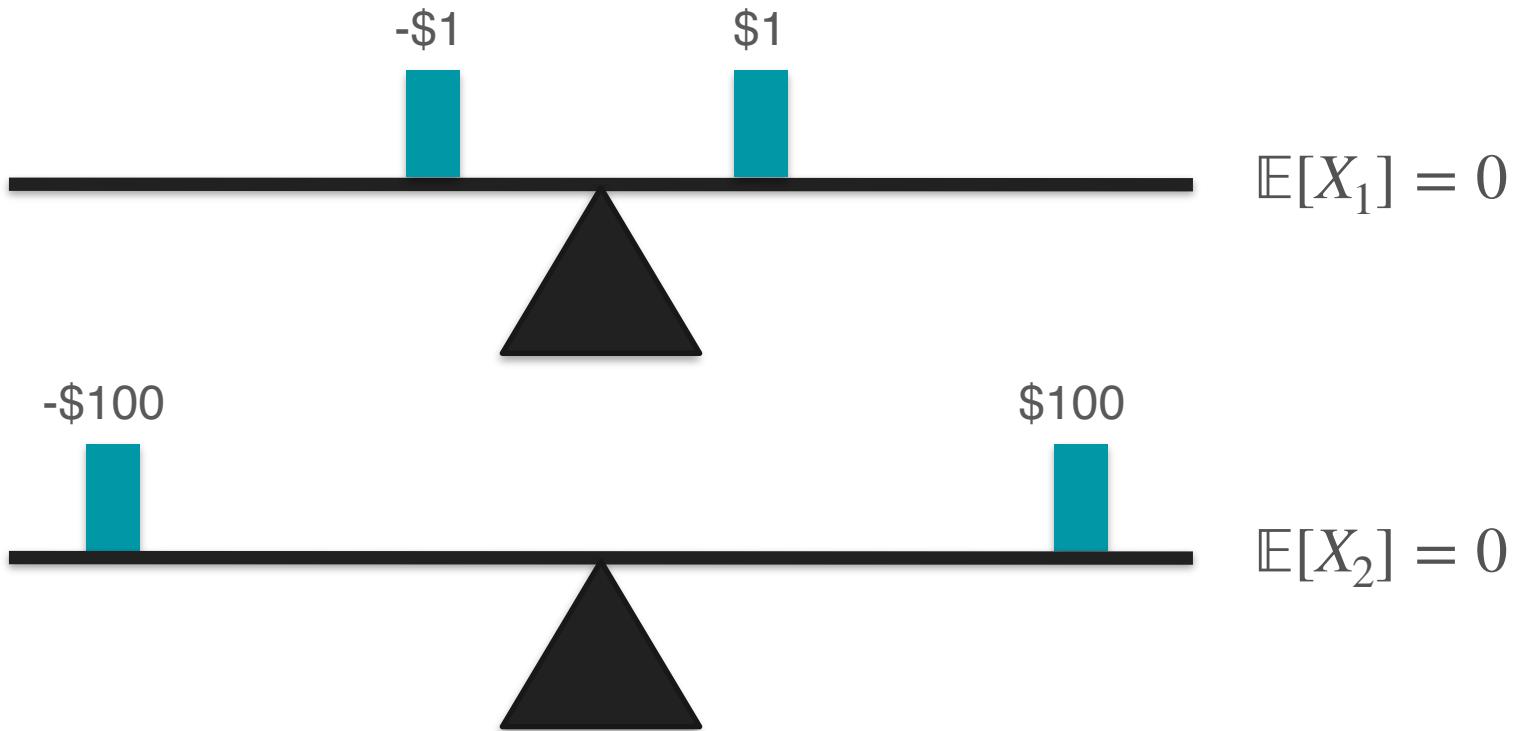


Variance Motivation: Measuring Spread

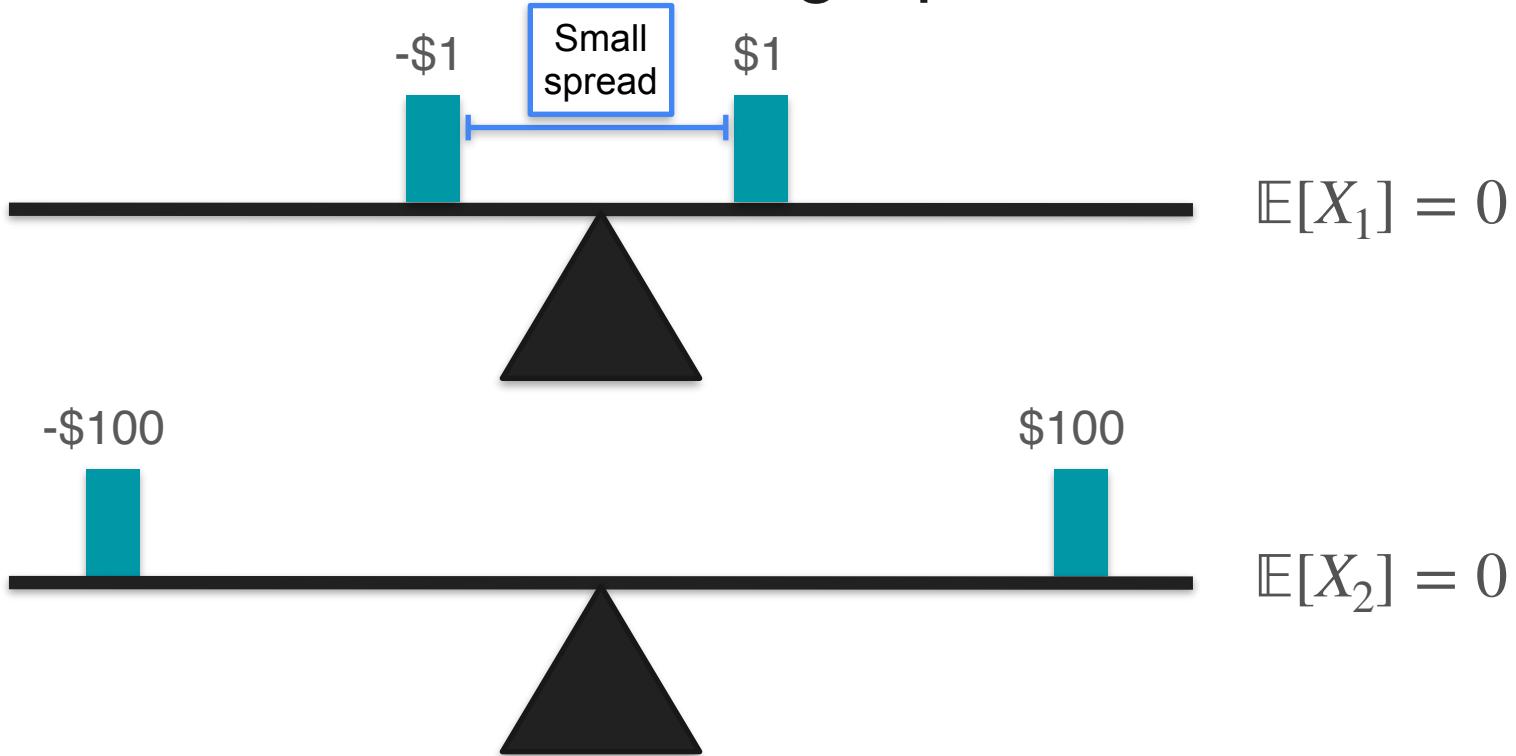
X_2 = expected amount of money gained in game 2



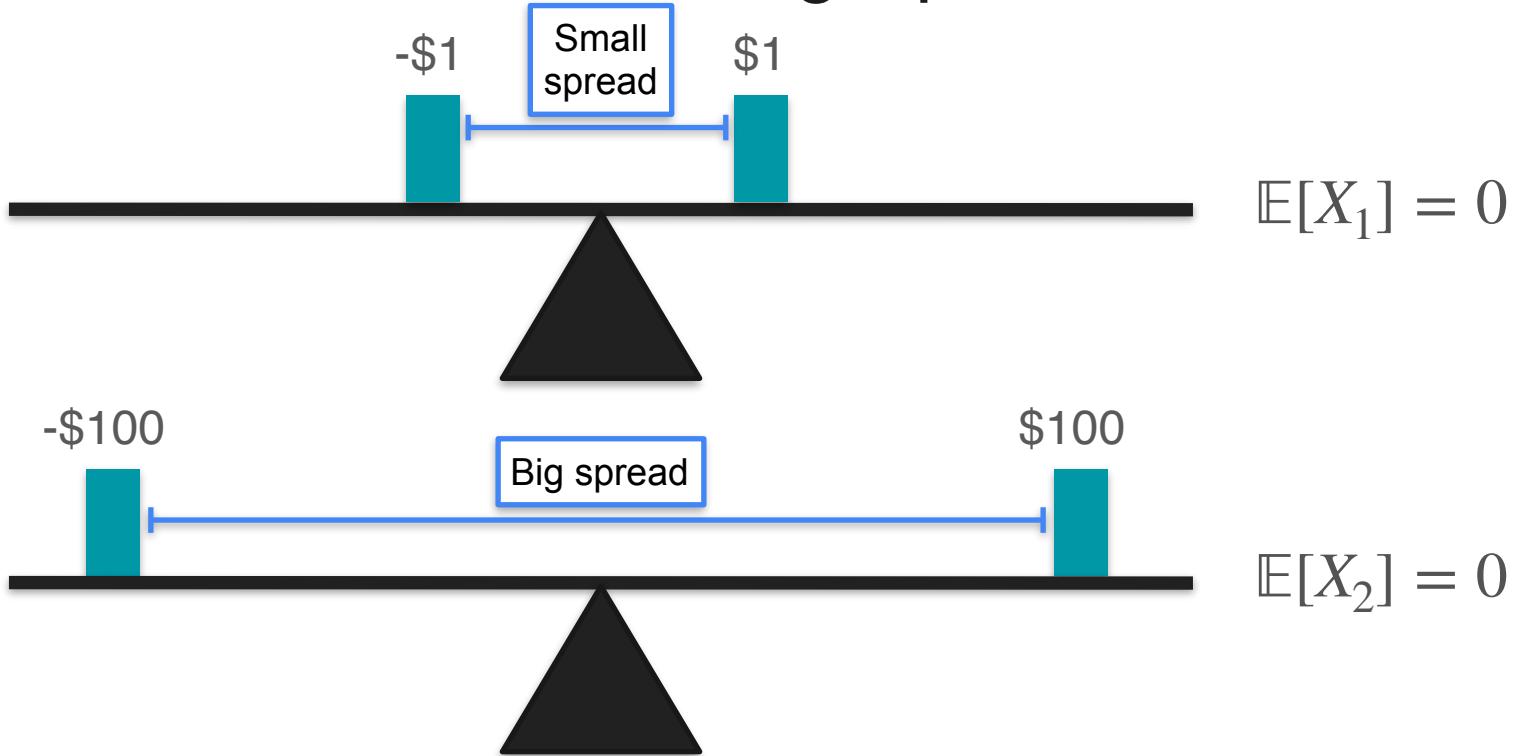
Variance Motivation: Measuring Spread



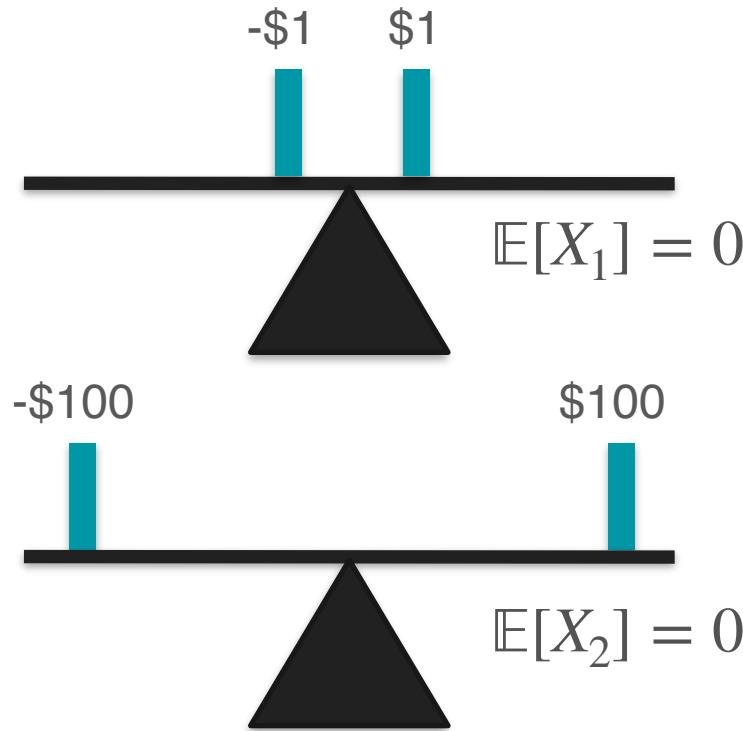
Variance Motivation: Measuring Spread



Variance Motivation: Measuring Spread



Variance Motivation: Measuring Spread



Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$



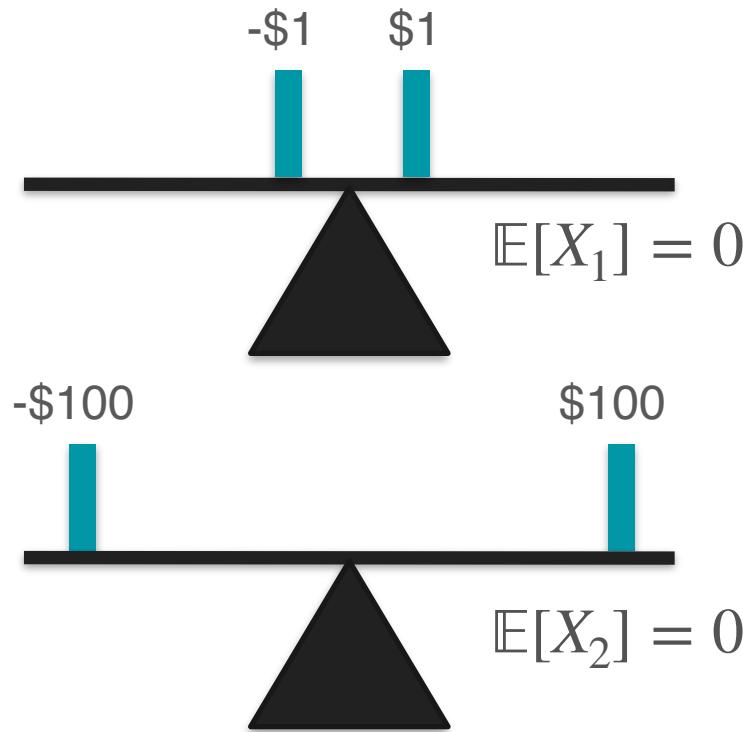
Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$

$$\mathbb{E}[X_2] = \frac{(100) + (-100)}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$

Turn into positive?

$$\mathbb{E}[X_2] = \frac{(100) + (-100)}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1)^2 + (-1)^2}{2} = 0$$

$$\mathbb{E}[X_2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1)^2 + (-1)^2}{2} = 0$$

$$\mathbb{E}[X_2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 0$$



$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

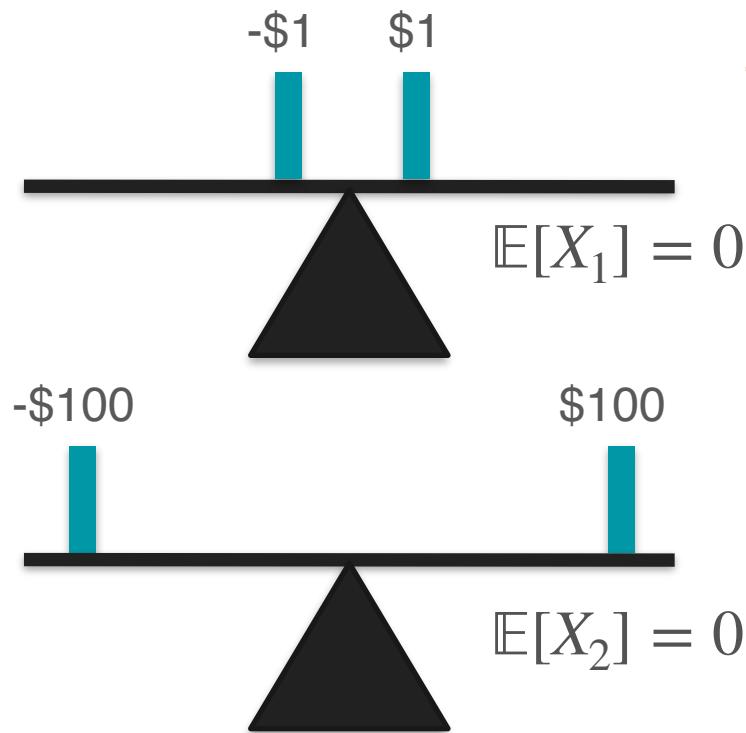
Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Variance Motivation: Measuring Spread

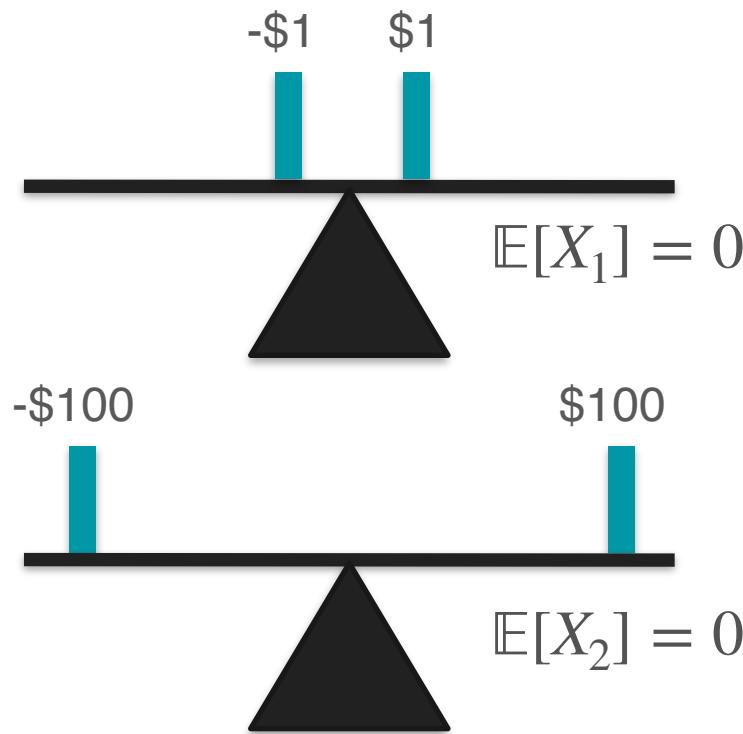


Key for telling game
1 and game 2 apart!

$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Variance Motivation: Measuring Spread



Key for telling game
1 and game 2 apart!

$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Measure of spread

Variance

$$\mathbb{E}[X^2]$$

Variance

$$\mathbb{E}[X^2]$$

Almost...

Variance Motivation: Centering With Mean

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

How risky are these two games in comparison?

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

How risky are these two games in comparison?

Hint: Think of the spread

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

They are equally risky

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1$$

Game 2



You win 6 dollars



You win 4 dollars

$$\mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1$$



Same risk?



$$\mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$

Game 2



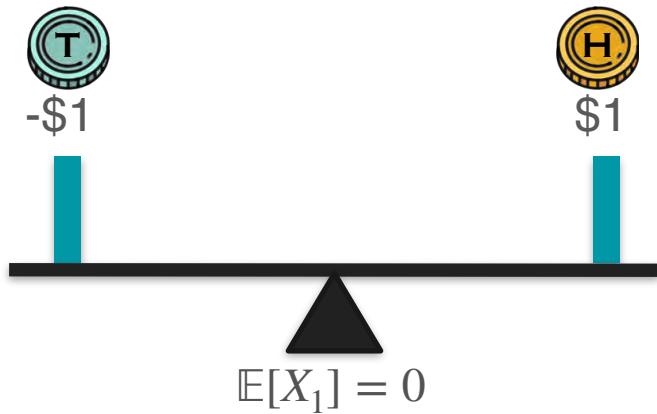
You win 6 dollars



You win 4 dollars

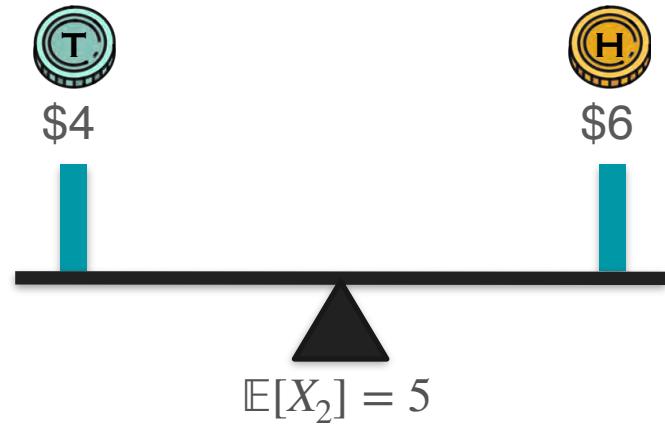
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

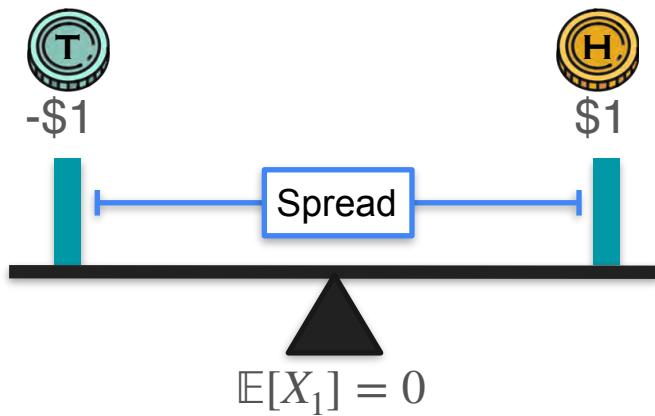
$$\mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$



Game 2

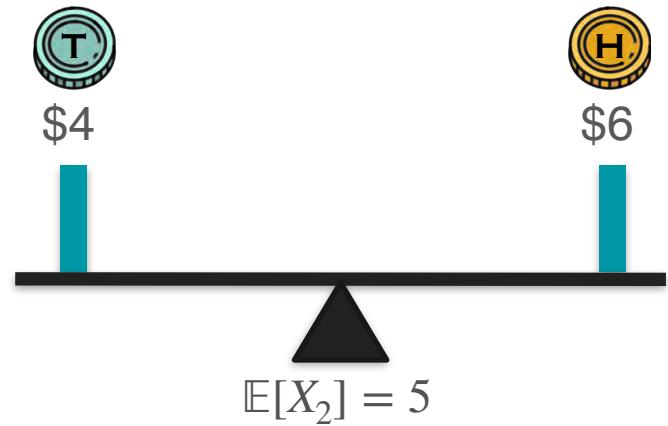
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

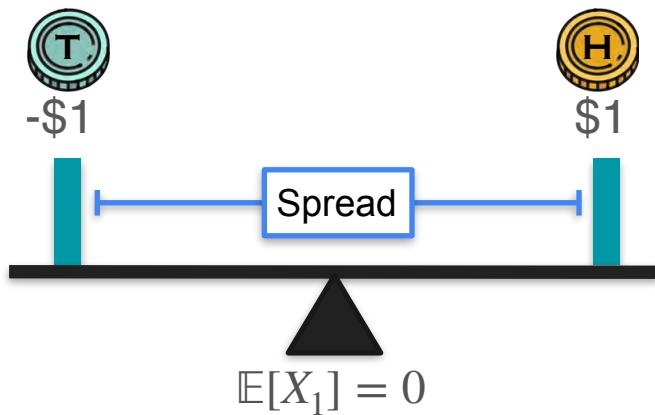
$$\mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$



Game 2

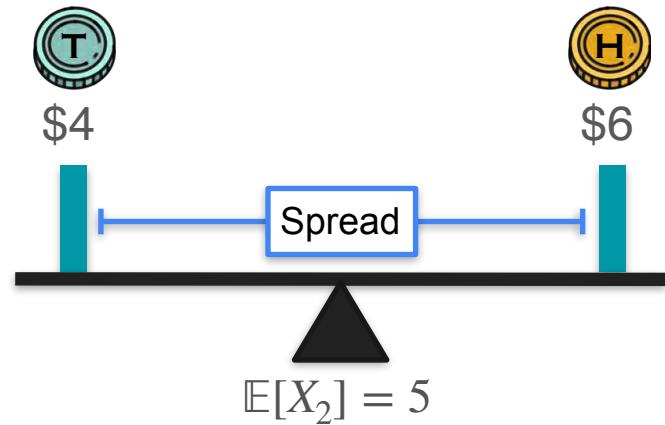
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

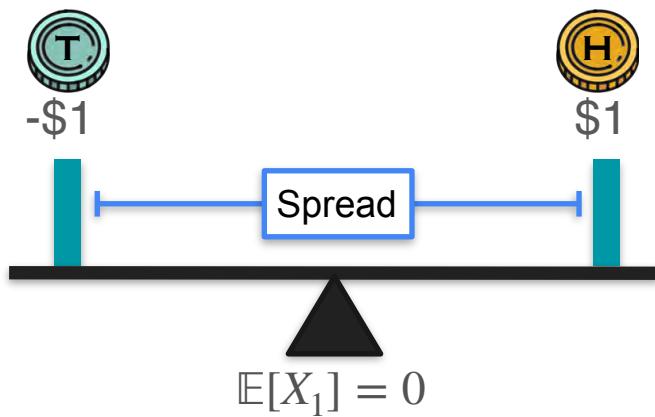
$$\mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$



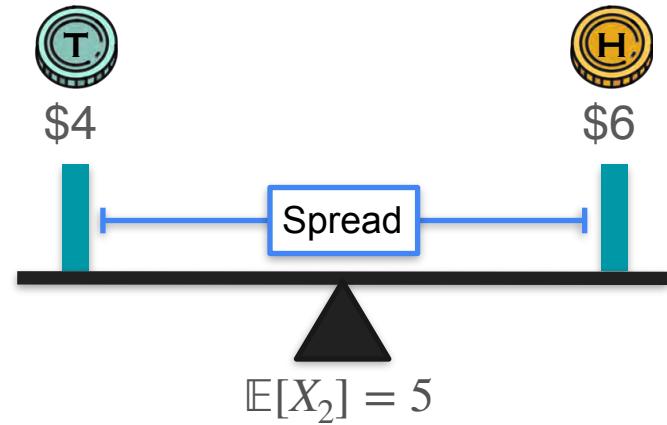
Game 2

Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1^2] = \frac{(-1)^2 + (1)^2}{2} = 1 \quad \leftarrow \text{Same spread?} \rightarrow \quad \mathbb{E}[X_2^2] = \frac{(4)^2 + (6)^2}{2} = 26$$



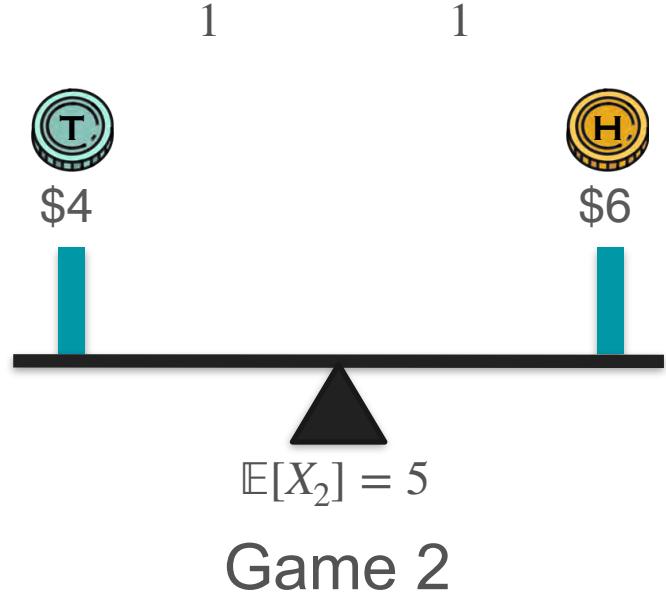
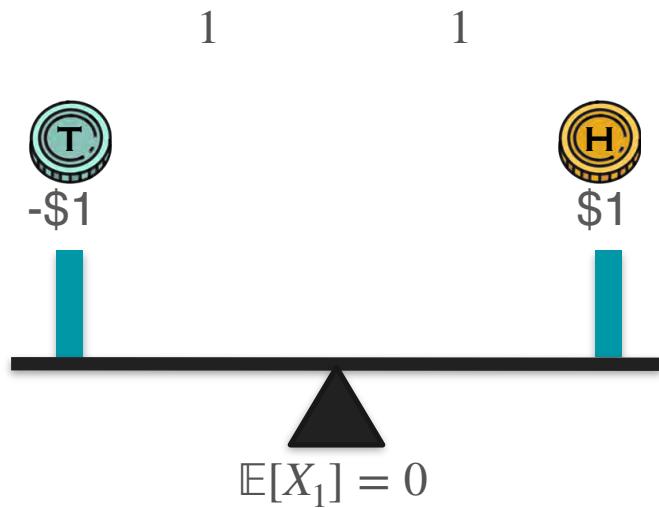
Game 1



Game 2

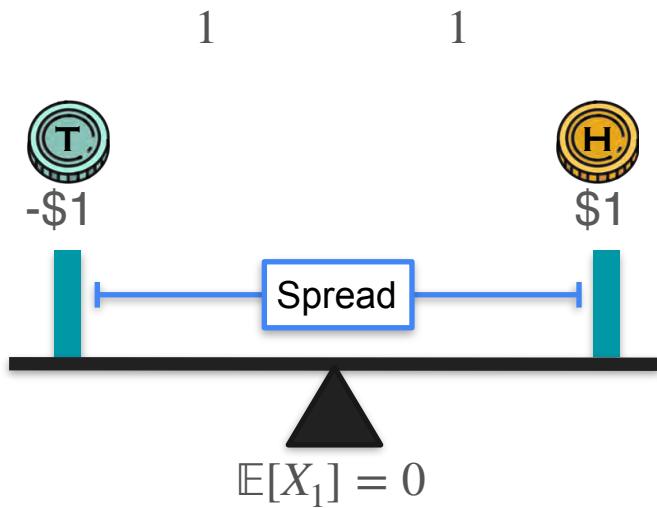
Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$

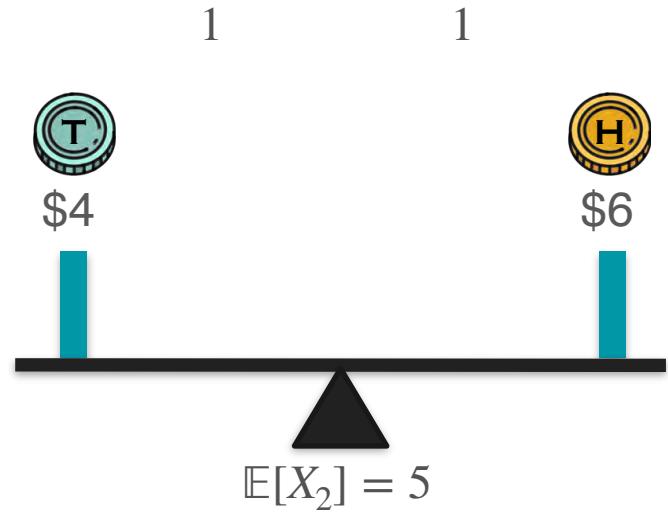


Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



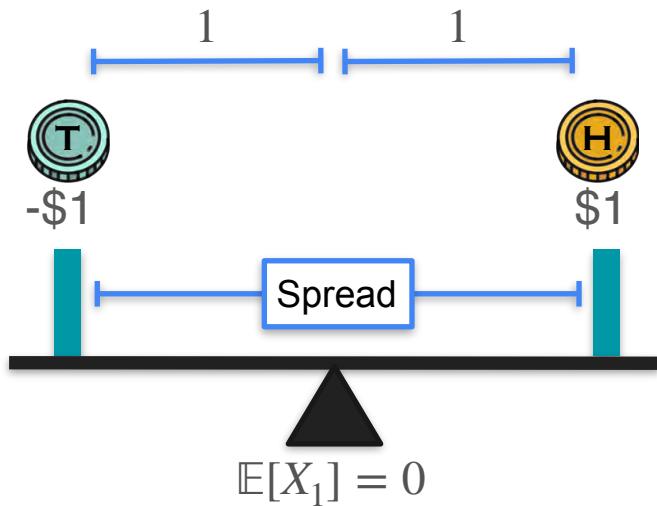
Game 1



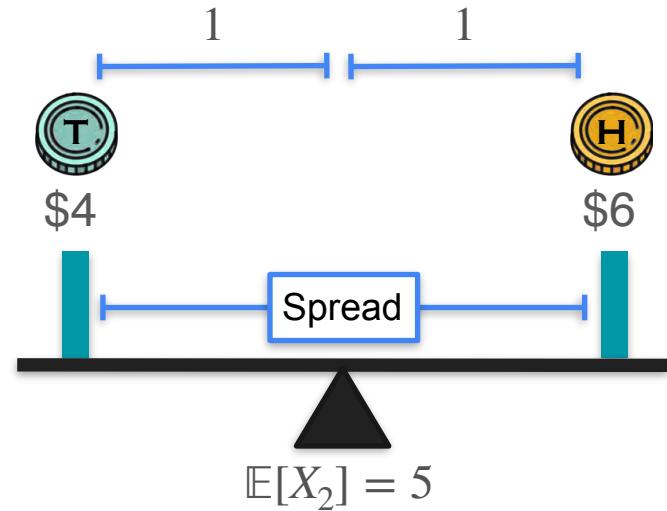
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



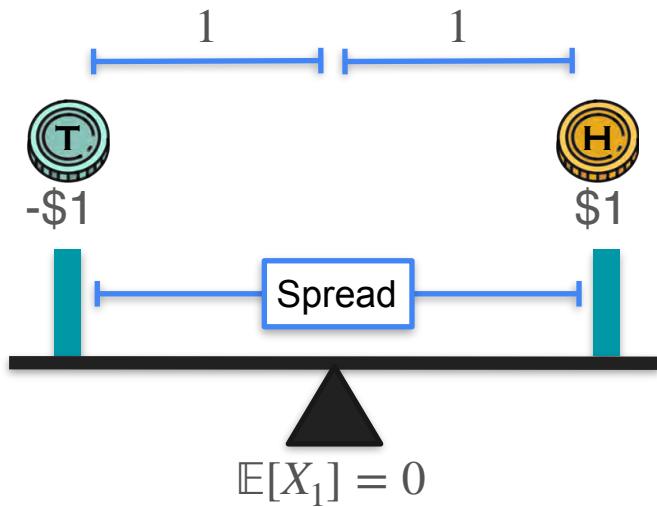
Game 1



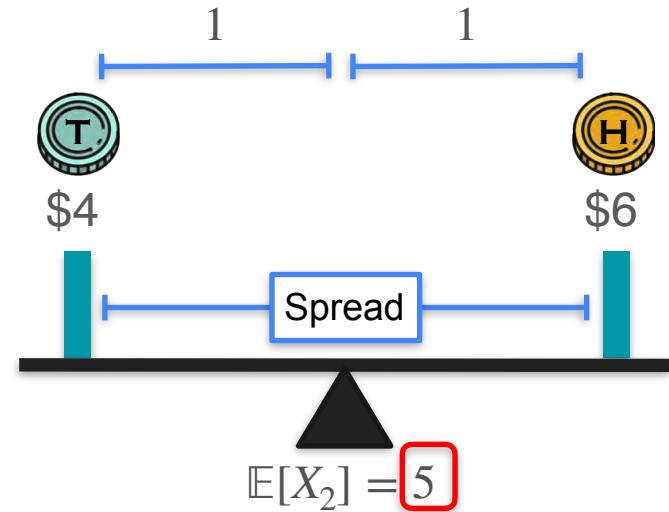
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



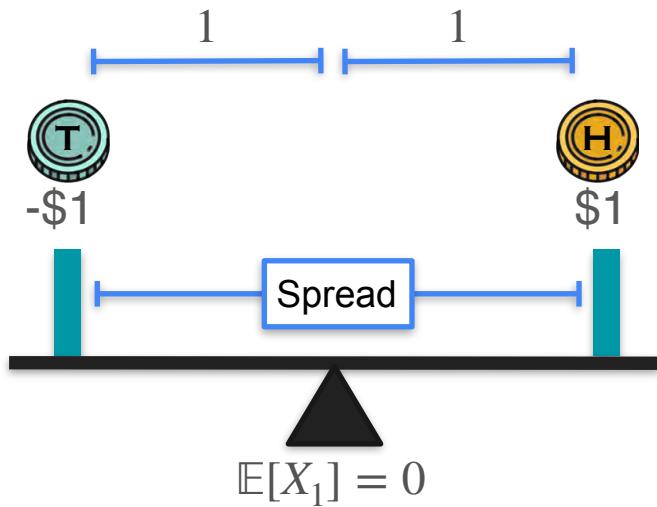
Game 1



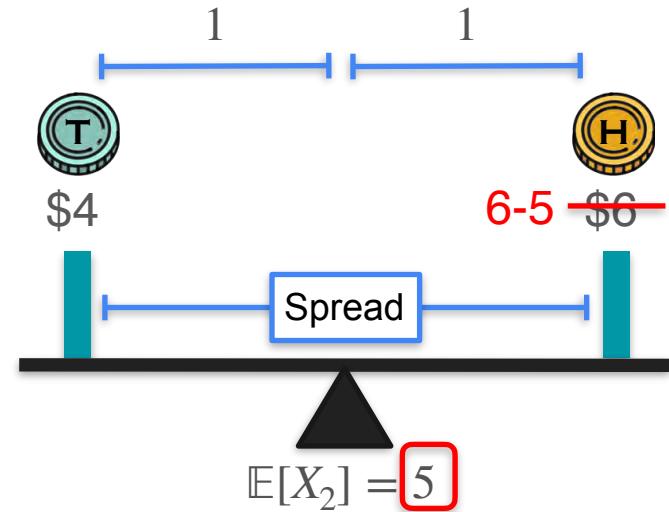
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



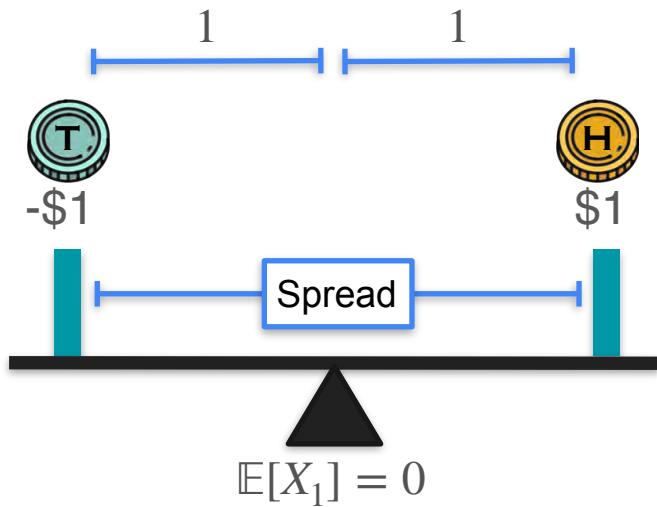
Game 1



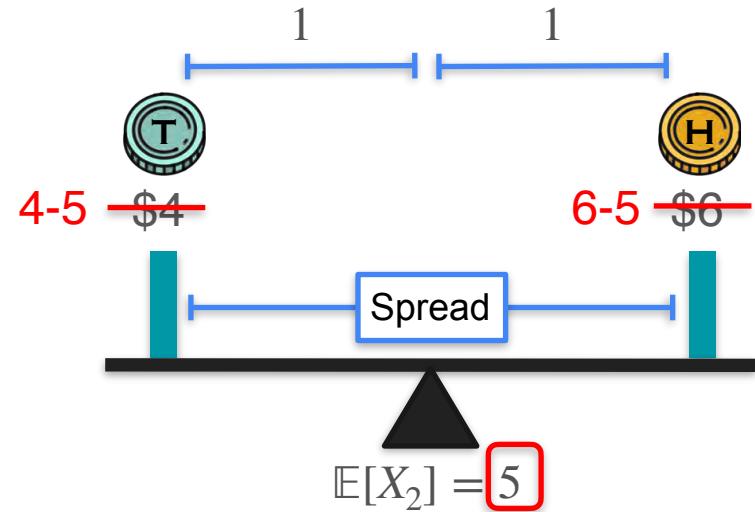
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



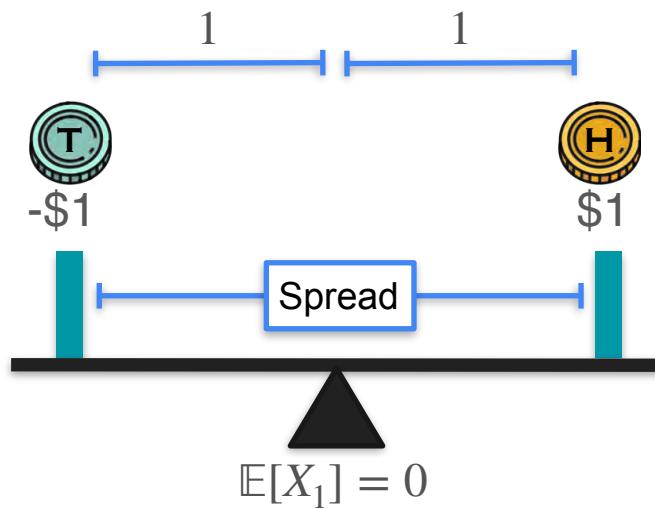
Game 1



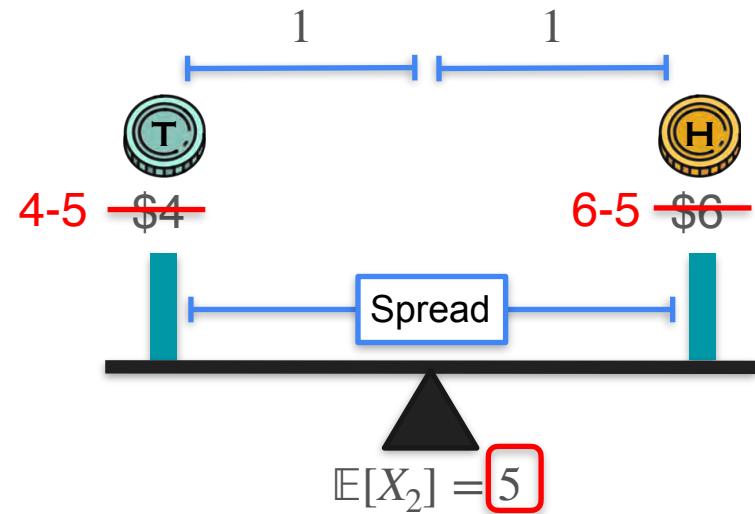
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$

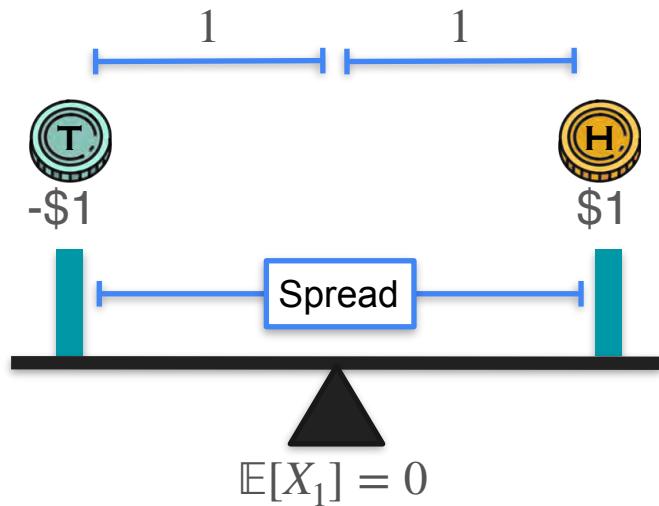


$$\mathbb{E}[(X_2 - 5)^2]$$



Variance: Spread and Shift

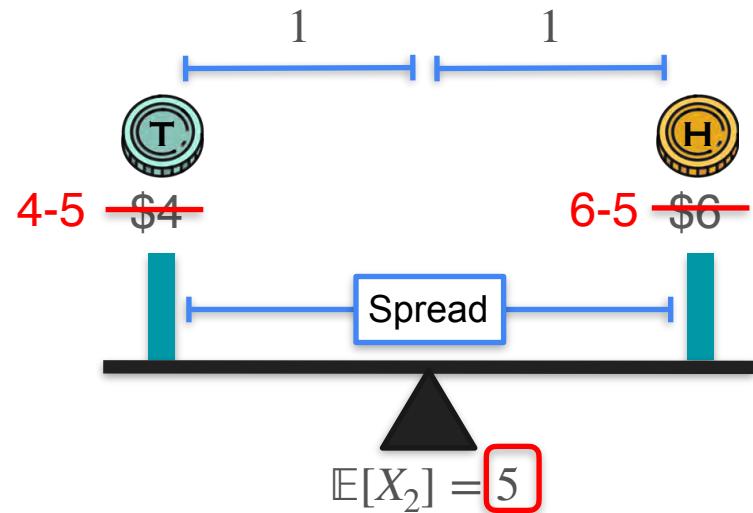
$$\mathbb{E}[X_1^2] = 1$$



Game 1

$$\text{Mean: } \mu$$

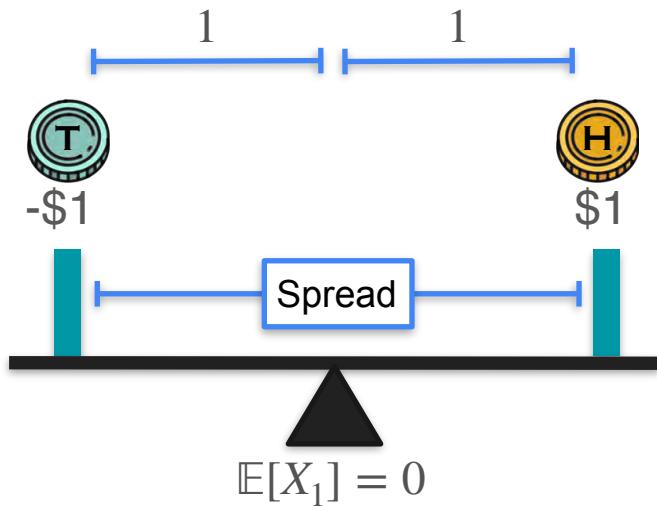
$$\mathbb{E}[(X_2 - 5)^2]$$



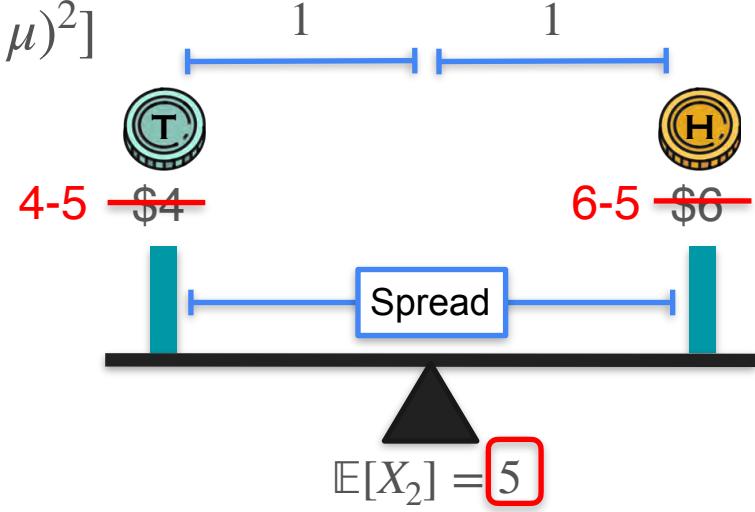
Game 2

Variance: Spread and Shift

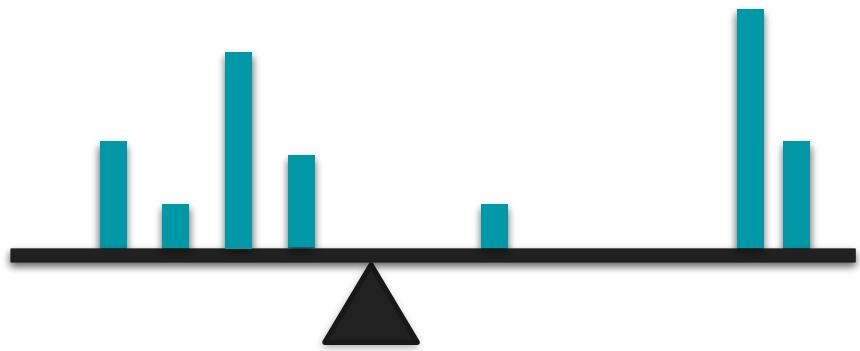
$$\mathbb{E}[X_1^2] = 1$$



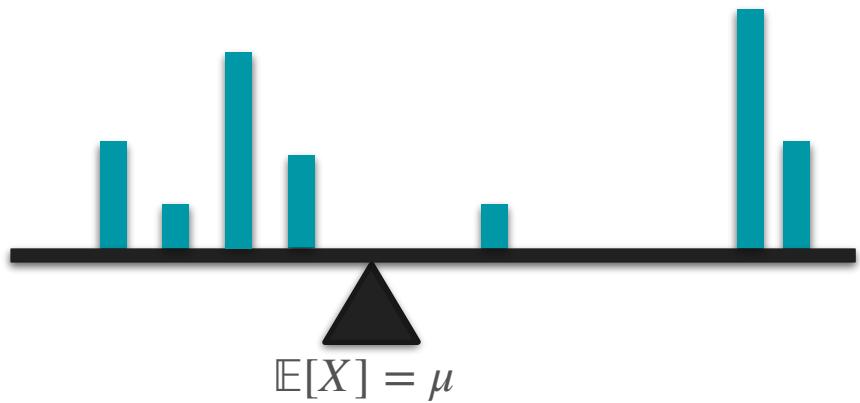
$$\text{Variance: } \mathbb{E}[(X - \mu)^2]$$



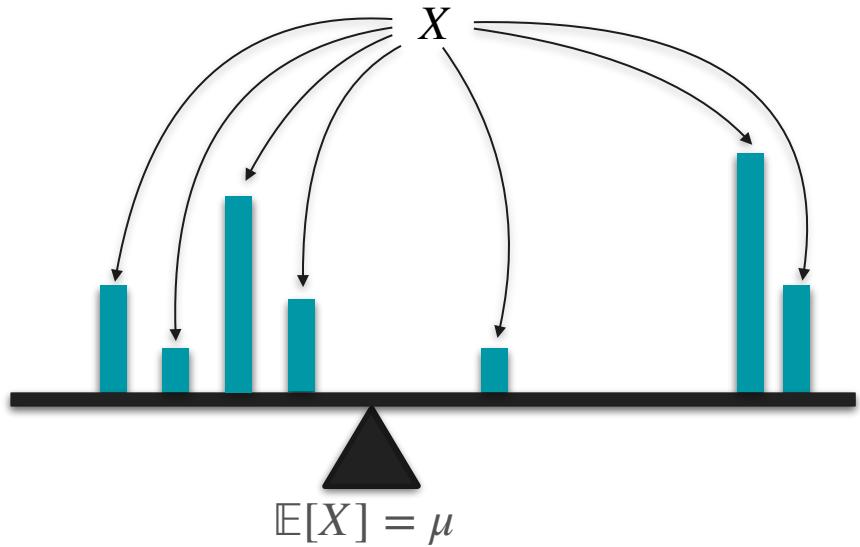
Variance Formula



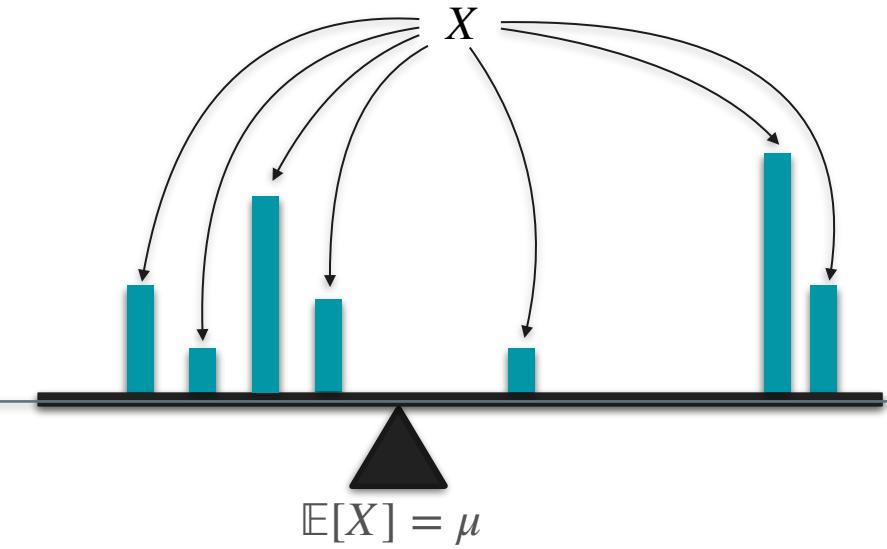
Variance Formula



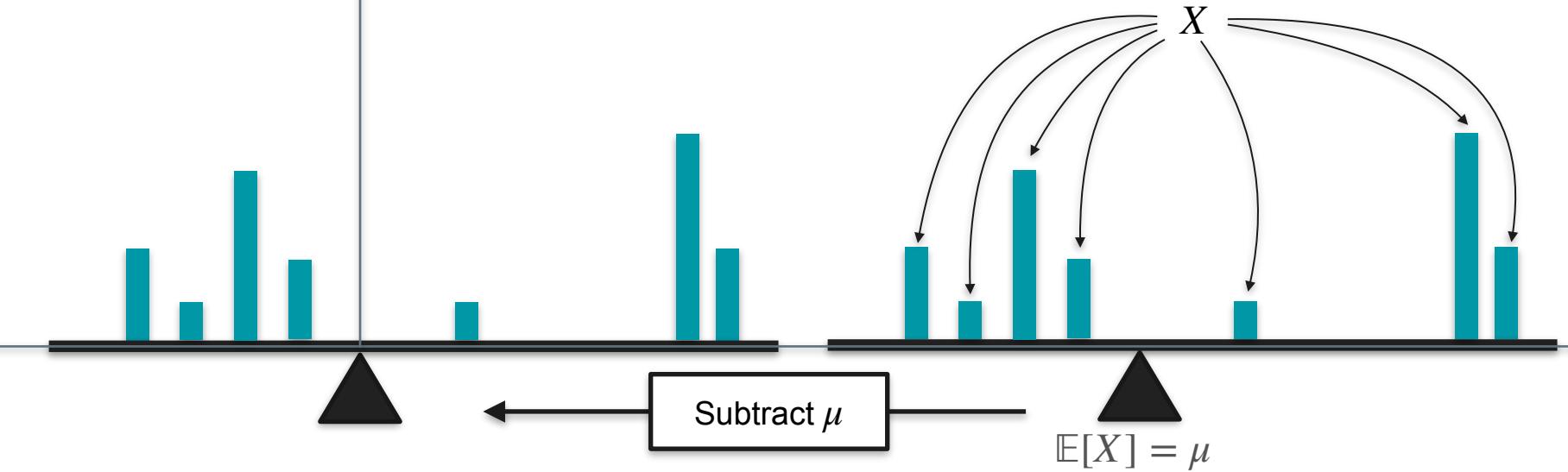
Variance Formula



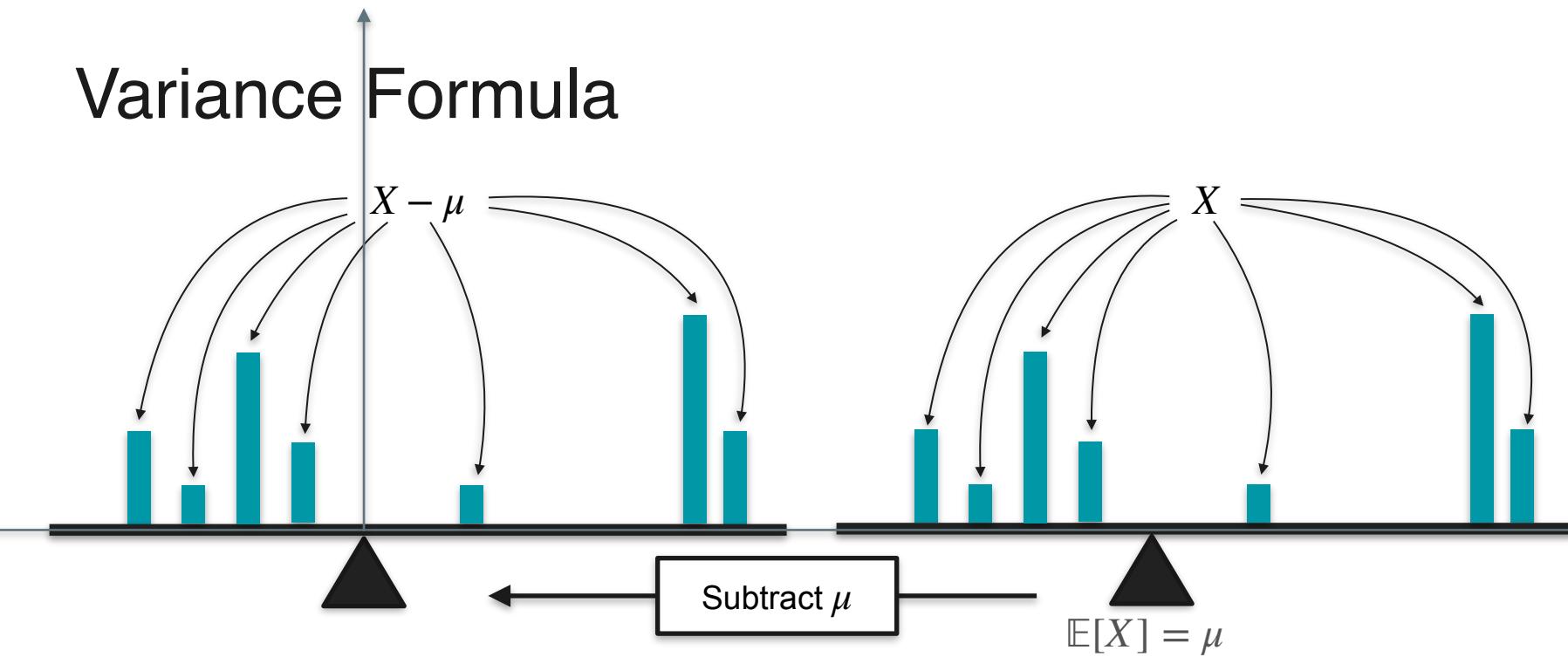
Variance Formula



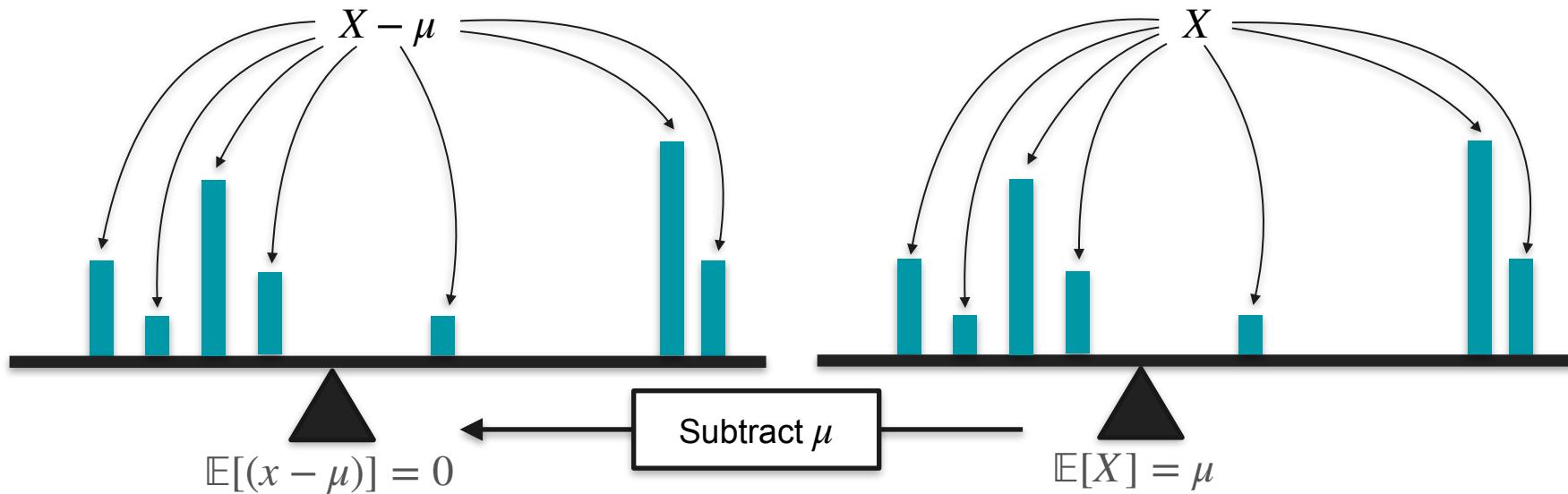
Variance Formula



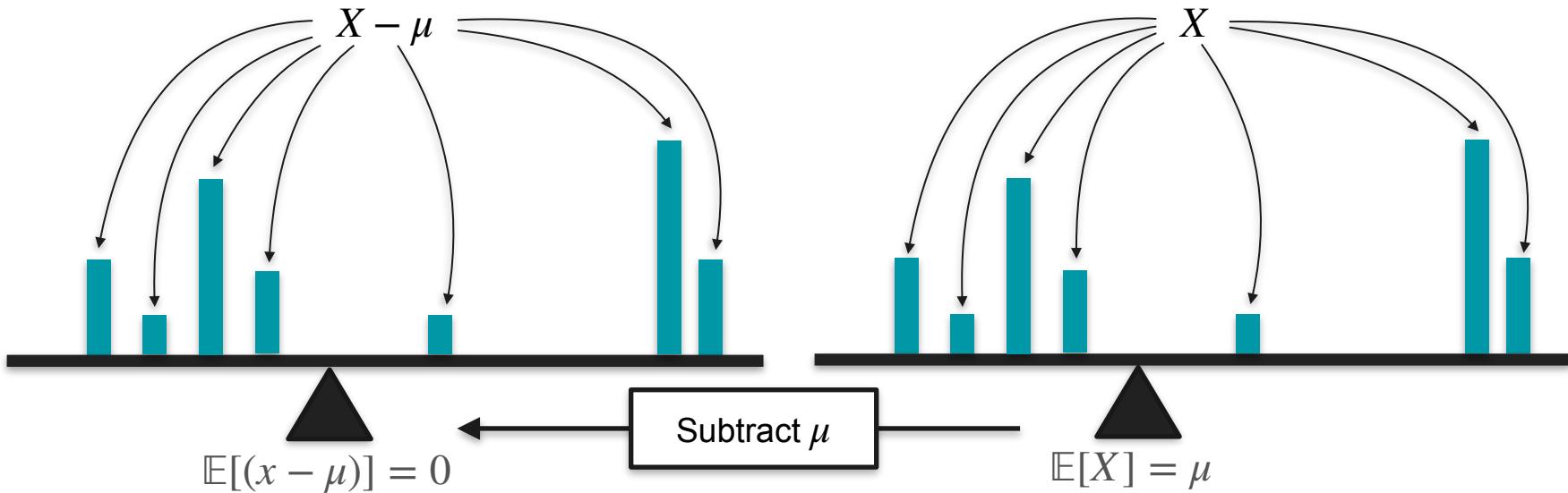
Variance Formula



Variance Formula



Variance Formula



$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

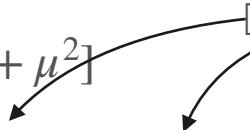
Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2]$$

Variance Formula

$$\begin{aligned} \text{Var}(X) = \mathbb{E}[(X - \mu)^2] &= \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \end{aligned}$$


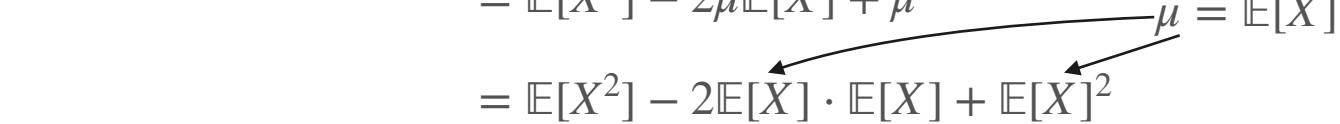
Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \quad \text{[constant} \cdot X \text{]} = \text{constant} \cdot \mathbb{E}[X] \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \end{aligned}$$

Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \quad \mathbb{E}[\text{constant}] = \text{constant} \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \end{aligned}$$

Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \end{aligned}$$


Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\&= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2\end{aligned}$$

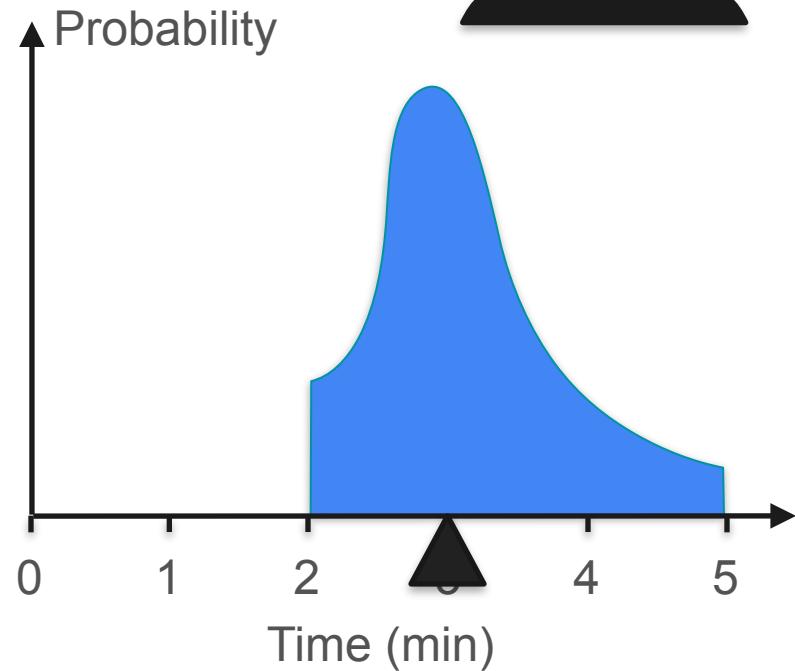
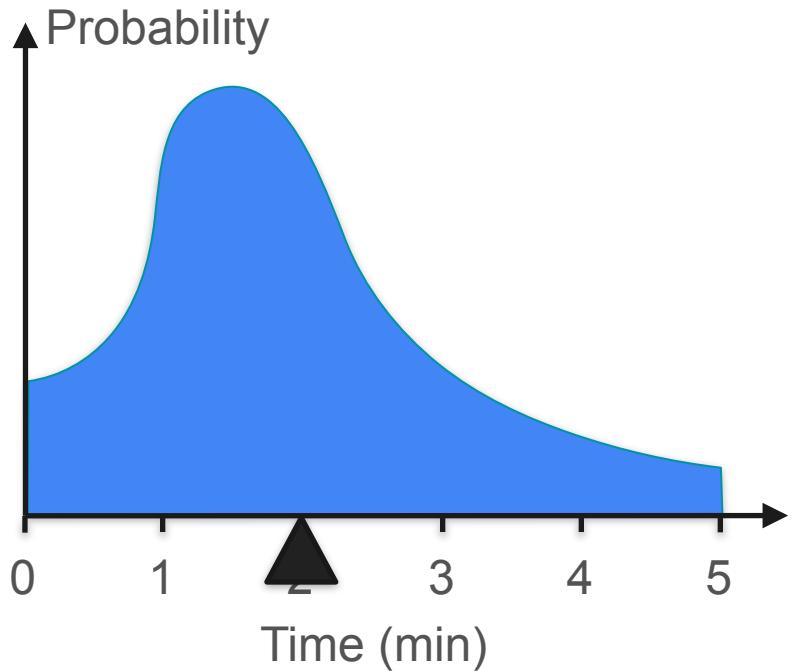
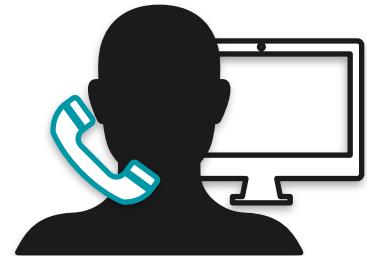
Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\&= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

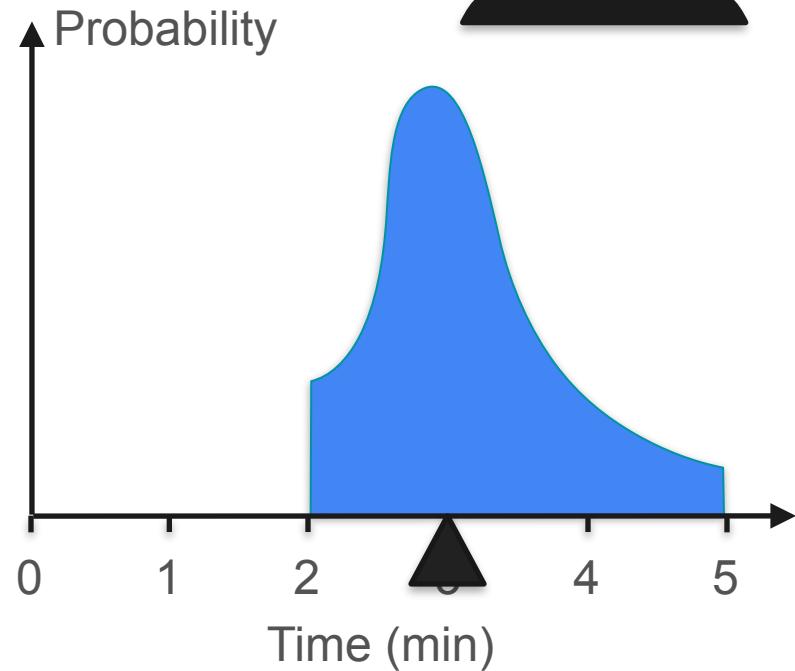
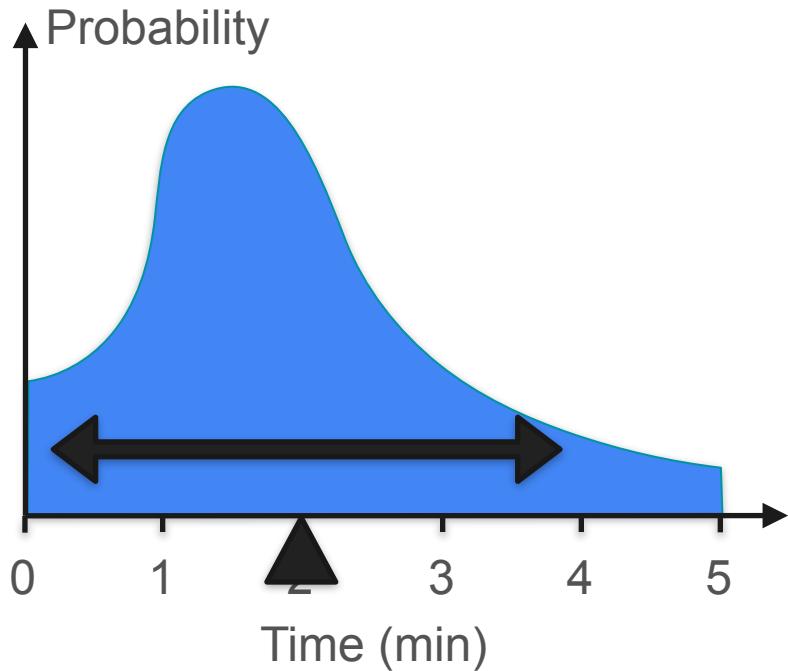
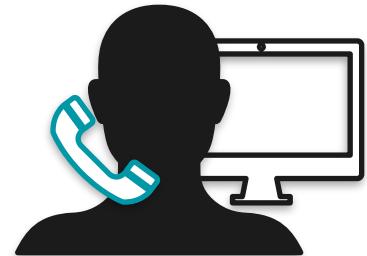
Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

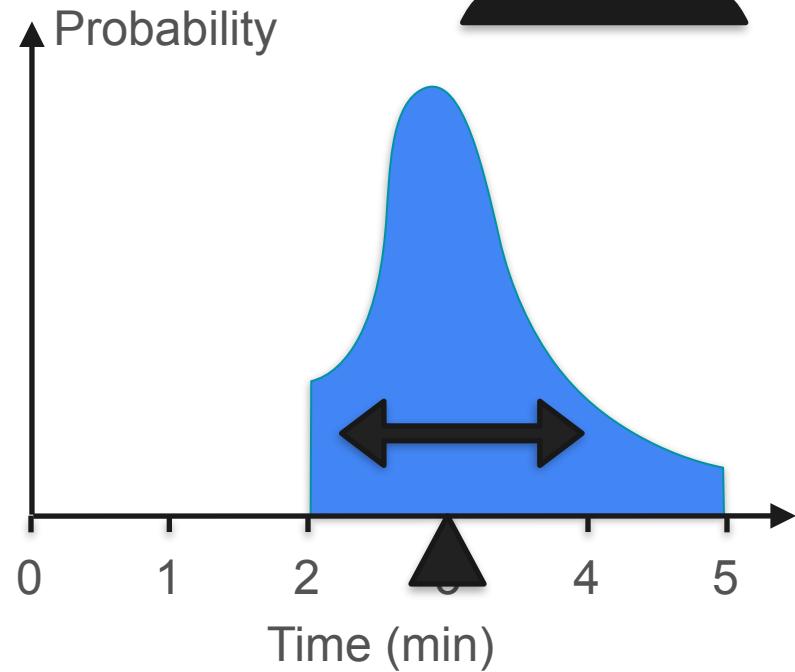
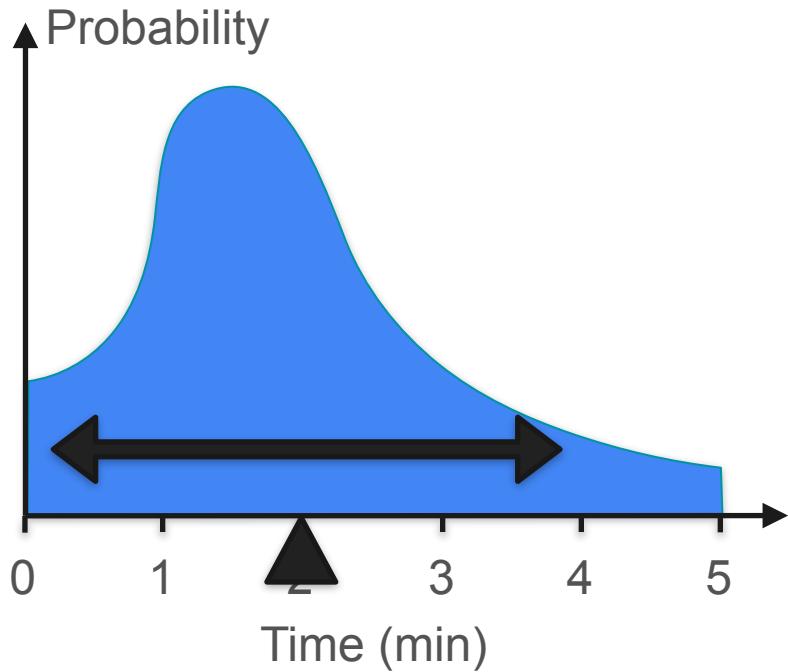
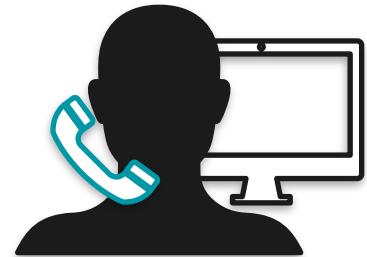
Variance for Continuous Distributions



Variance for Continuous Distributions

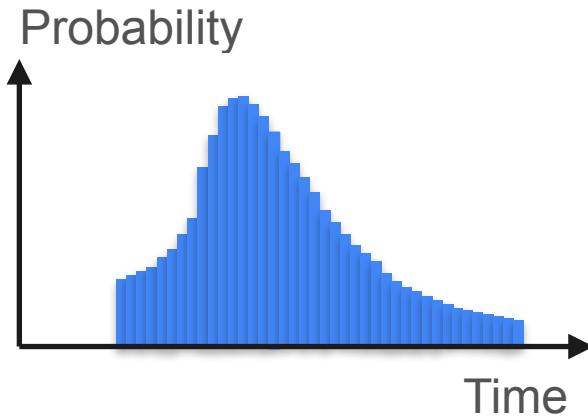


Variance for Continuous Distributions



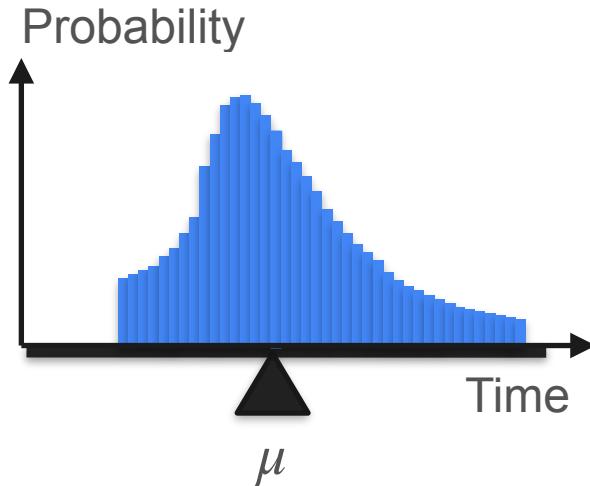
Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$



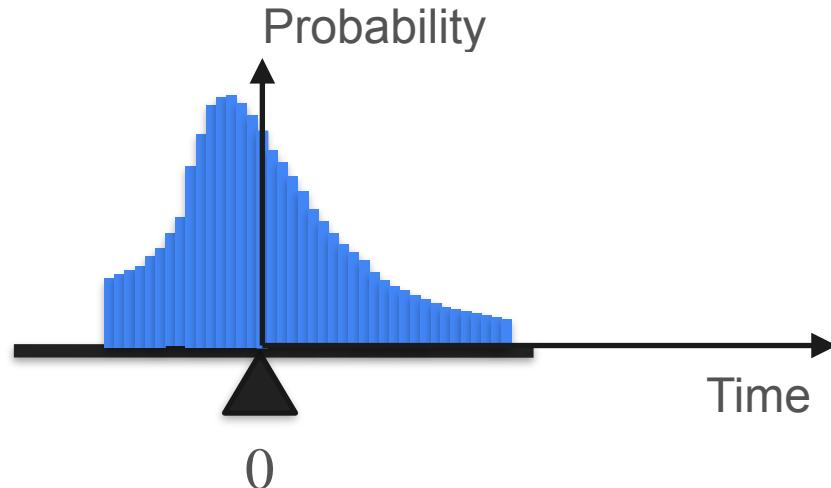
Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$



Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$





DeepLearning.AI

Describing Distributions

**Measures of
Central Tendency**

Mean: Example

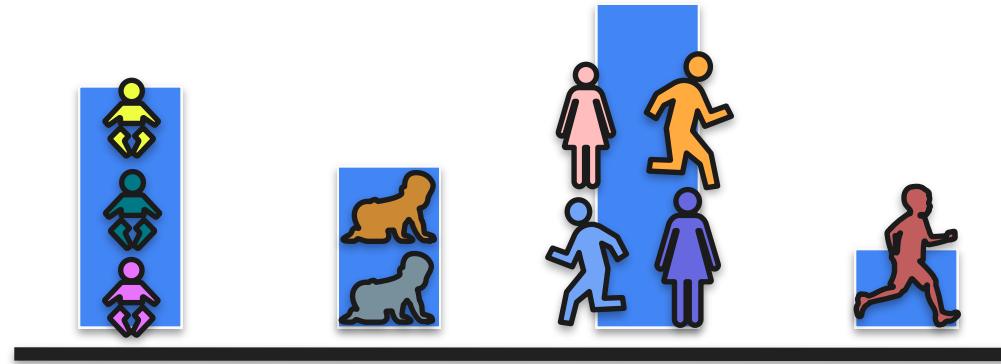
Age:

0

1

2

3



Mean: Example

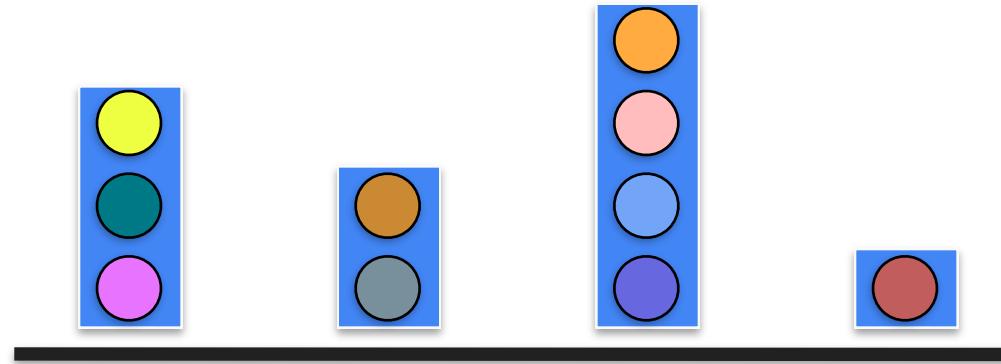
Age:

0

1

2

3



Mean: Example

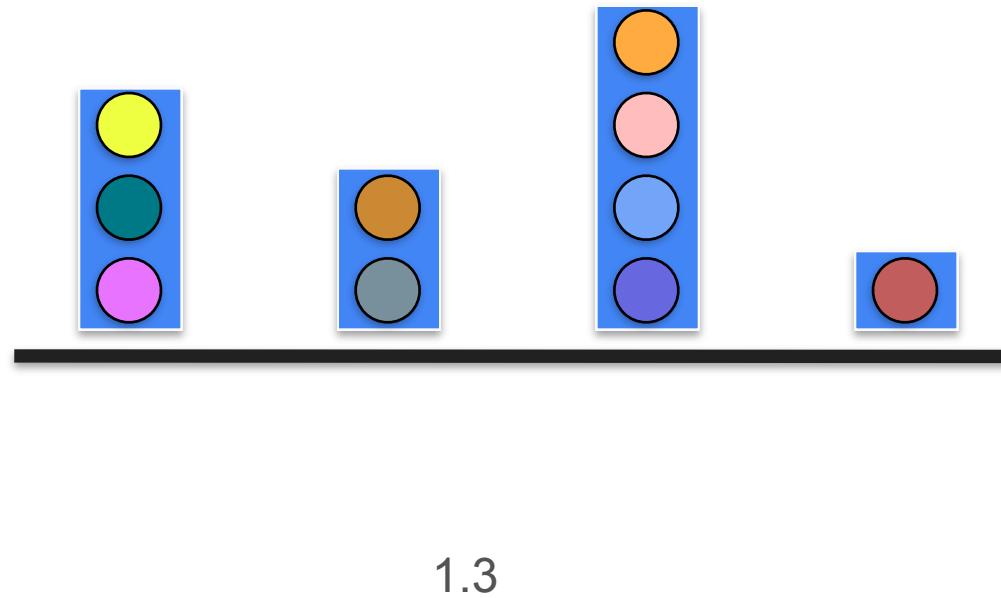
Age:

0

1

2

3



Mean: Example

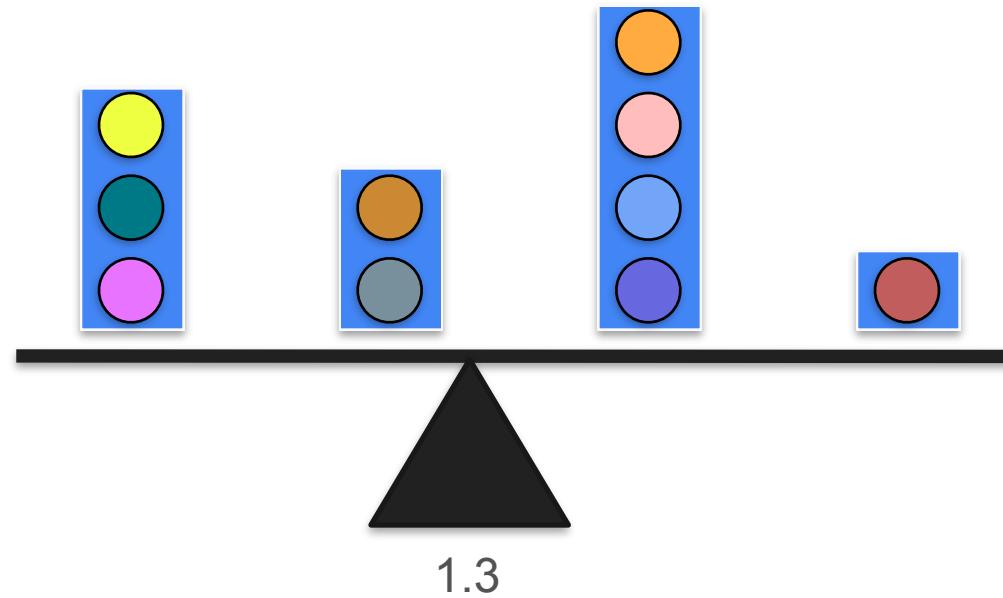
Age:

0

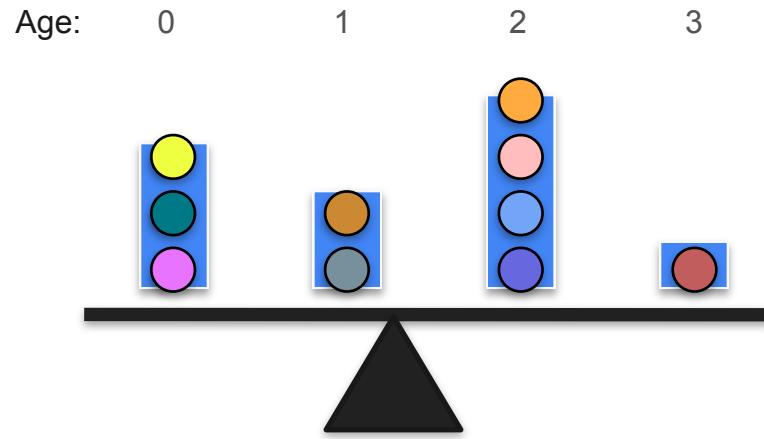
1

2

3

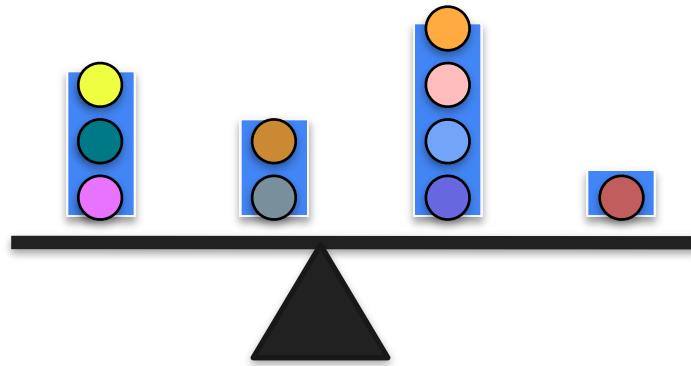


Mean: Example



Mean: Example

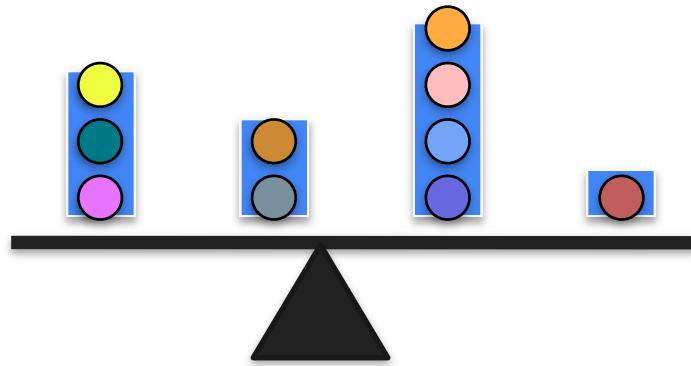
Age: 0 1 2 3 $0 + 0 + 0$



Mean: Example

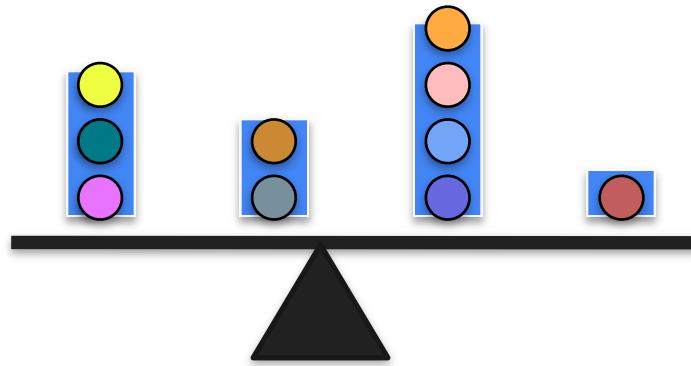
Age: 0 1 2 3

$0 + 0 + 0 + 1 + 1$



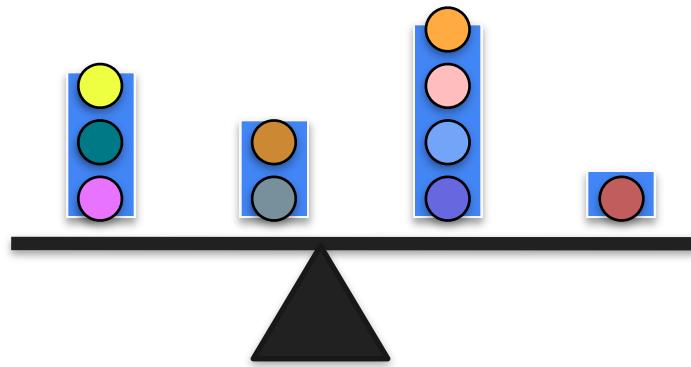
Mean: Example

Age: 0 1 2 3 $0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 2$



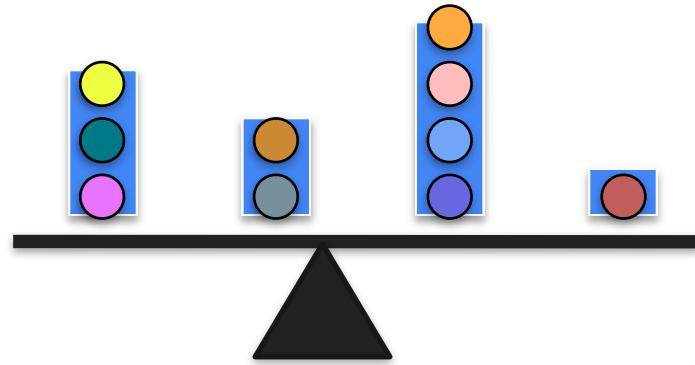
Mean: Example

Age: 0 1 2 3 $0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 2 + 3$



Mean: Example

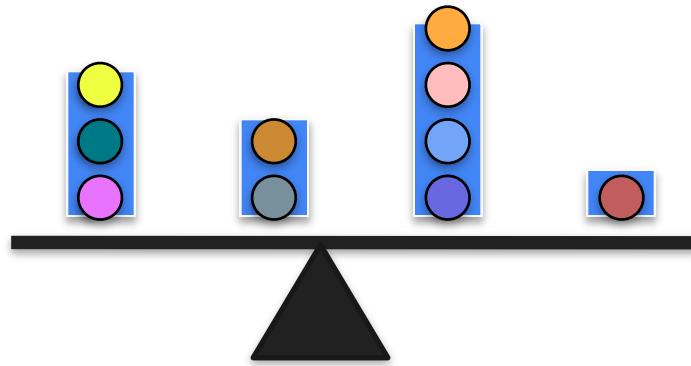
Age: 0 1 2 3



$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

Mean: Example

Age: 0 1 2 3

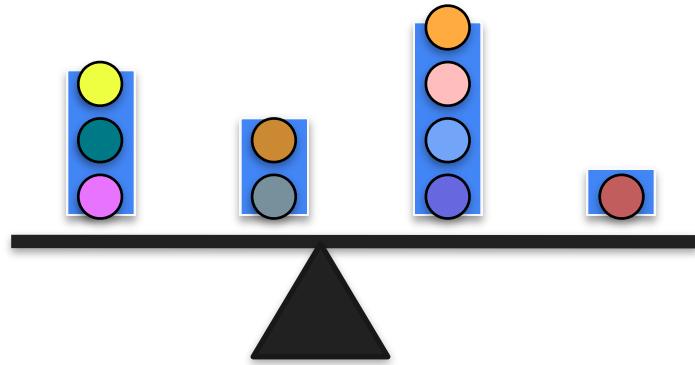


$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

$$= \frac{13}{10}$$

Mean: Example

Age: 0 1 2 3



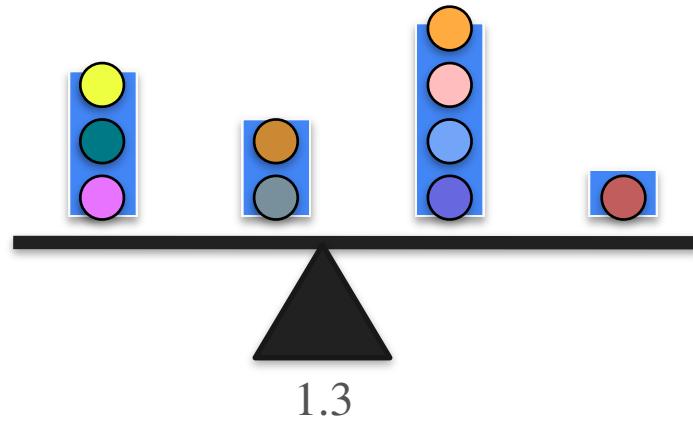
$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

$$= \frac{13}{10}$$

$$= 1.3$$

Mean: Example

Age: 0 1 2 3

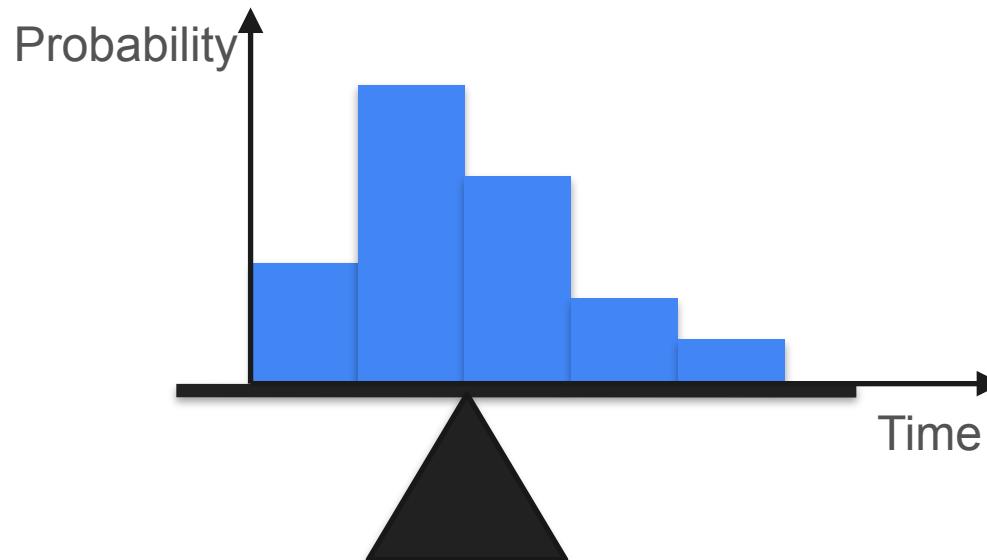


$$\begin{array}{r} 0 + 0 + 0 + 1 + 1 + 2 + 2 + 2 + 3 \\ \hline 10 \end{array}$$

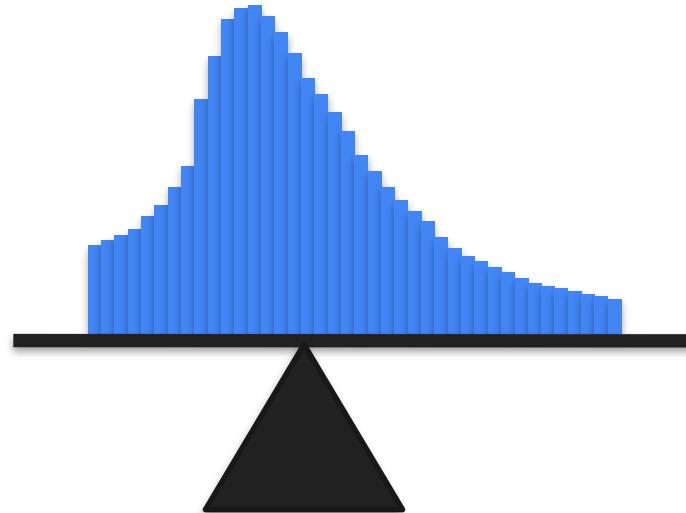
$$= \frac{13}{10}$$

$$= 1.3$$

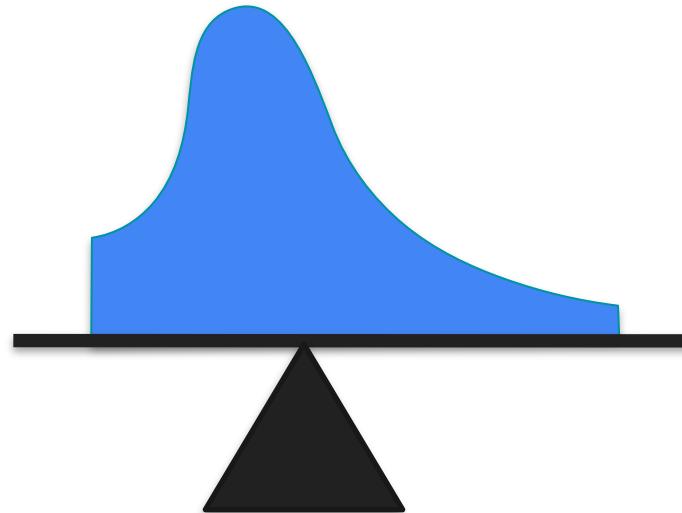
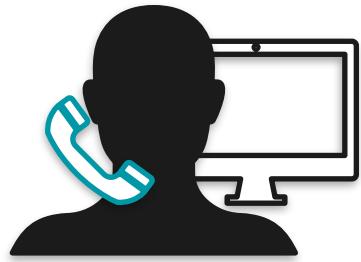
Mean



Mean



Mean



Median: Motivation

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections

Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money

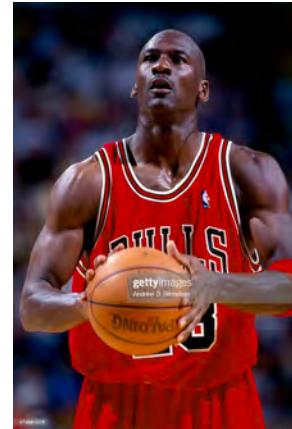
Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
- For the rest of the country, the starting salary for a geography graduate was \$22,000.
- Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money



Median: Motivation

- In the 1980s, the starting salary for a geography graduate at the University of North Carolina was \$250,000.
 - For the rest of the country, the starting salary for a geography graduate was \$22,000.
 - Why?
 1. The program was really good
 2. The university had great connections
 3. One student made lots of money
- 



Michael Jordan

Outliers

Outliers

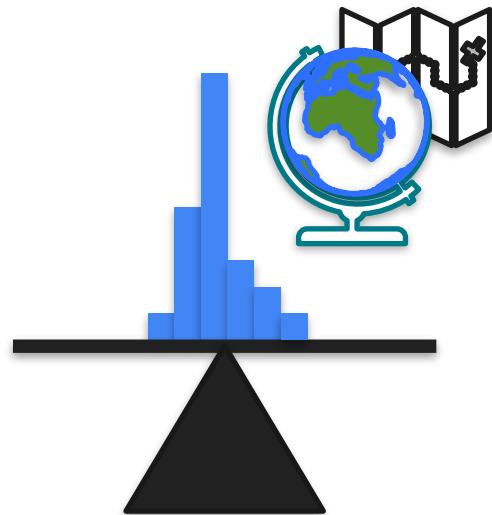


Outliers



Outliers

Graduates



Salary

Outliers

Graduates



Outliers

Graduates



Outliers

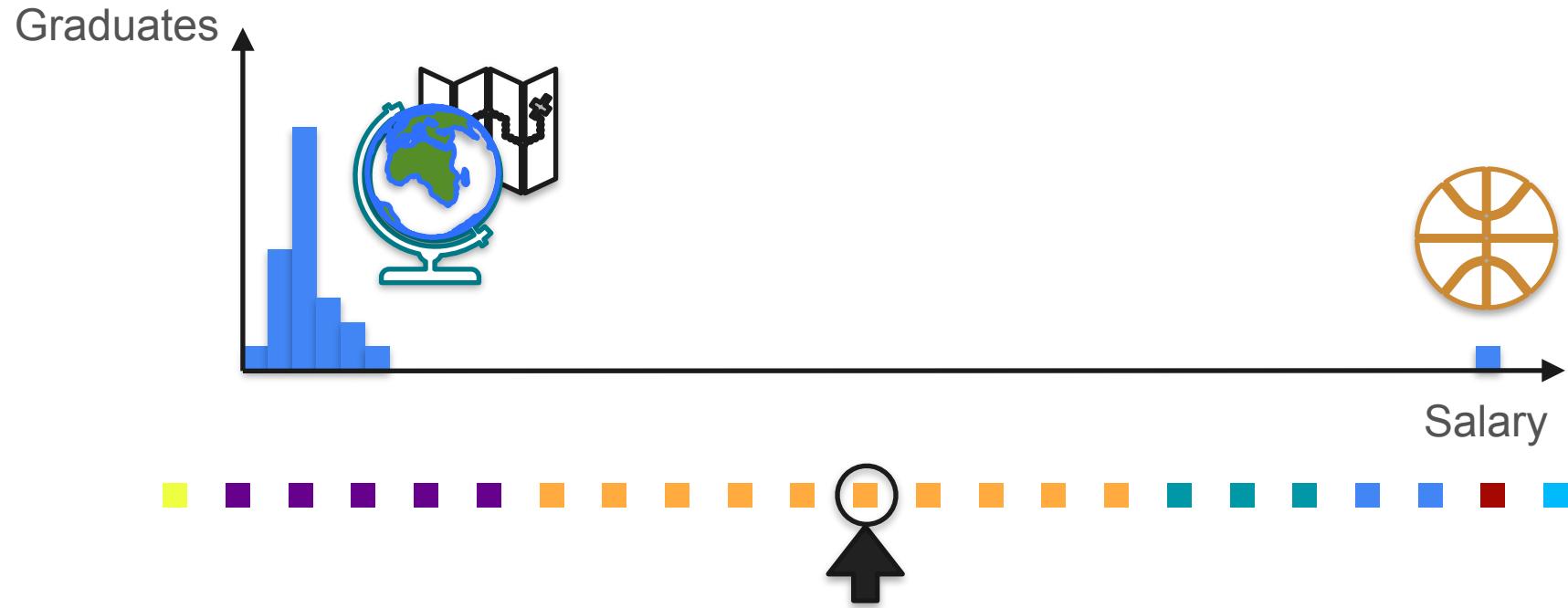


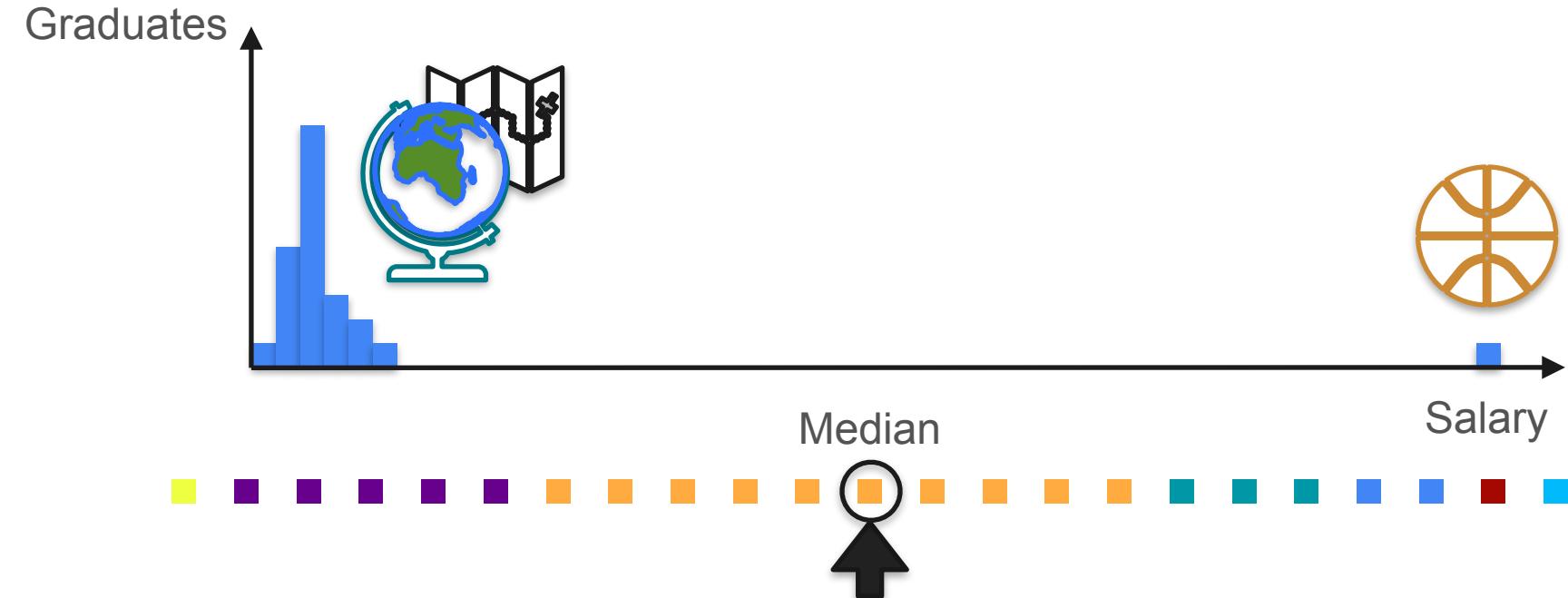
Median

Graduates



Median





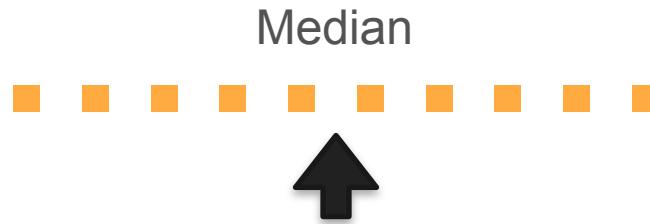
Median



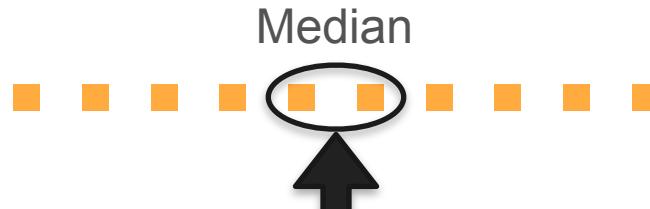
Median



Median

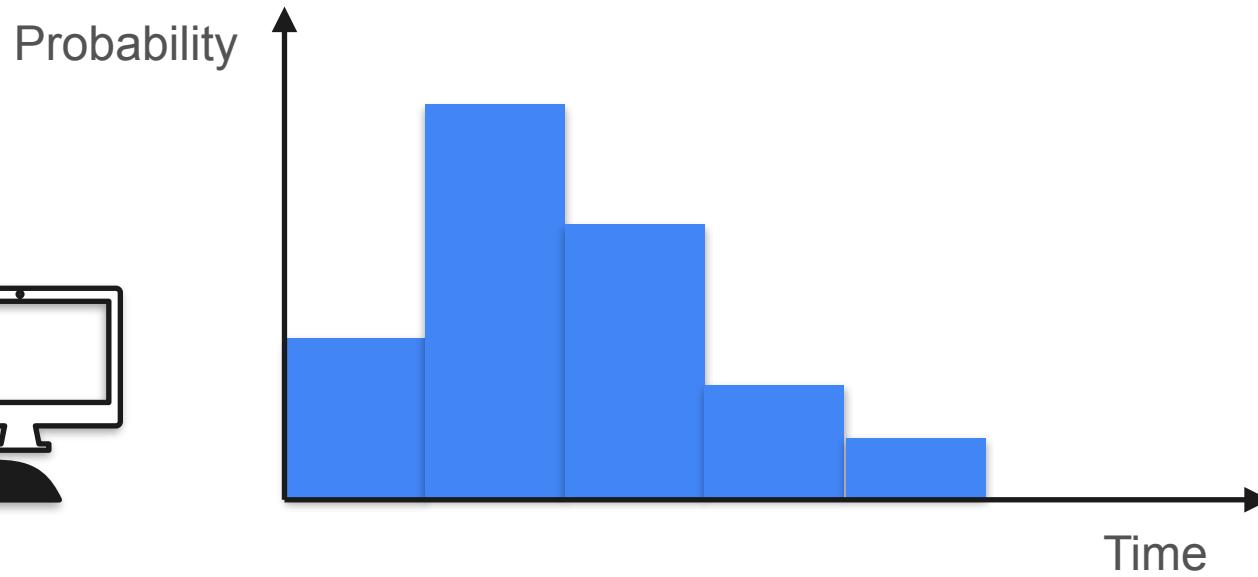


Median

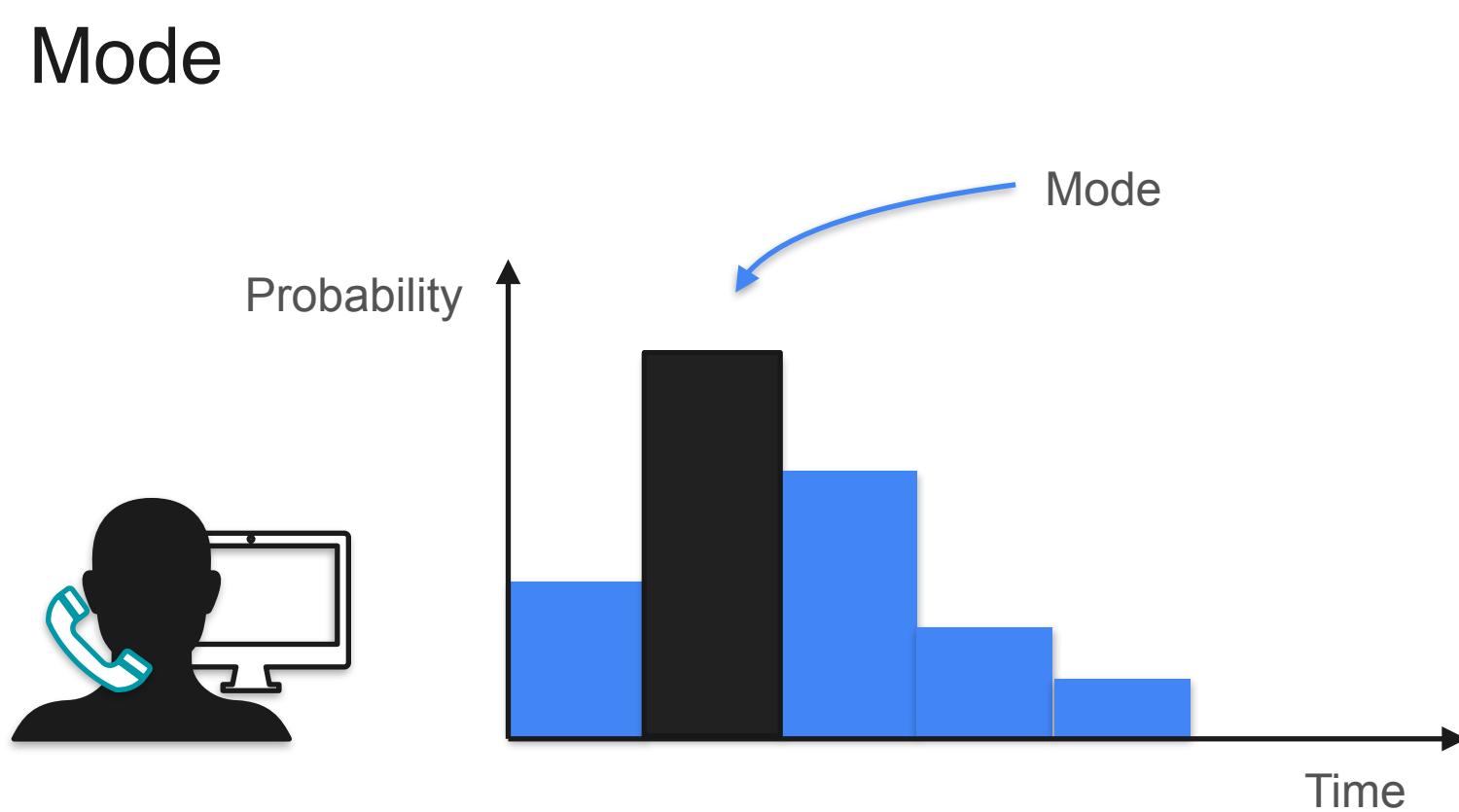


Median
Average of the
two middle ones

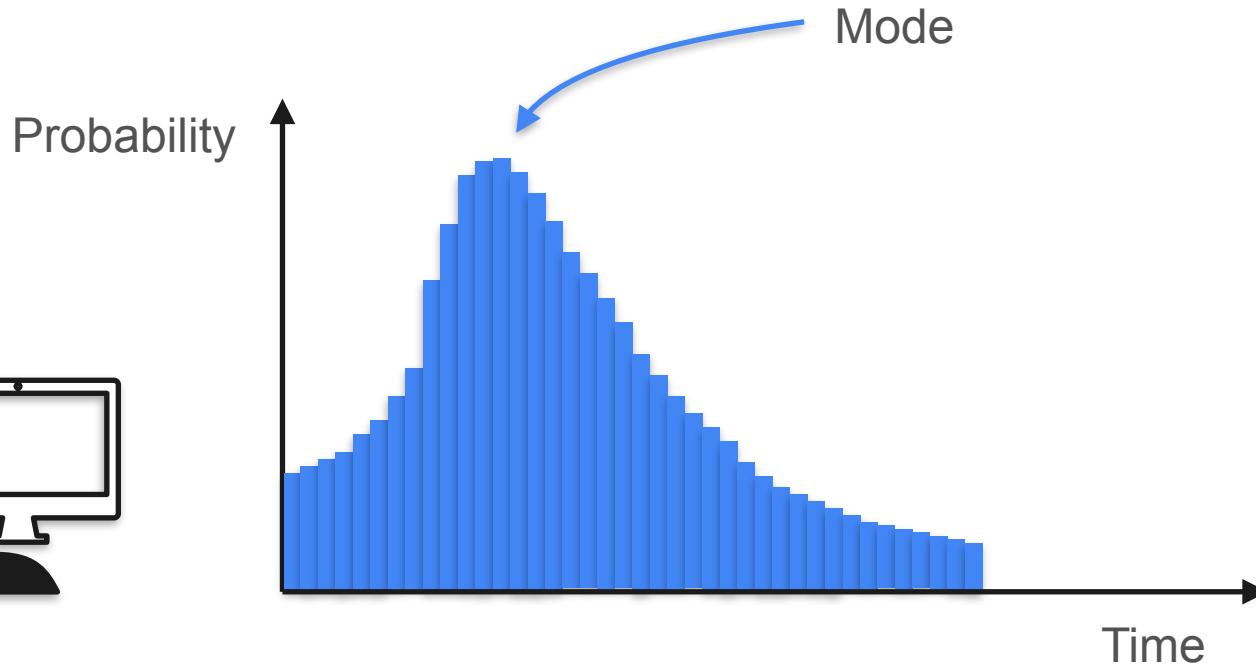
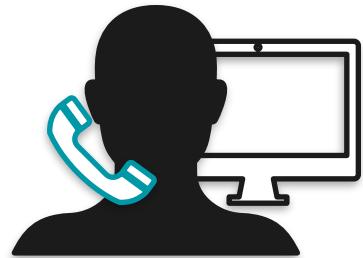
Mode



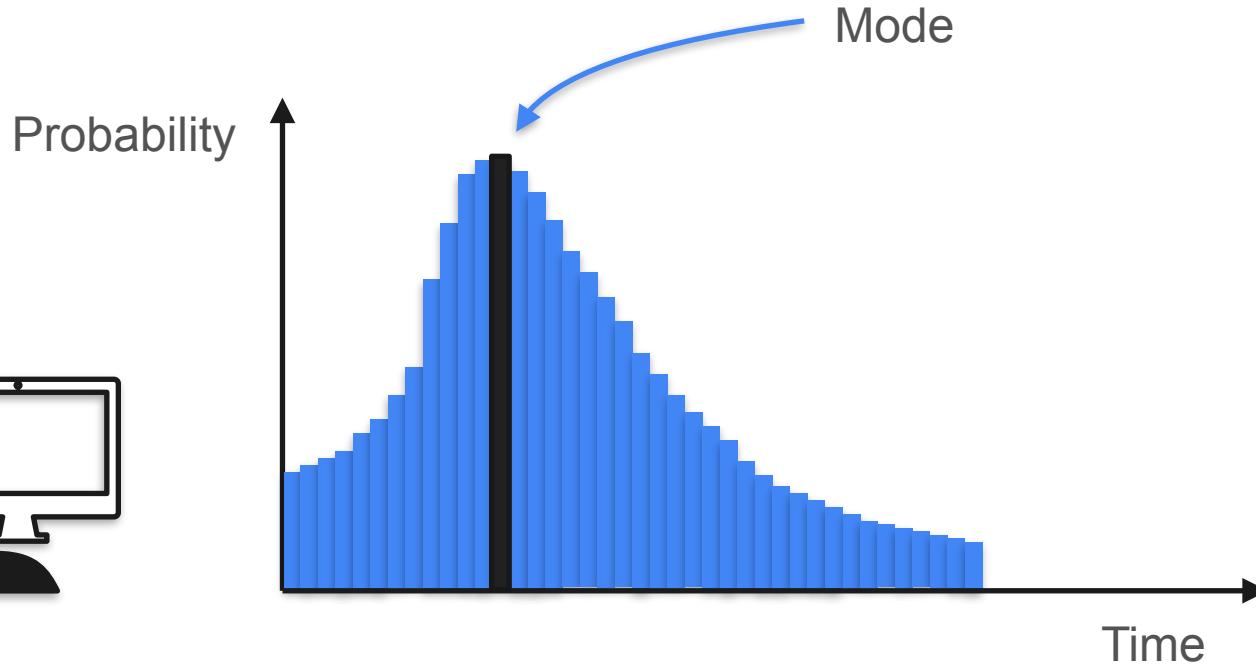
Mode



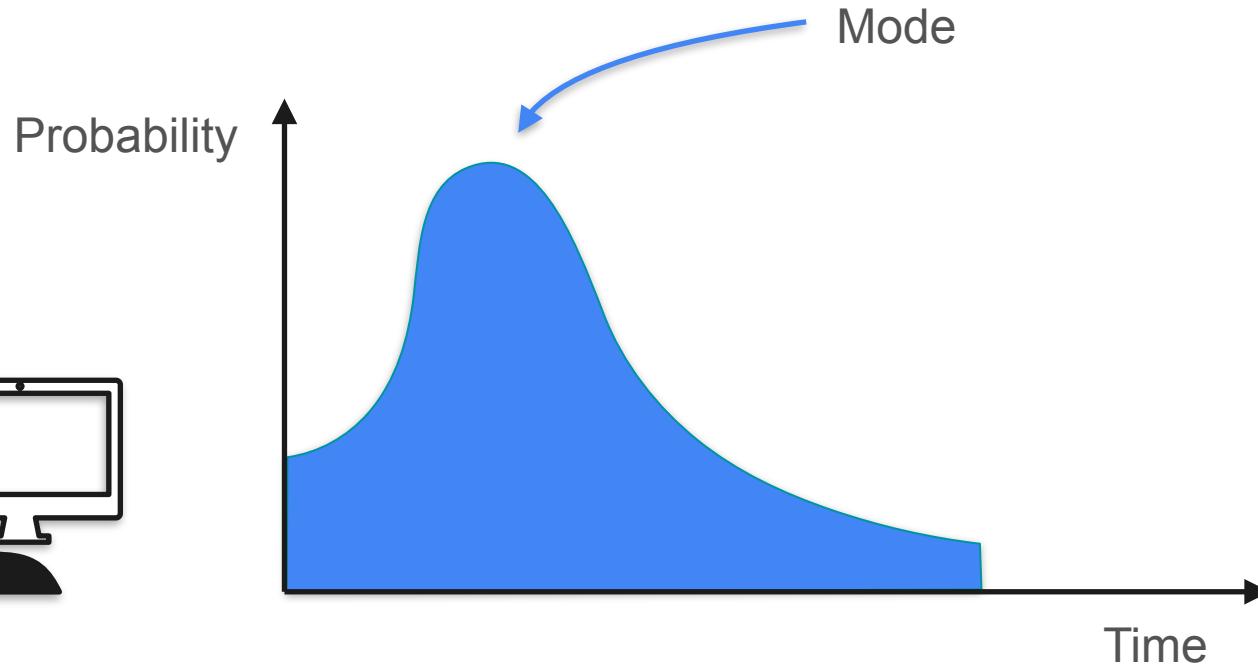
Mode



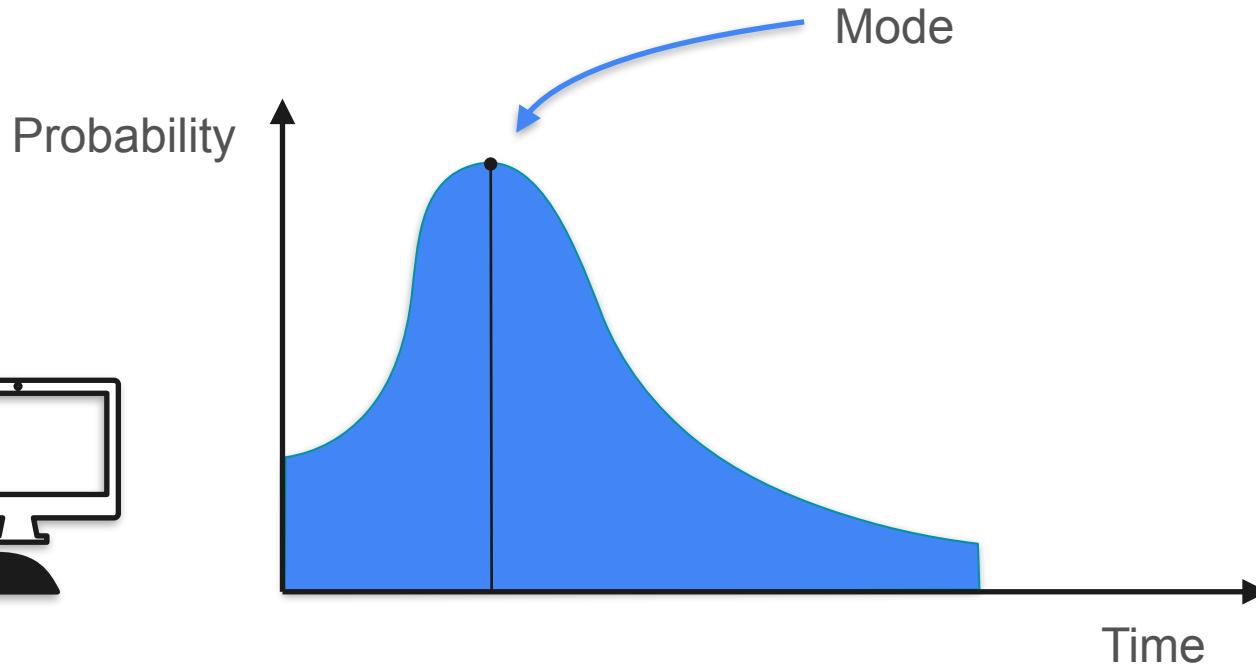
Mode



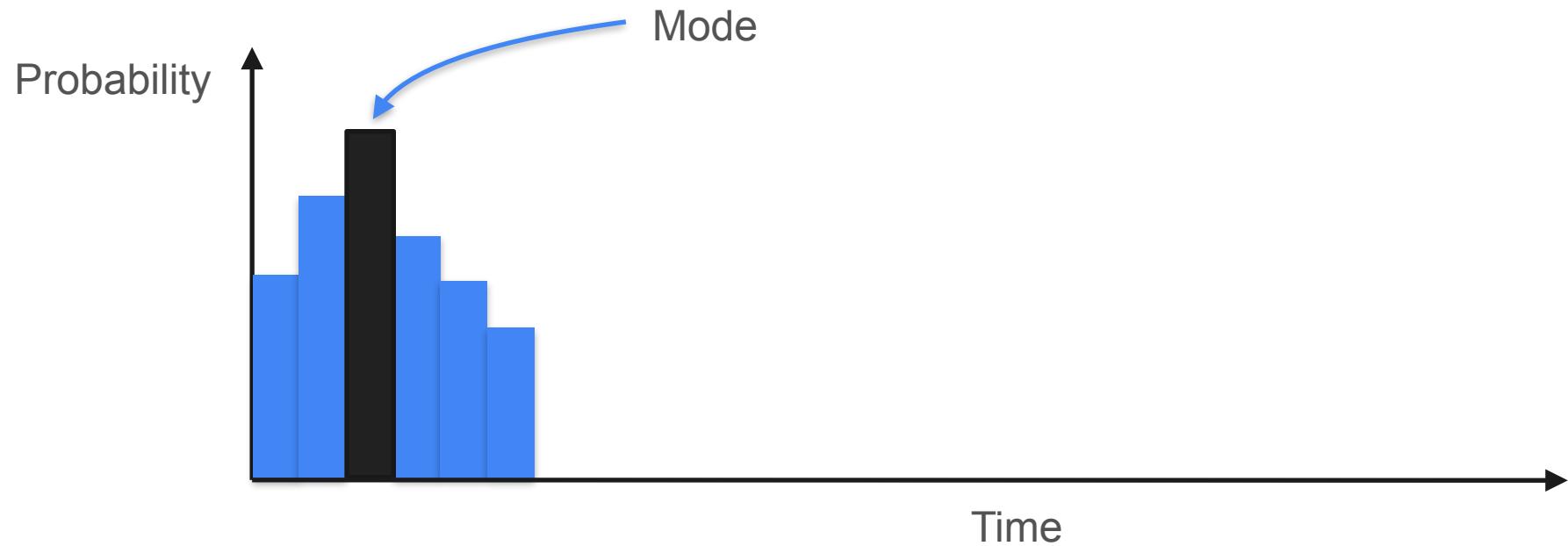
Mode



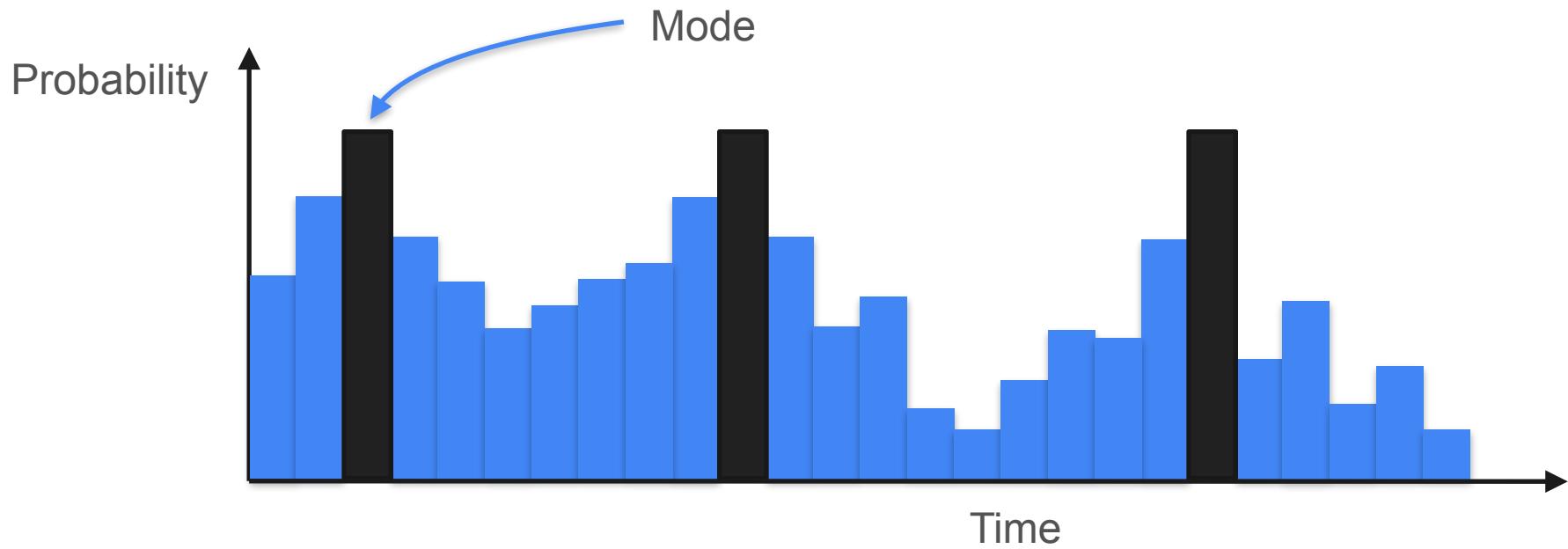
Mode



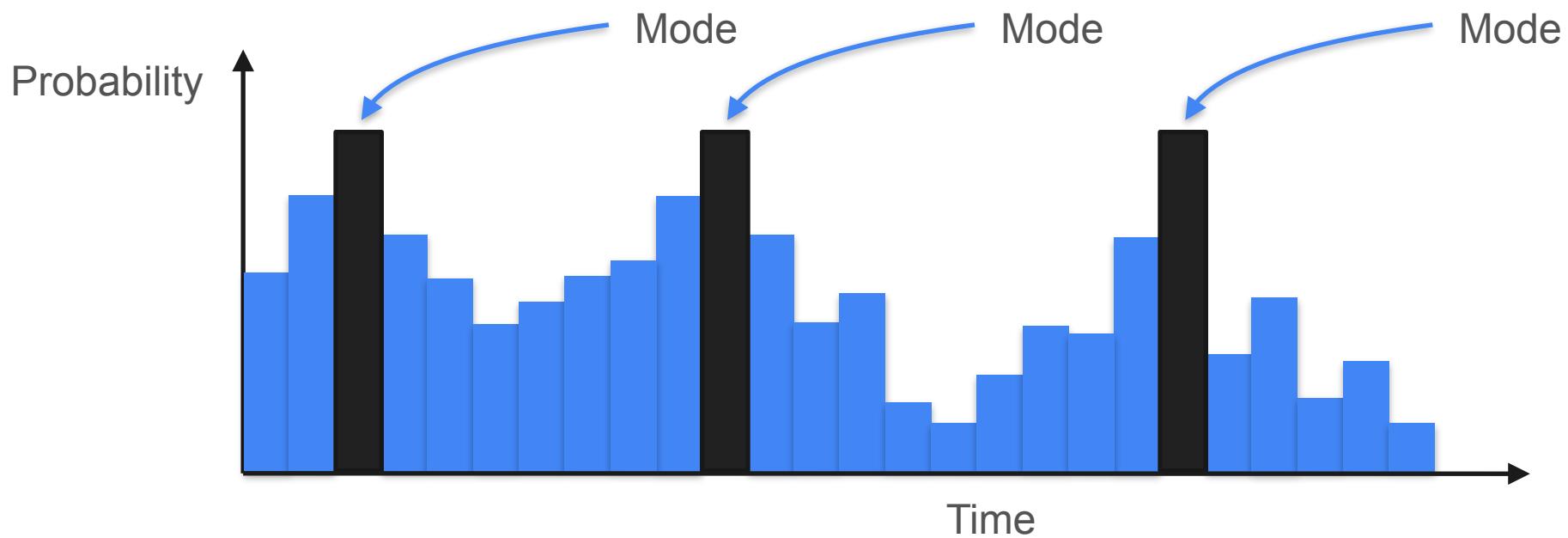
Mode: Multimodal Distribution



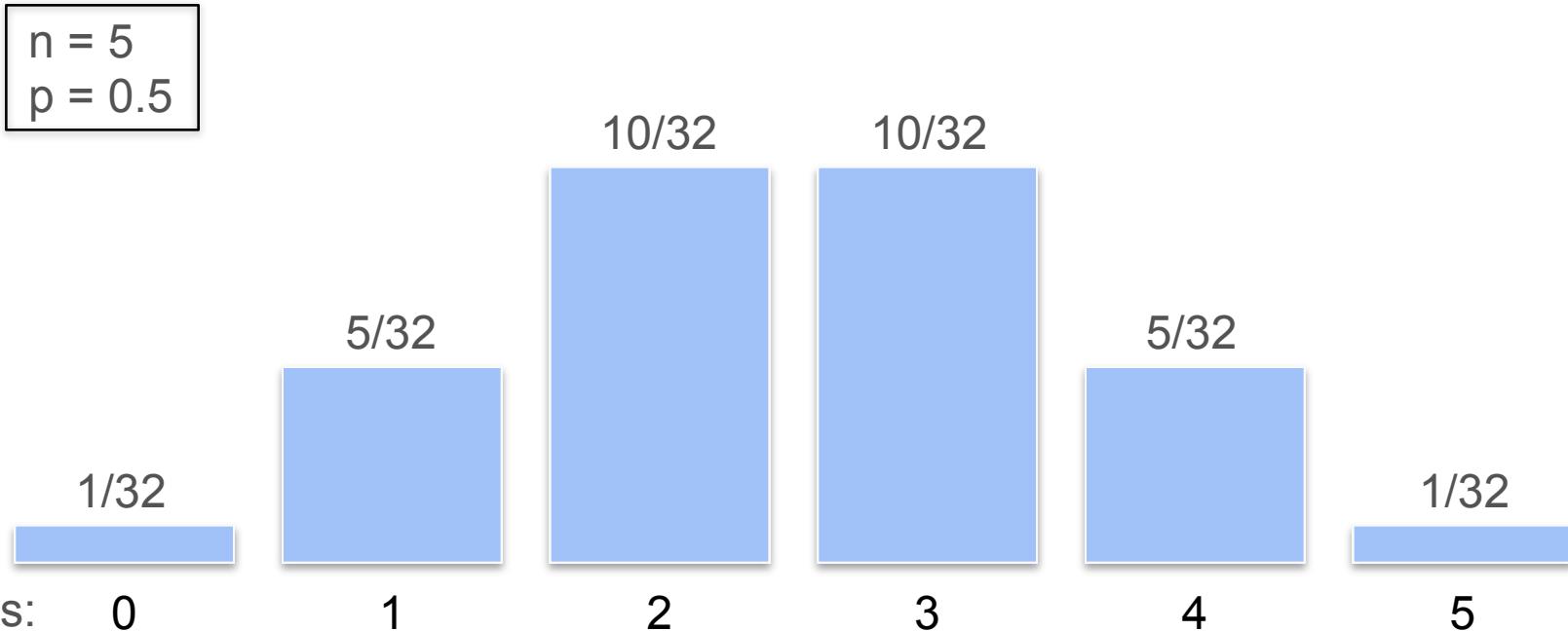
Mode: Multimodal Distribution



Mode: Multimodal Distribution

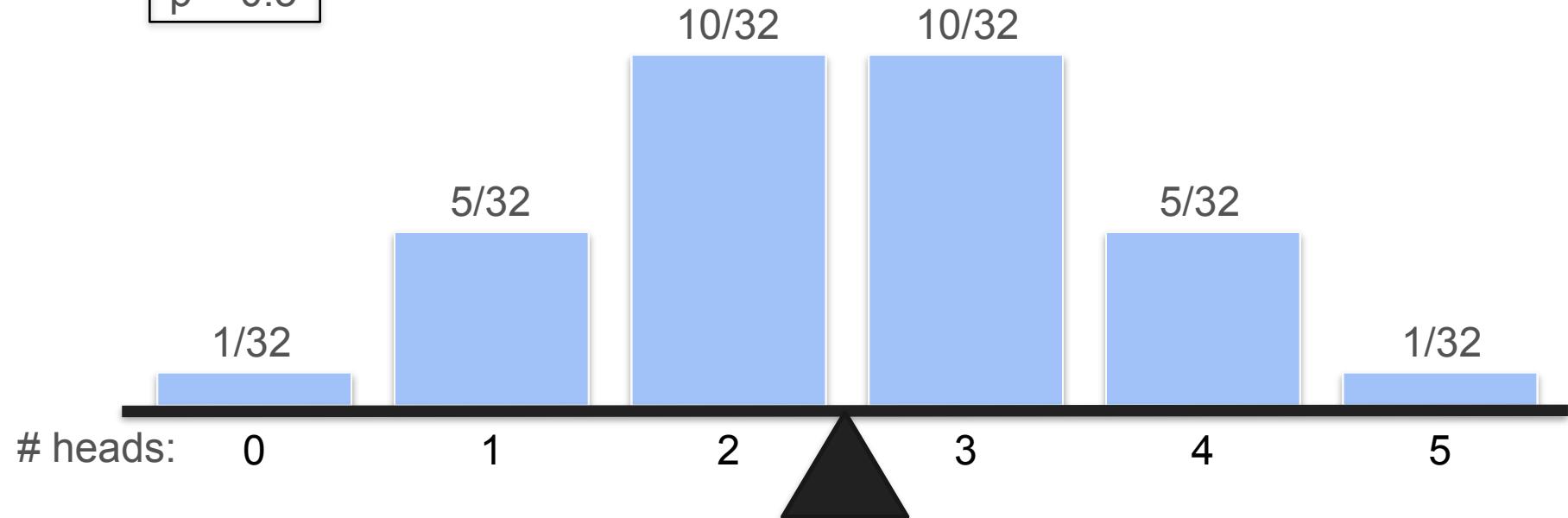


Mean, Median and Mode in Binomial Distribution

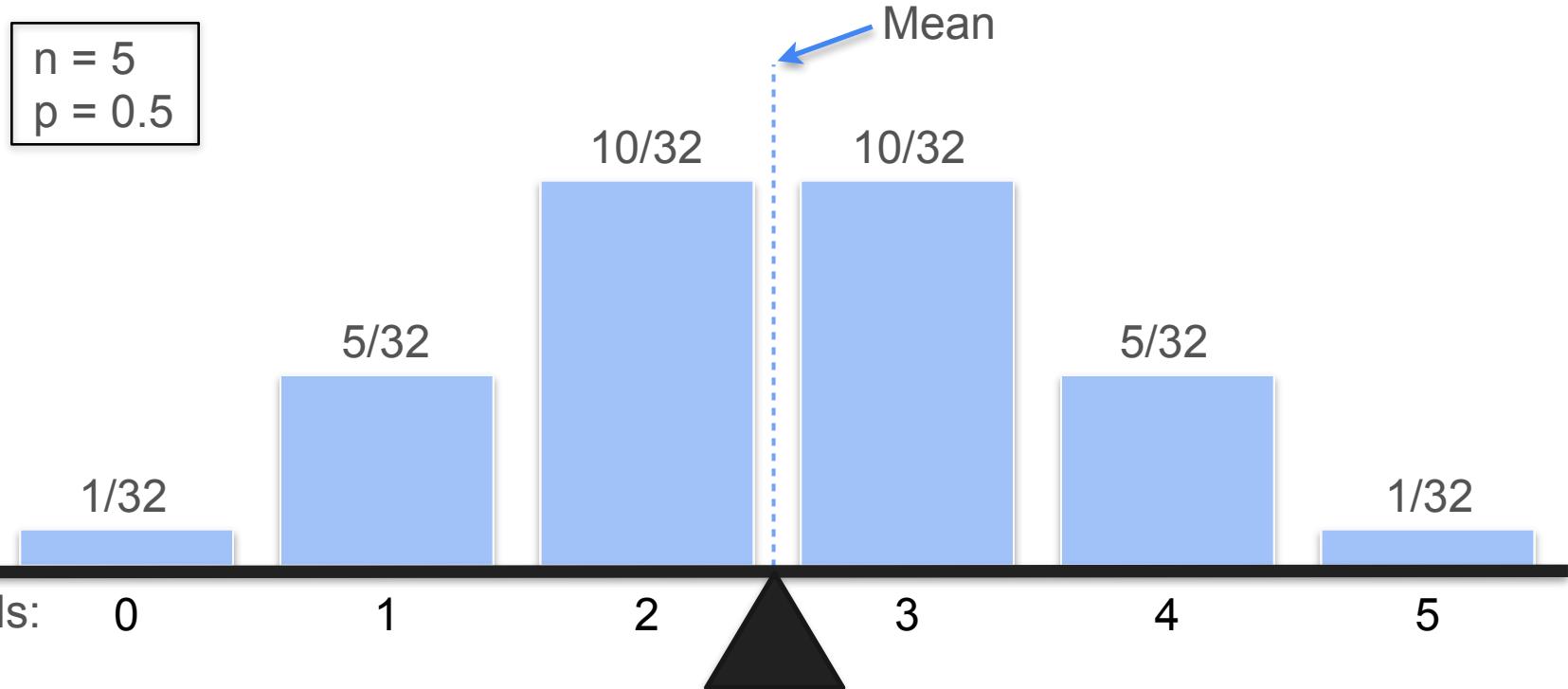


Mean, Median and Mode in Binomial Distribution

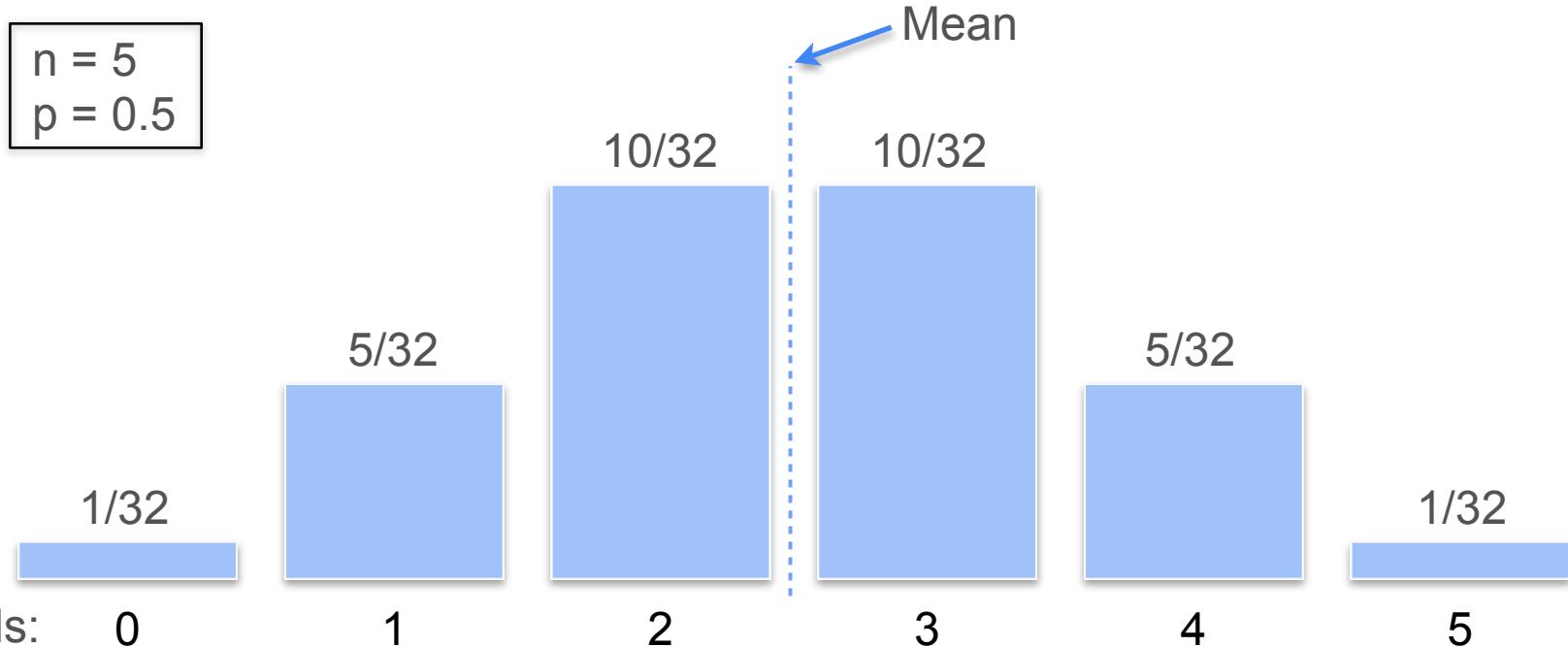
$n = 5$
 $p = 0.5$



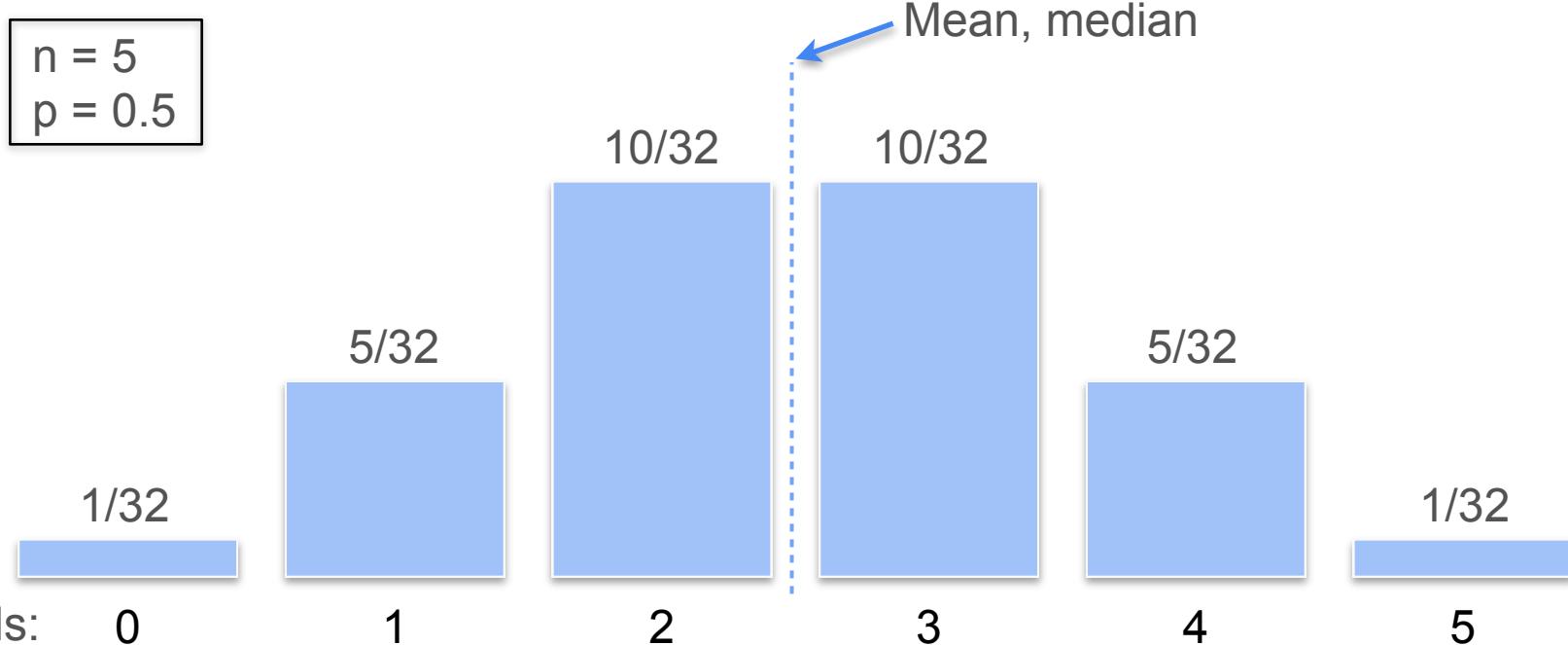
Mean, Median and Mode in Binomial Distribution



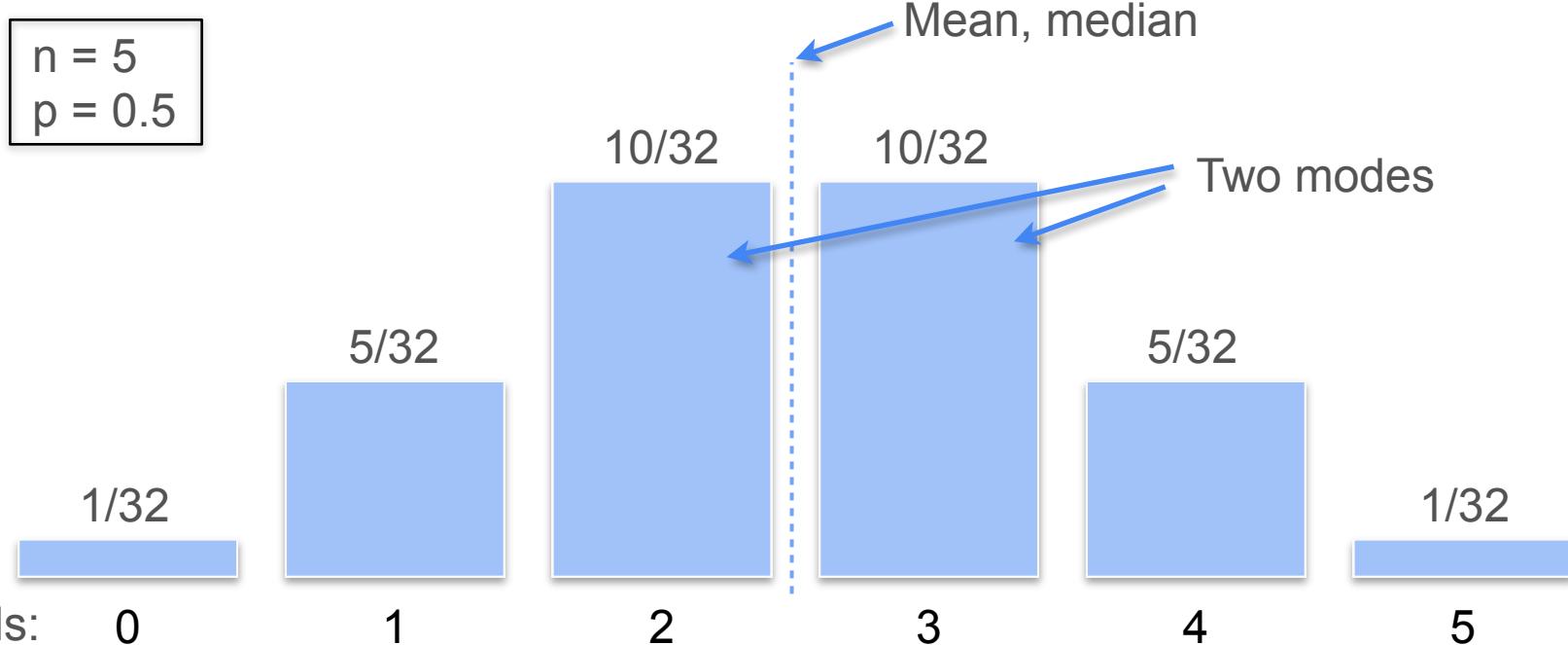
Mean, Median and Mode in Binomial Distribution



Mean, Median and Mode in Binomial Distribution

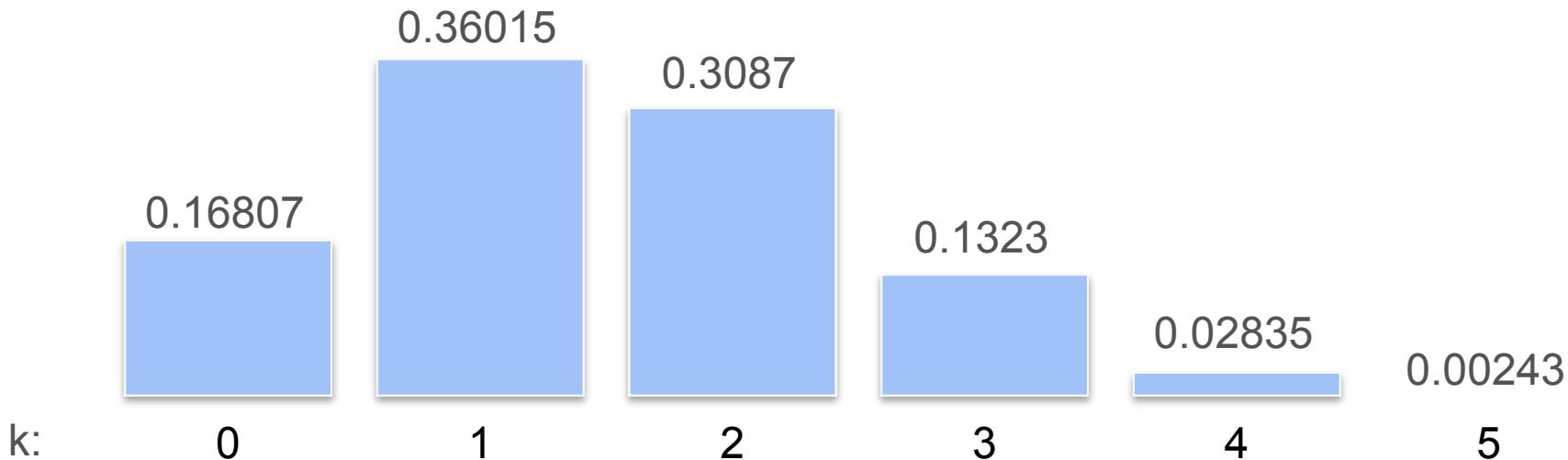


Mean, Median and Mode in Binomial Distribution



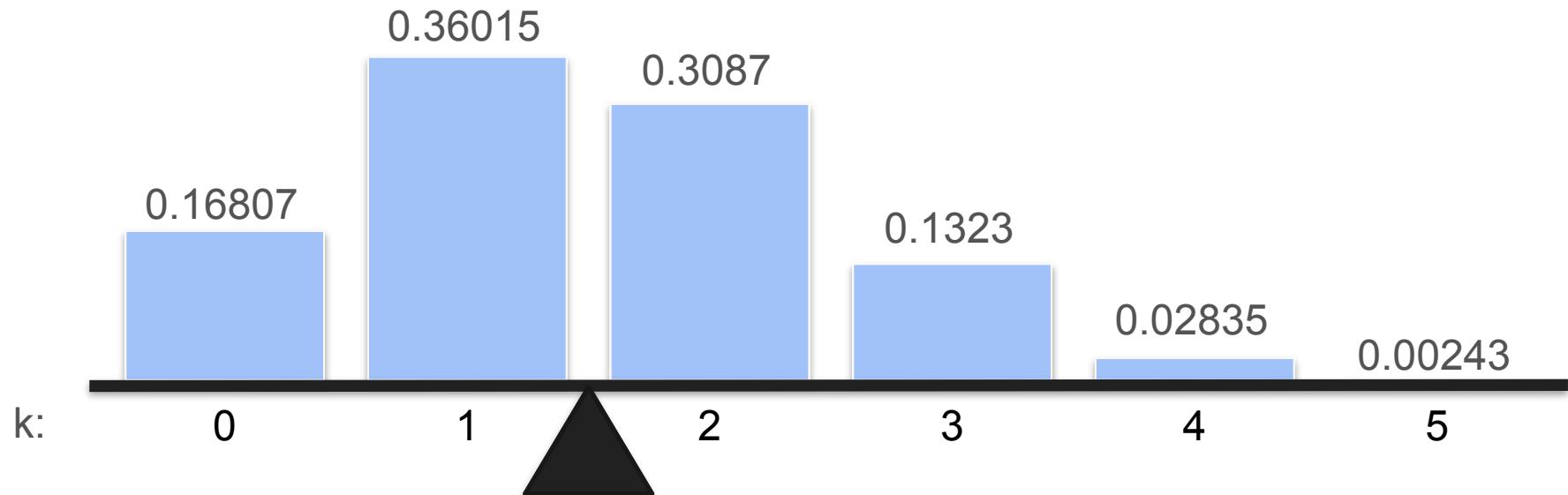
Mean, Median and Mode in Binomial Distribution

$n = 5$
 $p = 0.3$

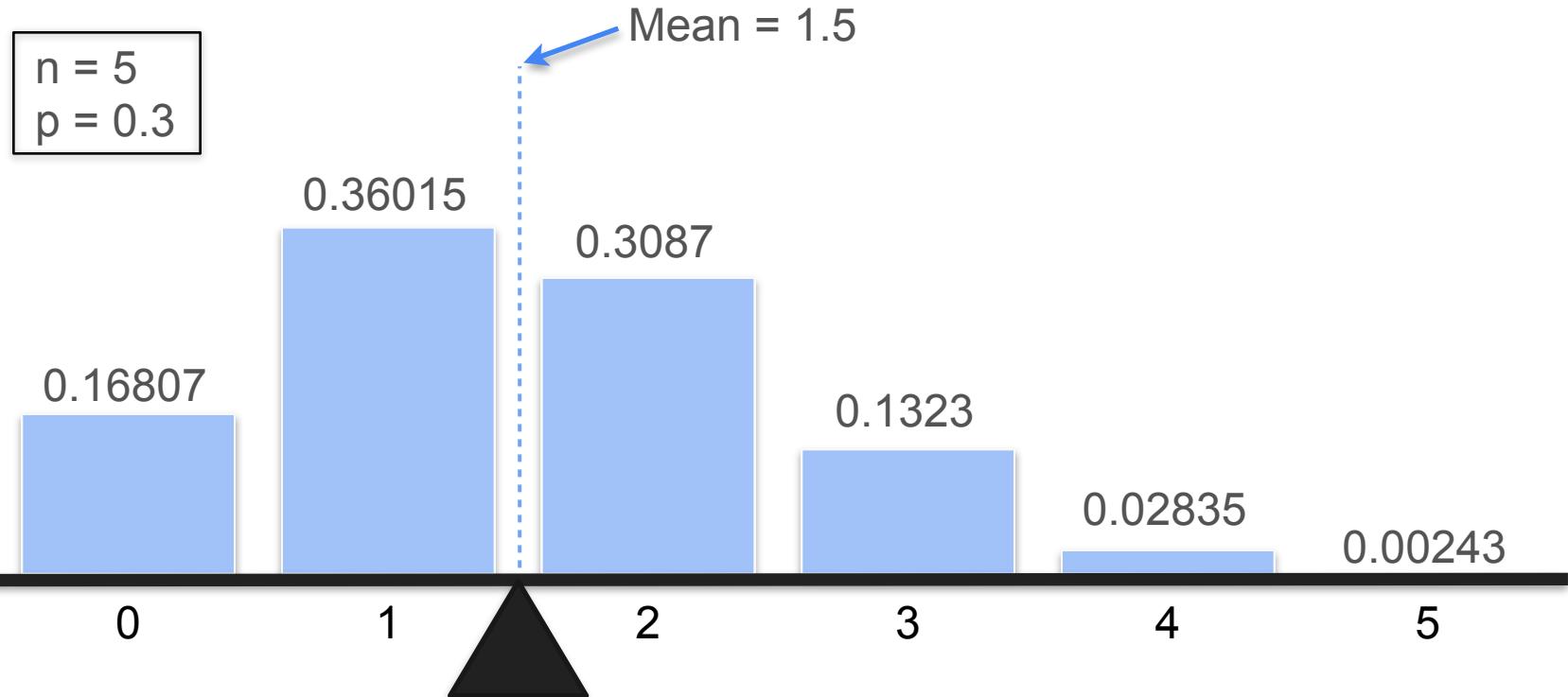


Mean, Median and Mode in Binomial Distribution

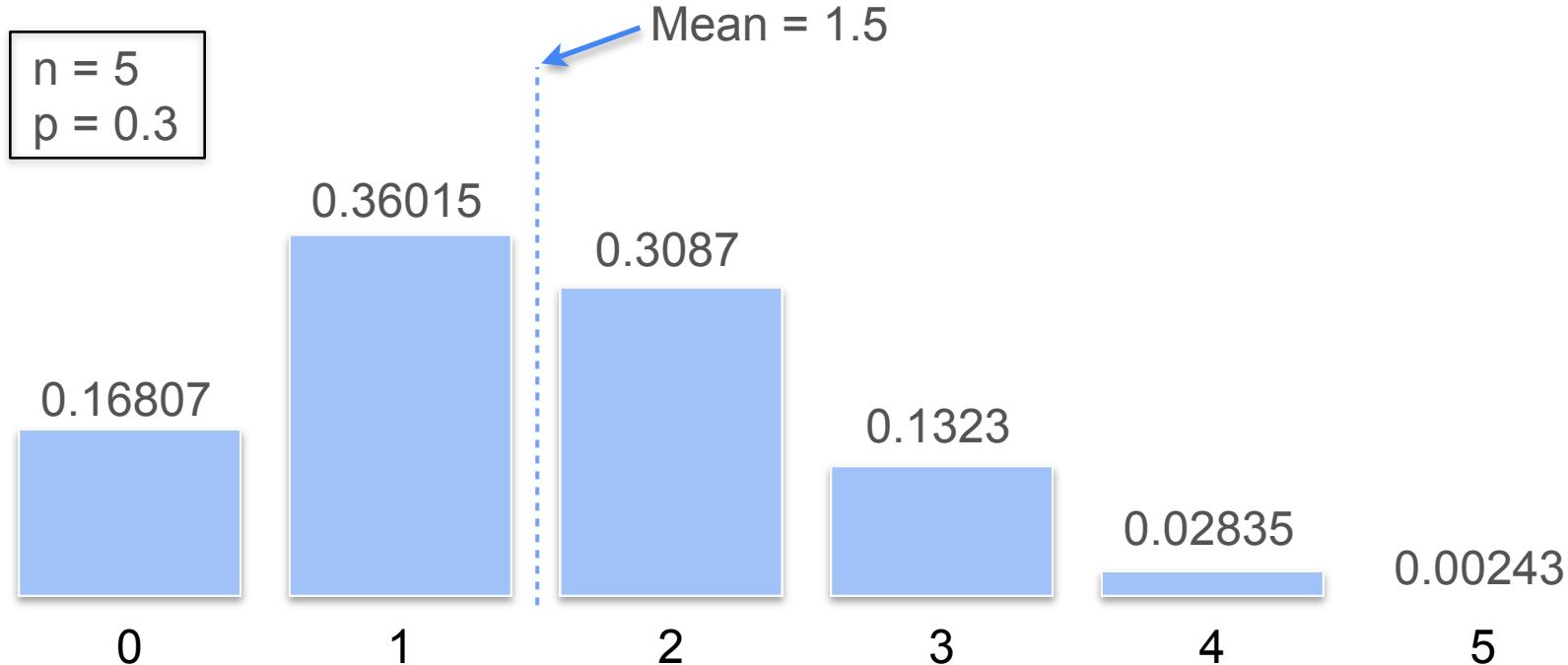
$n = 5$
 $p = 0.3$



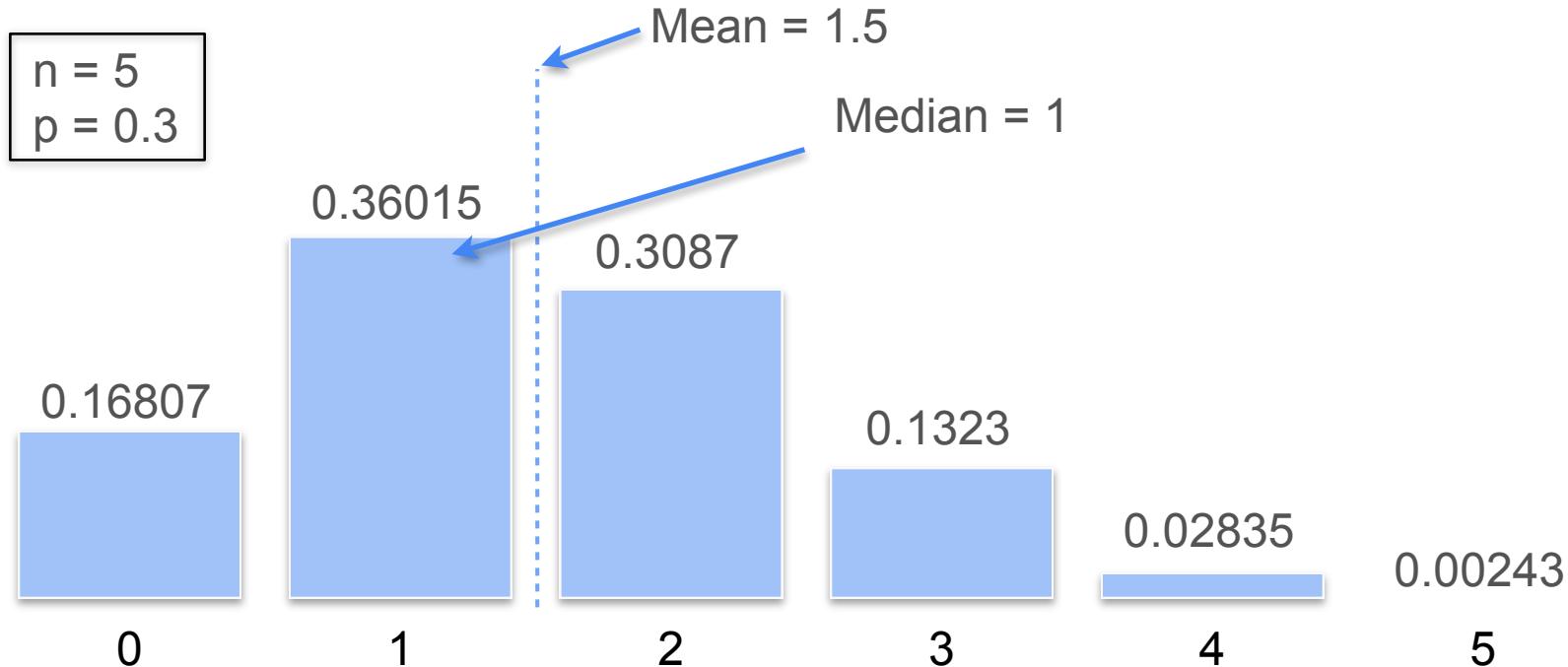
Mean, Median and Mode in Binomial Distribution



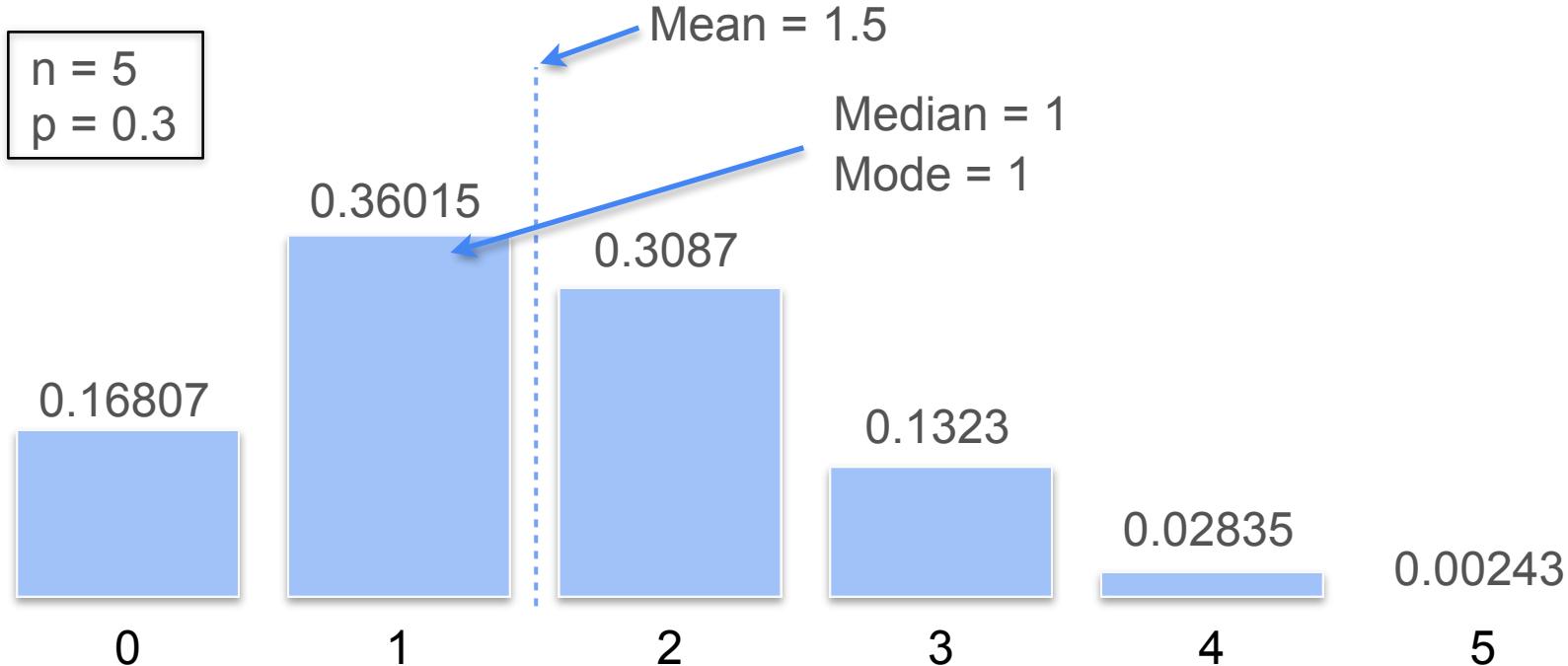
Mean, Median and Mode in Binomial Distribution



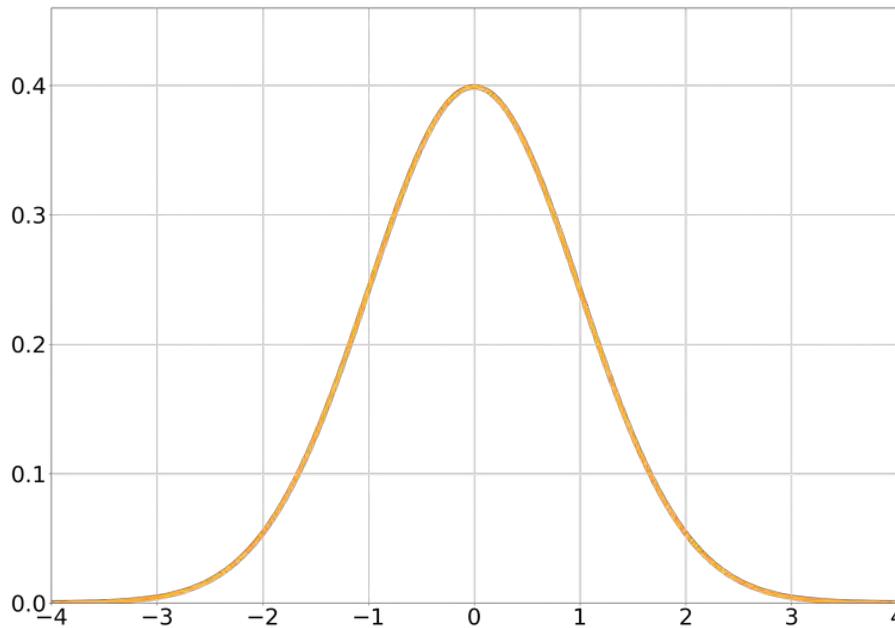
Mean, Median and Mode in Binomial Distribution



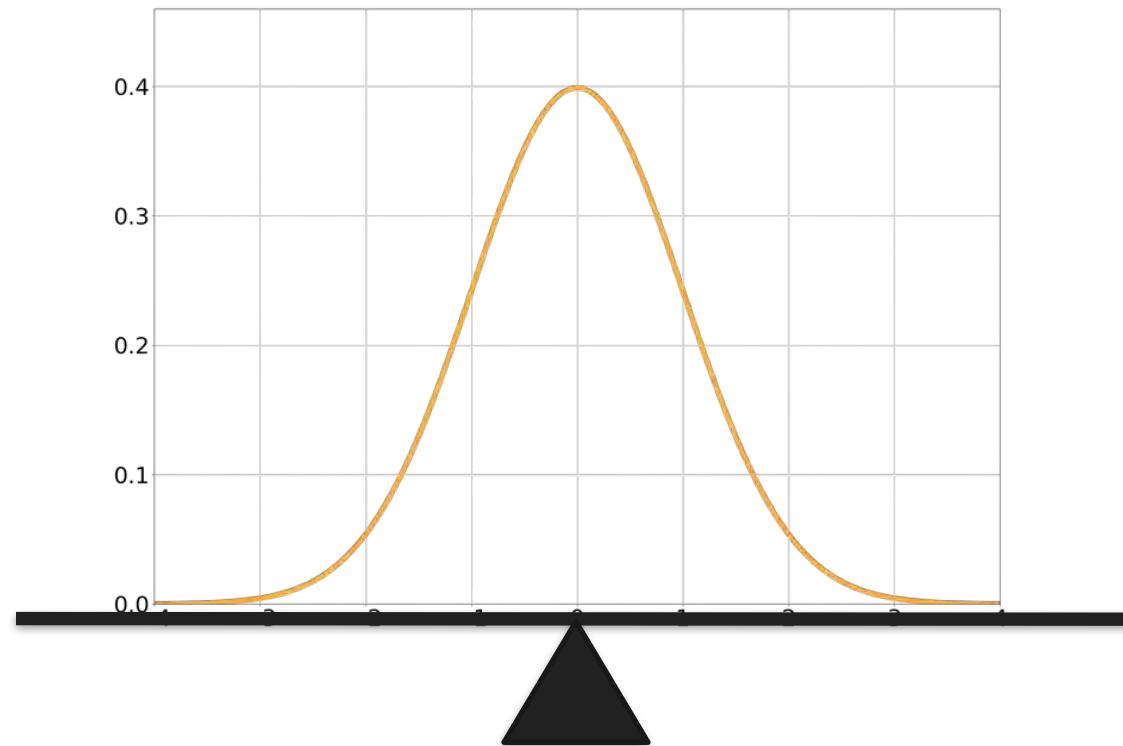
Mean, Median and Mode in Binomial Distribution



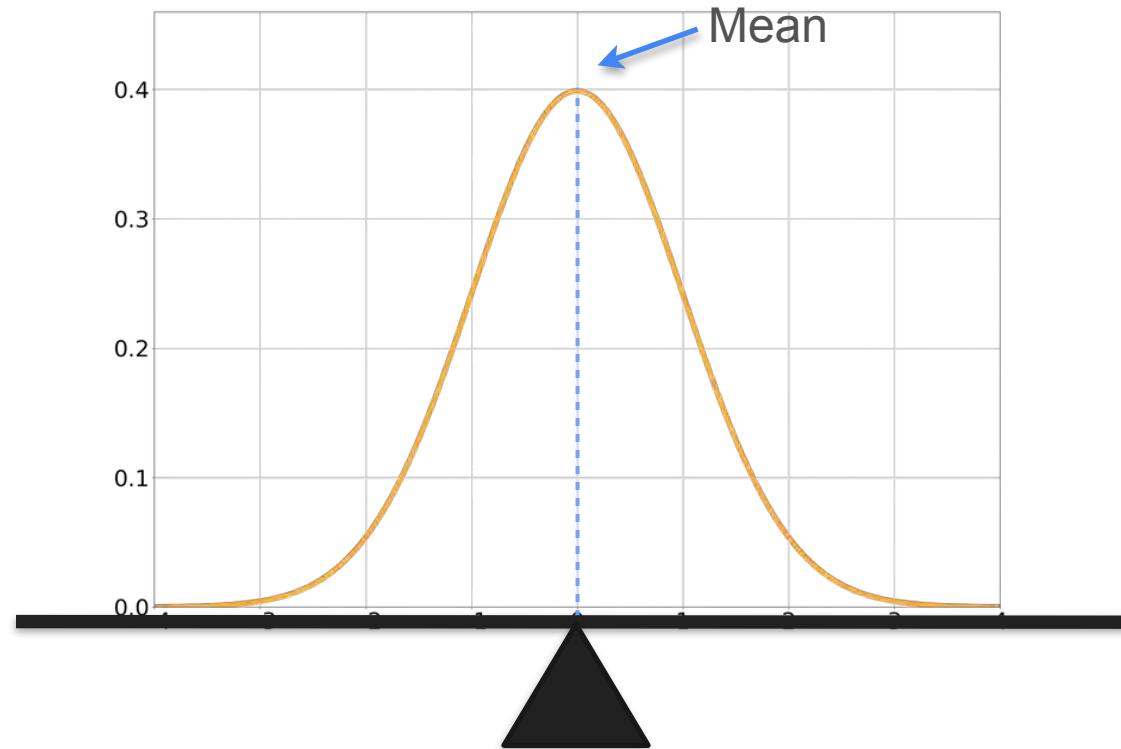
Mean, Median and Mode in Normal Distribution



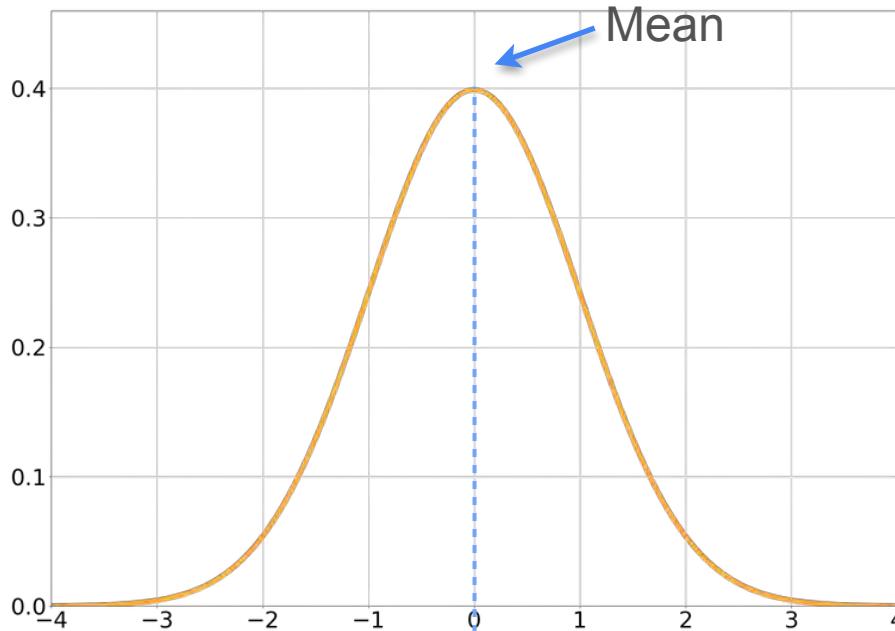
Mean, Median and Mode in Normal Distribution



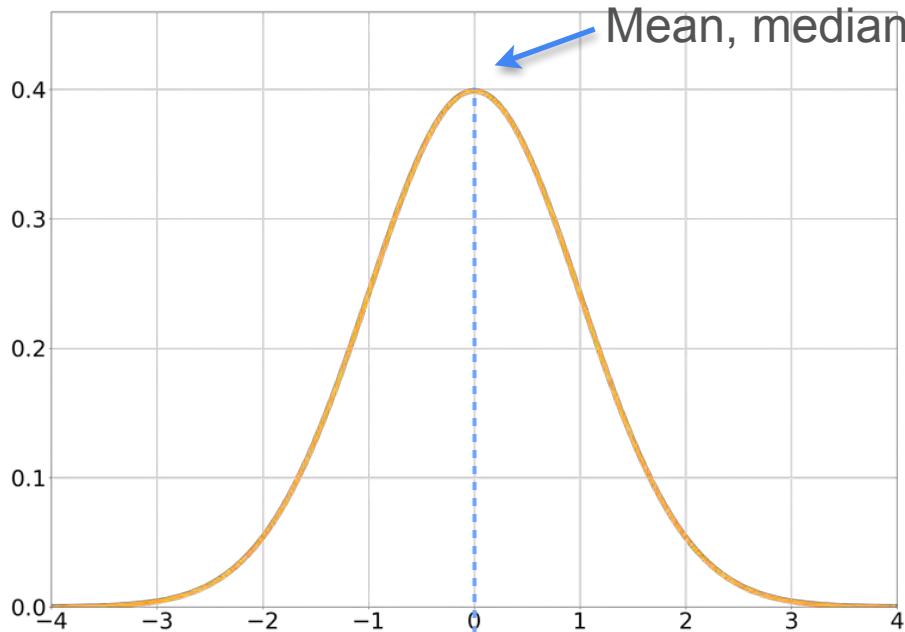
Mean, Median and Mode in Normal Distribution



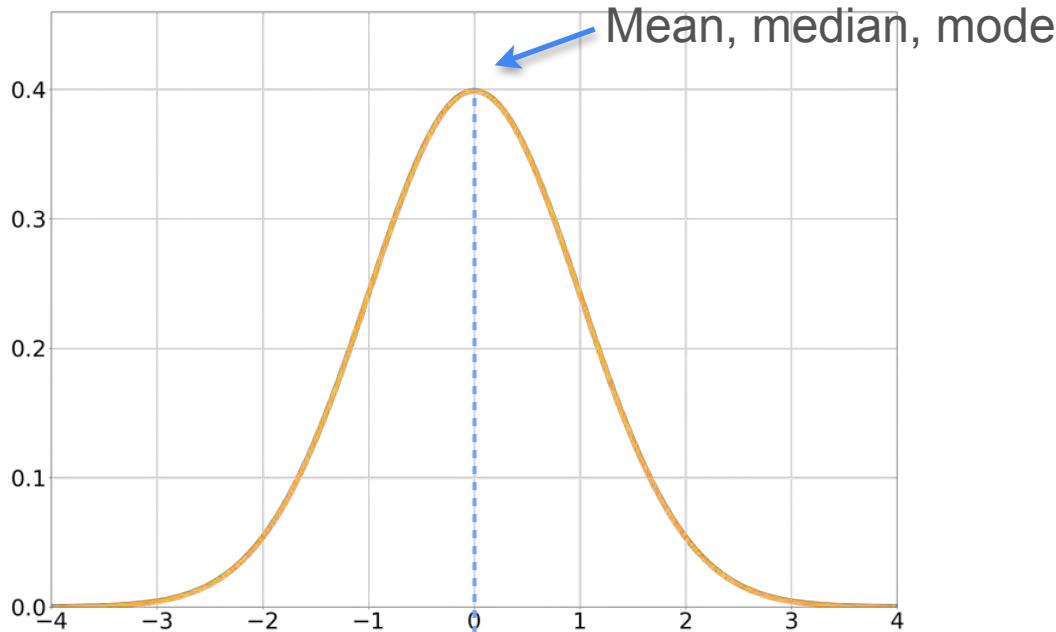
Mean, Median and Mode in Normal Distribution



Mean, Median and Mode in Normal Distribution



Mean, Median and Mode in Normal Distribution





DeepLearning.AI

Describing Distributions

Expected Value

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$6

Do you play the game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$4

Do you play the game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$4

Do you play the game?

What is the maximum amount of money you would pay to play this game?

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term:

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \0

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5 \rightarrow$ You expect to win \$5 on average
 $\mathbb{E}[X] = 5$

Expected Value: Motivation Example 1

You play a game with a friend



You win 10 dollars



You win nothing

Game cost:

\$5

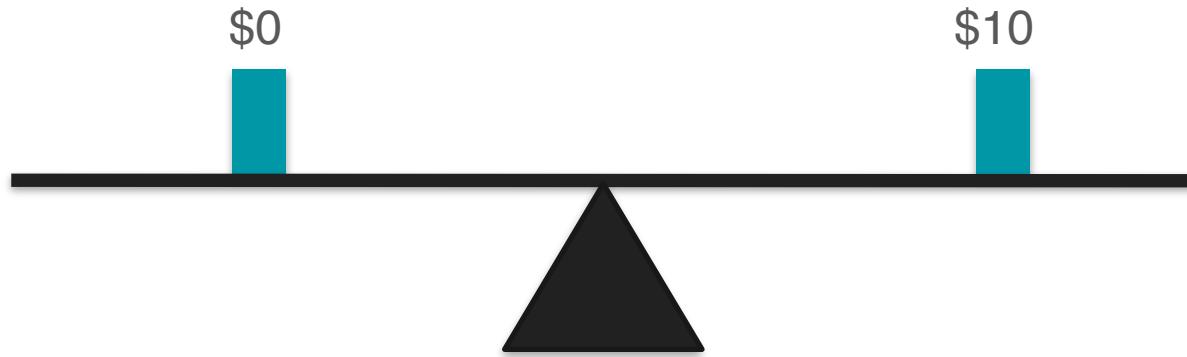
Long term: $0.5 \cdot \$10 + 0.5 \cdot \$0 = \$5$ 

You expect to win \$5 on average
 $E[X] = 5$

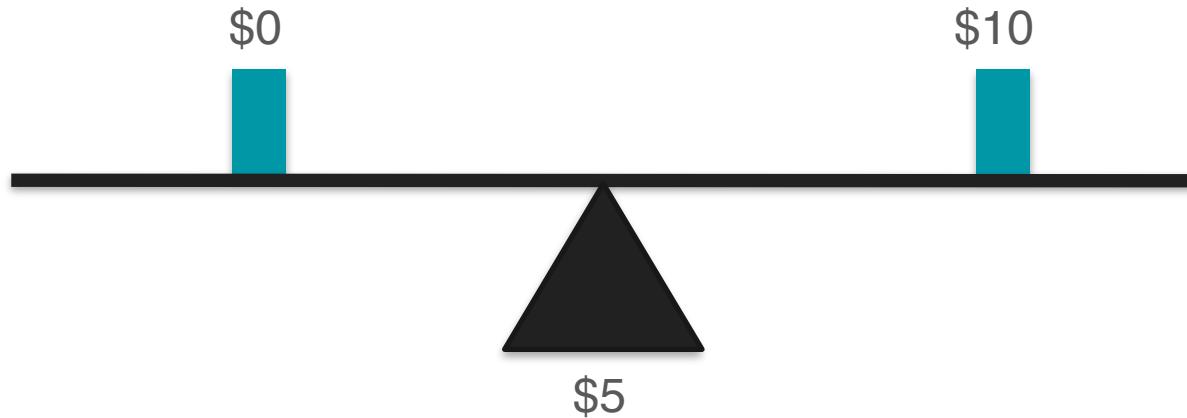
Expected Value: Motivation Example 1



Expected Value: Motivation Example 1



Expected Value: Motivation Example 1



Expected Value: Motivation Example 2

Play another game



Flip three coins. For each heads you win \$1

What is the maximum amount of money you would pay to play this game?

Expected Value: Motivation Example 2

Number of heads:

0



1



2



3



Expected Value: Motivation Example 2

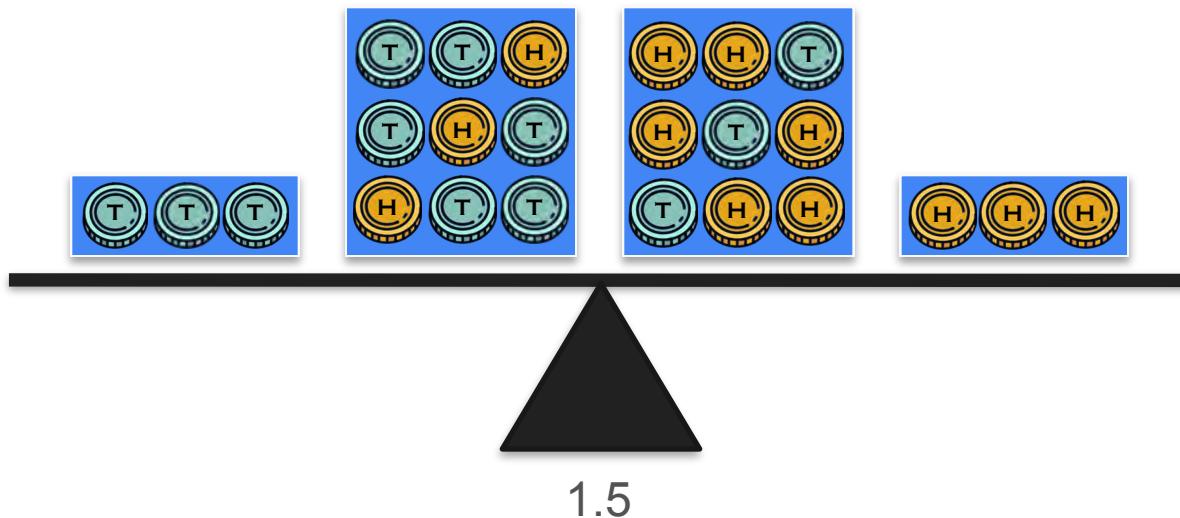
Number of heads:

0

1

2

3



Expected Value: Motivation Example 2

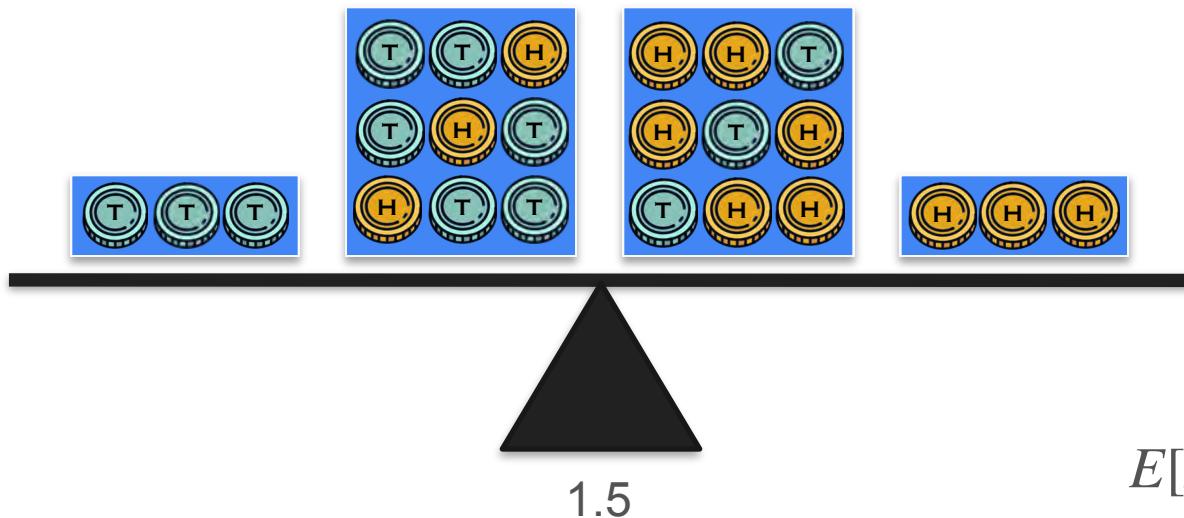
Number of heads:

0

1

2

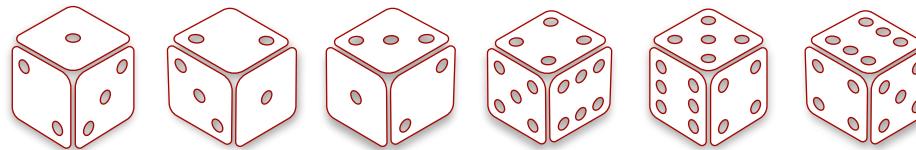
3



Expected Value: Motivation Example 3

Probability: 1/6 1/6 1/6 1/6 1/6 1/6

Roll: 1 2 3 4 5 6

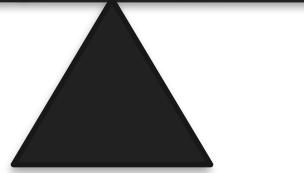
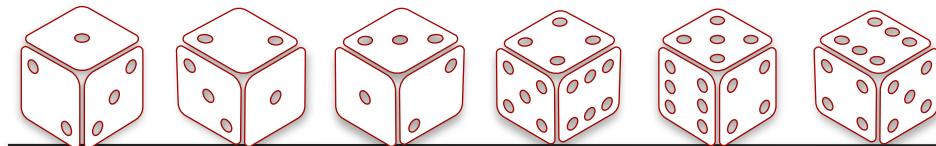


$$\frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

Expected Value: Motivation Example 3

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6

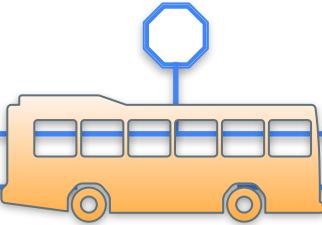


3.5

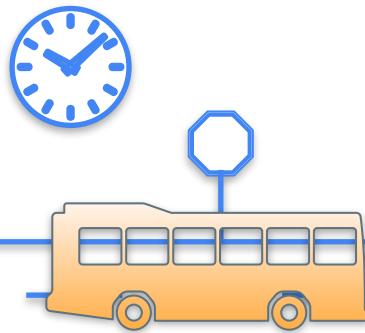
Expected Value



Expected Value



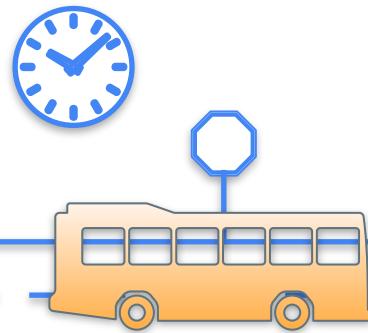
Expected Value



Expected Value

Waiting Time

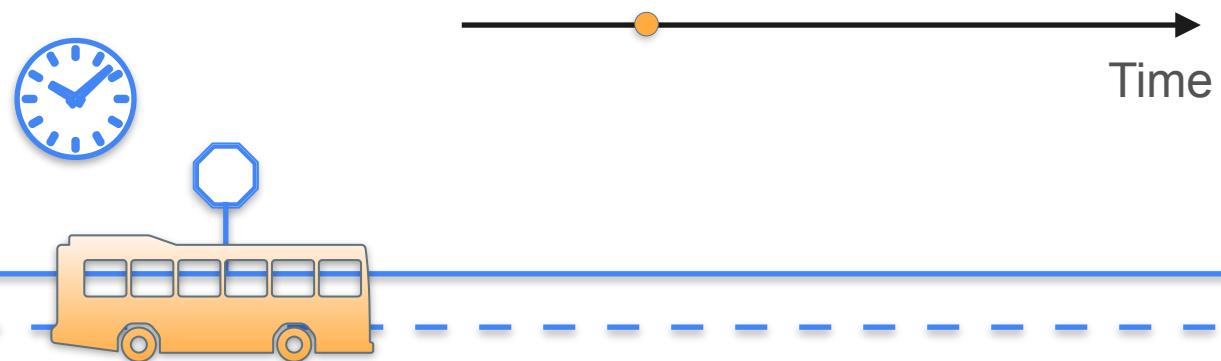
15 min



Expected Value

Waiting Time

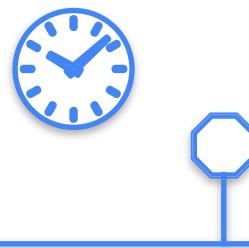
15 min



Expected Value

Waiting Time

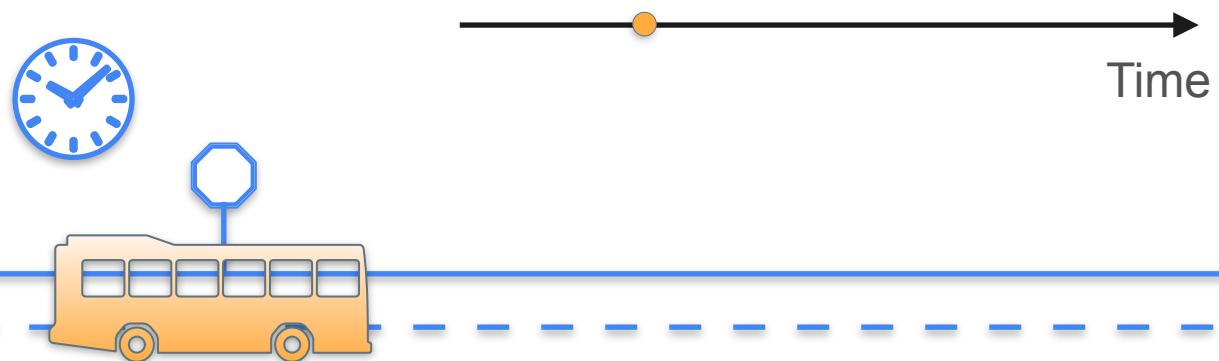
15 min



Expected Value

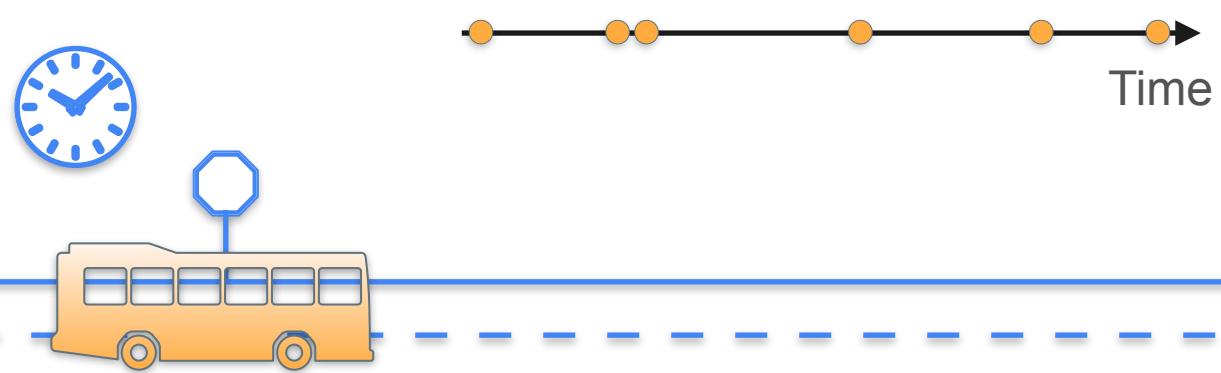
Waiting Time

15 min



Expected Value

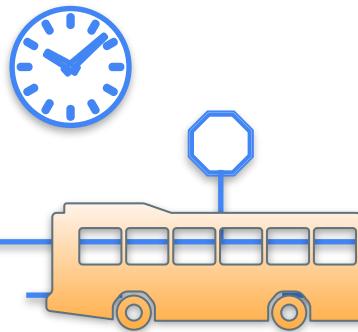
Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min

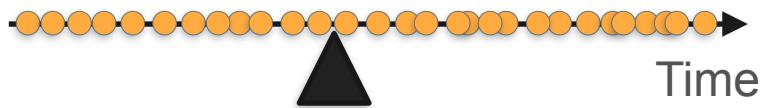
Average = 27.833



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

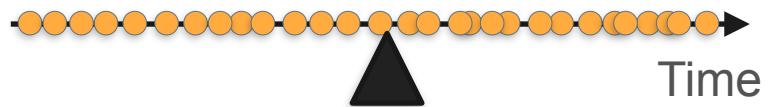
Average = 27.833



Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

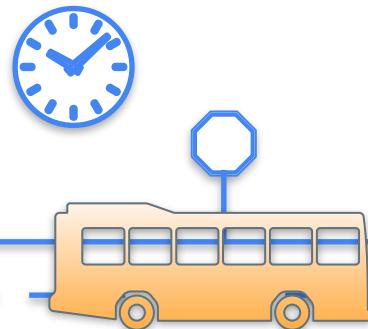
Average = 30



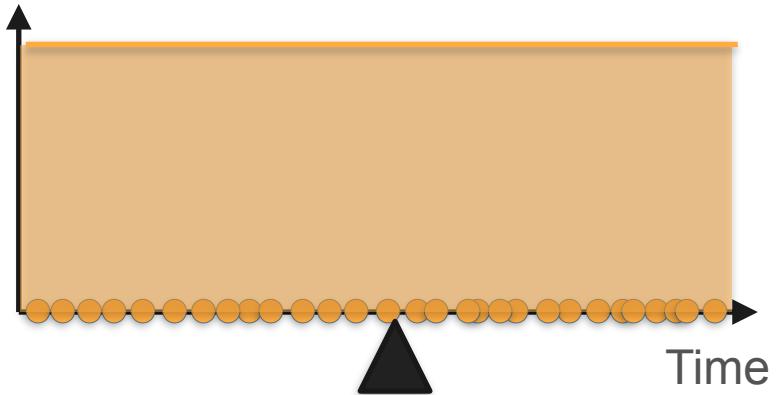
Expected Value

Waiting Time
15 min
32 min
58 min
1 min
47 min
14 min
37 min
8 min
29 min
55 min
...

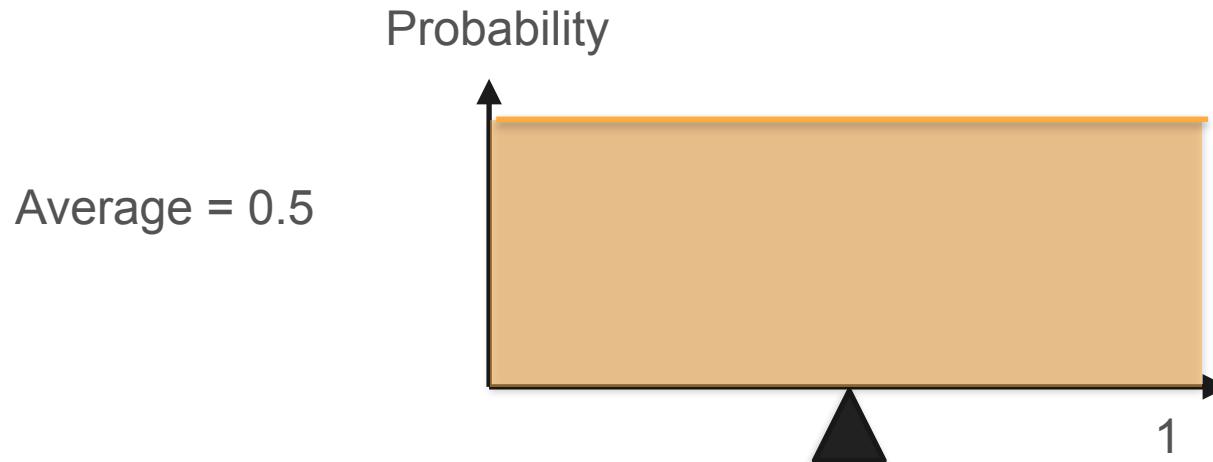
Average = 30



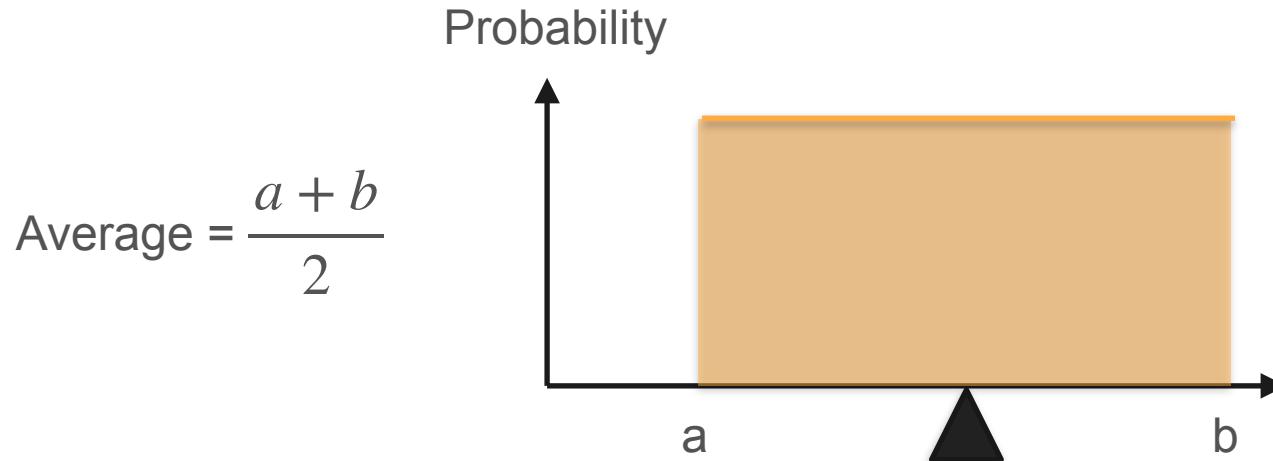
Probability



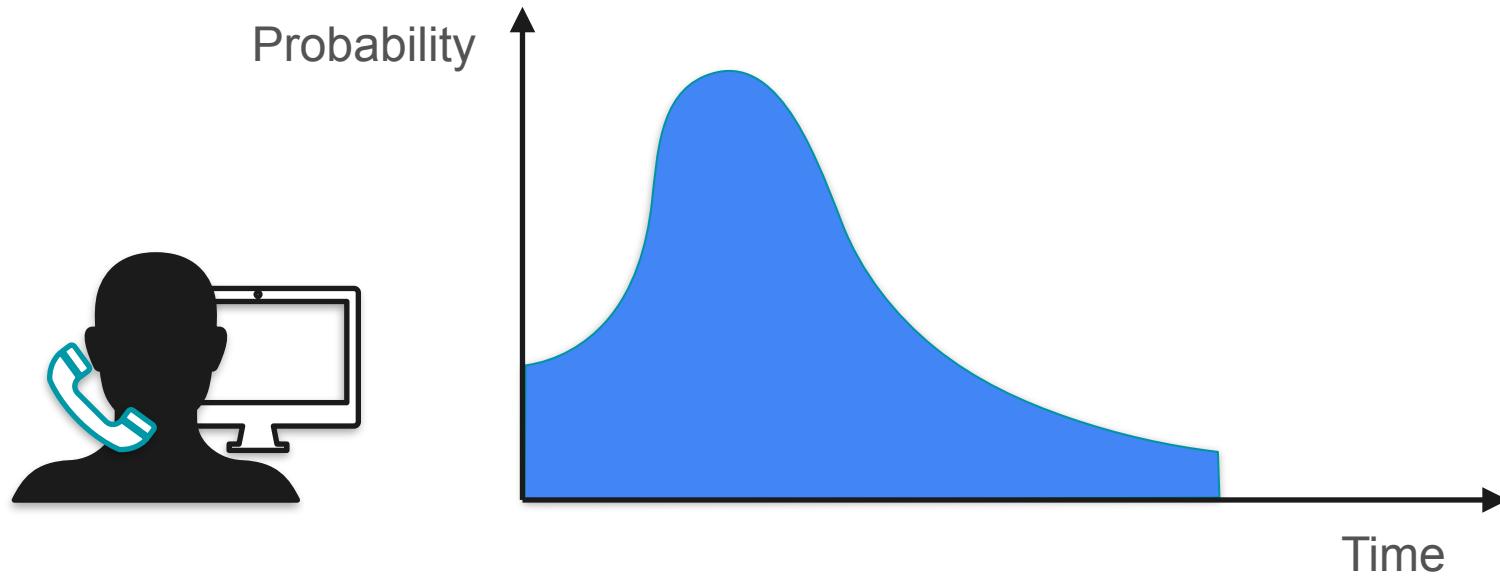
Expected Value: Uniform Distribution



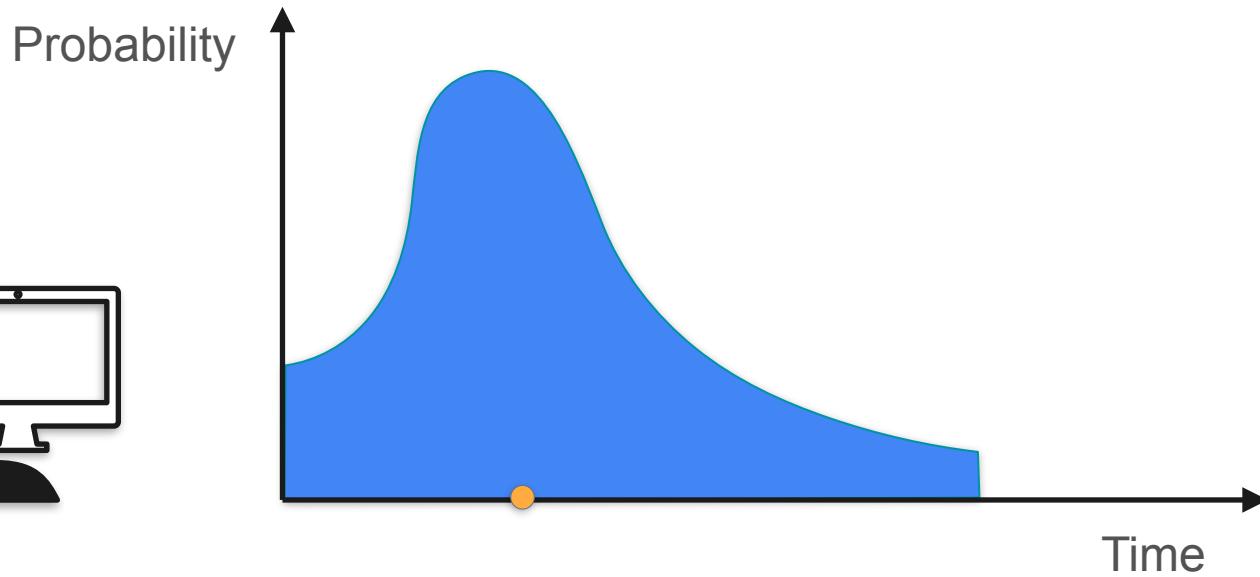
Expected Value: Uniform Distribution



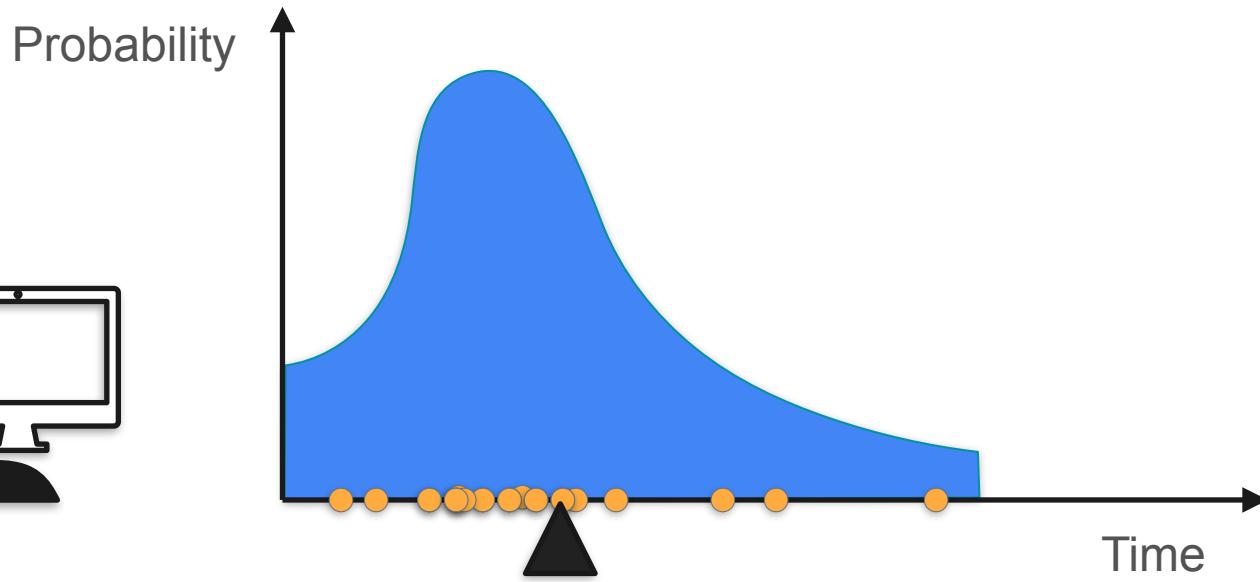
Expected Value



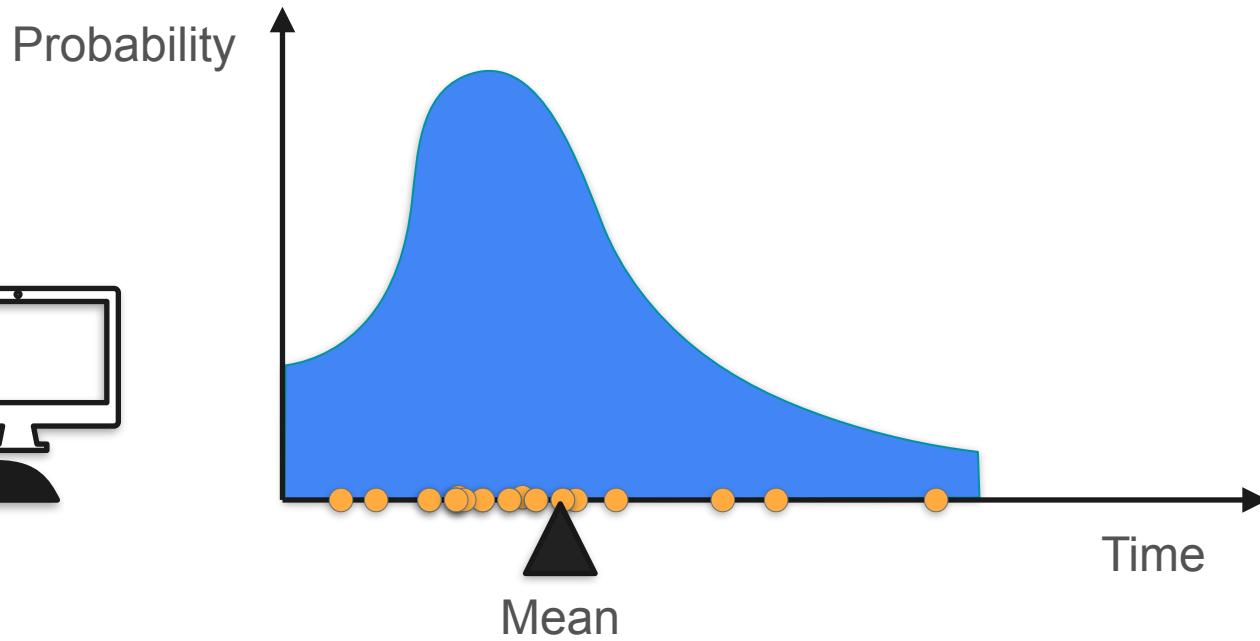
Expected Value



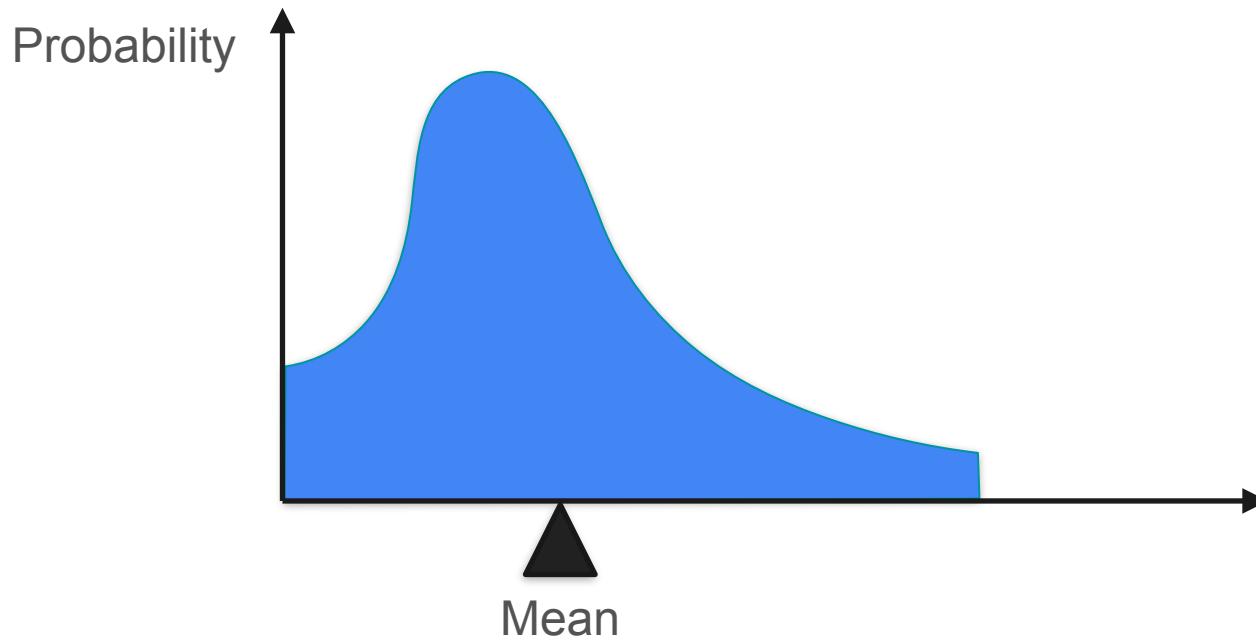
Expected Value



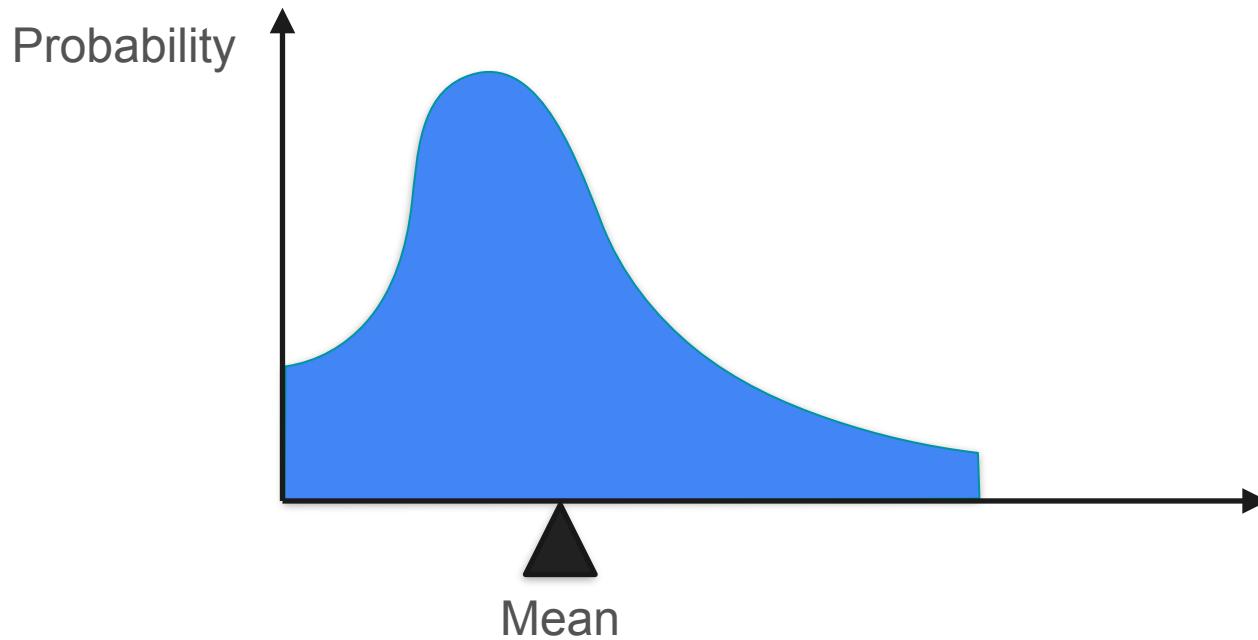
Expected Value



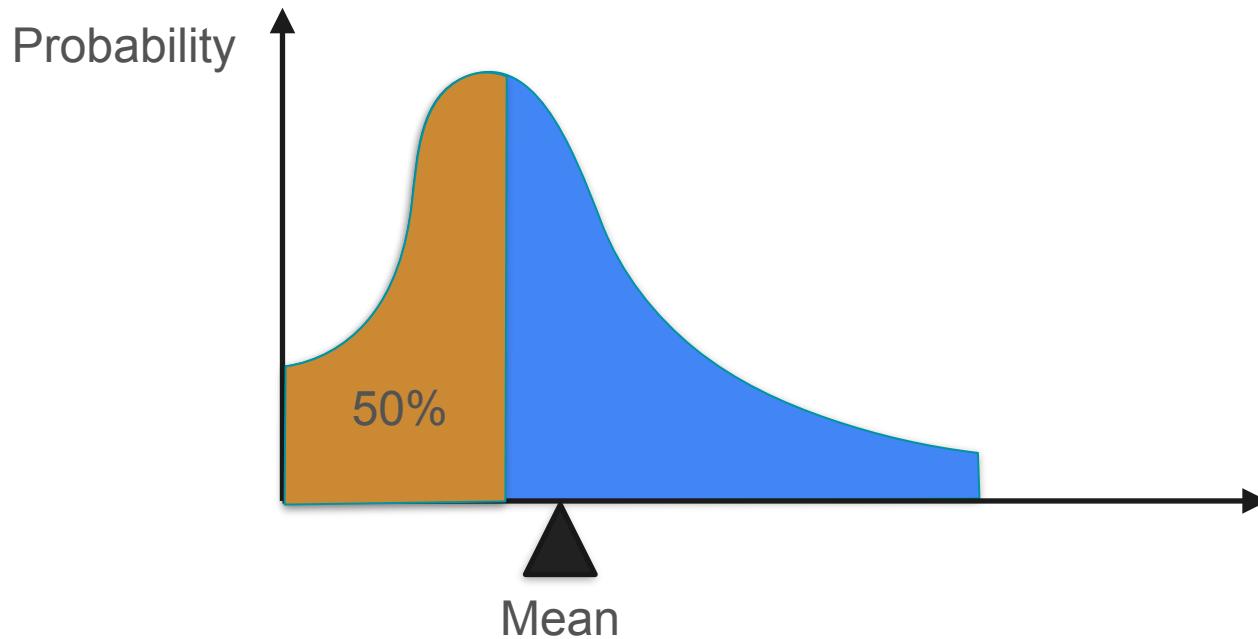
Expected Value: General Case



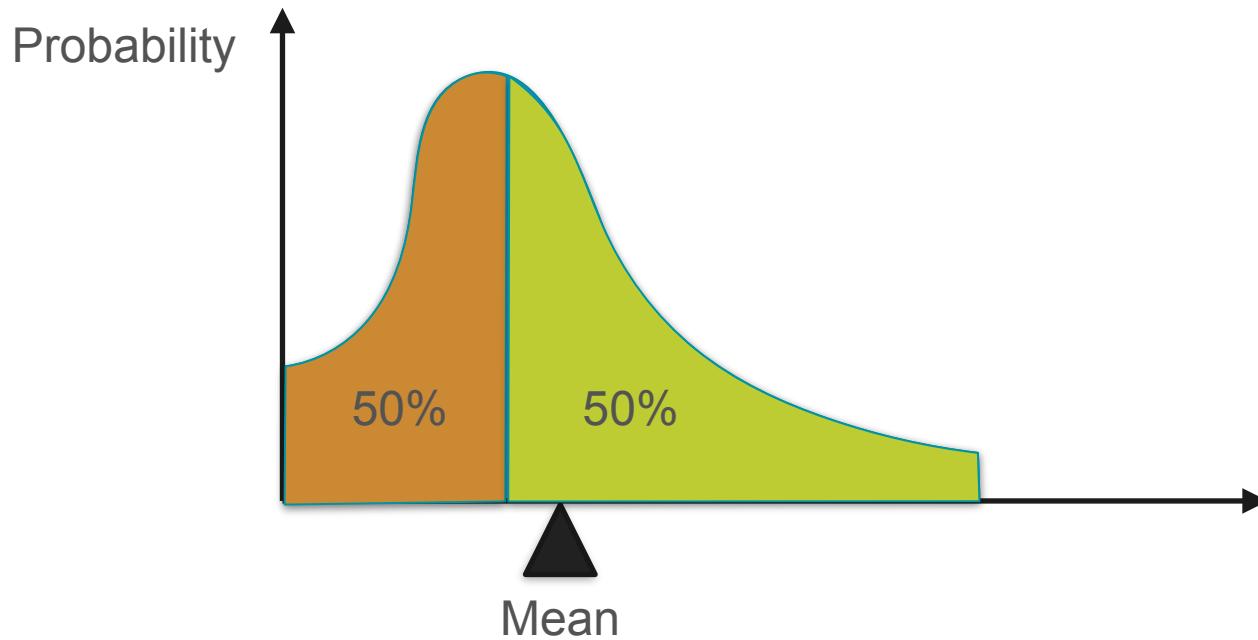
Expected Value: Common Misconception



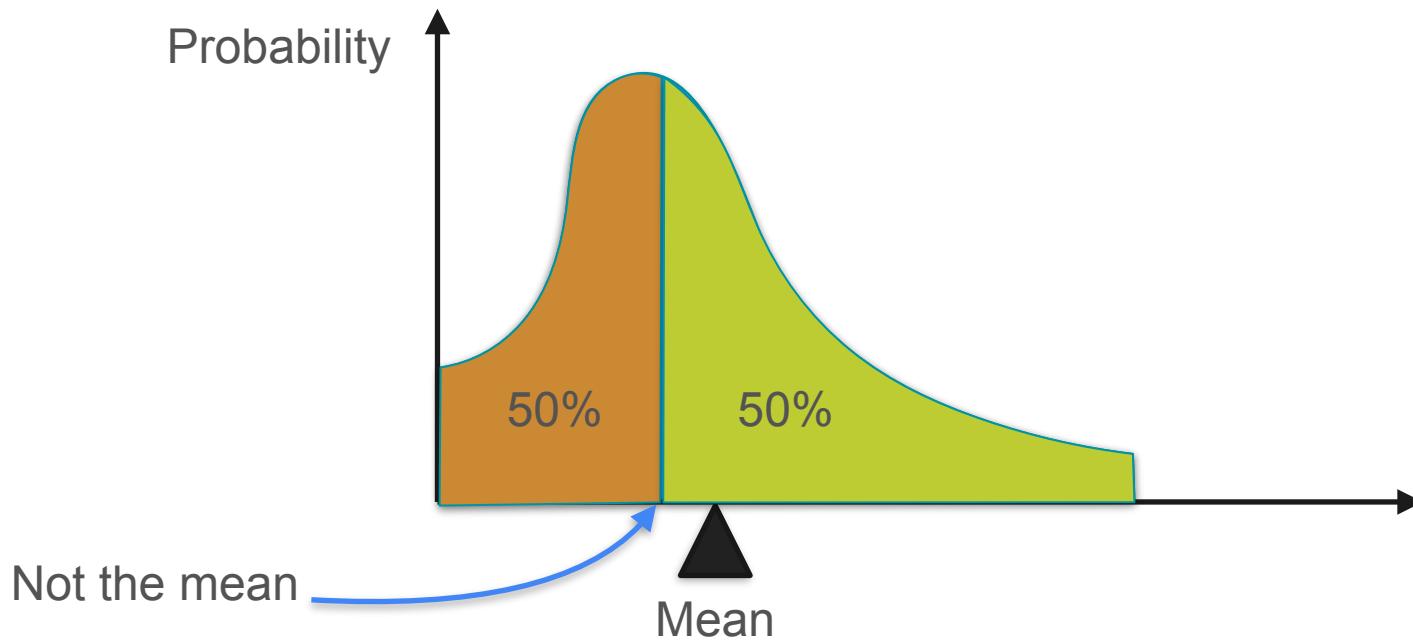
Expected Value: Common Misconception



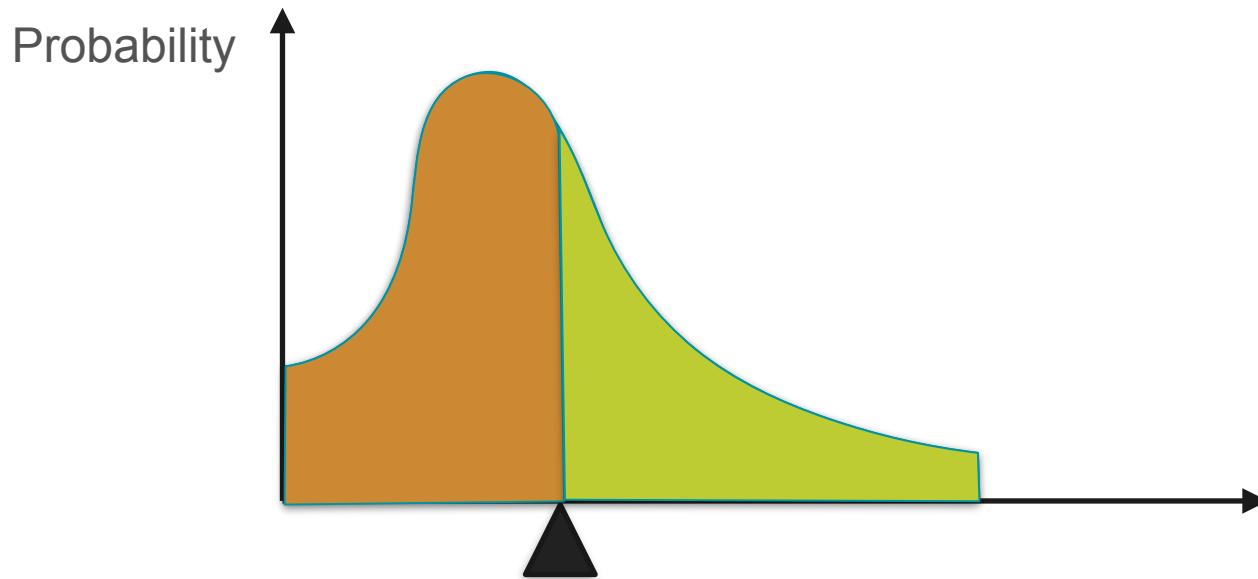
Expected Value: Common Misconception



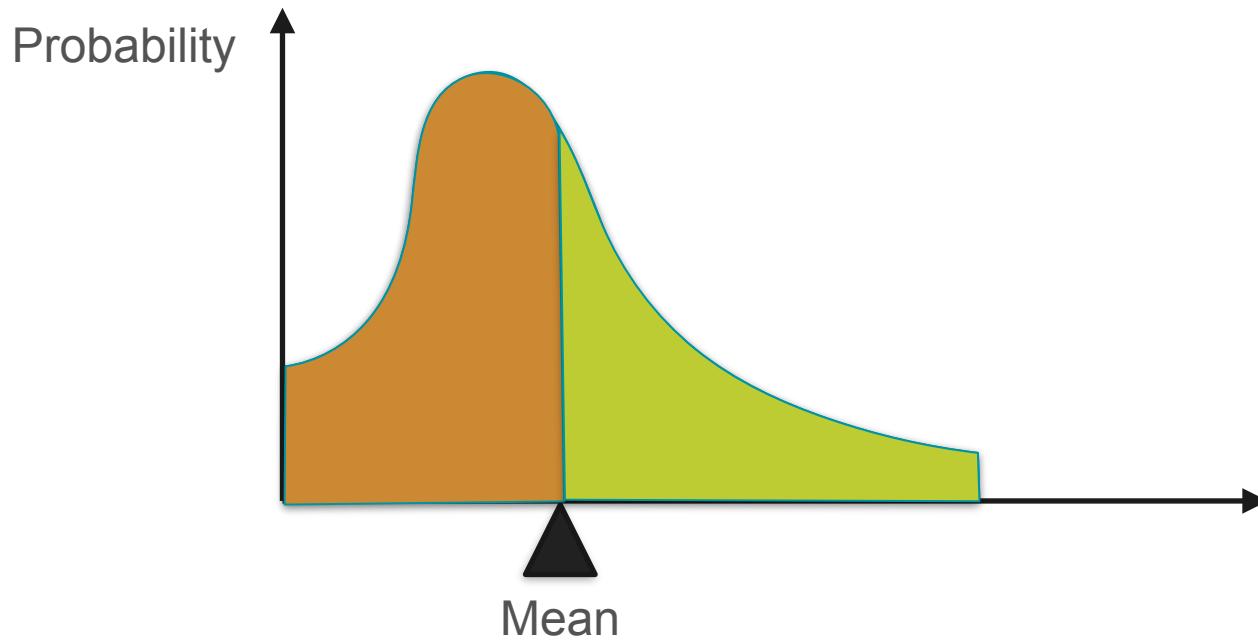
Expected Value: Common Misconception



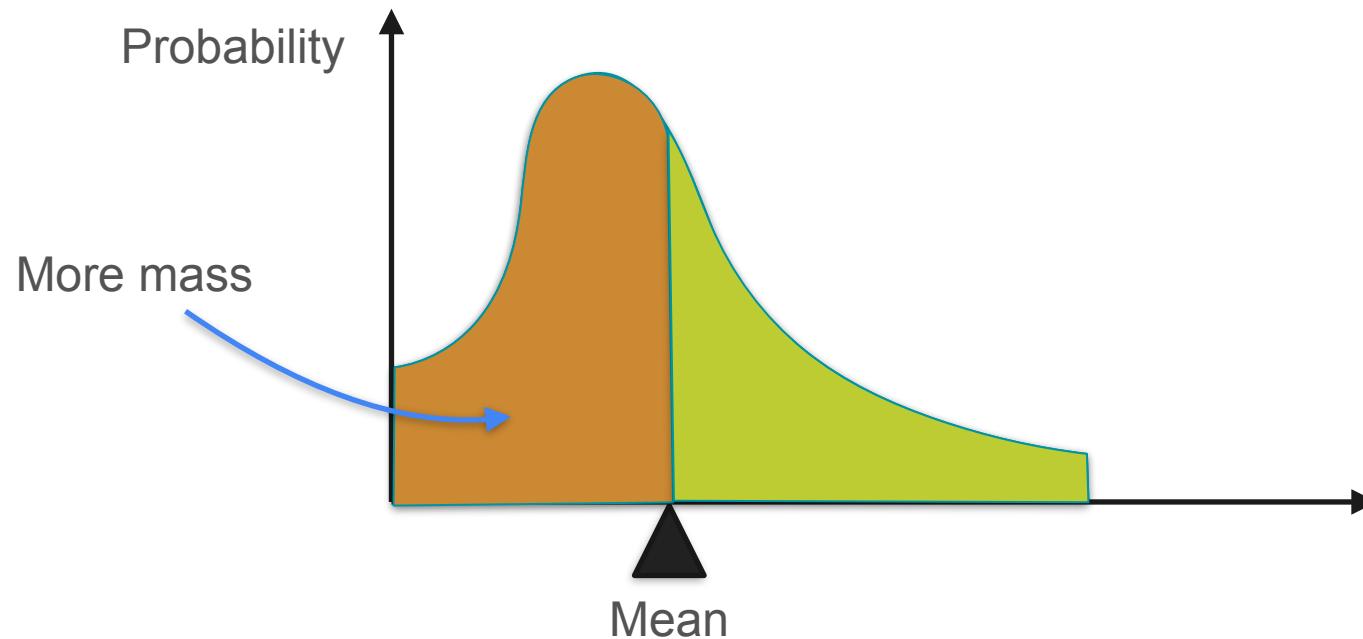
Expected Value: Common Misconception



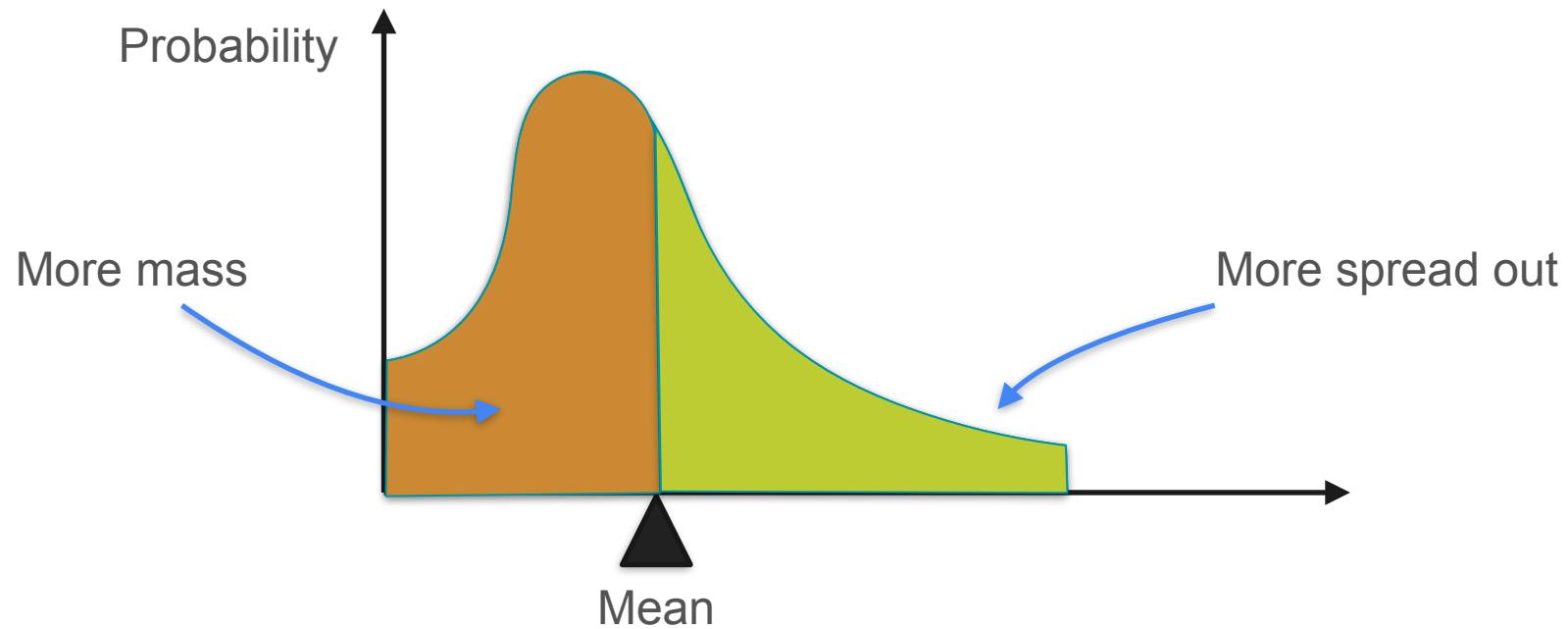
Expected Value: Common Misconception



Expected Value: Common Misconception



Expected Value: Common Misconception



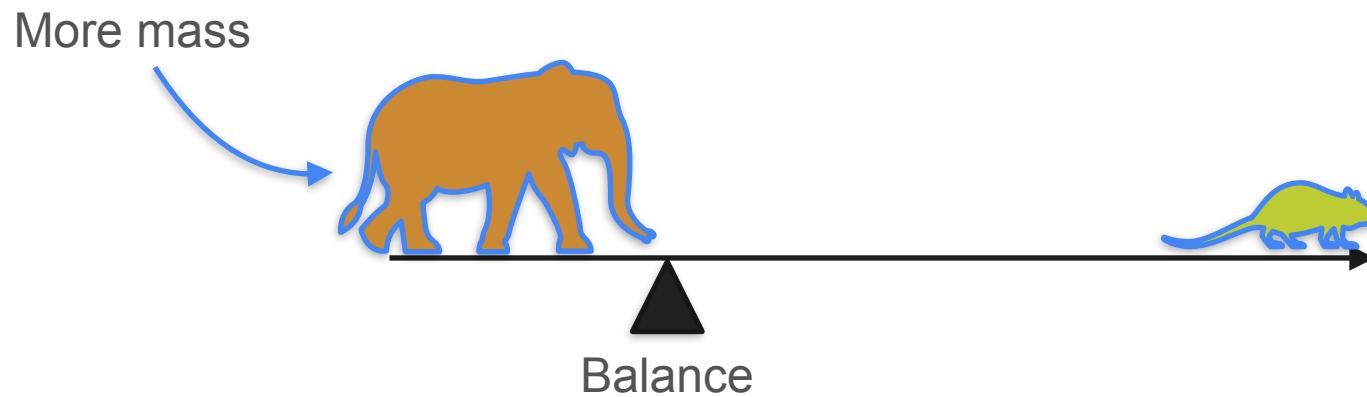
Expected Value: Common Misconception



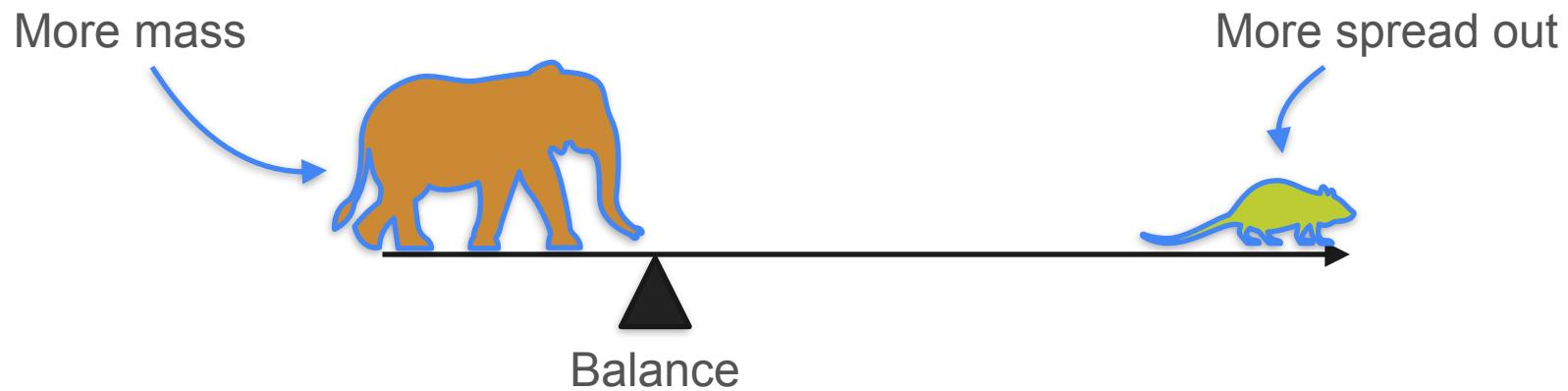
Expected Value: Common Misconception



Expected Value: Common Misconception



Expected Value: Common Misconception



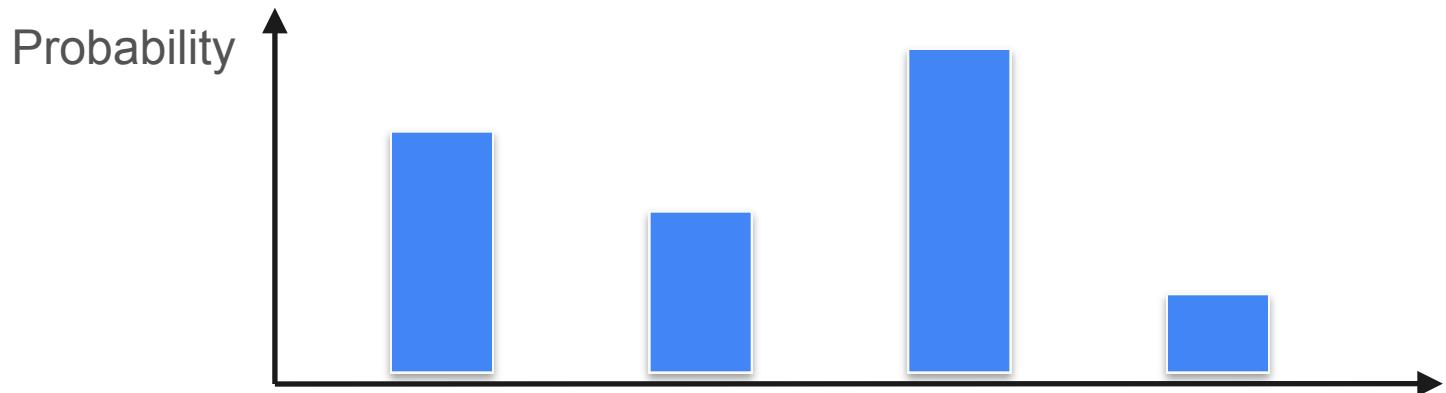


DeepLearning.AI

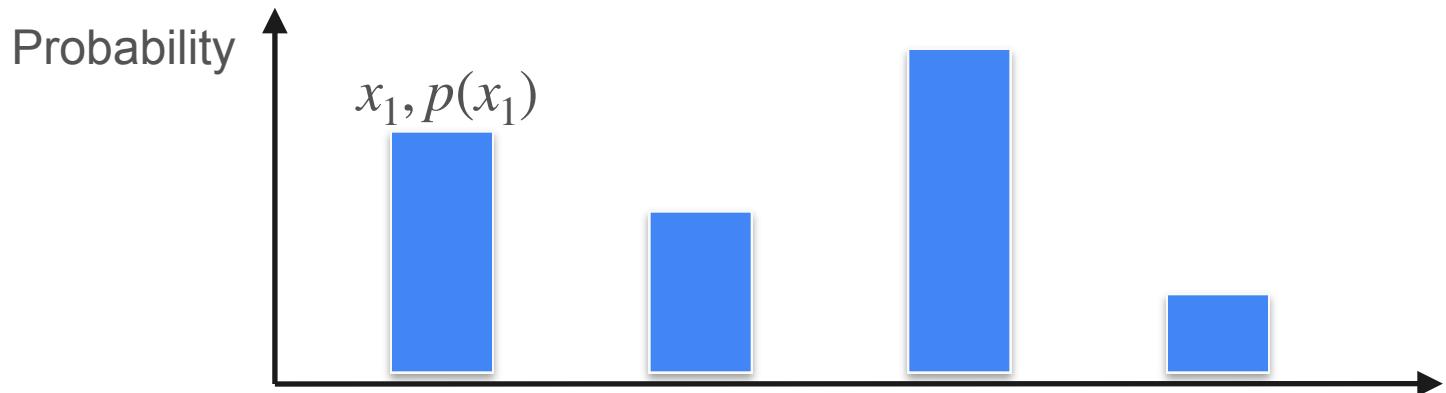
Describing Distributions

Expected value of a function

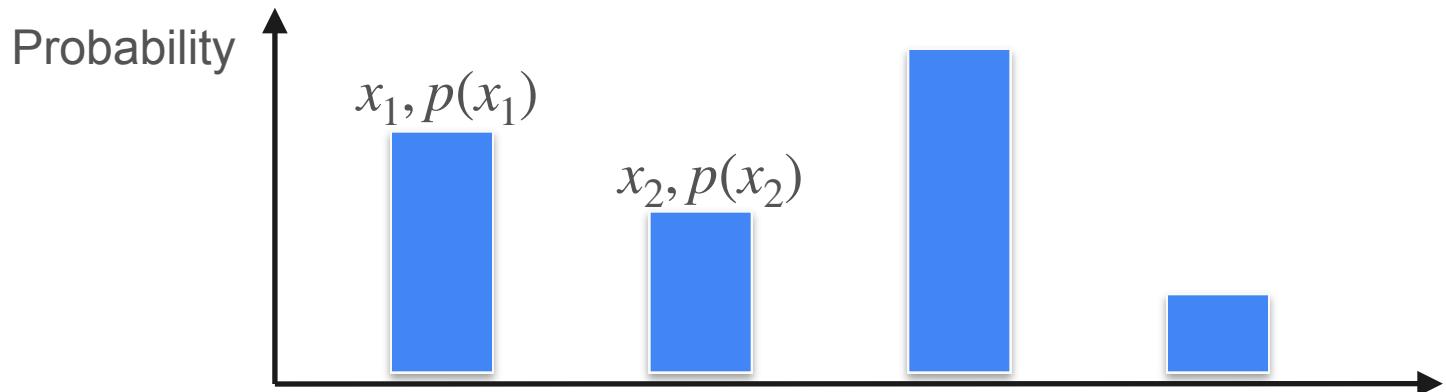
Expected Value of a Function



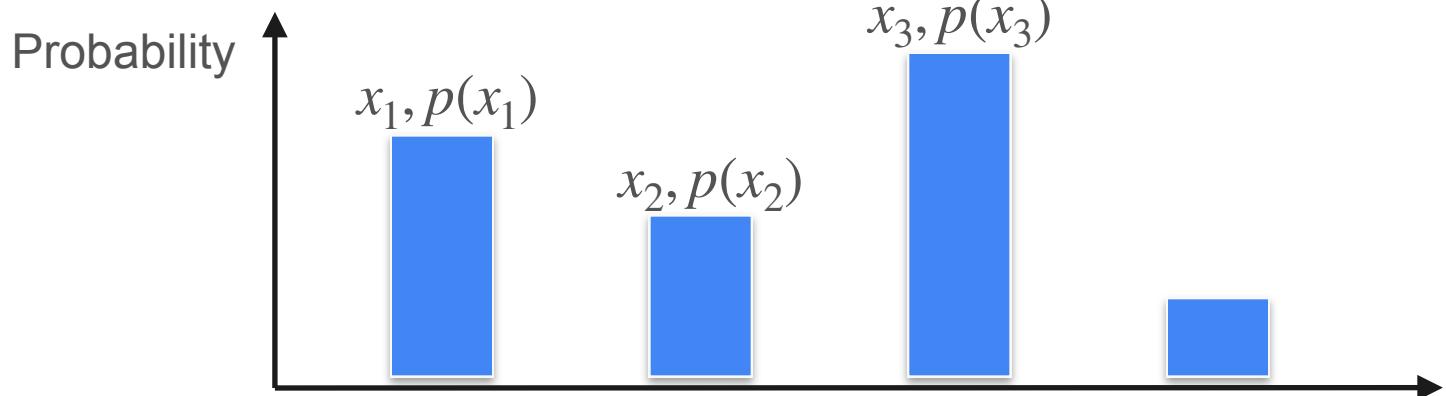
Expected Value of a Function



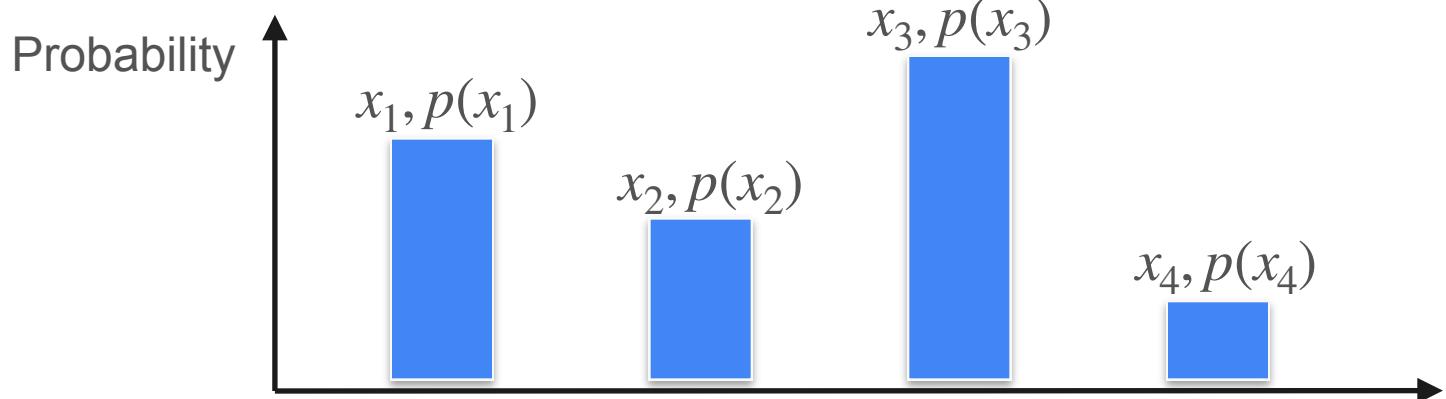
Expected Value of a Function



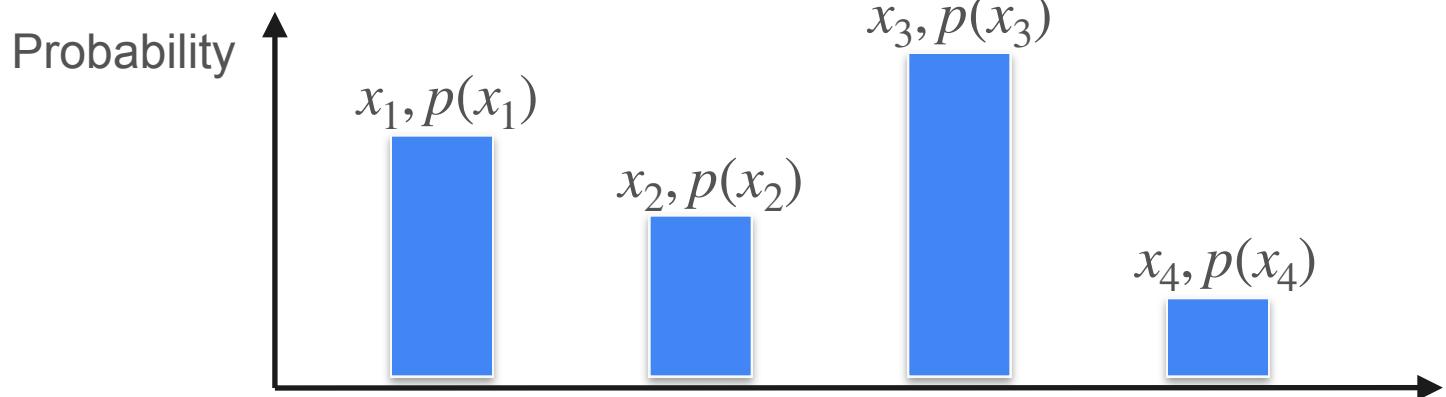
Expected Value of a Function



Expected Value of a Function

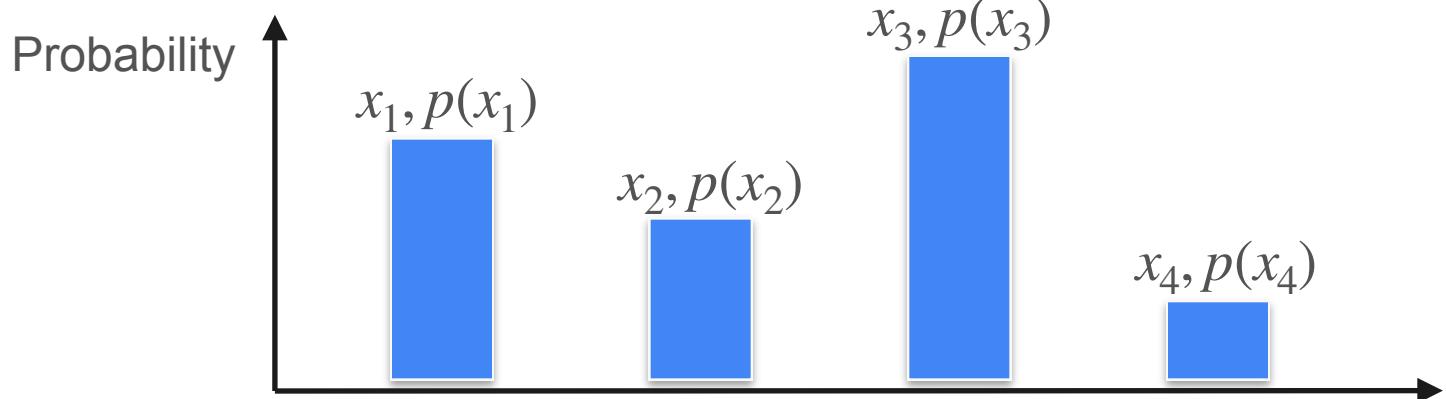


Expected Value of a Function



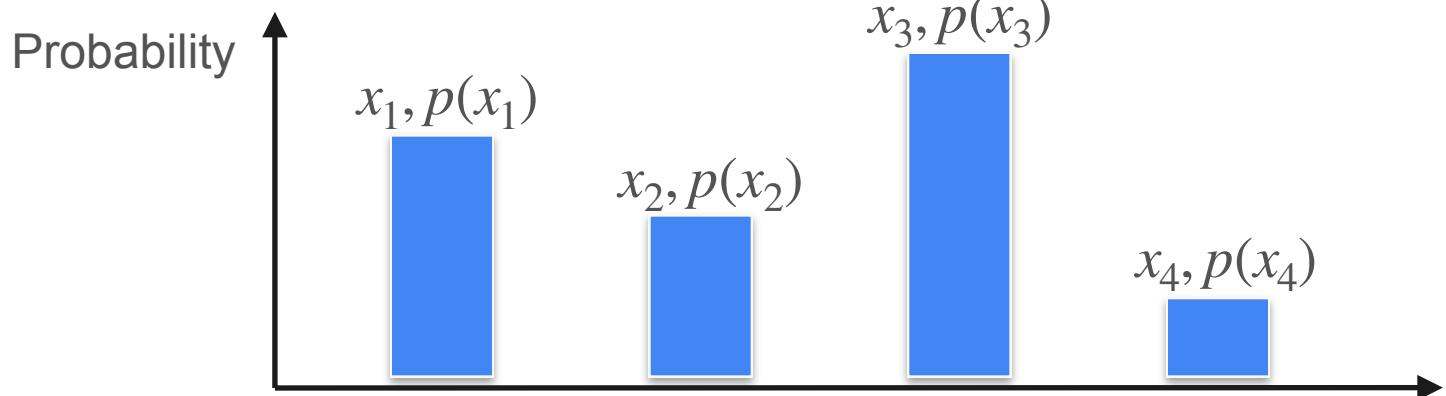
$$\mathbb{E}[X] =$$

Expected Value of a Function



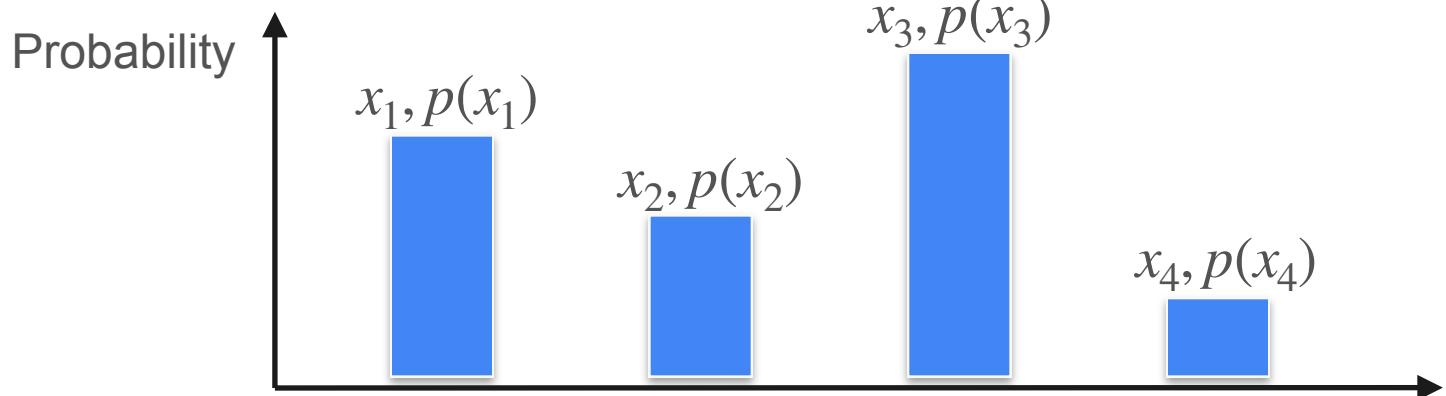
$$\mathbb{E}[X] = x_1 p(x_1)$$

Expected Value of a Function



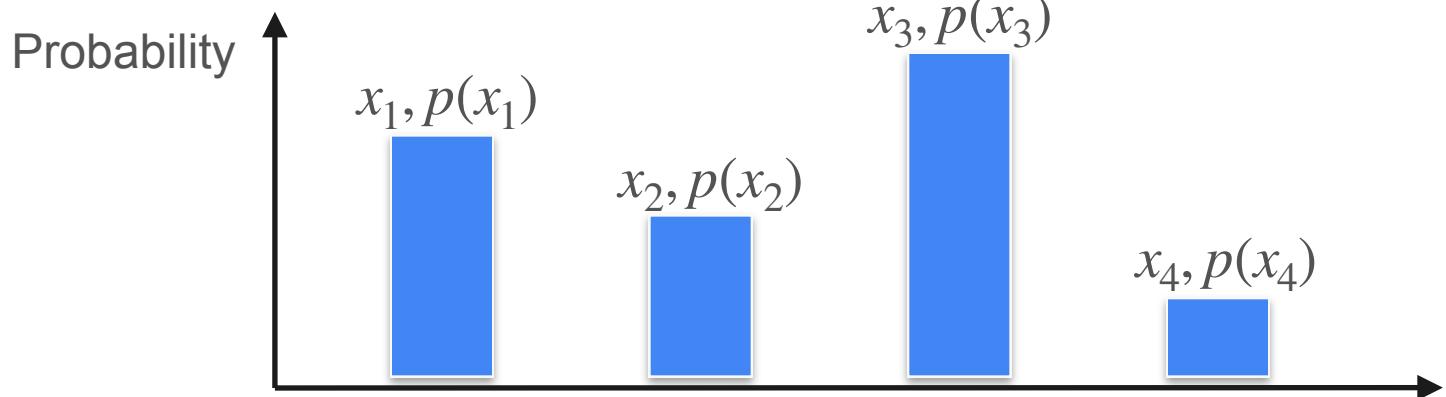
$$\mathbb{E}[X] = x_1 p(x_1) + x_2 p(x_2)$$

Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

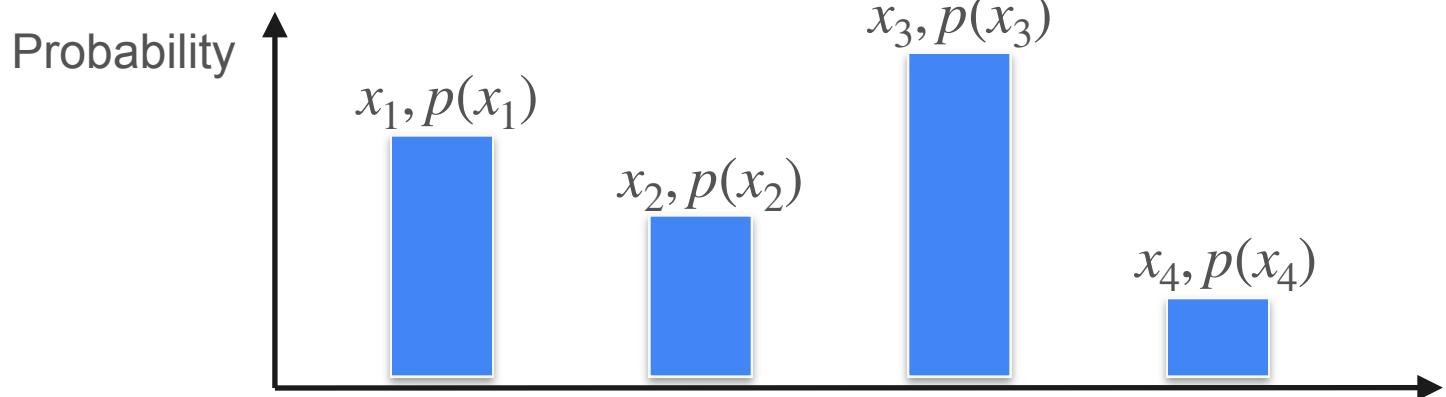
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] =$$

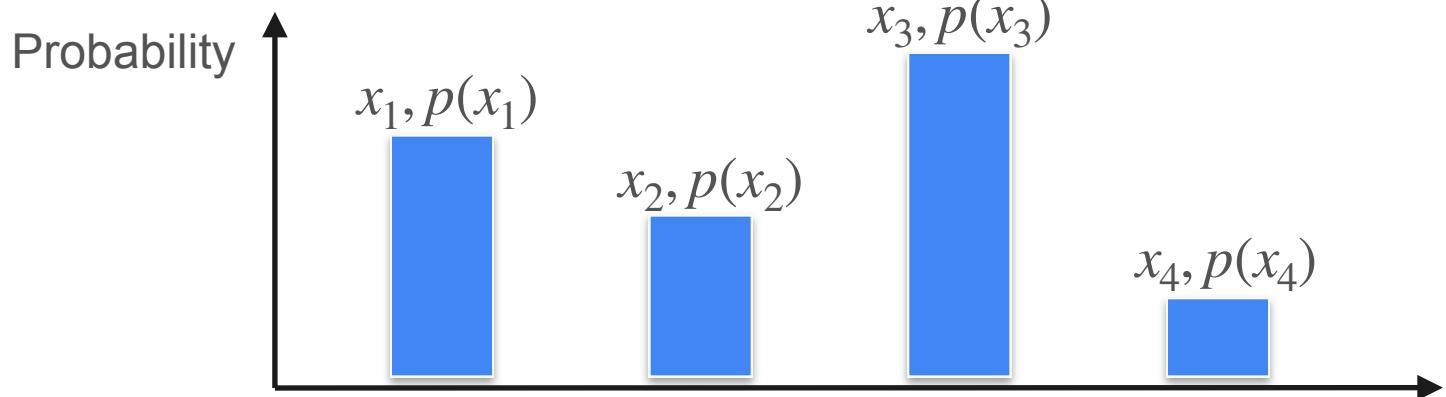
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1)$$

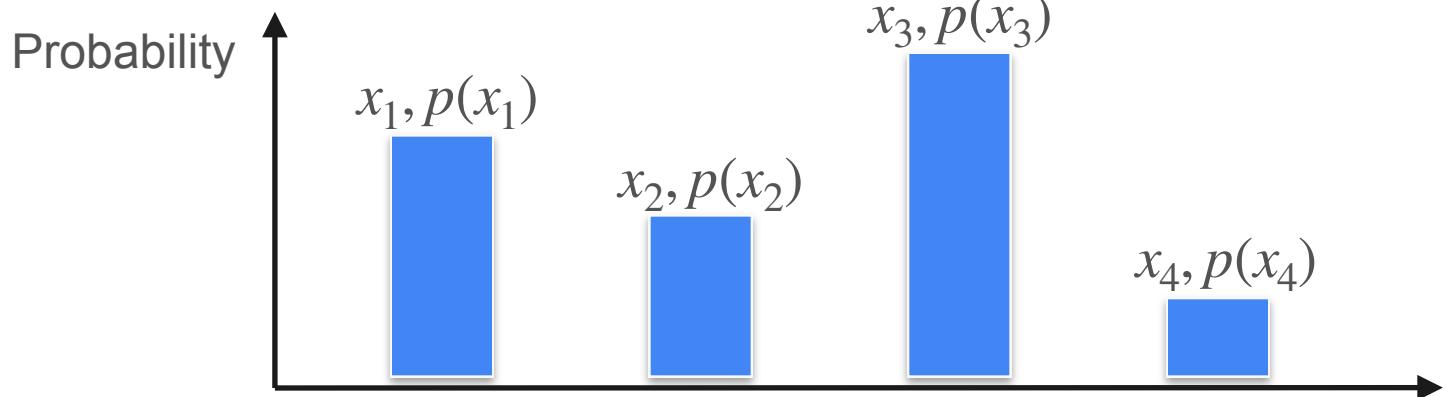
Expected Value of a Function



$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1) + f(x_2)p(x_2)$$

Expected Value of a Function



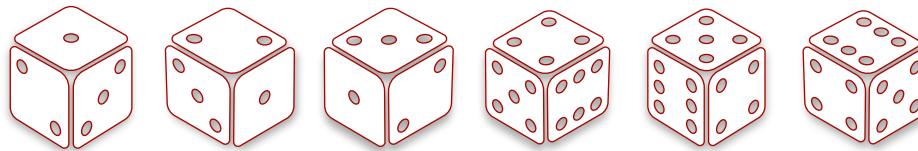
$$\mathbb{E}[X] = x_1p(x_1) + x_2p(x_2) + x_3p(x_3) + x_4p(x_4)$$

$$E[f(X)] = f(x_1)p(x_1) + f(x_2)p(x_2) + f(x_3)p(x_3) + f(x_4)p(x_4)$$

Expected Value of a Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

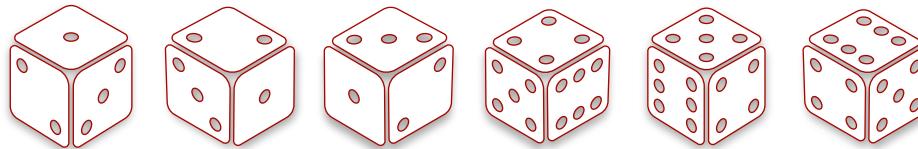
Roll: 1 2 3 4 5 6



Expected Value of a Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

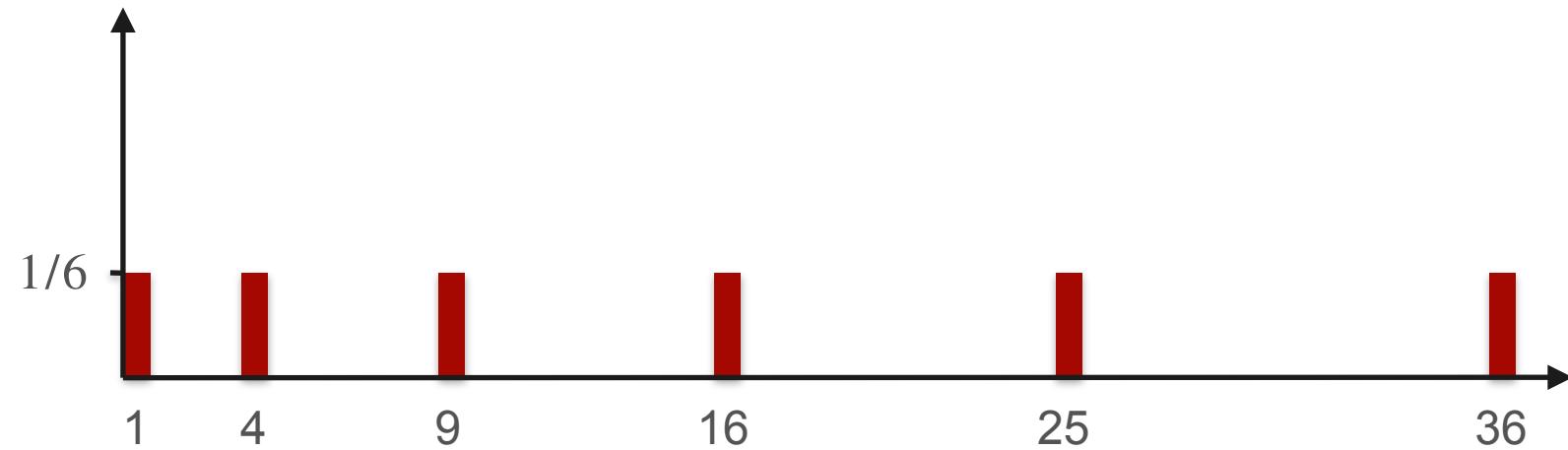
Roll: 1 2 3 4 5 6



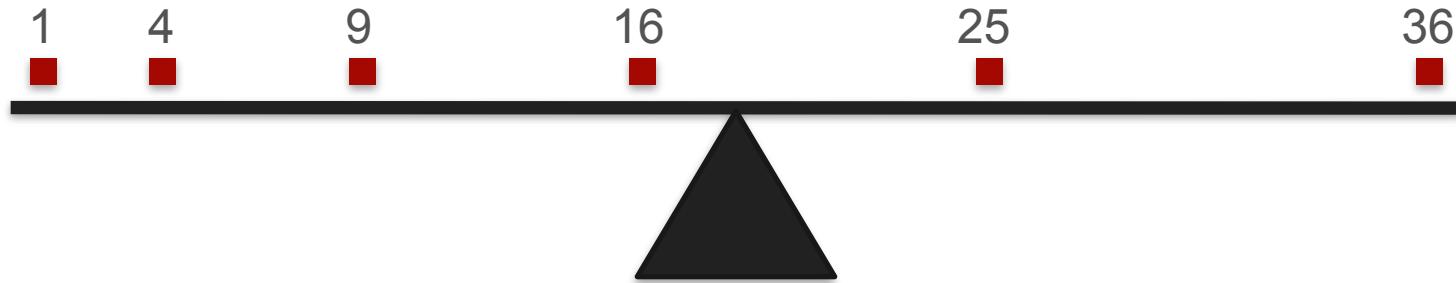
Square: 1 4 9 16 25 36

Expected Value of a Function

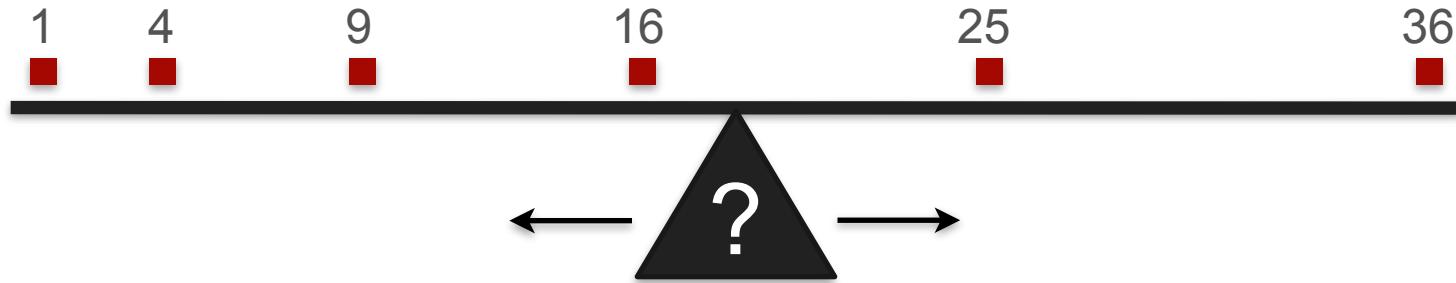
Probability



Expected Value of a Function

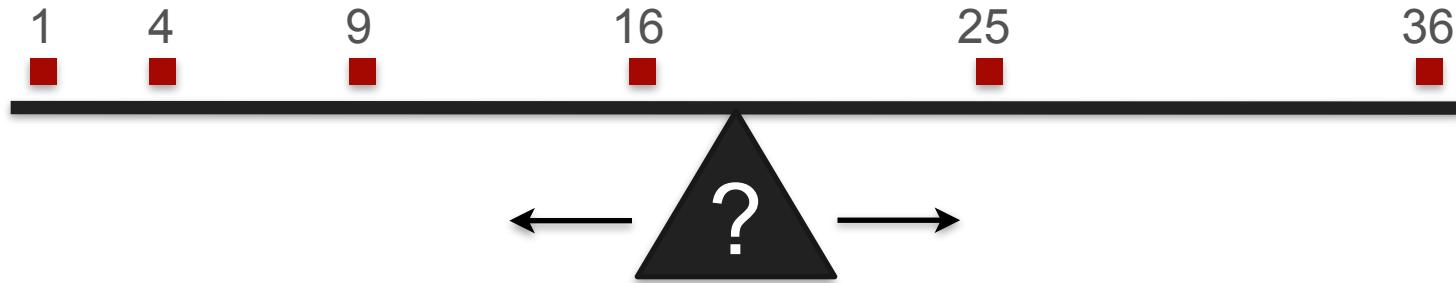


Expected Value of a Function



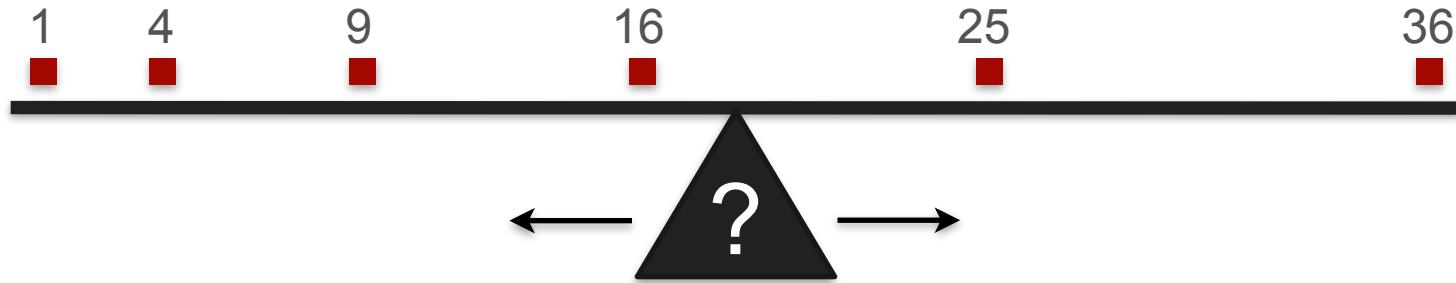
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6}$$



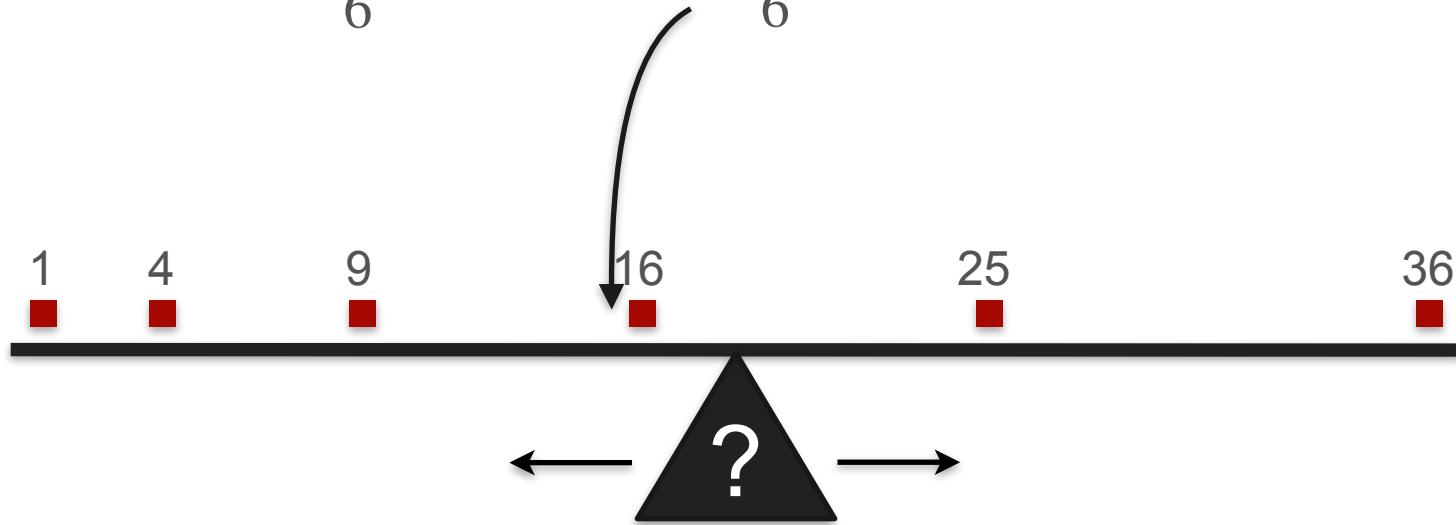
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



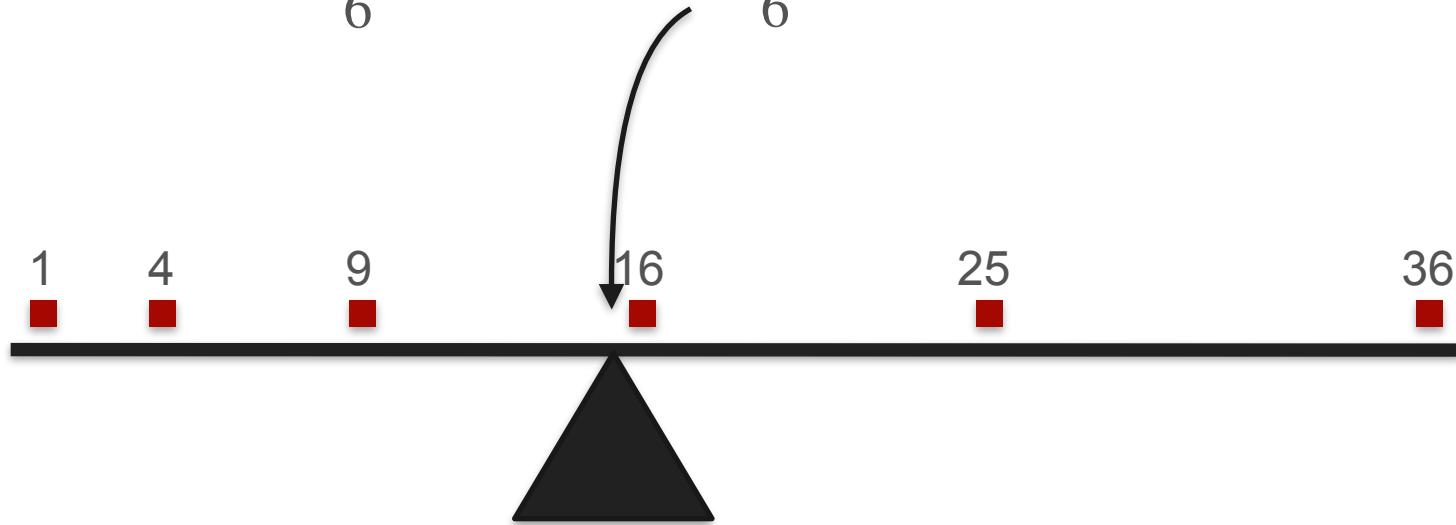
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



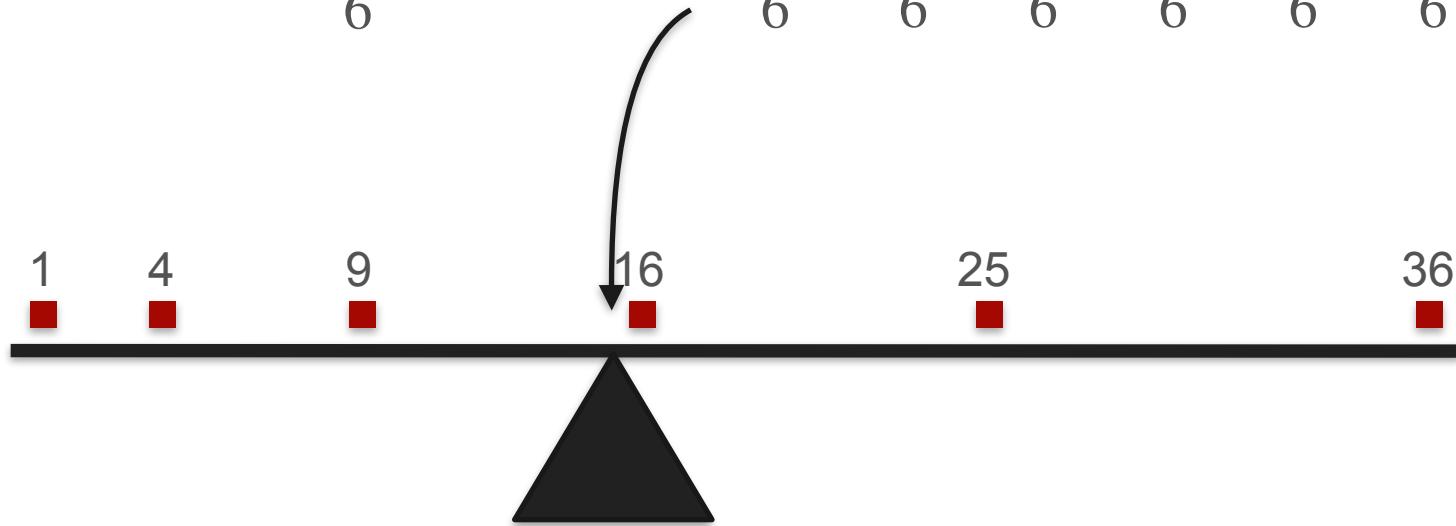
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$



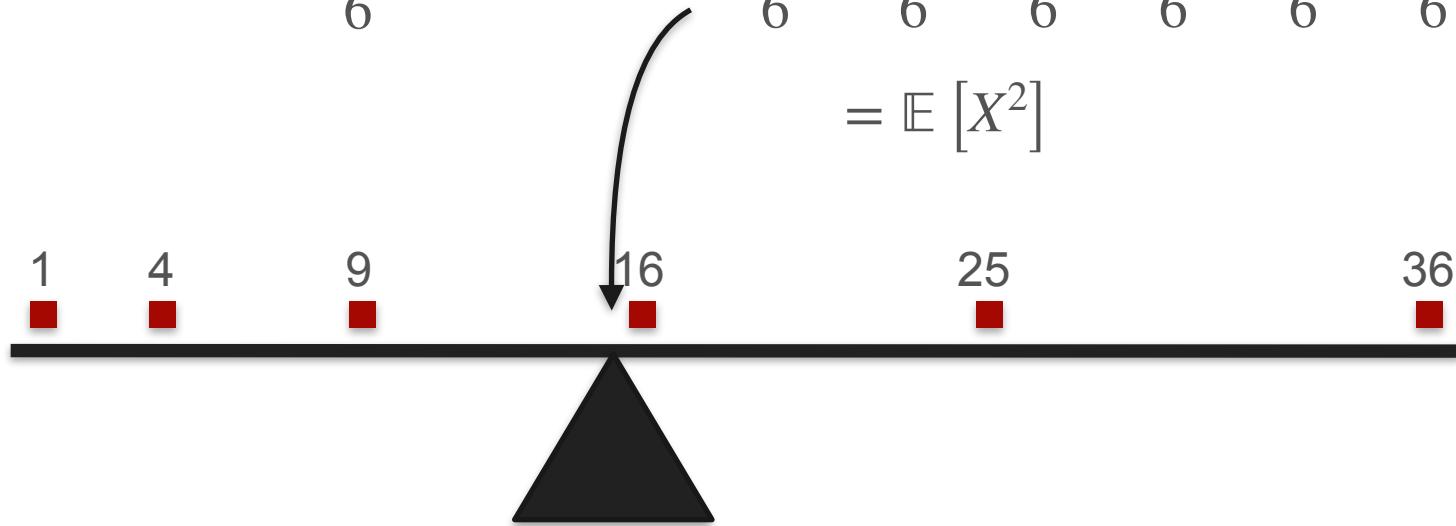
Expected Value of a Function

$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6} = \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6}$$



Expected Value of a Function

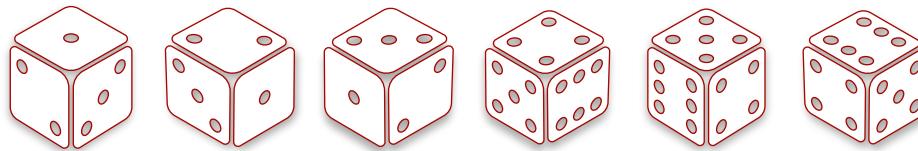
$$\frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6} = \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6}$$
$$= \mathbb{E}[X^2]$$



Expectation of Linear Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6

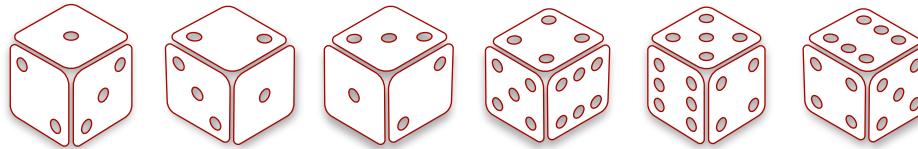


Wins

Expectation of Linear Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6



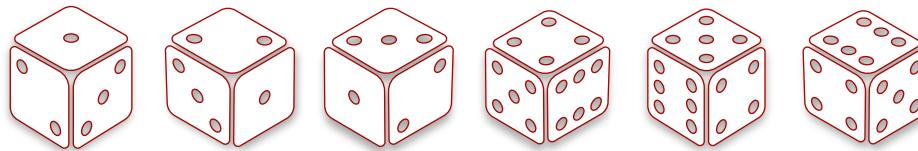
Double: 2 4 6 8 10 12

Wins 2 - 5 4 - 5 6 - 5 8 - 5 10 - 5 12 - 5

Expectation of Linear Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6

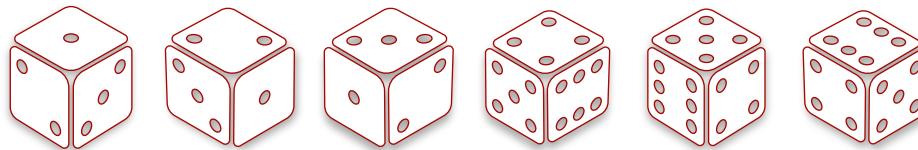


Wins

Expectation of Linear Function

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6

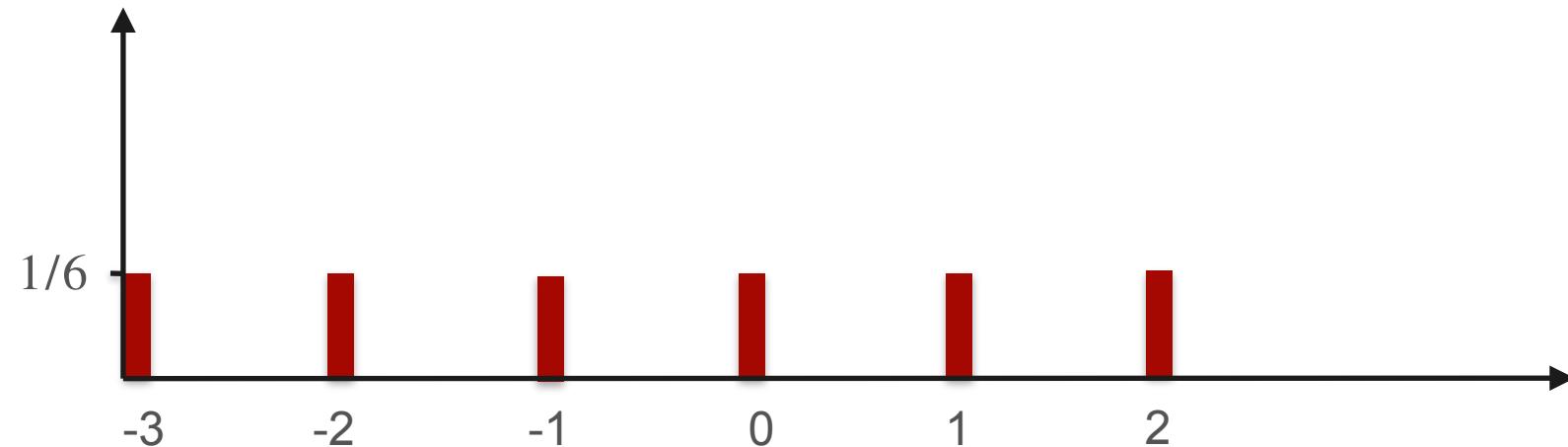


Double: 2 4 6 8 10 12

Wins -3 -2 -1 0 1 2

Expected Value of a Function

Probability



Expected Value of a Function



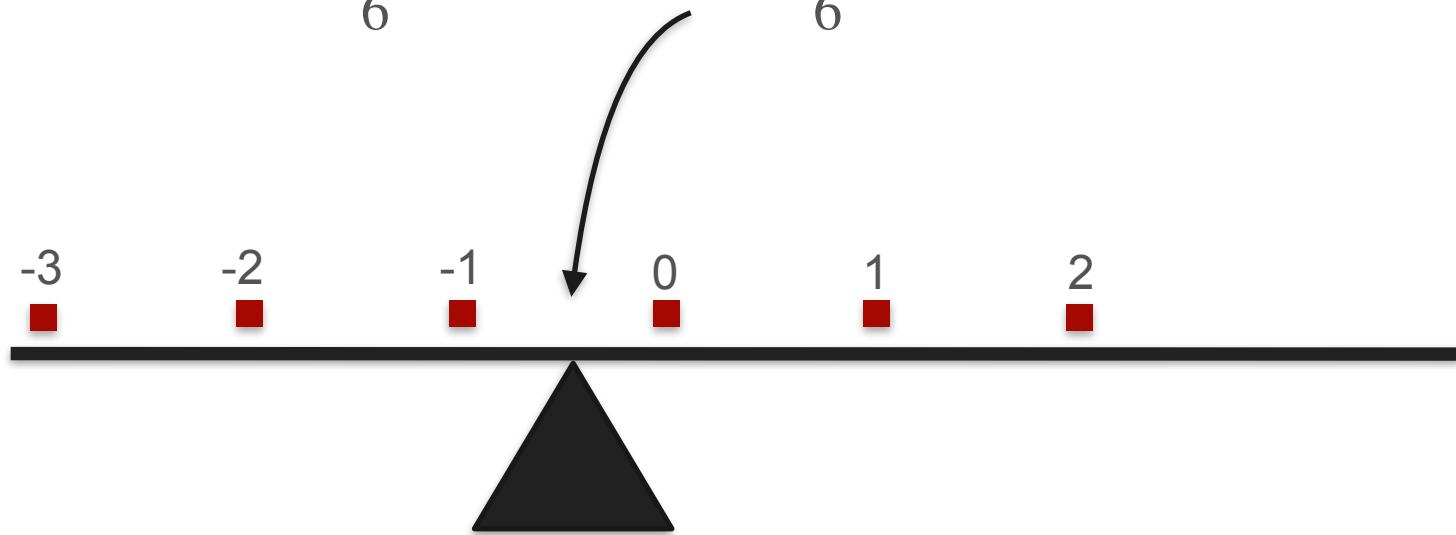
Expected Value of a Function

$$\frac{-3 + -2 + -1 + 0 + 1 + 2}{6}$$



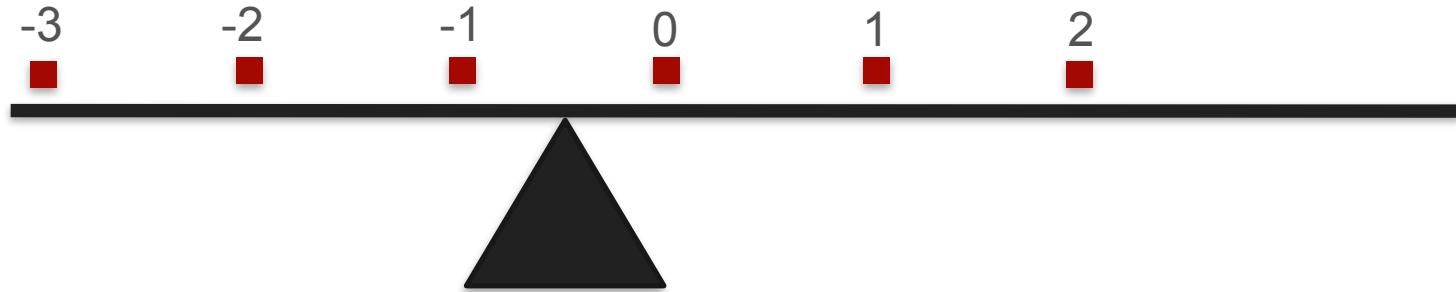
Expected Value of a Function

$$\frac{-3 + -2 + -1 + 0 + 1 + 2}{6} = \frac{-3}{6} = -0.5$$



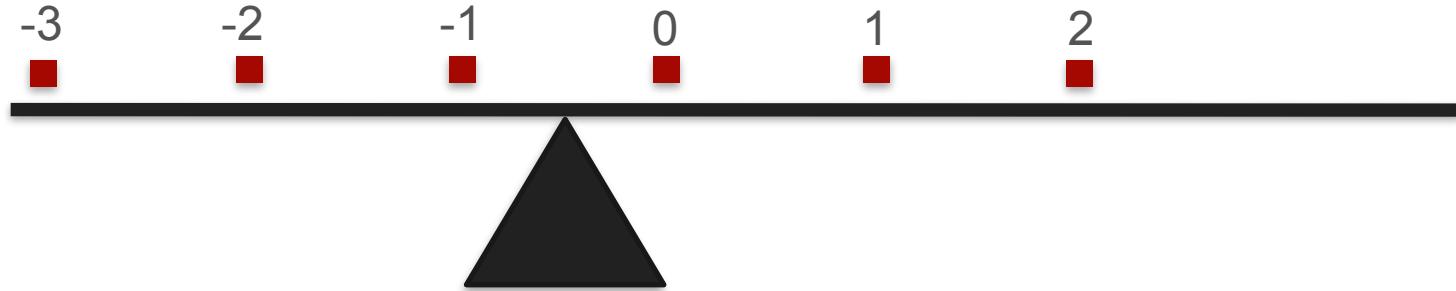
Expected Value of a Function

$$\frac{(2 \cdot 1 - 5) + (2 \cdot 2 - 5) + (2 \cdot 3 - 5) + (2 \cdot 4 - 5) + (2 \cdot 5 - 5) + (2 \cdot 6 - 5)}{6} = \frac{-3}{6} = -0.5$$



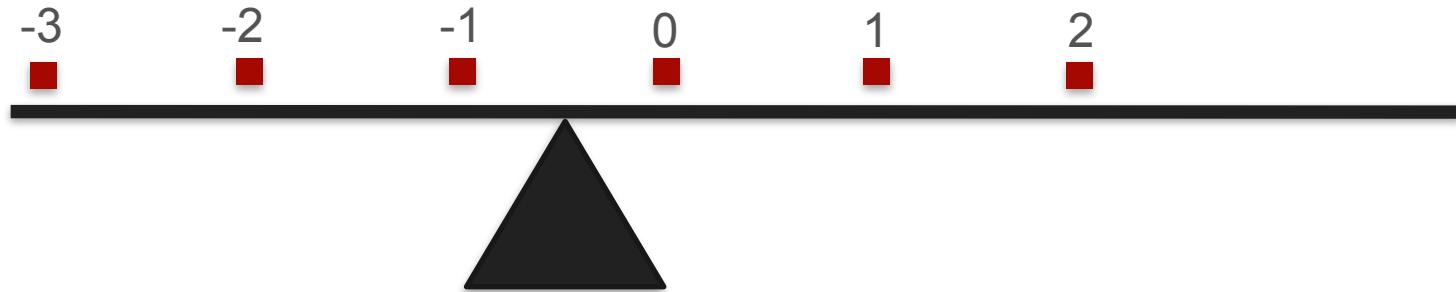
Expected Value of a Function

$$\frac{(2 \cdot 1) + (2 \cdot 2) + (2 \cdot 3) + (2 \cdot 4) + (2 \cdot 5) + (2 \cdot 6) + 6 \cdot (-5)}{6} = \frac{-3}{6} = -0.5$$



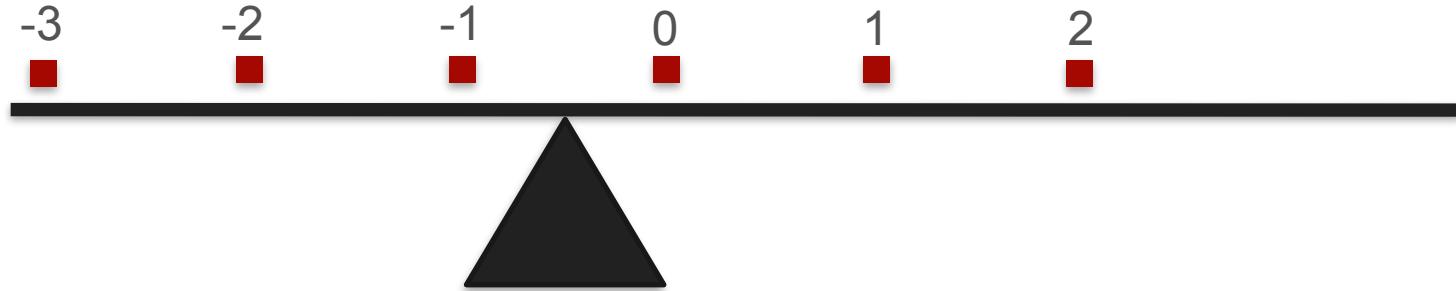
Expected Value of a Function

$$\frac{(2 \cdot 1) + (2 \cdot 2) + (2 \cdot 3) + (2 \cdot 4) + (2 \cdot 5) + (2 \cdot 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$

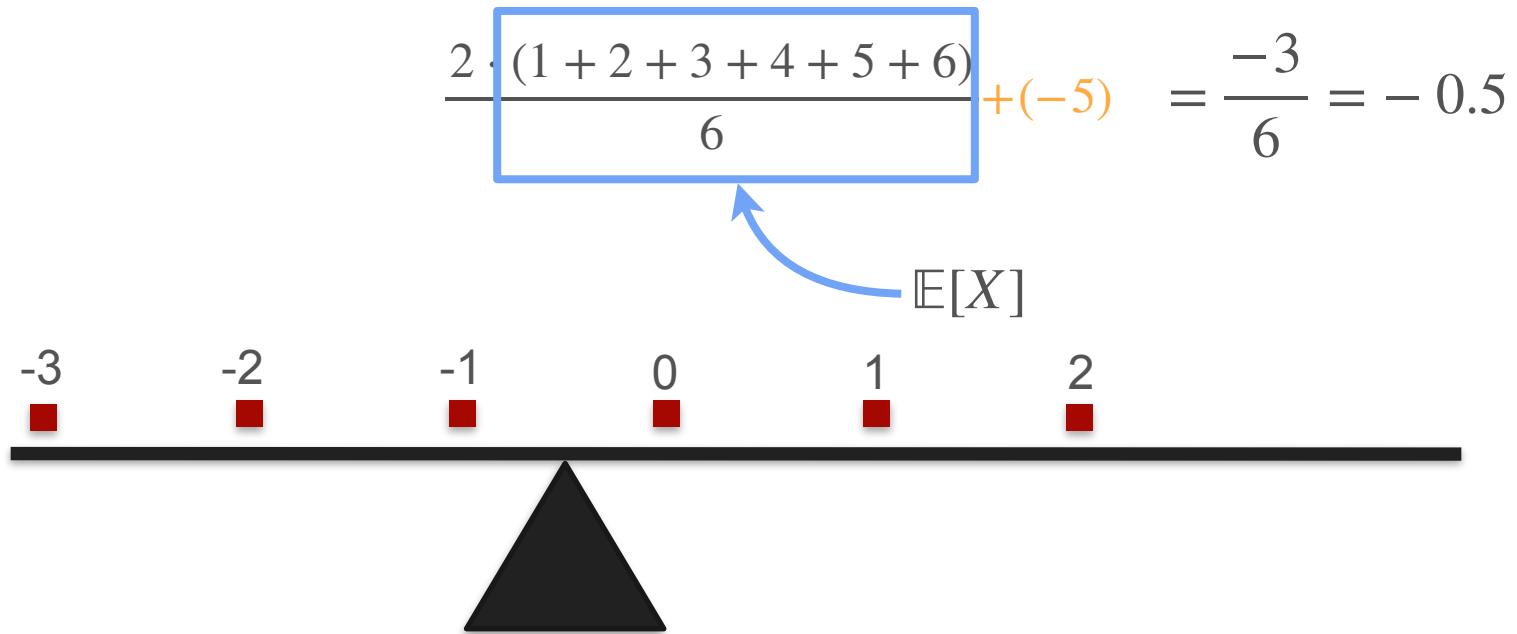


Expected Value of a Function

$$\frac{2 \cdot (1 + 2 + 3 + 4 + 5 + 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$



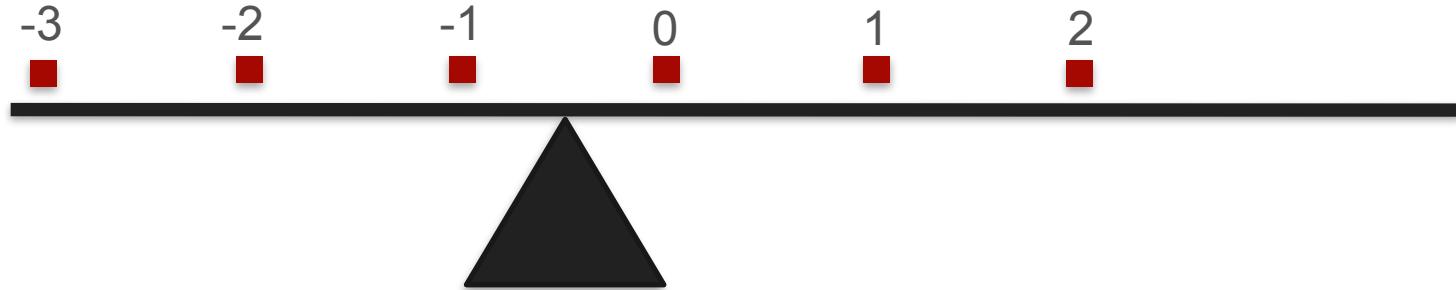
Expected Value of a Function



Expected Value of a Function

$$\mathbb{E}[2 \cdot X + (-5)] = \frac{2 \cdot (1 + 2 + 3 + 4 + 5 + 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$

$= 2 \cdot \mathbb{E}[X] + (-5)$



Expected Value of a Function

$$\mathbb{E}[2 \cdot X + (-5)] = \frac{2 \cdot (1 + 2 + 3 + 4 + 5 + 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$

$= 2 \cdot \mathbb{E}[X] + (-5)$



In general:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Expected Value of a Function

$$\mathbb{E}[2 \cdot X + (-5)] = \frac{2 \cdot (1 + 2 + 3 + 4 + 5 + 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$

$= 2 \cdot \mathbb{E}[X] + (-5)$



In general:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

Expected Value of a Function

$$\mathbb{E}[2 \cdot X + (-5)] = \frac{2 \cdot (1 + 2 + 3 + 4 + 5 + 6)}{6} + (-5) = \frac{-3}{6} = -0.5$$

$= 2 \cdot \mathbb{E}[X] + (-5)$



In general:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

$$\mathbb{E}[b] = b$$



DeepLearning.AI

Describing Distributions

Sum of expectations

Sum of Expectation: Expected Winnings

You play a game:

Sum of Expectation: Expected Winnings

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Sum of Expectation: Expected Winnings

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.



Win \$1



Win nothing

Sum of Expectation: Expected Winnings

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then throw a dice. You win the amount you roll.



Win \$1



Win nothing

Sum of Expectation: Expected Winnings

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then throw a dice. You win the amount you roll.

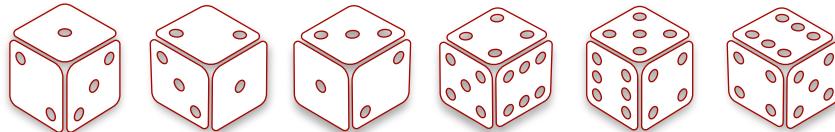


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



Sum of Expectation: Expected Winnings

You play a game:

Flip a coin. If heads you win \$ 1, else you win nothing.

Then throw a dice. You win the amount you roll.

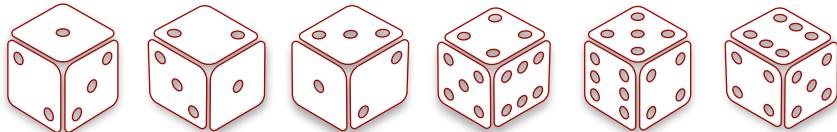


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



What are your expected winnings for the game?

Sum of Expectations: Expected Winnings

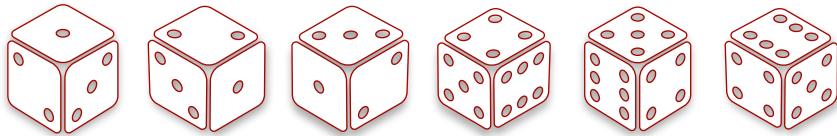


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



Sum of Expectations: Expected Winnings

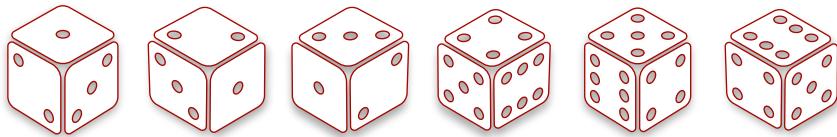


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E} [X_{coin}] = \$0.5$$

Sum of Expectations: Expected Winnings



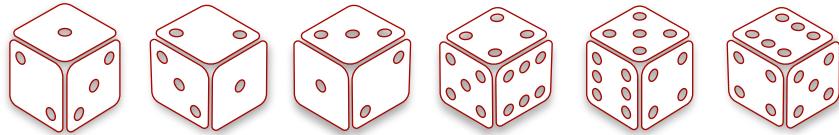
Win \$1



Win nothing

$$\mathbb{E}[X_{coin}] = \$0.5$$

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E}[X_{dice}] = \$3.5$$

Sum of Expectations: Expected Winnings

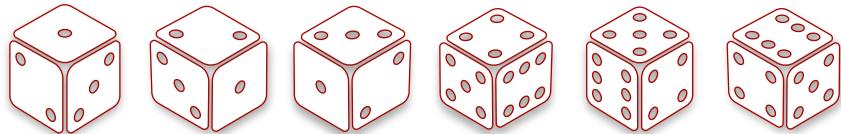


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E}[X_{coin}] = \$0.5$$

$$\mathbb{E}[X_{dice}] = \$3.5$$

$$\mathbb{E}[X] = \mathbb{E}[X_{coin}] + \mathbb{E}[X_{dice}] = \$0.5 + \$3.5 = \$4$$

Sum of Expectations: Expected Winnings

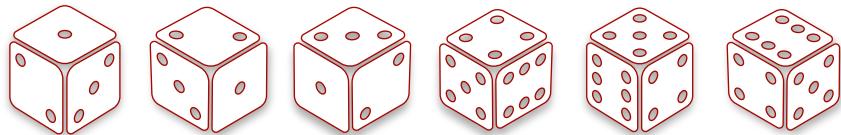


Win \$1



Win nothing

Win \$1 \$2 \$3 \$4 \$5 \$6



$$\mathbb{E}[X_{coin}] = \$0.5$$

$$\mathbb{E}[X_{dice}] = \$3.5$$

$$\mathbb{E}[X] = \mathbb{E}[X_{coin}] + \mathbb{E}[X_{dice}] = \$0.5 + \$3.5 = \$4$$

In general: $\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$

Sum of Expectations: Bag With Names



8 billion people

Sum of Expectations: Bag With Names



8 billion people

Sum of Expectations: Bag With Names



8 billion people

Sum of Expectations: Bag With Names



Expected number of
correct assignments?



8 billion people

Sum of Expectations: Bag With Names



1

Expected number of
correct assignments?



8 billion people

Sum of Expectations: Bag With Names



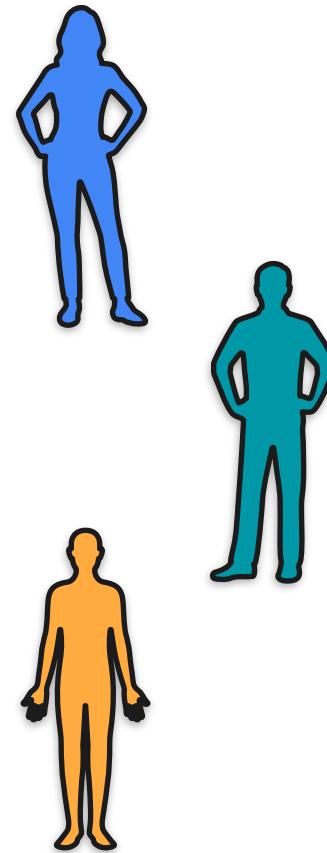
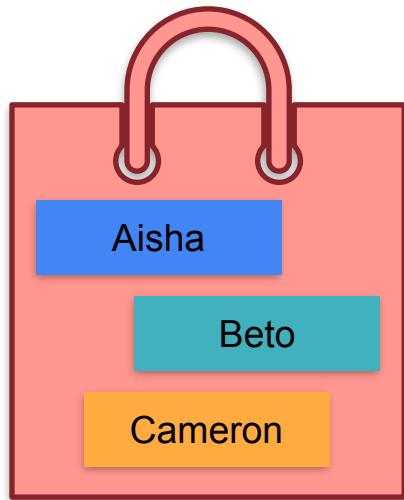
1

Expected number of
correct assignments?

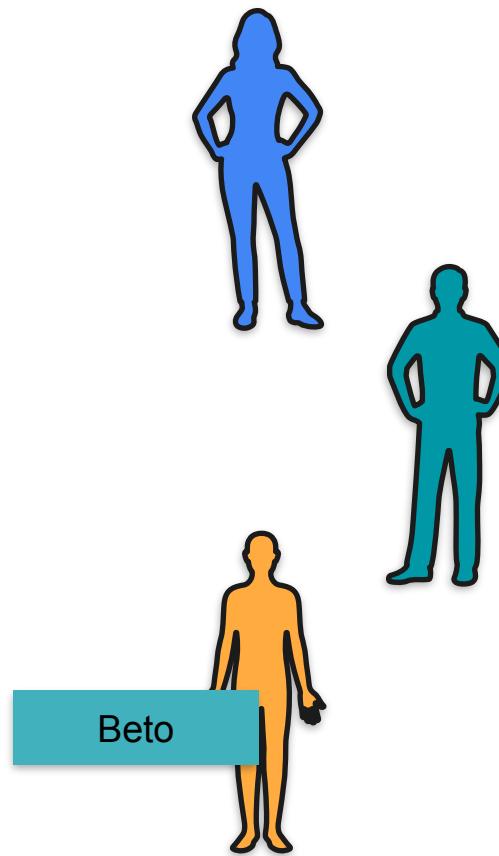
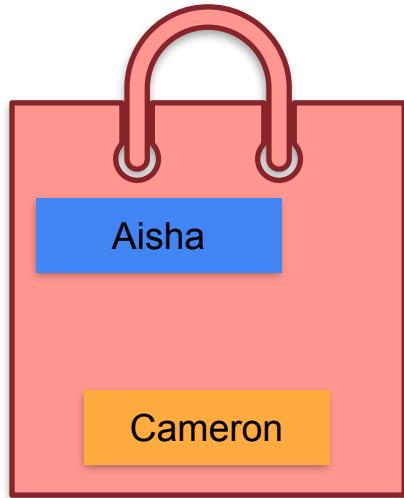


8 billion people

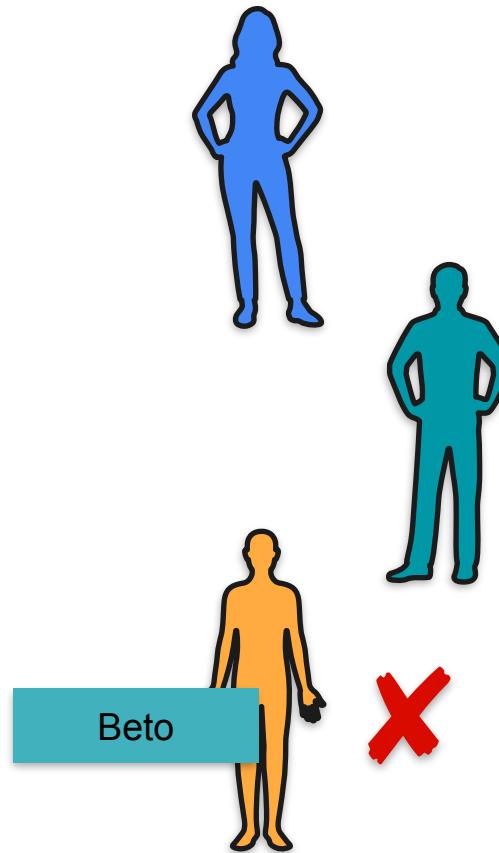
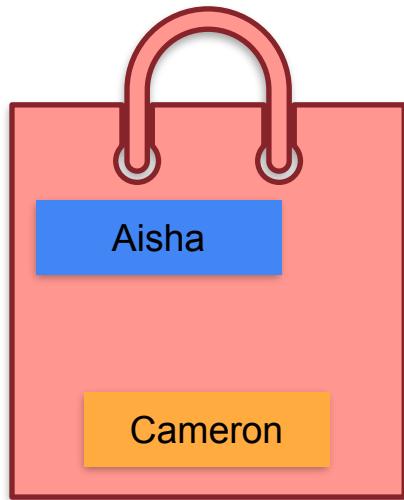
Bag With 3 Names



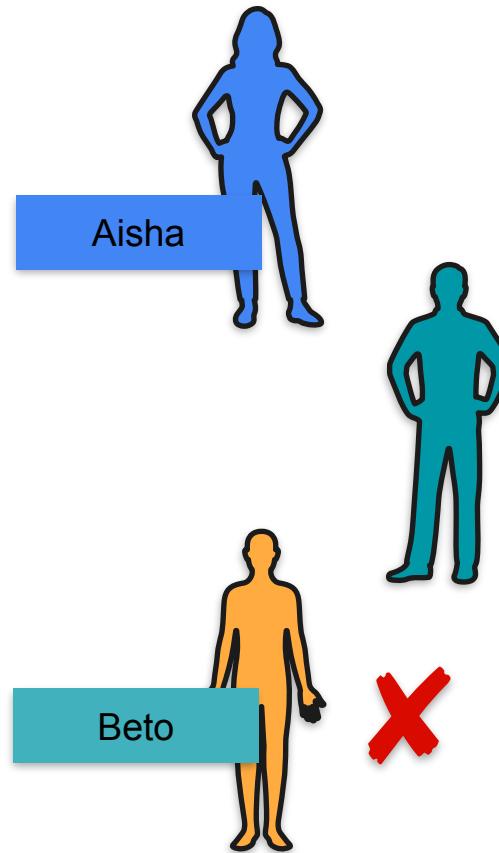
Bag With 3 Names



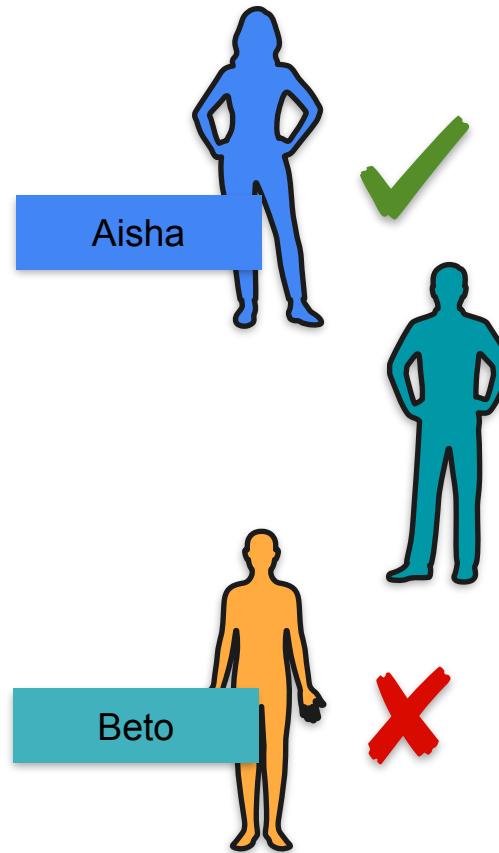
Bag With 3 Names



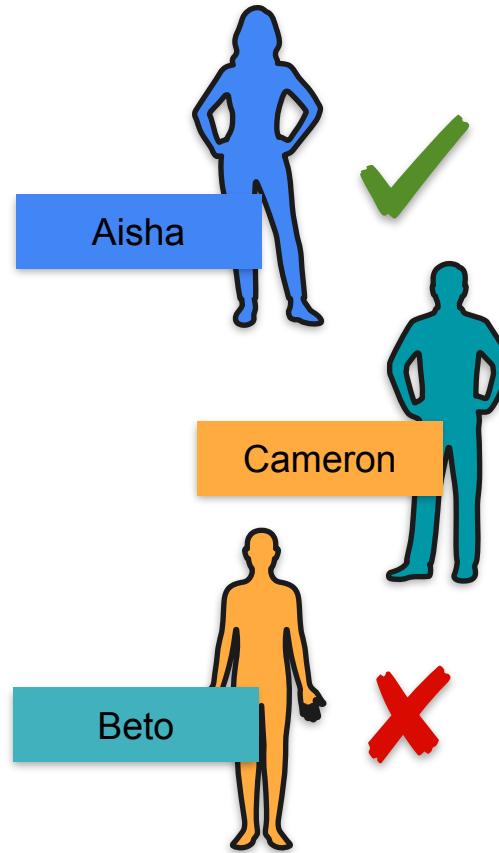
Bag With 3 Names



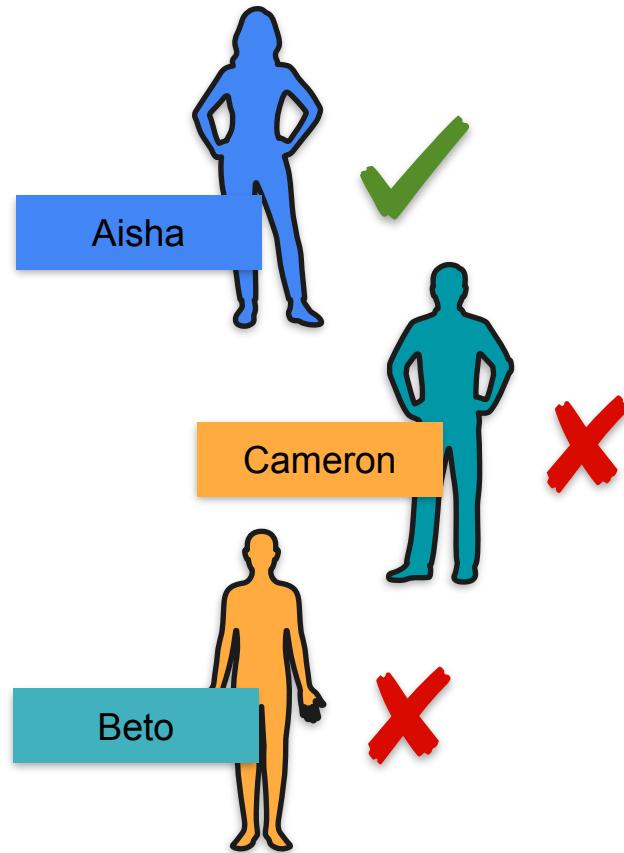
Bag With 3 Names



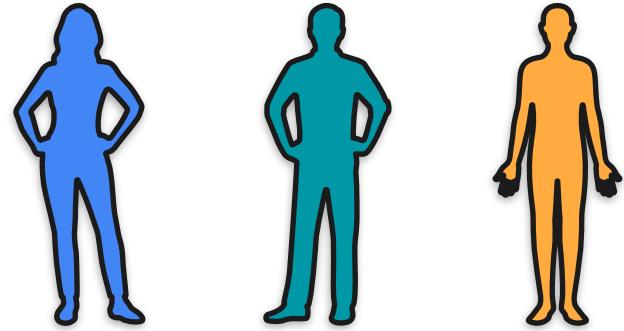
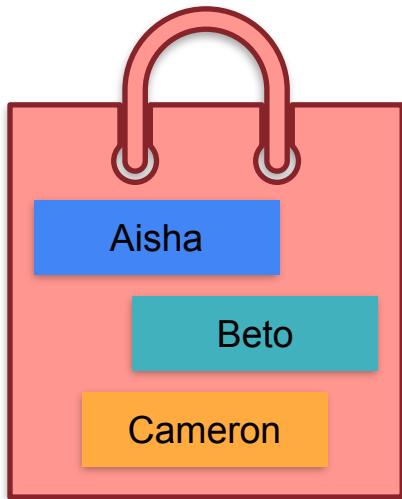
Bag With 3 Names



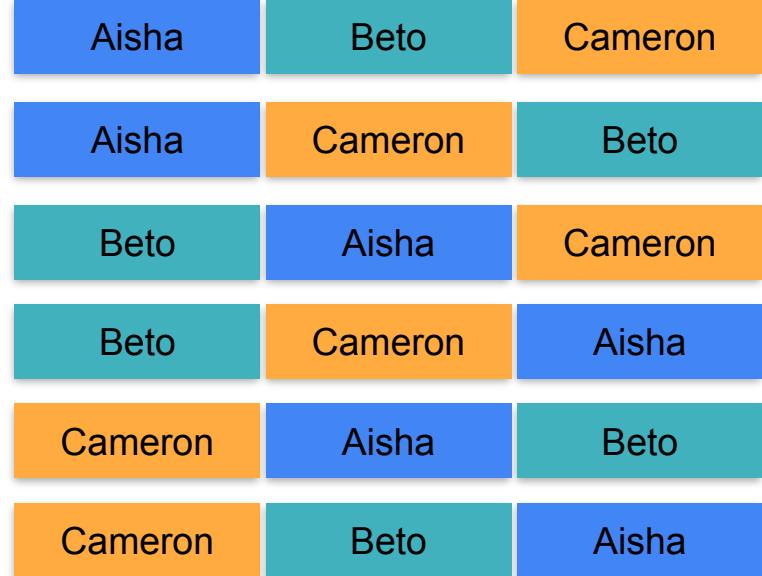
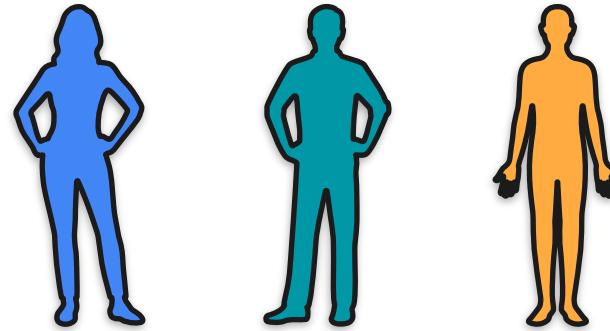
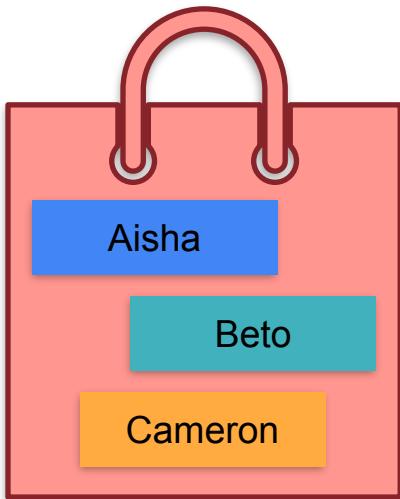
Bag With 3 Names



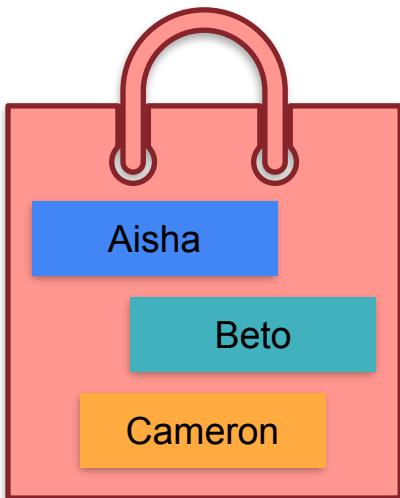
Bag With 3 Names



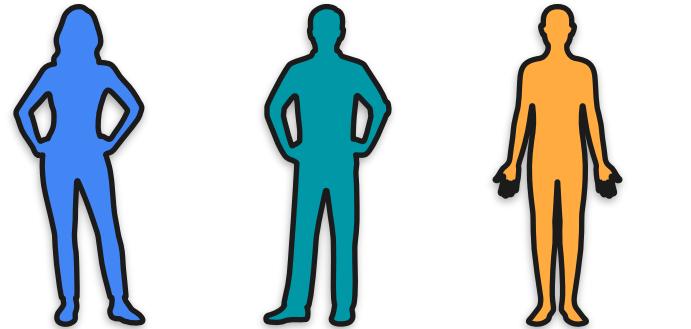
Bag With 3 Names



Bag With 3 Names

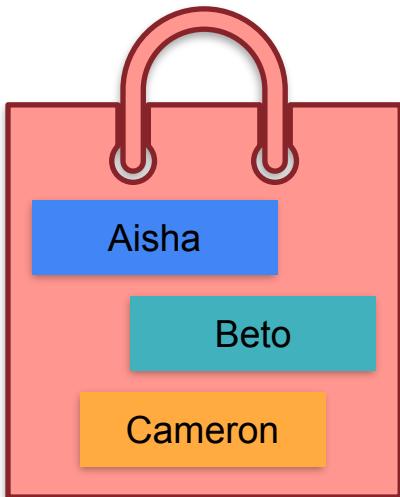


Correct

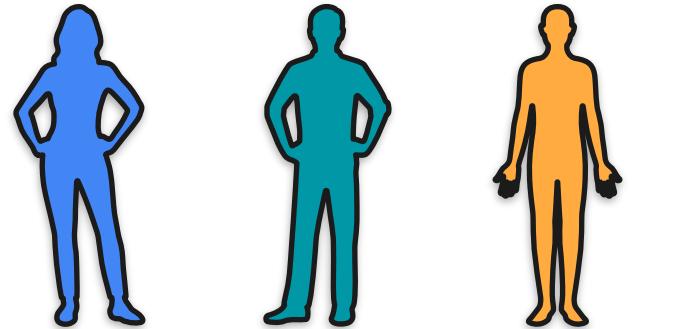


Aisha	Beto	Cameron
Aisha	Cameron	Beto
Beto	Aisha	Cameron
Beto	Cameron	Aisha
Cameron	Aisha	Beto
Cameron	Beto	Aisha

Bag With 3 Names

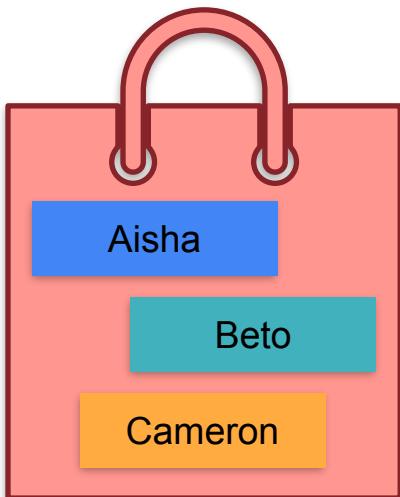


Correct



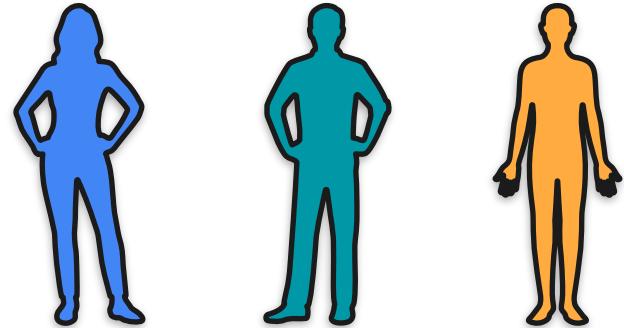
Aisha	Beto	Cameron
Aisha	Cameron	Beto
Beto	Aisha	Cameron
Beto	Cameron	Aisha
Cameron	Aisha	Beto
Cameron	Beto	Aisha

Bag With 3 Names

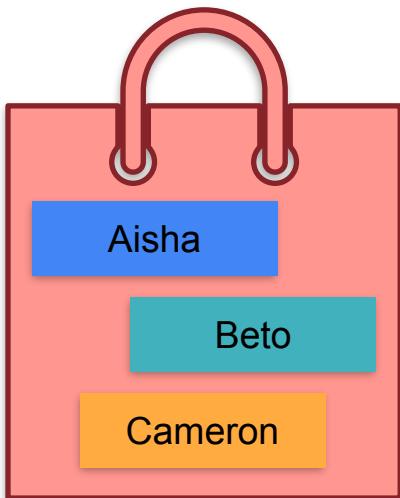


Correct

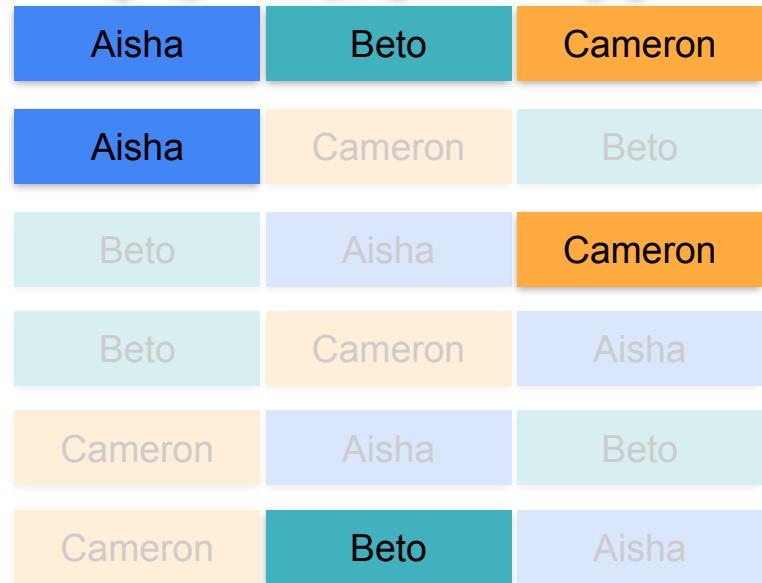
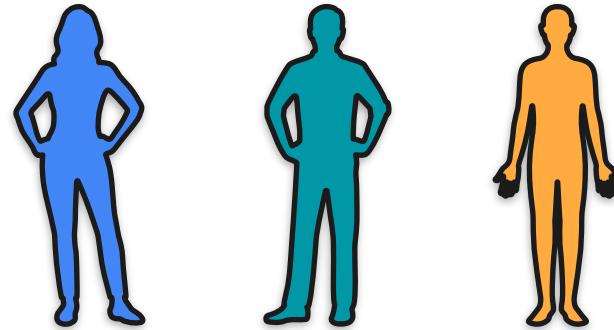
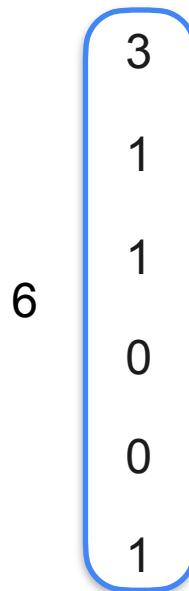
3	Aisha	Beto	Cameron
1	Aisha	Cameron	Beto
1	Beto	Aisha	Cameron
0	Beto	Cameron	Aisha
0	Cameron	Aisha	Beto
1	Cameron	Beto	Aisha



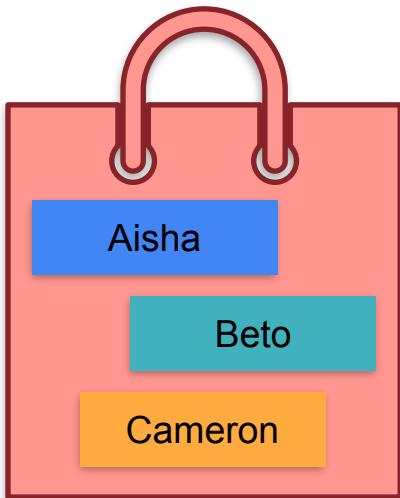
Bag With 3 Names



Correct



Bag With 3 Names

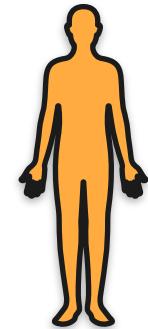
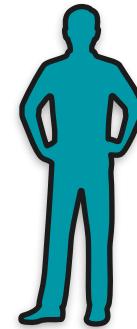
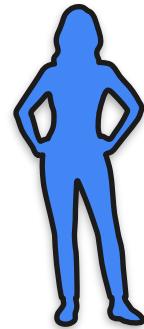
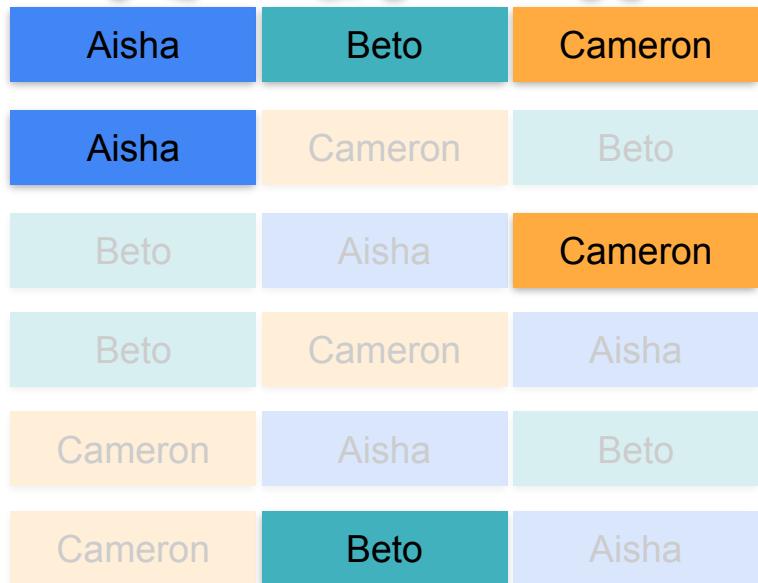


Average
1

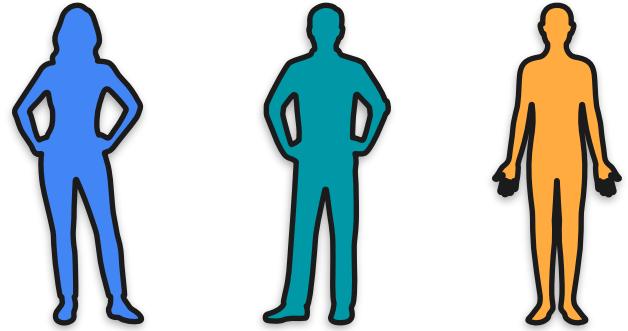
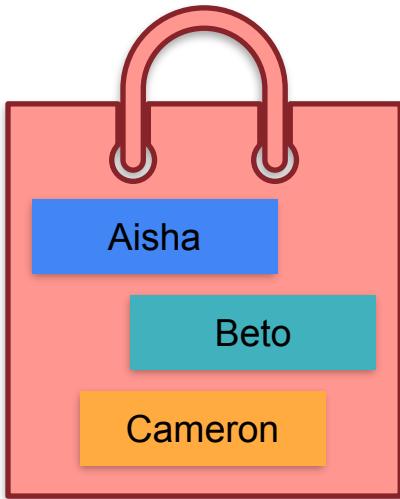
Correct

3
1
1
0
0
1

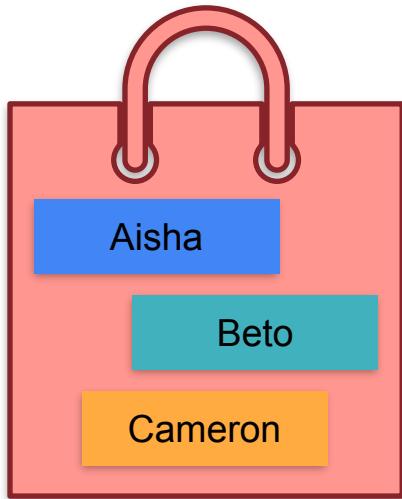
6



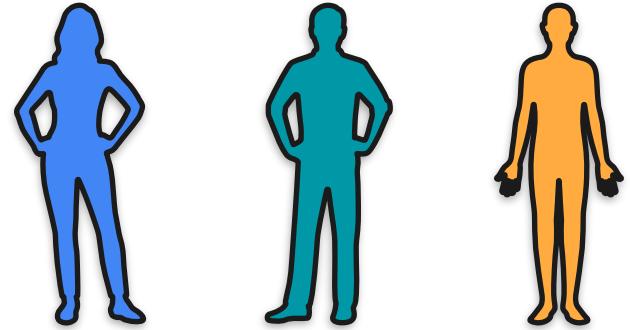
Adding Expectations



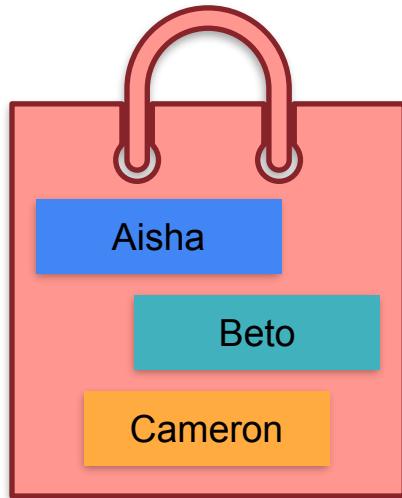
Adding Expectations



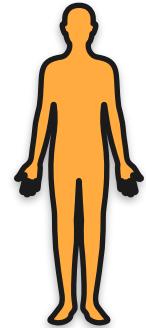
$\mathbb{E}[\text{Matches}]$



Adding Expectations



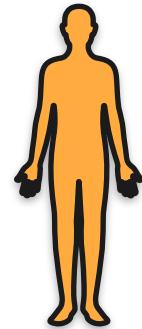
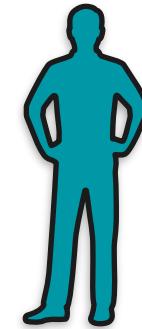
$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A]$$



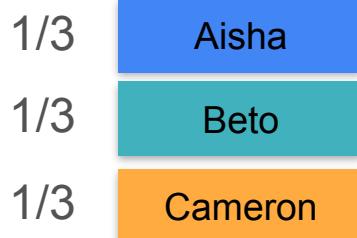
Adding Expectations



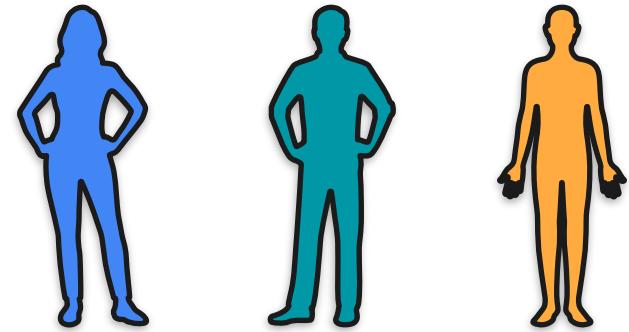
$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A]$$



Adding Expectations



$\mathbb{E}[\text{Matches}]$



$= \mathbb{E}[A]$

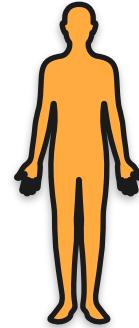
Adding Expectations



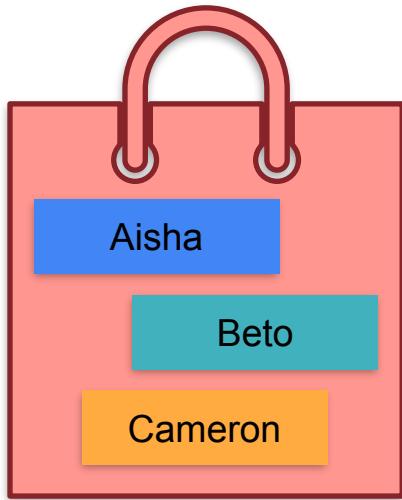
$\mathbb{E}[\text{Matches}]$

$= \mathbb{E}[A]$

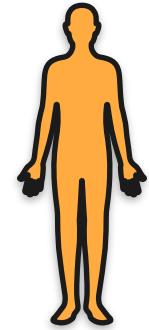
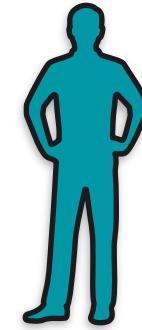
$= 1/3$



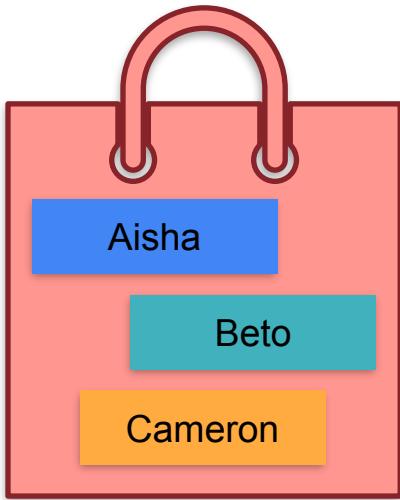
Adding Expectations



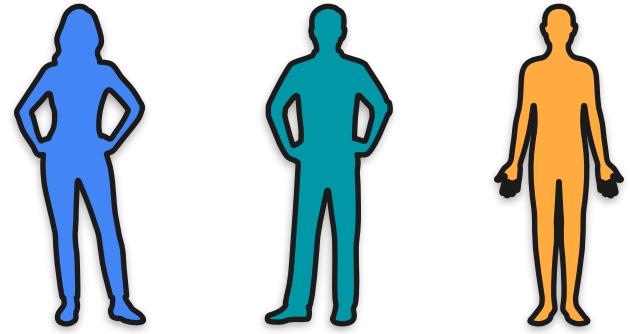
$$\mathbb{E}[\text{Matches}] = \mathbb{E}[A] = 1/3$$



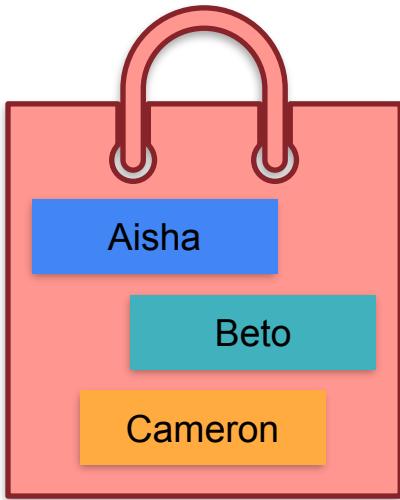
Adding Expectations



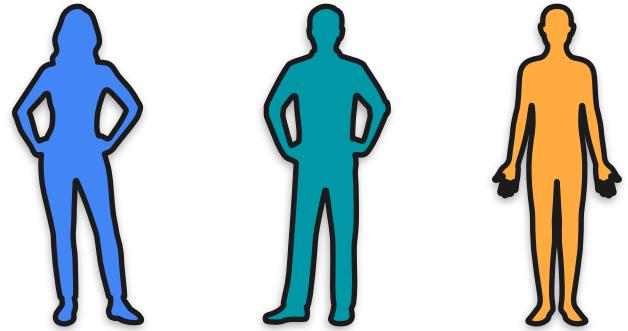
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] \\ &= 1/3 + 1/3\end{aligned}$$



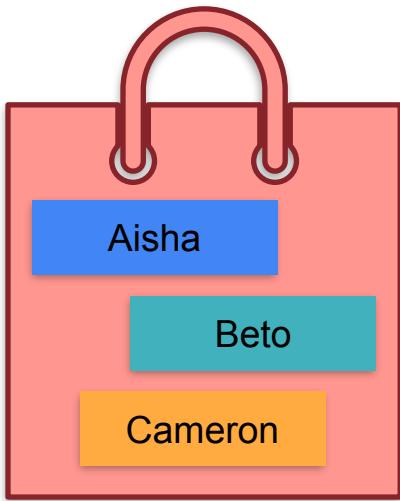
Adding Expectations



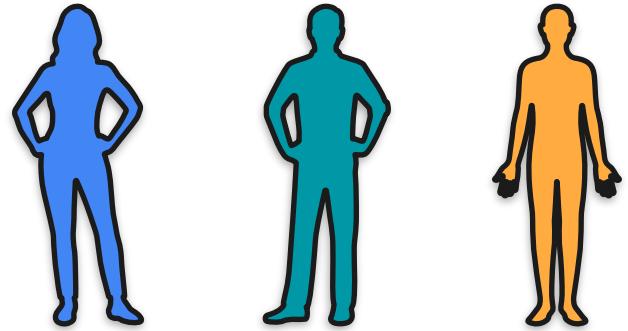
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3\end{aligned}$$



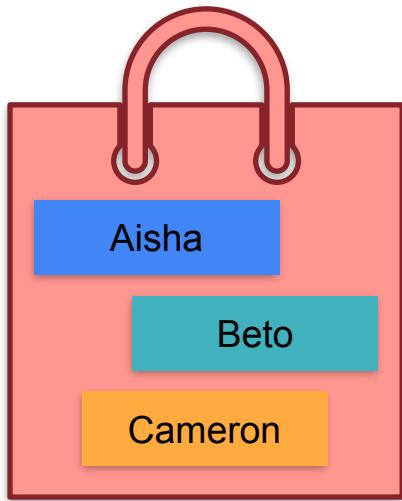
Adding Expectations



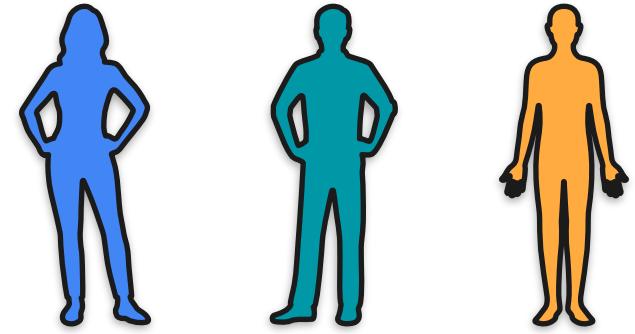
$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3 \\ &= 1\end{aligned}$$



Adding Expectations



$$\begin{aligned}\mathbb{E}[\text{Matches}] &= \mathbb{E}[A] + \mathbb{E}[B] + \mathbb{E}[C] \\ &= 1/3 + 1/3 + 1/3 \\ &= 1\end{aligned}$$



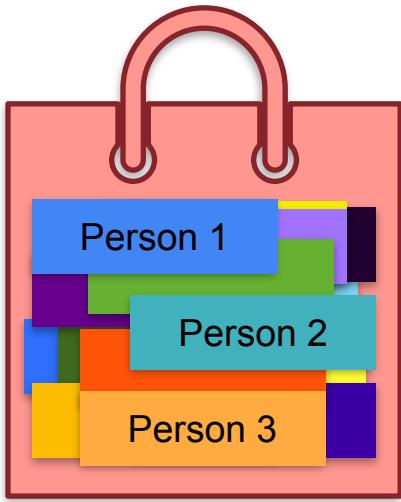
Average
1

Adding Expectations



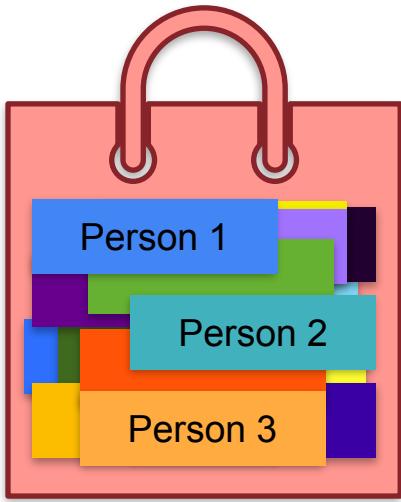
8 billion people

Adding Expectations



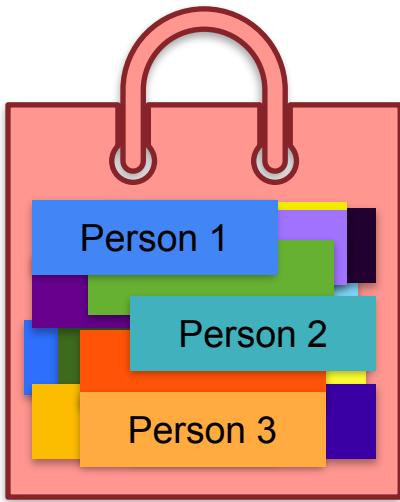
8 billion people

Adding Expectations



8 billion people

Adding Expectations



Expected number = ?



Adding Expectations



Adding Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}]$$

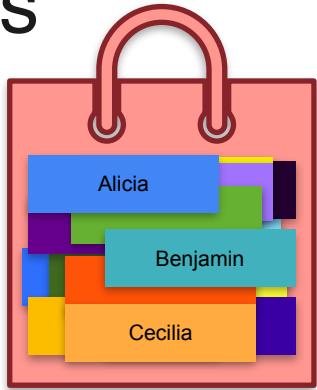
Adding Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}]$$

n people ($n = 8$ billion)

Adding Expectations



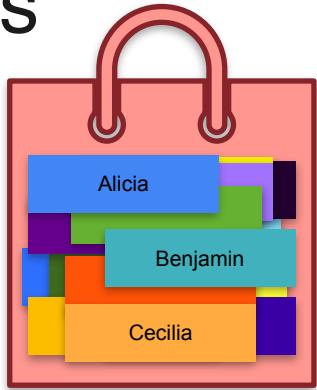
$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \quad \overbrace{\qquad \qquad \qquad}^{\text{n people (n = 8 billion)}} = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}$$

Adding Expectations



$$\begin{aligned}\mathbb{E} [\text{Matches}] &= \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \\ &\quad \overbrace{\hspace{10em}}^{\text{n people (n = 8 billion)}} \\ &= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} \\ &= n \cdot \frac{1}{n}\end{aligned}$$

Adding Expectations



$$\begin{aligned}\mathbb{E} [\text{Matches}] &= \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] \\ &\quad \overbrace{\hspace{10em}}^{\text{n people (n = 8 billion)}} \\ &= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} \\ &= n \cdot \frac{1}{n} = 1\end{aligned}$$

Adding Expectations



$$\mathbb{E} [\text{Matches}] = \mathbb{E} [X_{\text{person}_1}] + \mathbb{E} [X_{\text{person}_2}] + \dots + \mathbb{E} [X_{\text{person}_n}] = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}$$

$n \text{ people } (n = 8 \text{ billion})$

In general:

$$\mathbb{E} [X_1 + X_2 + \dots + X_n] = \mathbb{E} [X_1] + \mathbb{E} [X_2] + \dots + \mathbb{E} [X_n]$$

$$= n \cdot \frac{1}{n} = 1$$



DeepLearning.AI

Describing Distributions

Variance

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend

Game cost:



You win 1 dollar



You lose 1 dollar

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 1 dollar



You lose 1 dollar

Game cost:

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend

Game cost:



You win 100 dollars



You lose 100 dollars

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

You play a game with a friend



You win 100 dollars



You lose 100 dollars

Game cost:

\$0

What is the fair amount of money to pay to play this game?

Variance Motivation: Fair Price To Play the Game

How do you tell these two games apart?

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 100 dollars



You lose 100 dollars

Variance Motivation: Fair Price To Play the Game

How do you tell these two games apart?

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



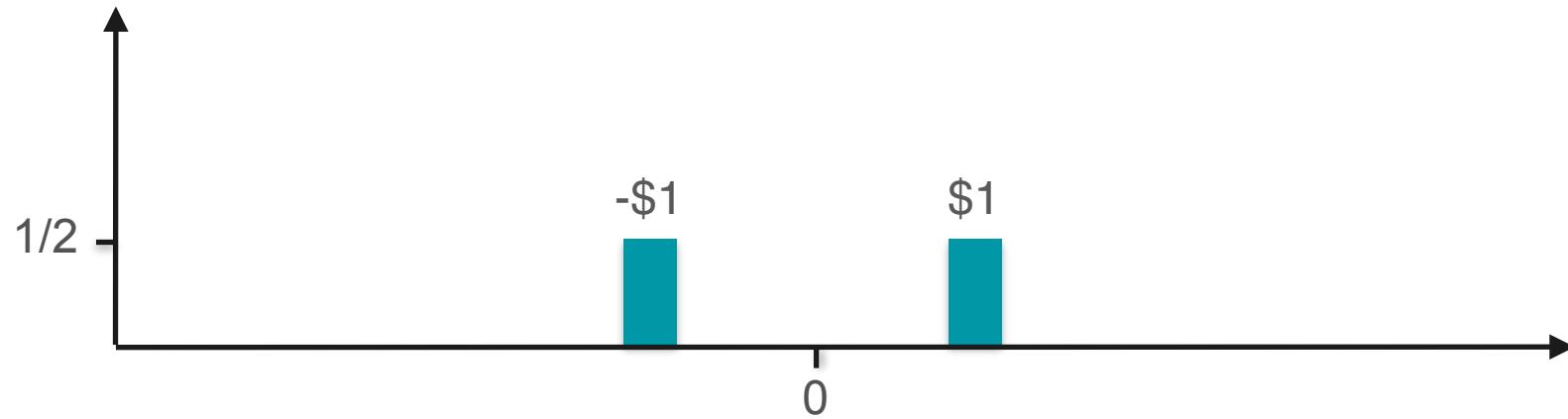
You win 100 dollars

You lose 100 dollars

Variance!

Variance Motivation: Measuring Spread

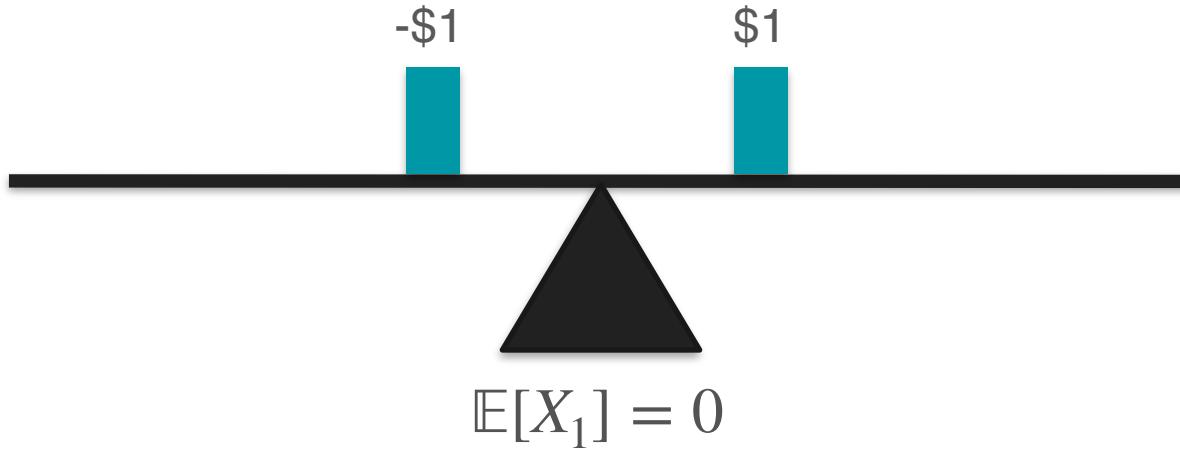
Probability



Variance Motivation: Measuring Spread

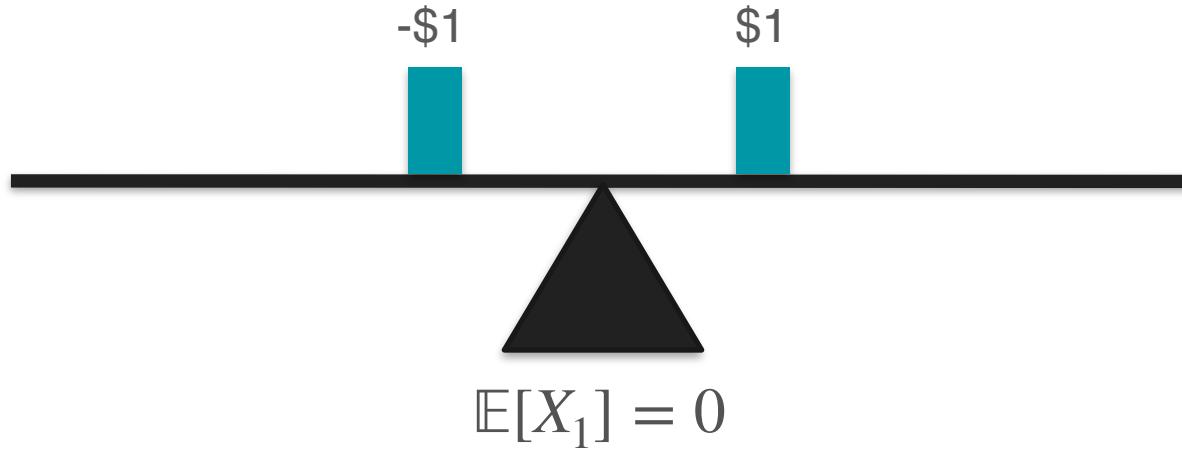


Variance Motivation: Measuring Spread

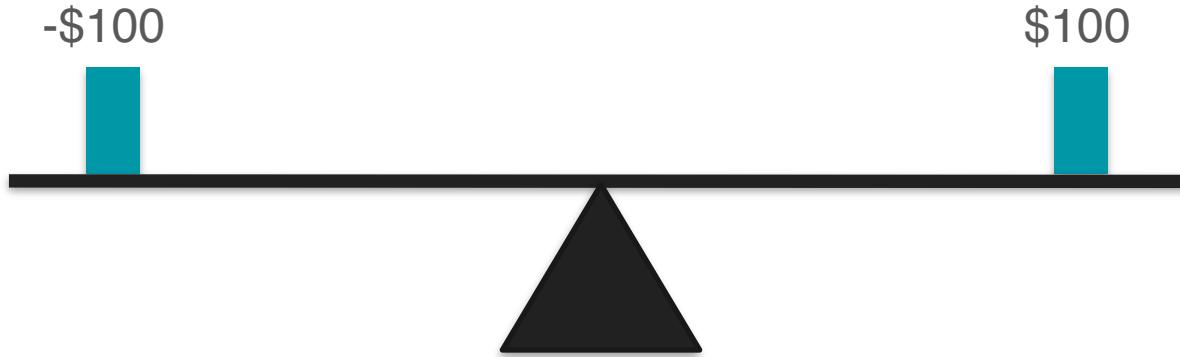


Variance Motivation: Measuring Spread

X_1 = expected amount of money gained in game 1

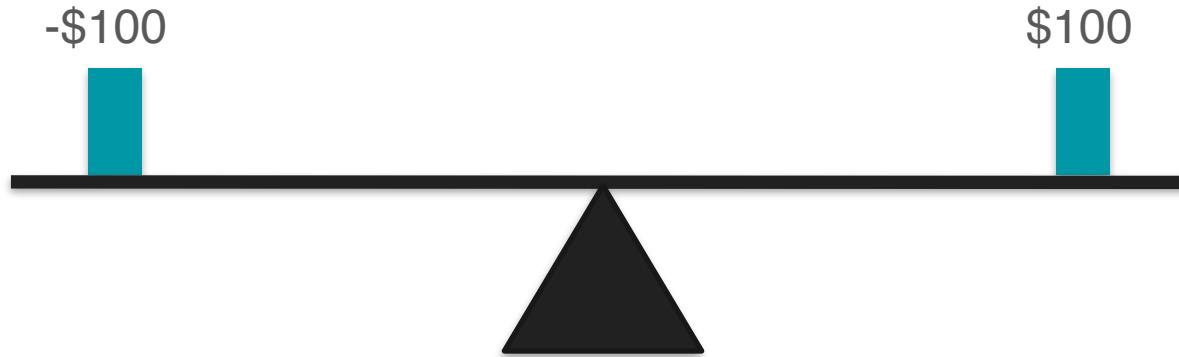


Variance Motivation: Measuring Spread



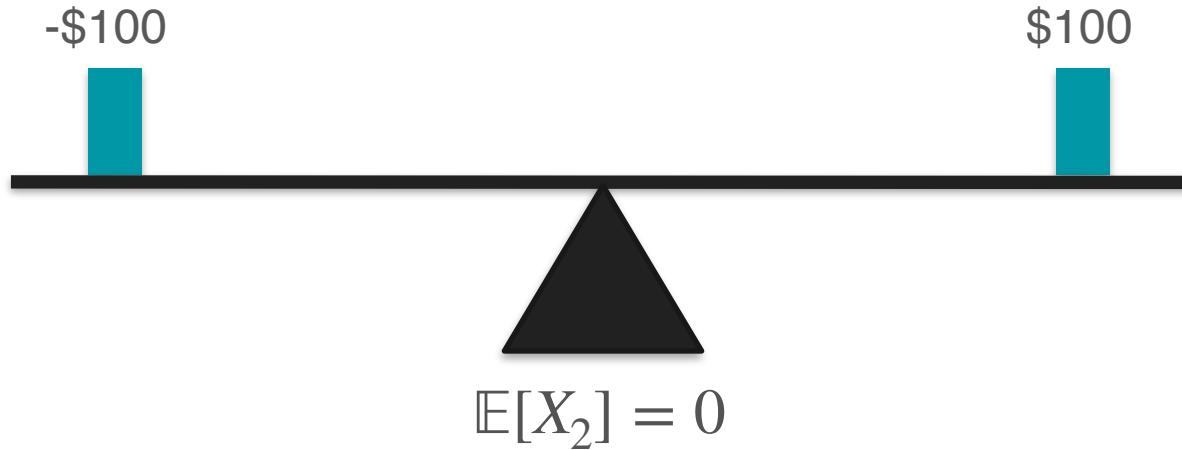
Variance Motivation: Measuring Spread

X_2 = expected amount of money gained in game 2

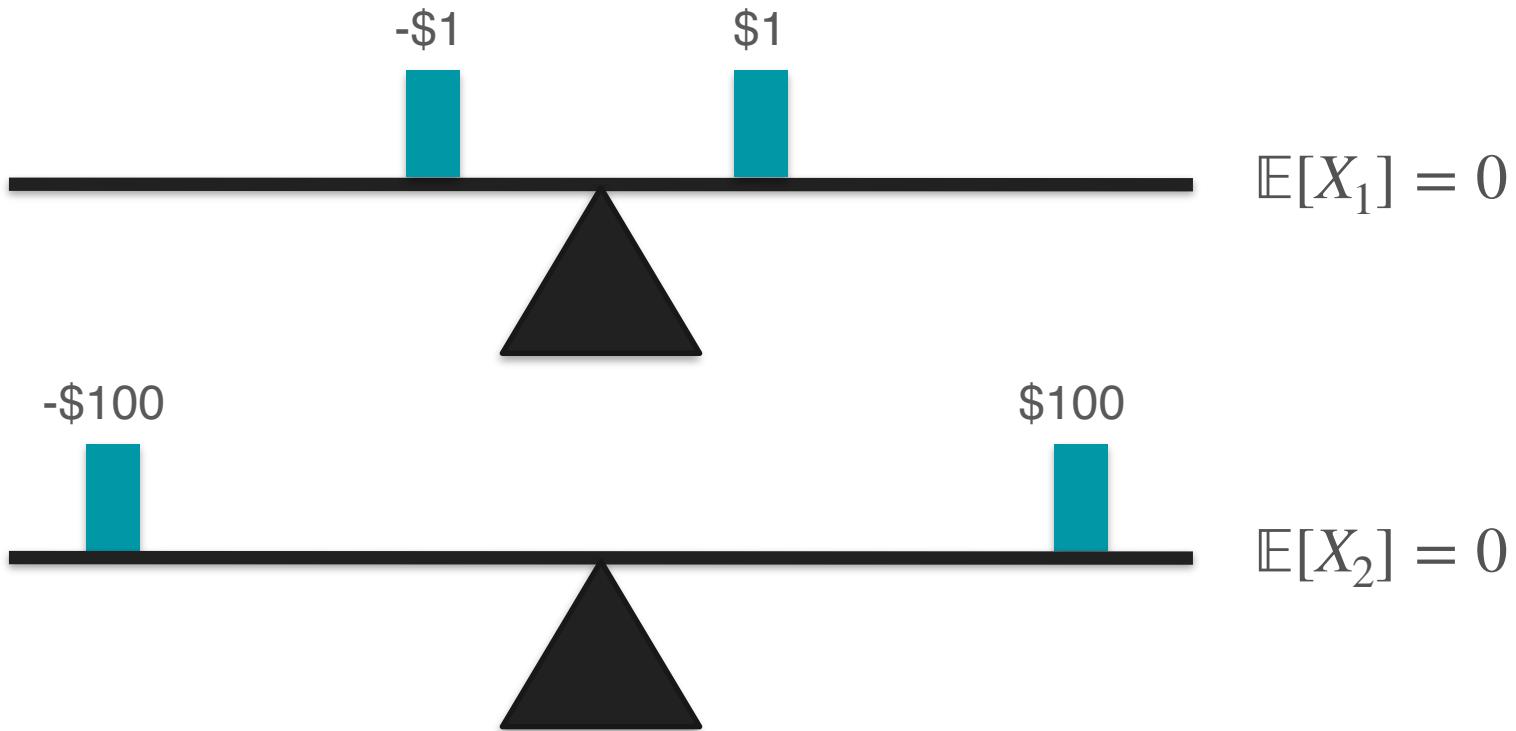


Variance Motivation: Measuring Spread

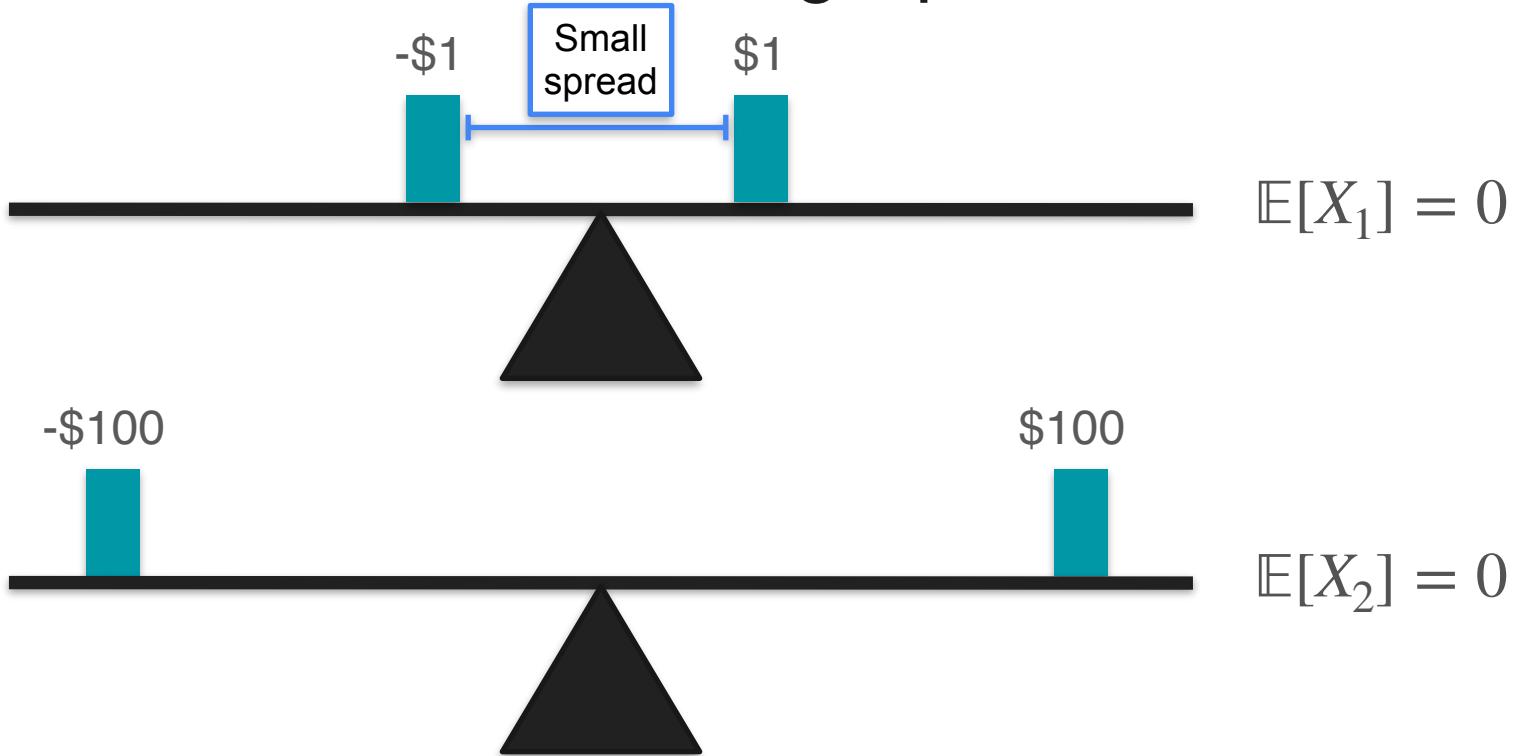
X_2 = expected amount of money gained in game 2



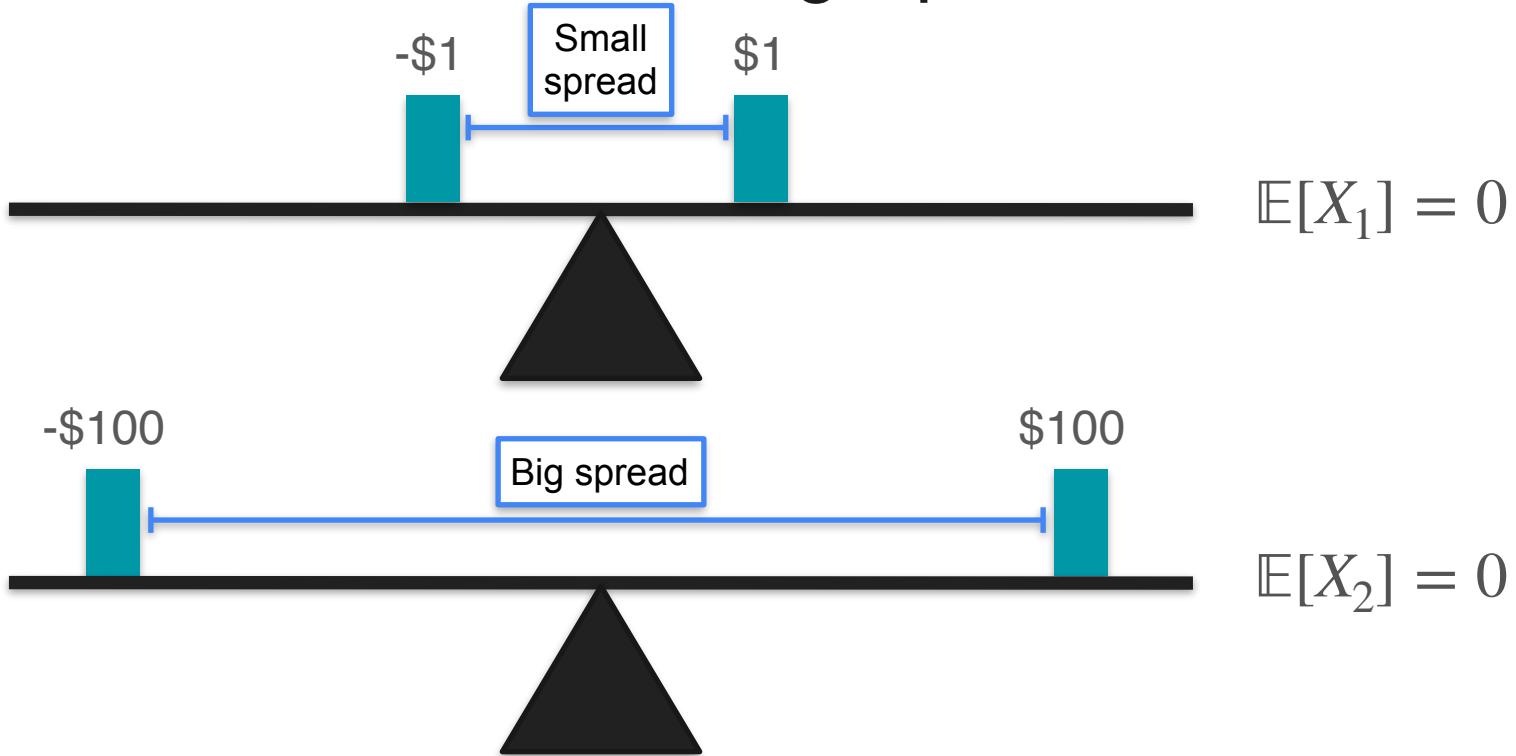
Variance Motivation: Measuring Spread



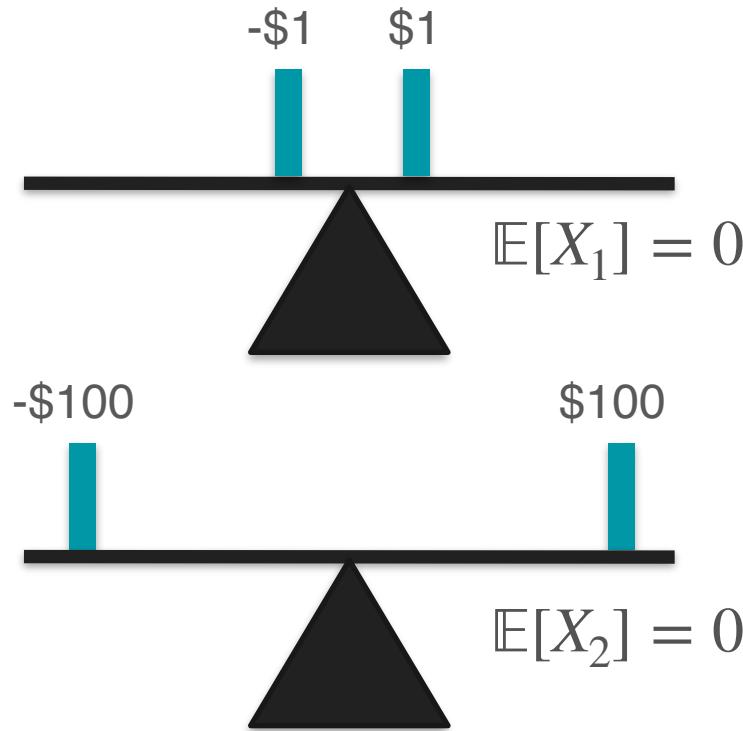
Variance Motivation: Measuring Spread



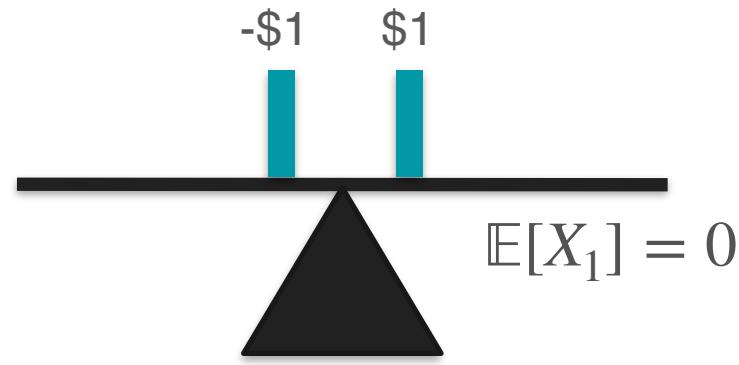
Variance Motivation: Measuring Spread



Variance Motivation: Measuring Spread



Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$



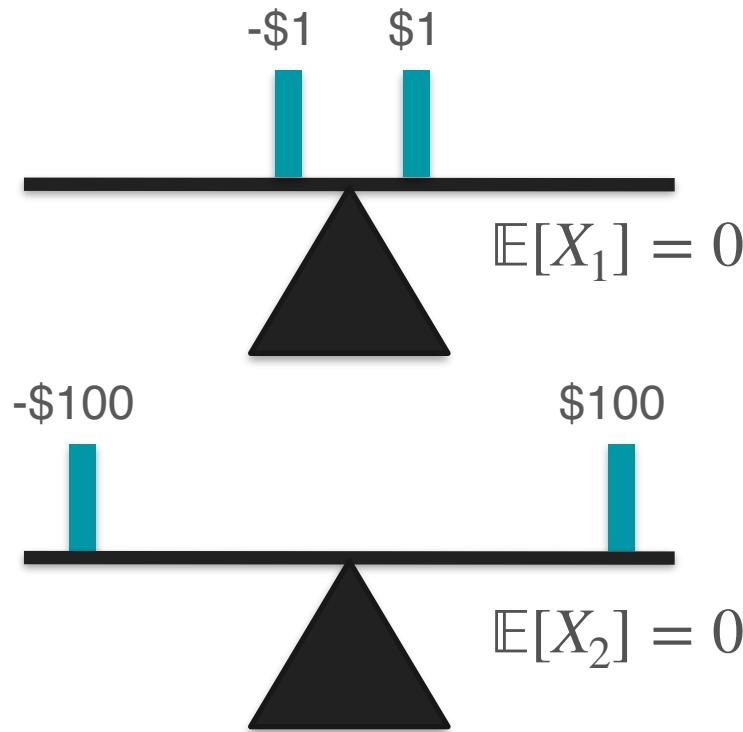
Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$

$$\mathbb{E}[X_2] = \frac{(100) + (-100)}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1) + (-1)}{2} = 0$$

Turn into positive?

$$\mathbb{E}[X_2] = \frac{(100) + (-100)}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1)^2 + (-1)^2}{2} = 0$$

$$\mathbb{E}[X_2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1] = \frac{(1)^2 + (-1)^2}{2} = 0$$



$$\mathbb{E}[X_2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 0$$



$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 0$$

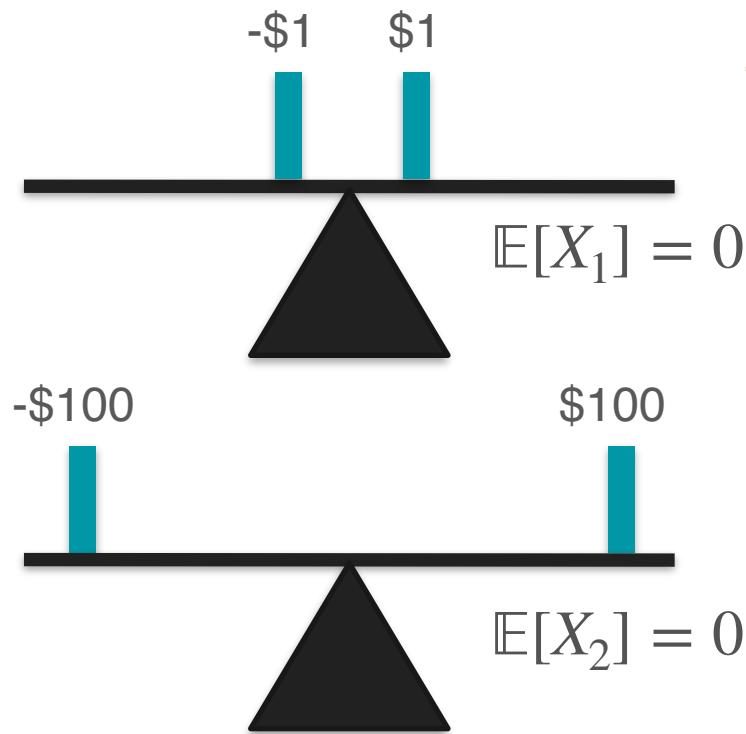
Variance Motivation: Measuring Spread



$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Variance Motivation: Measuring Spread

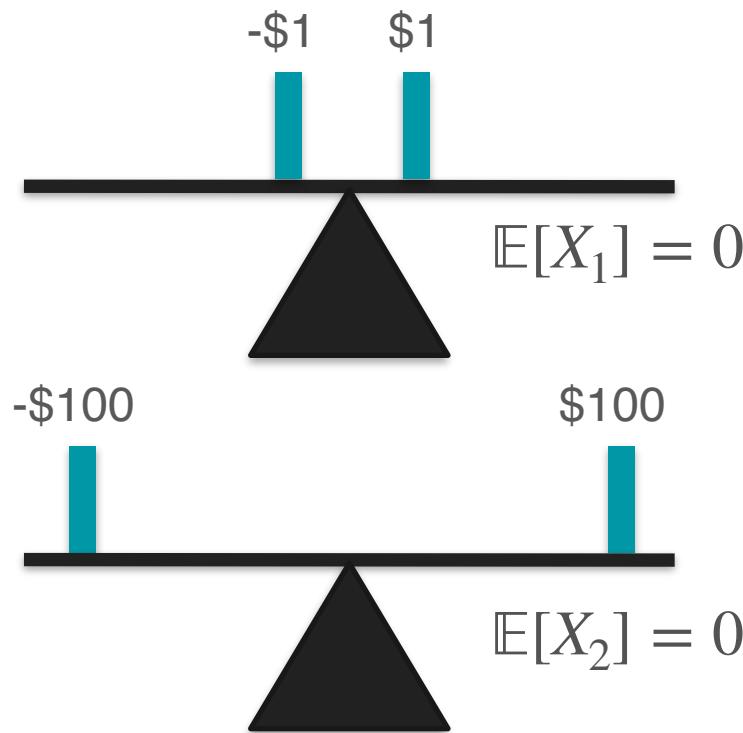


Key for telling game
1 and game 2 apart!

$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Variance Motivation: Measuring Spread



Key for telling game
1 and game 2 apart!

$$\mathbb{E}[X_1^2] = \frac{(1)^2 + (-1)^2}{2} = 1$$

$$\mathbb{E}[X_2^2] = \frac{(100)^2 + (-100)^2}{2} = 10,000$$

Measure of spread

Variance Motivation: Measuring Spread

$$\mathbb{E}[X^2]$$

Variance Motivation: Measuring Spread

$$\mathbb{E}[X^2]$$

Almost...

Variance Motivation: Centering With Mean

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

How risky are these two games in comparison?

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

How risky are these two games in comparison?

Hint: Think of the spread

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

Game 2



You win 6 dollars



You win 4 dollars

They are equally risky

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1$$

Game 2



You win 6 dollars



You win 4 dollars

$$\mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$

Variance Motivation: Centering With Mean

Game 1



You win 1 dollar



You lose 1 dollar

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1$$



Same risk?

$$\mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$



Game 2



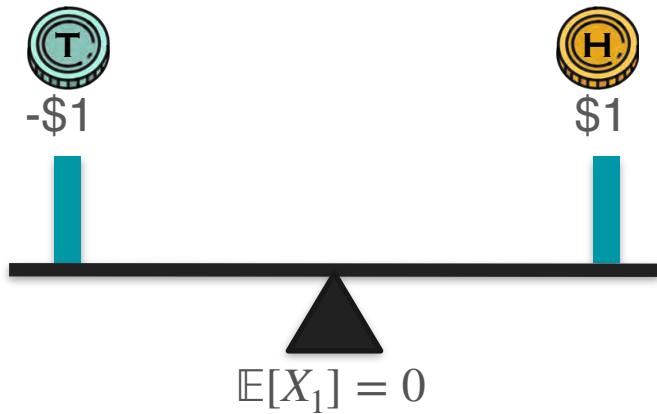
You win 6 dollars



You win 4 dollars

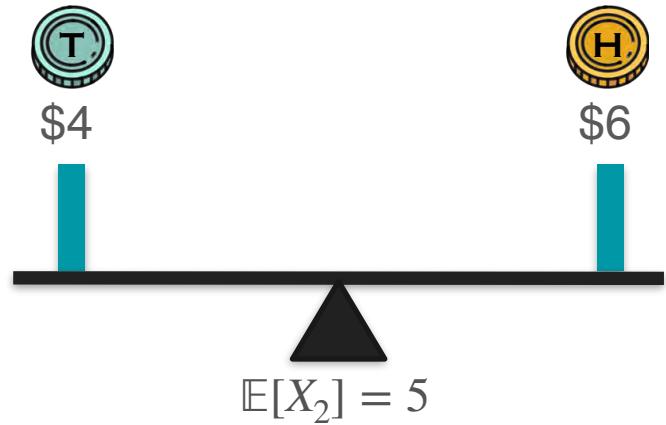
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

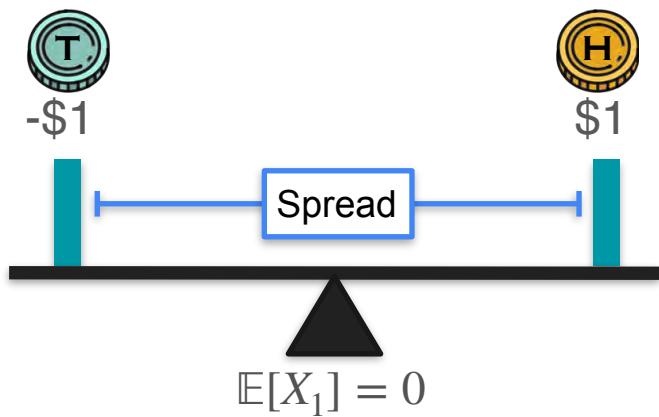
$$\mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$



Game 2

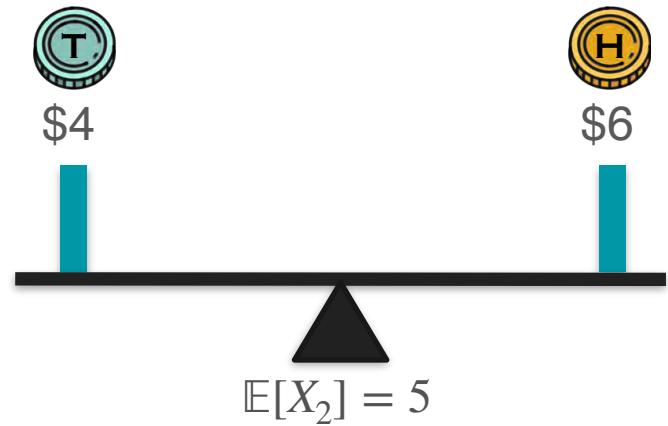
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

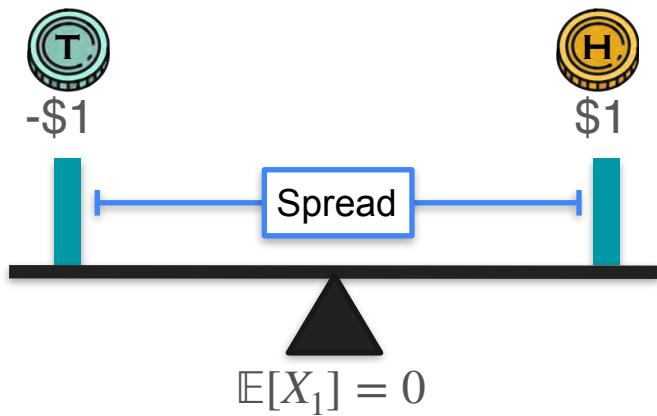
$$\mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$



Game 2

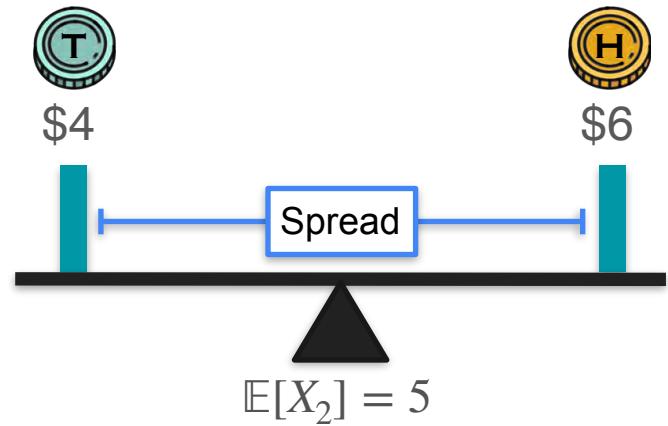
Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1$$



Game 1

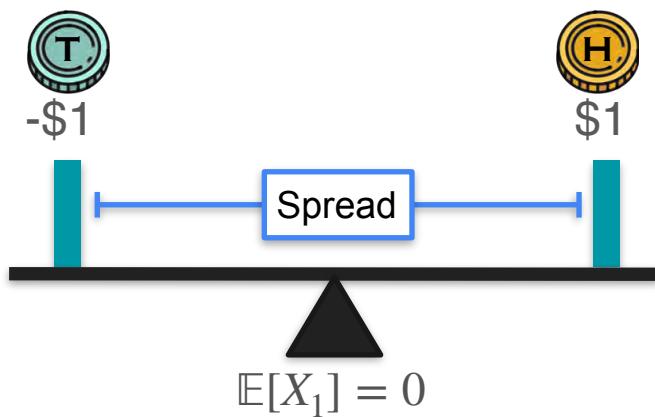
$$\mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$



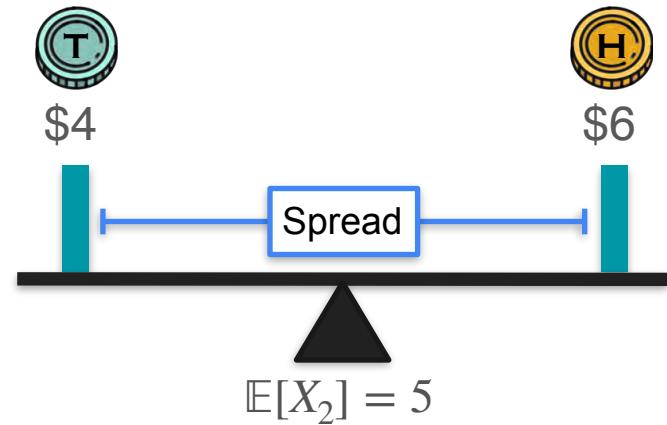
Game 2

Variance Motivation: Centering With Mean

$$\mathbb{E}[X_1]^2 = \frac{(-1)^2 + (1)^2}{2} = 1 \quad \leftarrow \text{Same spread?} \rightarrow \quad \mathbb{E}[X_2]^2 = \frac{(4)^2 + (6)^2}{2} = 26$$



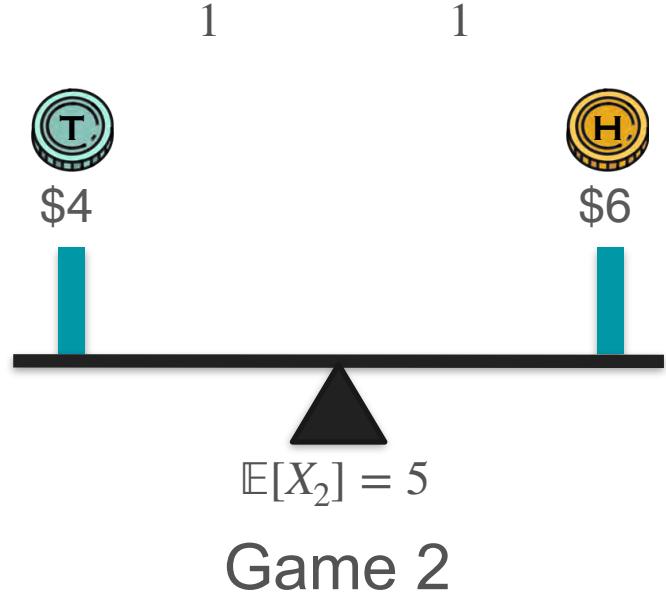
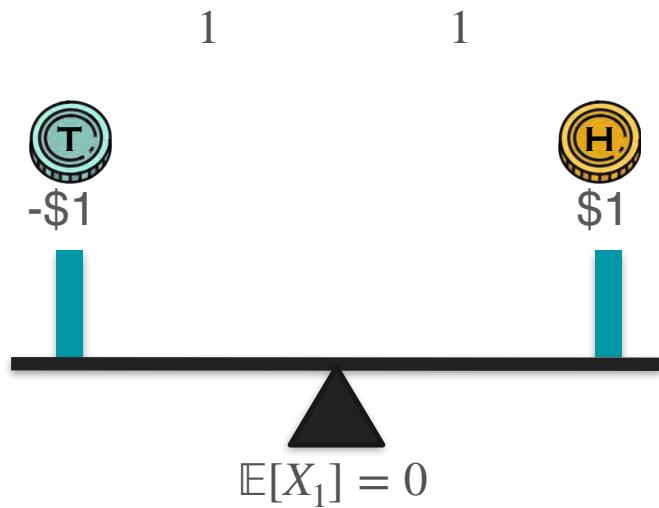
Game 1



Game 2

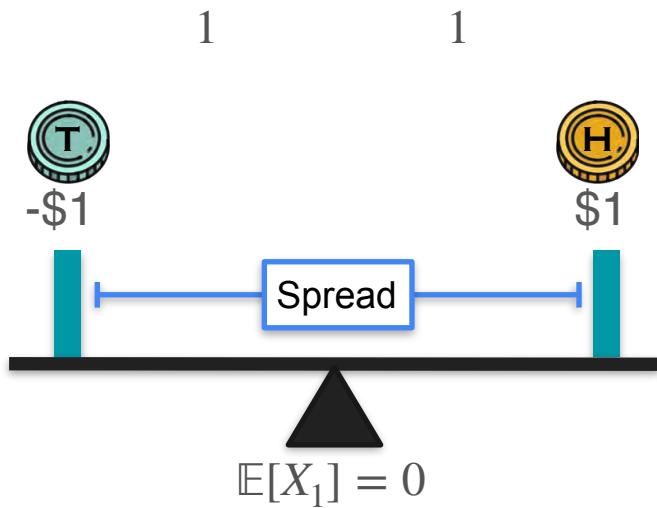
Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$

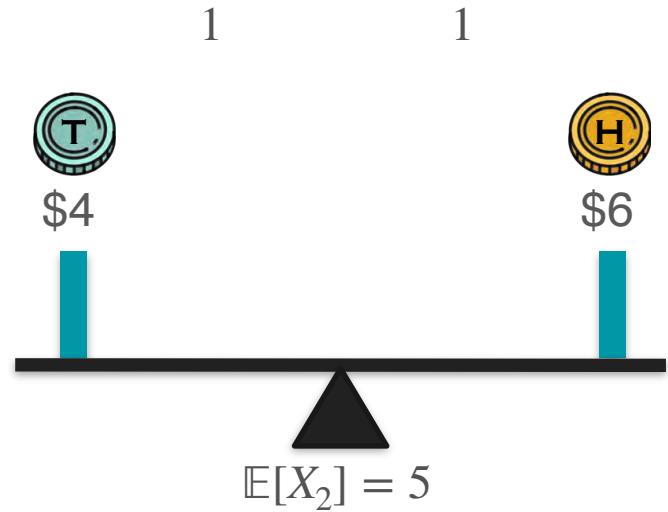


Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



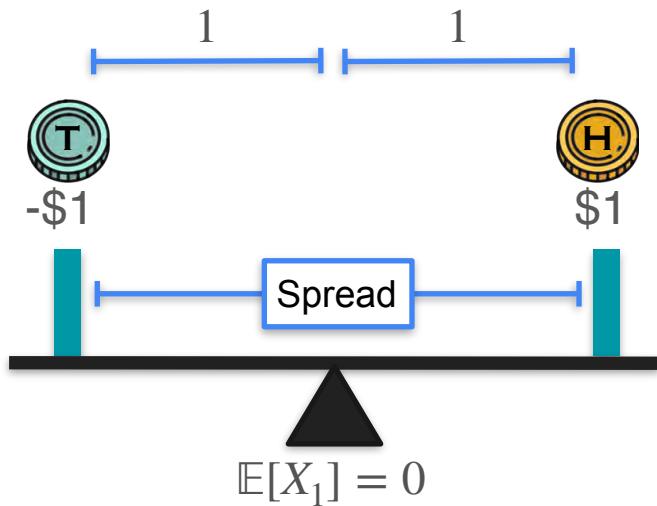
Game 1



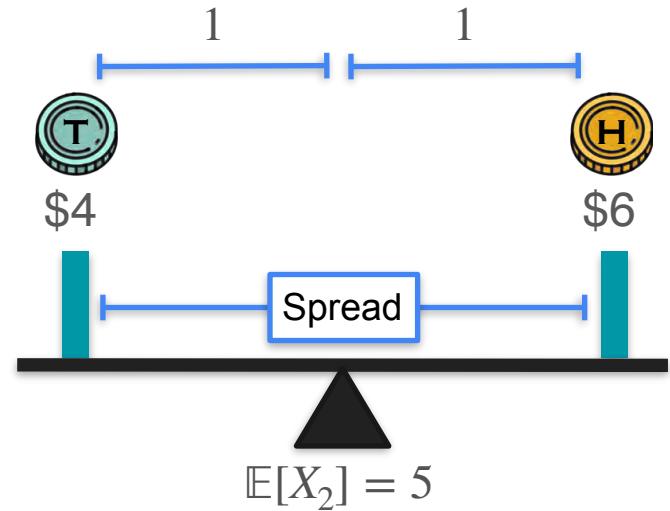
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



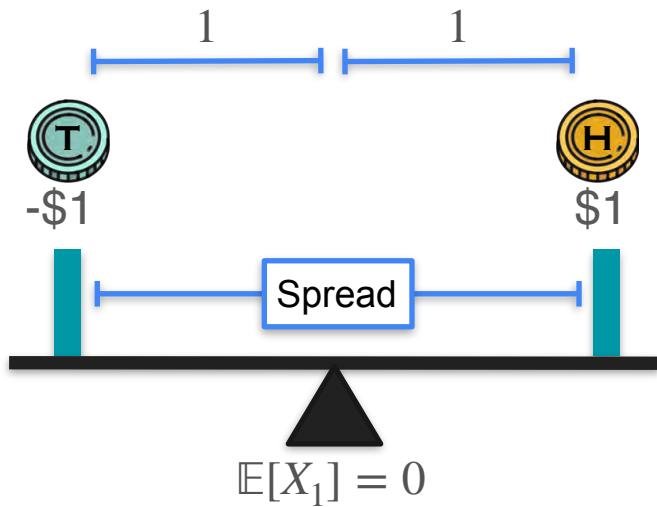
Game 1



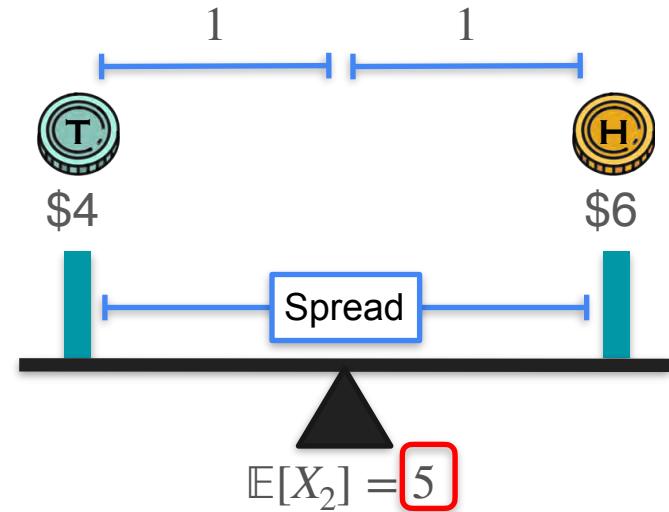
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



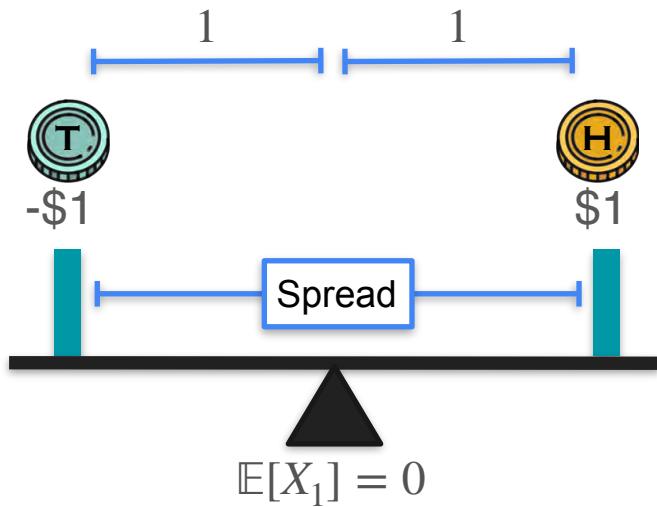
Game 1



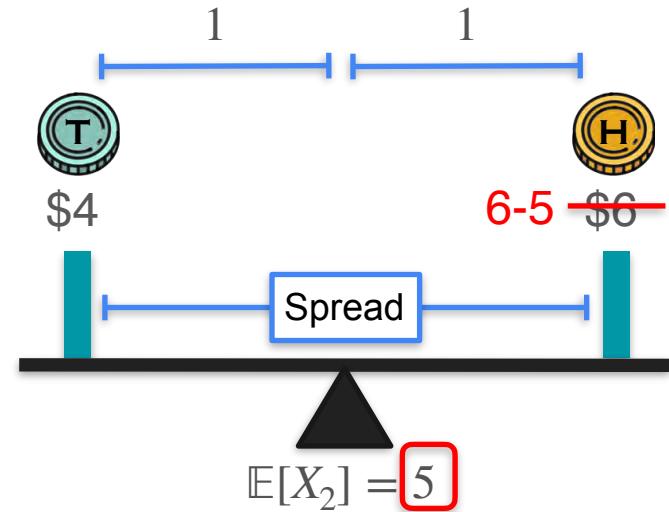
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



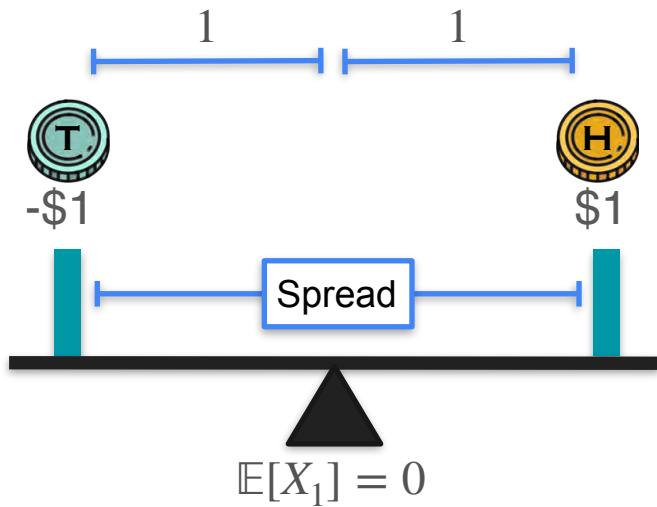
Game 1



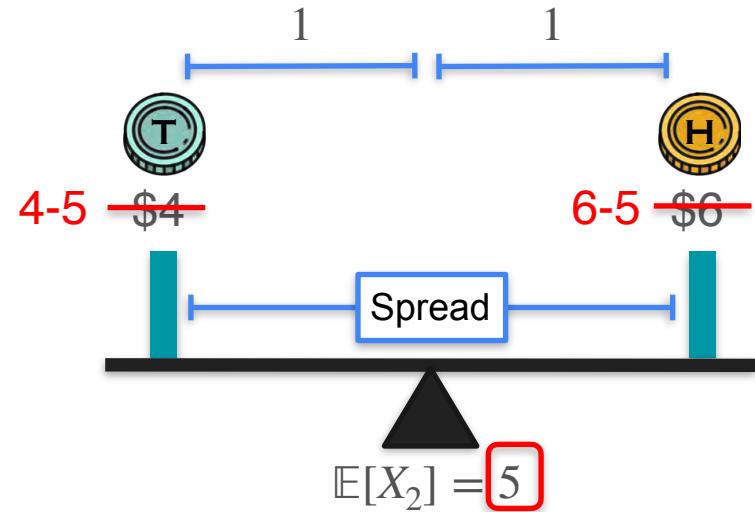
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



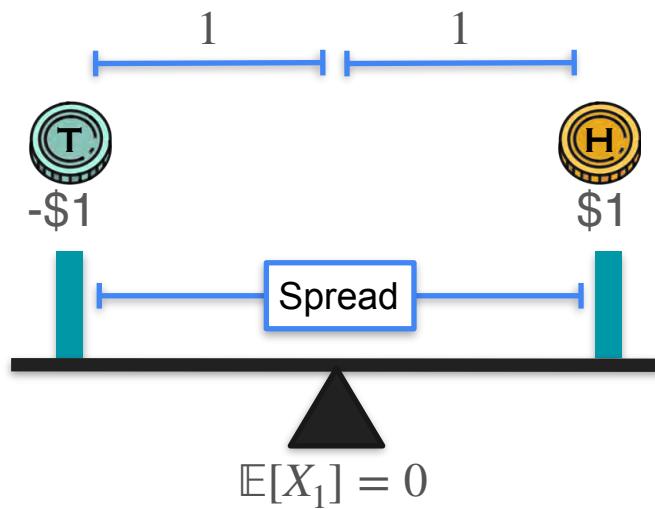
Game 1



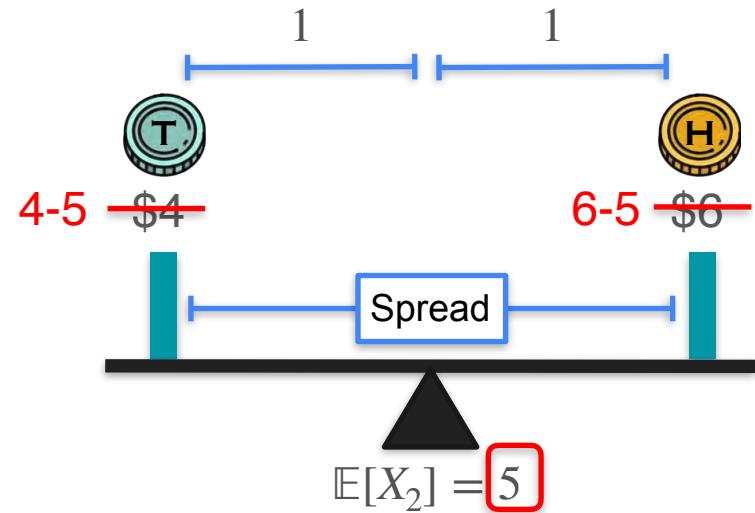
Game 2

Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$

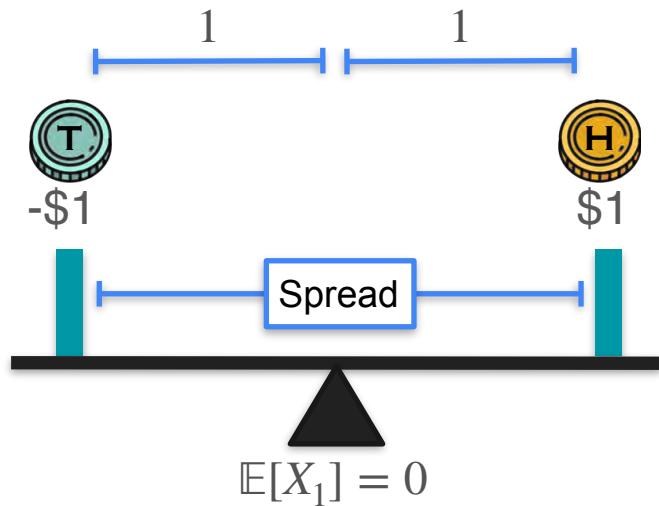


$$\mathbb{E}[(X_2 - 5)^2]$$



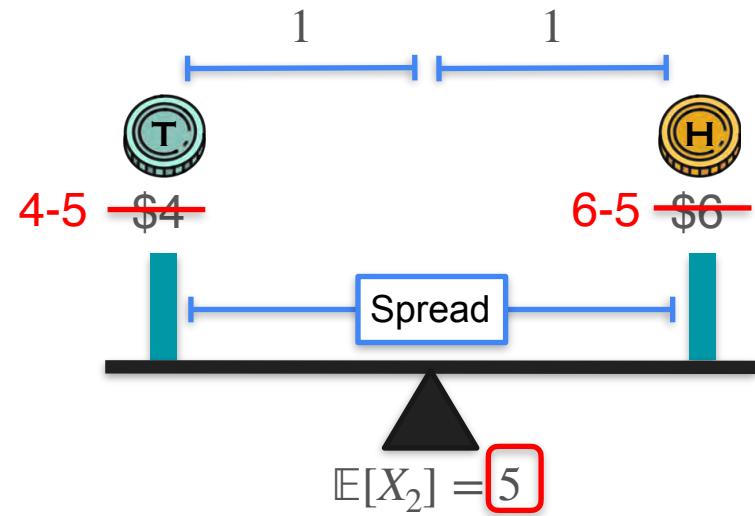
Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$



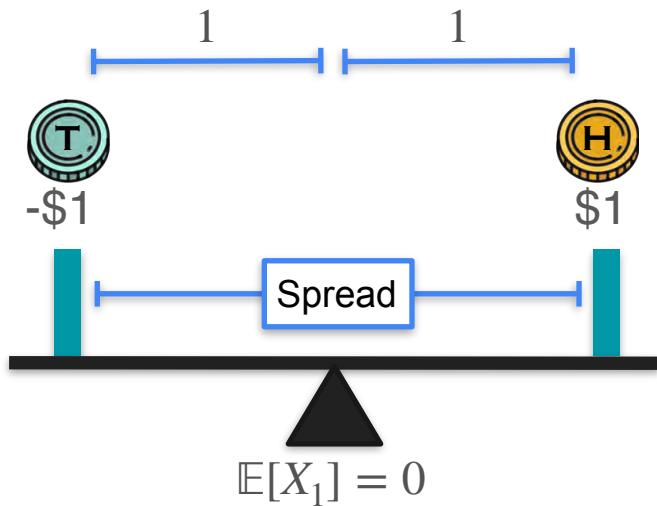
$$\text{Mean: } \mu$$

$$\mathbb{E}[(X_2 - 5)^2]$$



Variance: Spread and Shift

$$\mathbb{E}[X_1^2] = 1$$

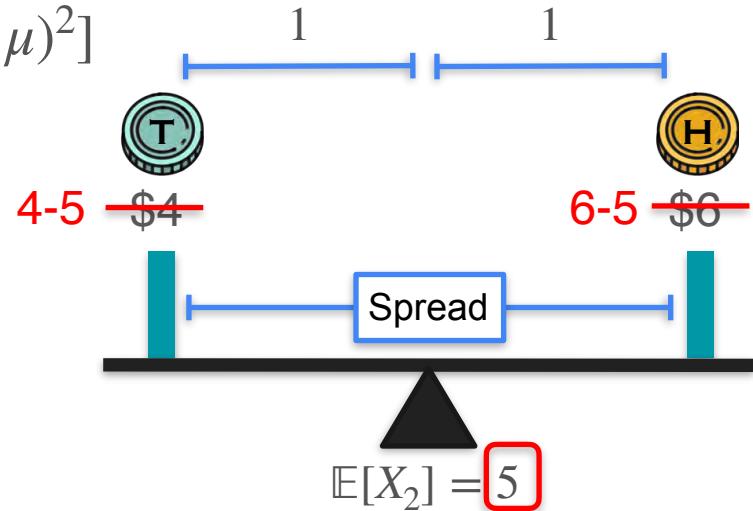


Game 1

$$\text{Mean: } \mu$$

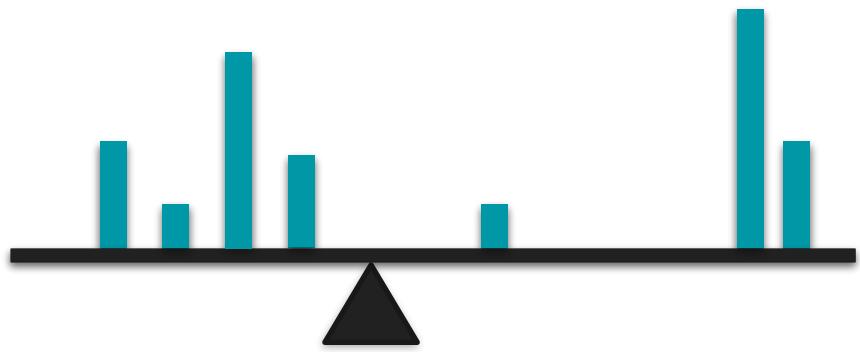
$$\text{Variance: } \mathbb{E}[(X - \mu)^2]$$

$$\mathbb{E}[(X_2 - 5)^2]$$

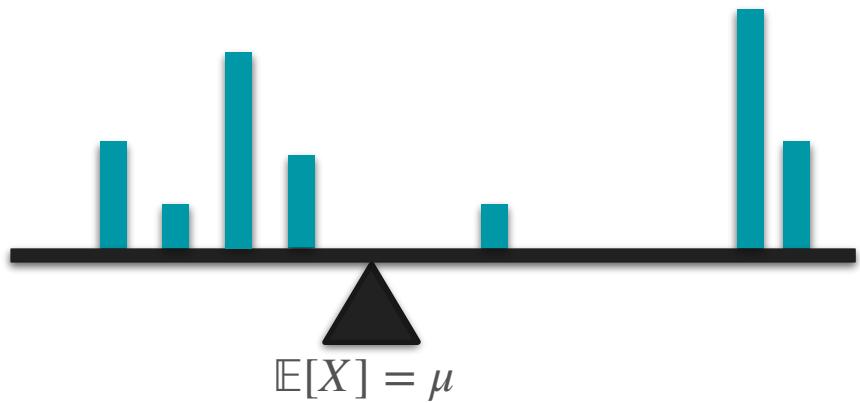


Game 2

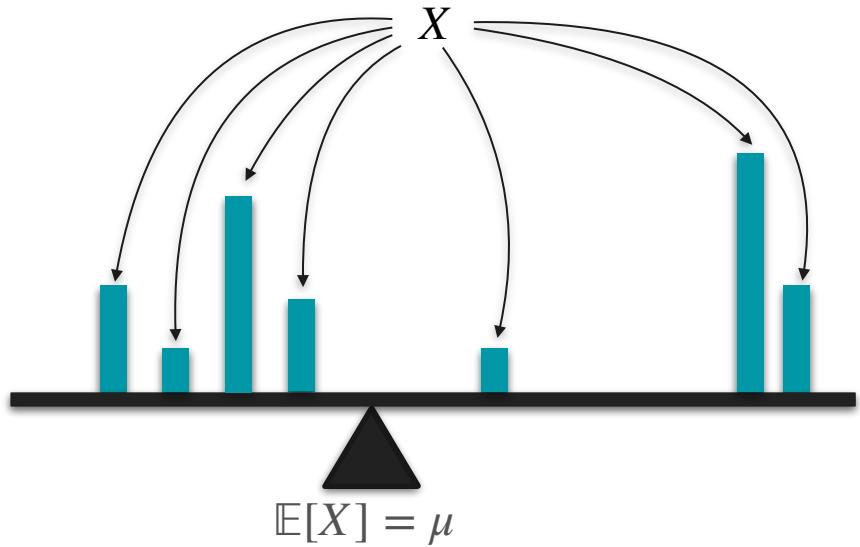
Variance Formula



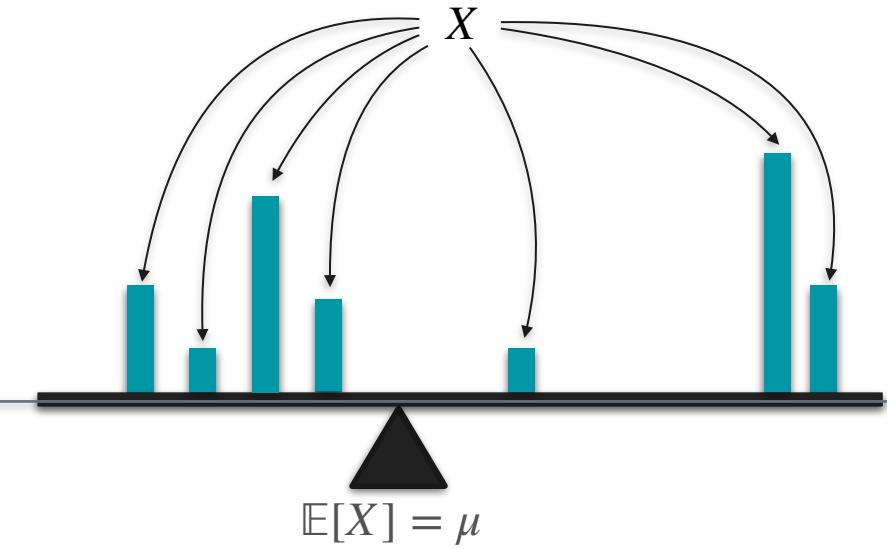
Variance Formula



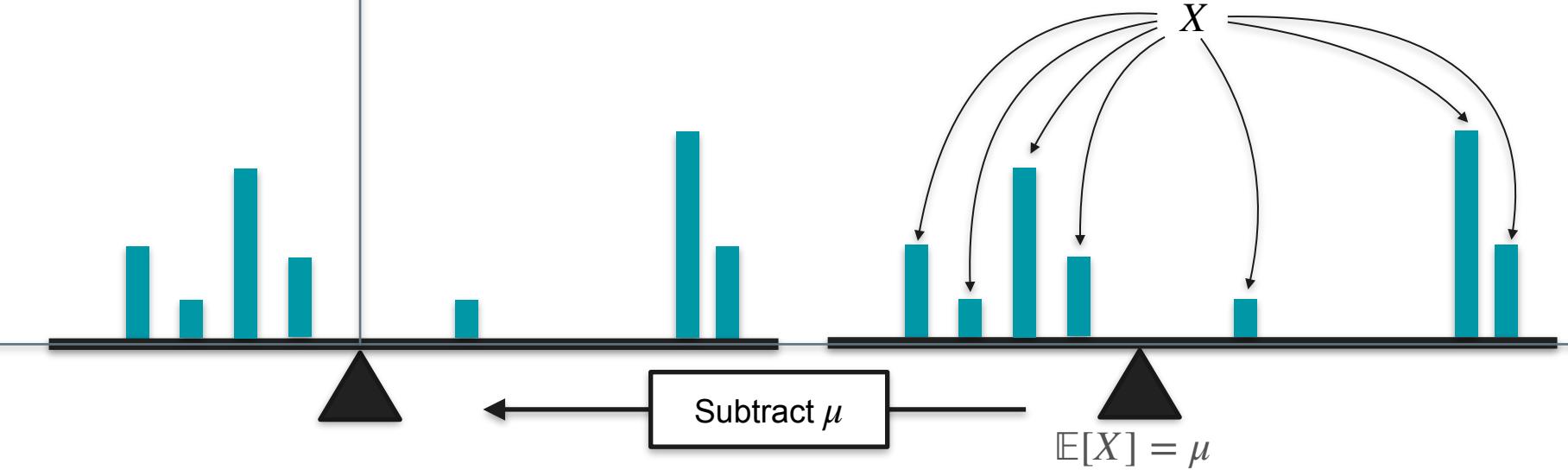
Variance Formula



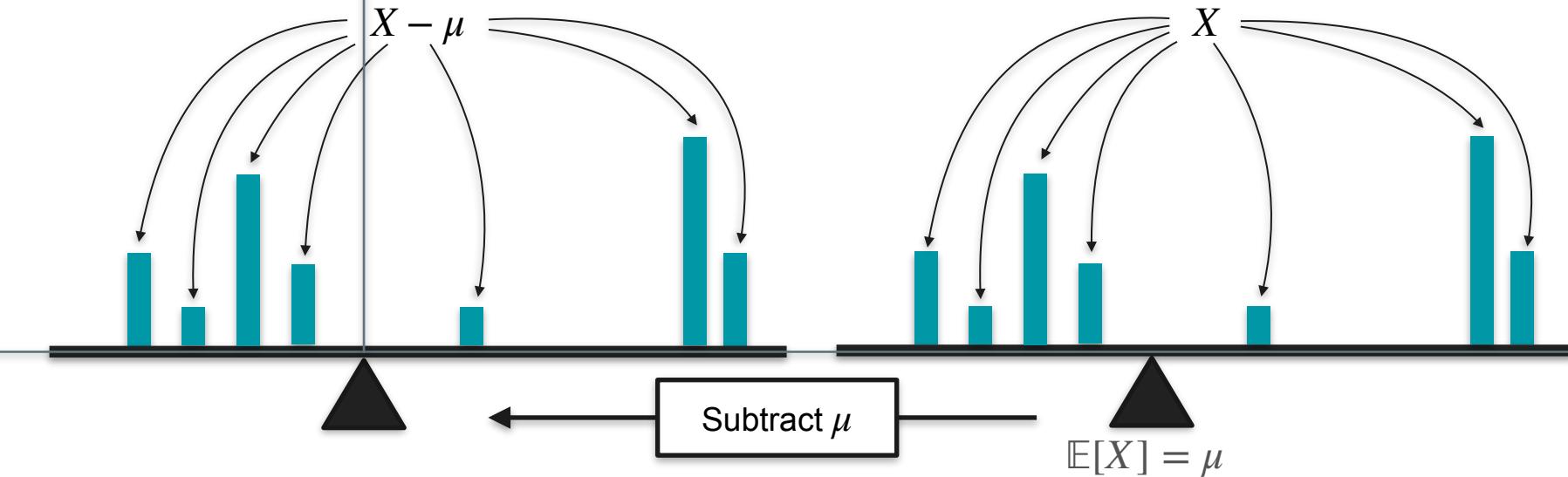
Variance Formula



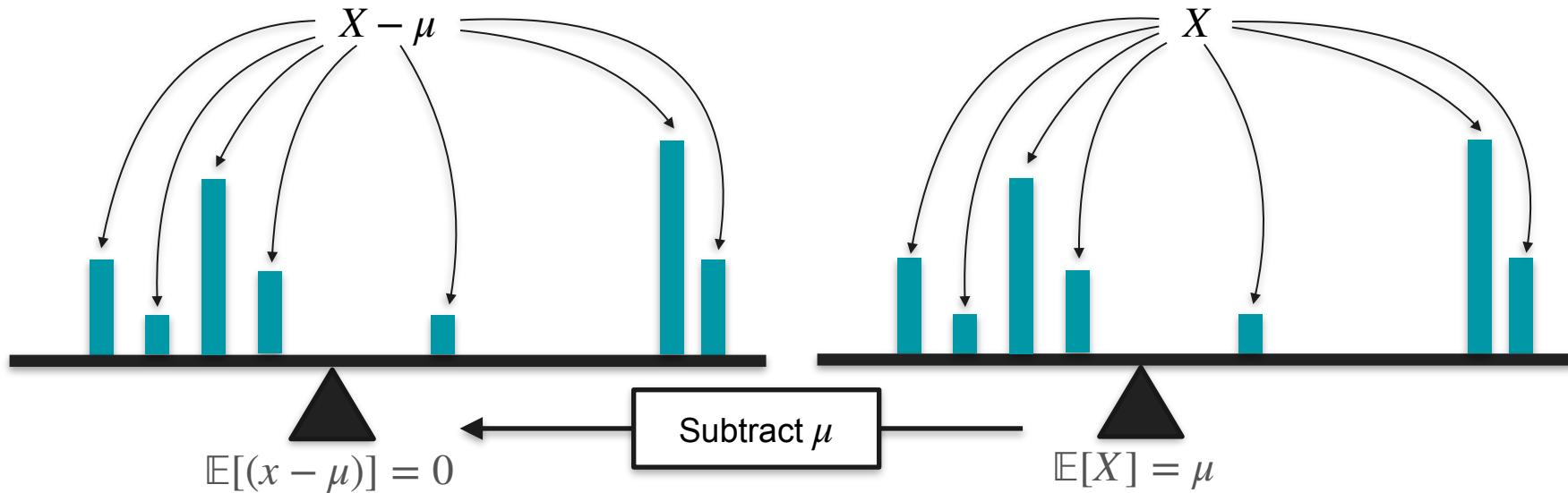
Variance Formula



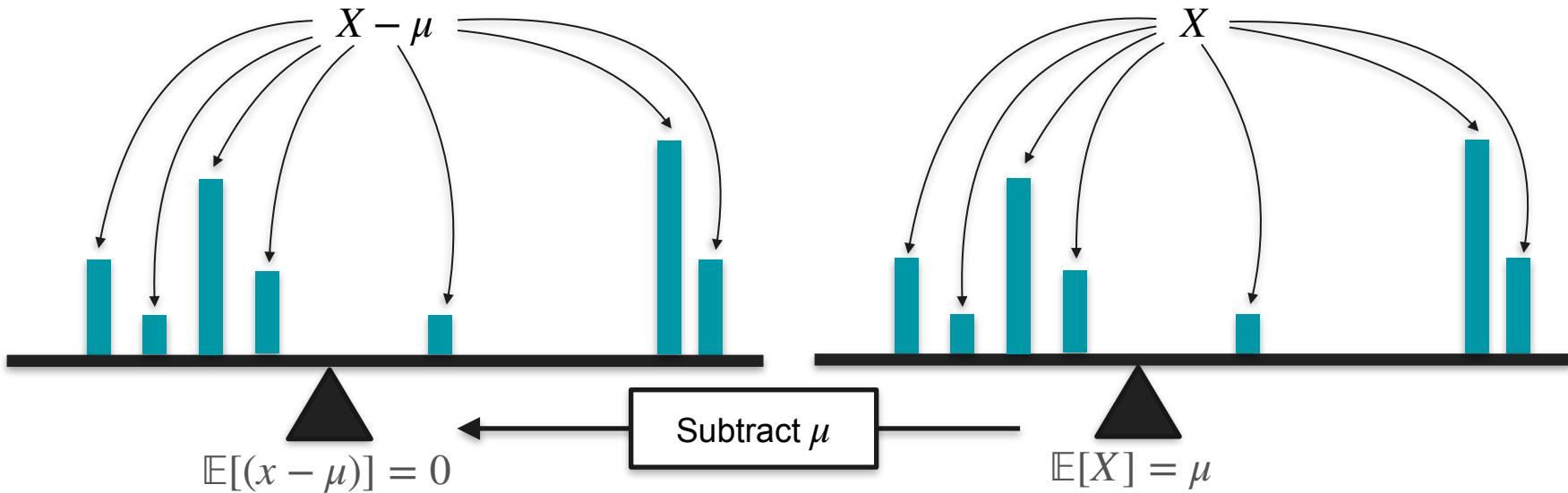
Variance Formula



Variance Formula



Variance Formula



$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

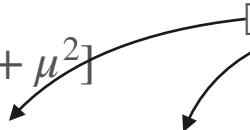
Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

Variance Formula

$$Var(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2]$$

Variance Formula

$$\begin{aligned} \text{Var}(X) = \mathbb{E}[(X - \mu)^2] &= \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \end{aligned}$$


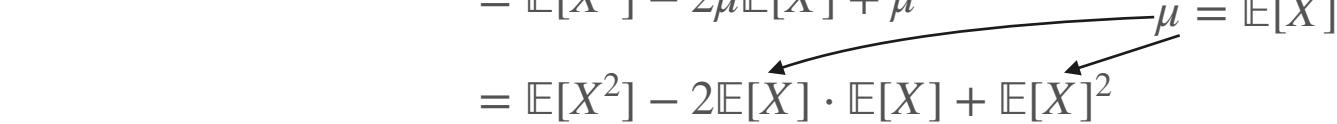
Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \quad \text{[constant} \cdot X \text{]} = \text{constant} \cdot \mathbb{E}[X] \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \end{aligned}$$

Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \quad \mathbb{E}[\text{constant}] = \text{constant} \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \end{aligned}$$

Variance Formula

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \end{aligned}$$


Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\&= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2\end{aligned}$$

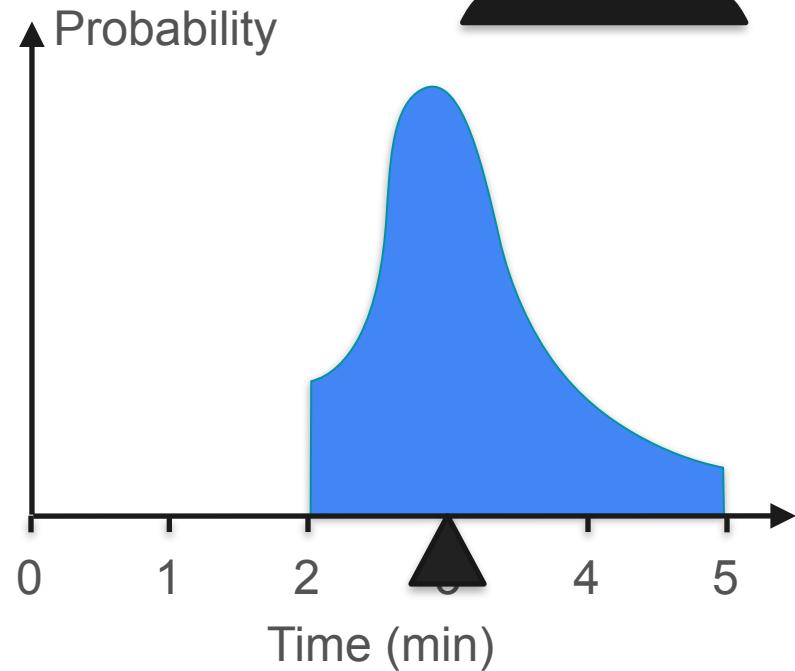
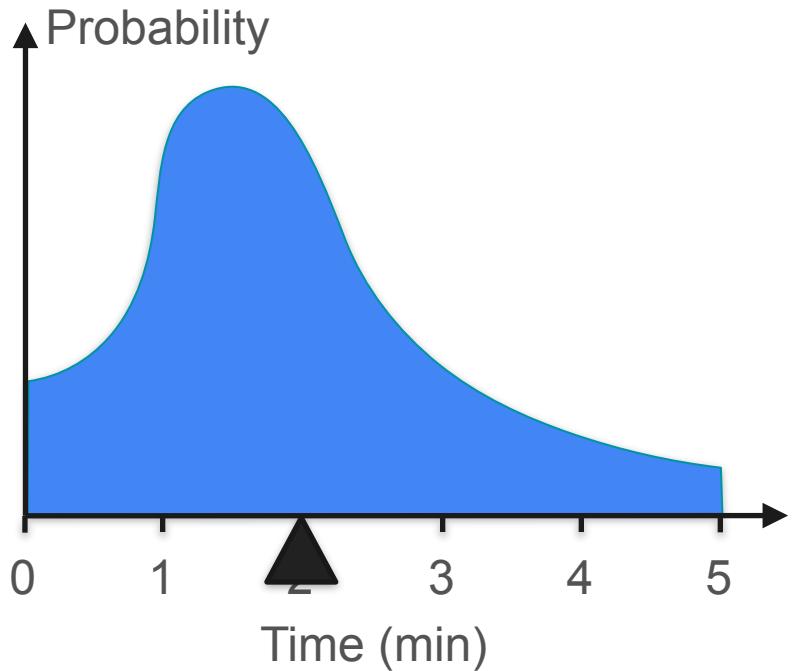
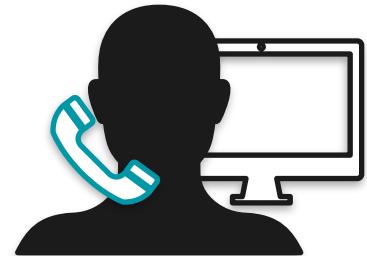
Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2] \\&= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

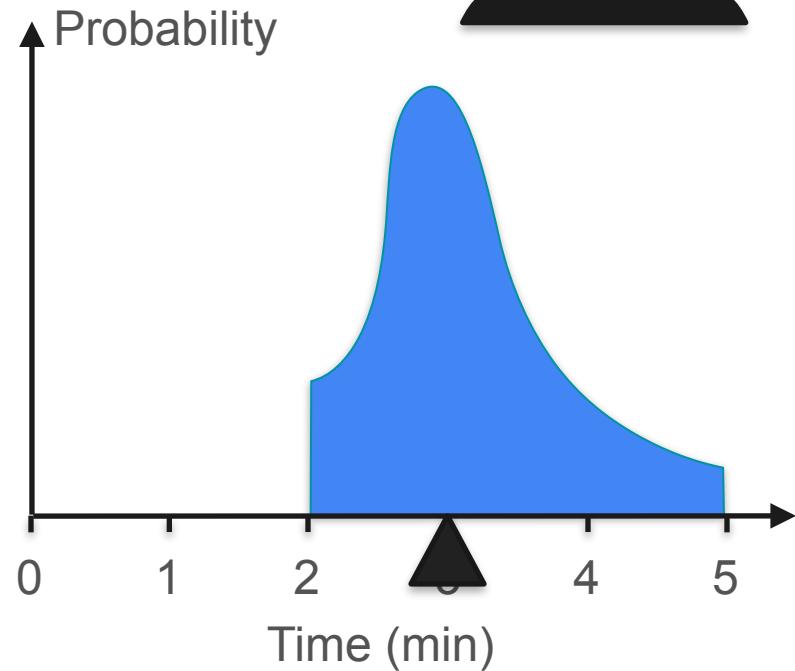
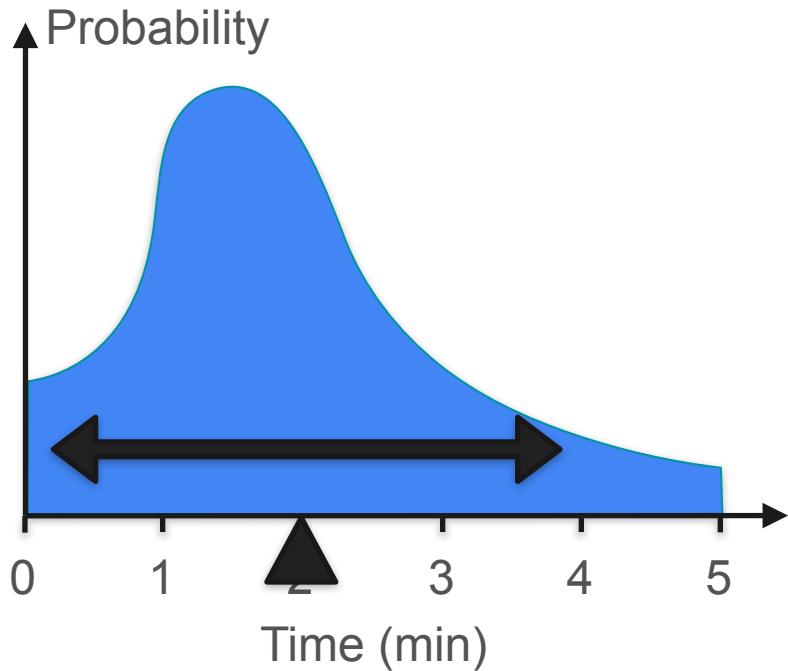
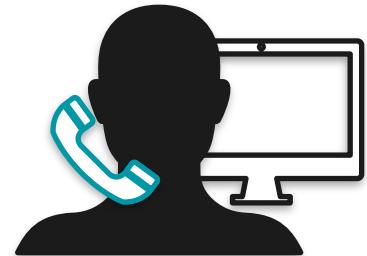
Variance Formula

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mu)^2] \\&= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

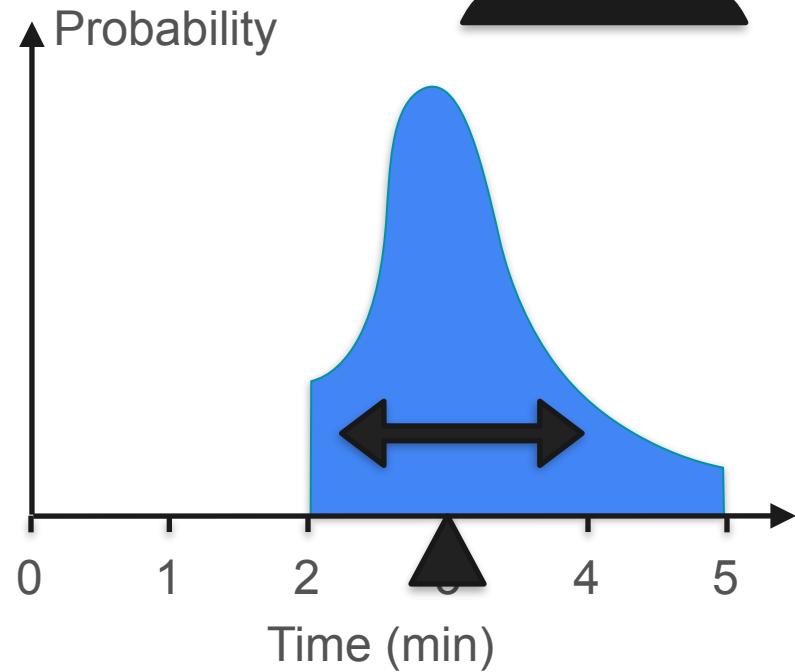
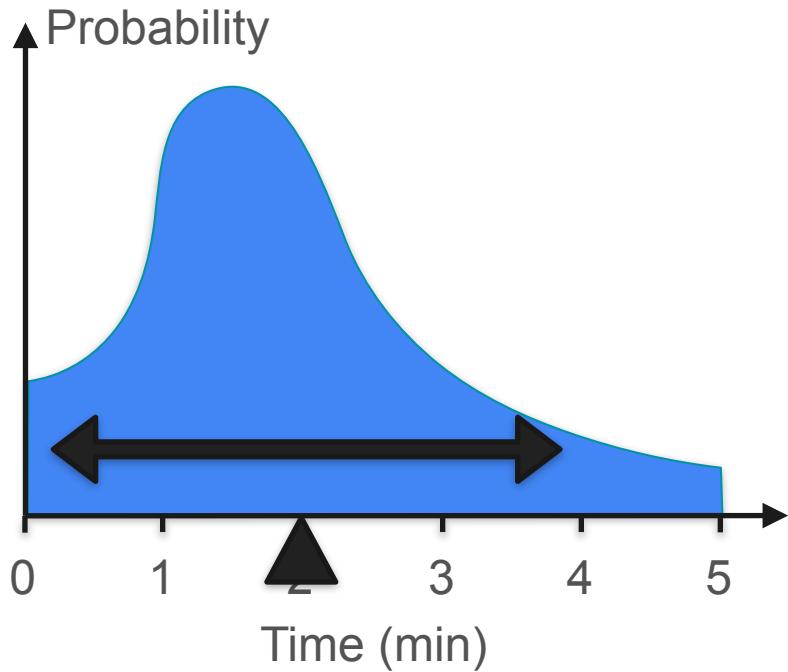
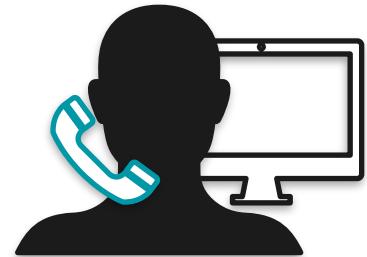
Variance for Continuous Distributions



Variance for Continuous Distributions

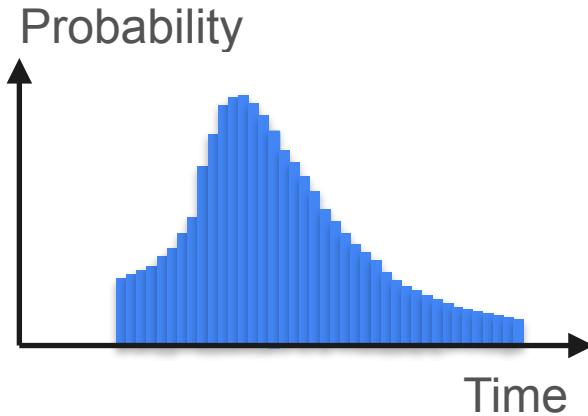


Variance for Continuous Distributions



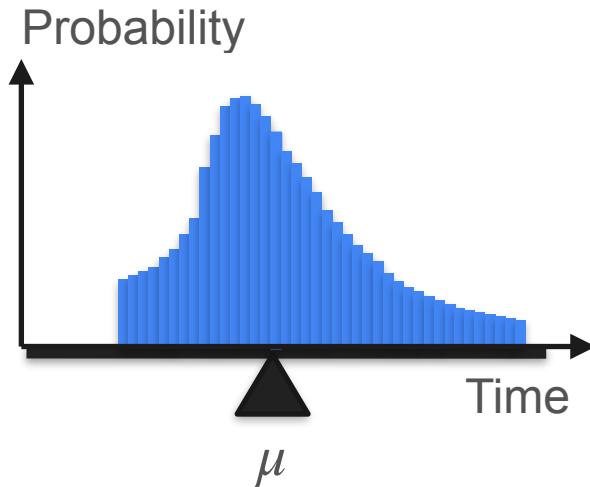
Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$



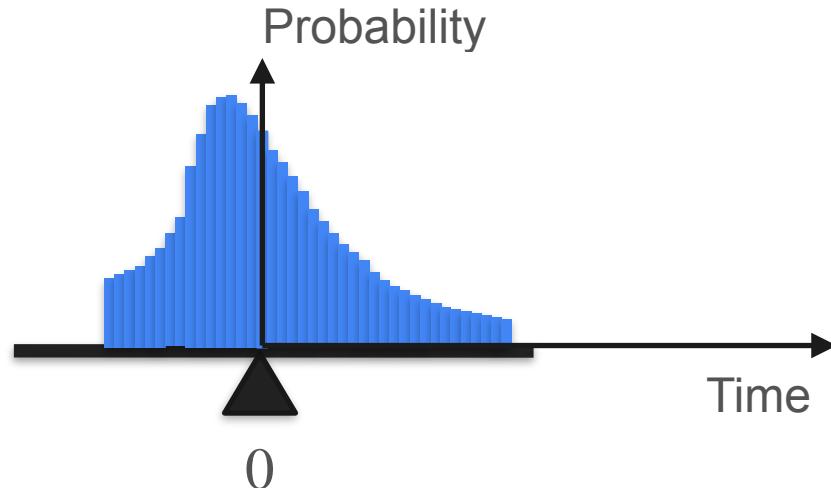
Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$



Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

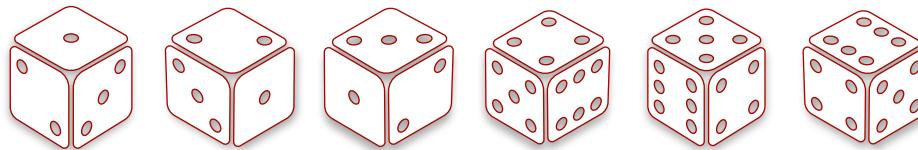


Properties of the Variance

Properties of the Variance

Probability: $1/6$ $1/6$ $1/6$ $1/6$ $1/6$ $1/6$

Roll: 1 2 3 4 5 6



Double: 2 4 6 8 10 12

Wins -3 -2 -1 0 1 2

Properties of the Variance

$$= 2.92$$

Properties of the Variance

$$= 2.92$$



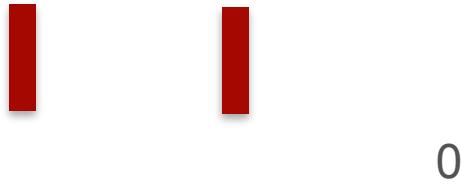
Properties of the Variance

$$= 2.92$$



Properties of the Variance

$$= 2.92$$



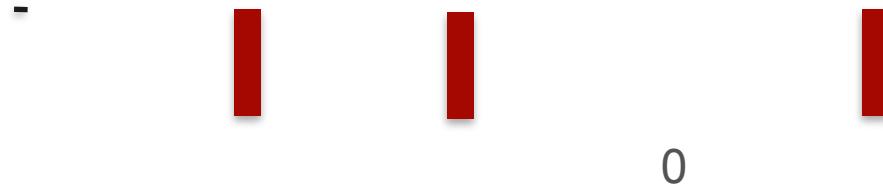
Properties of the Variance

$$= 2.92$$

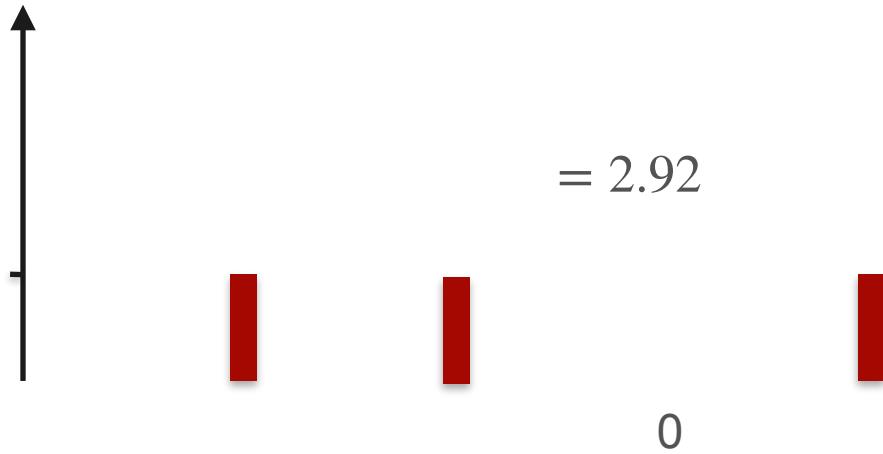


Properties of the Variance

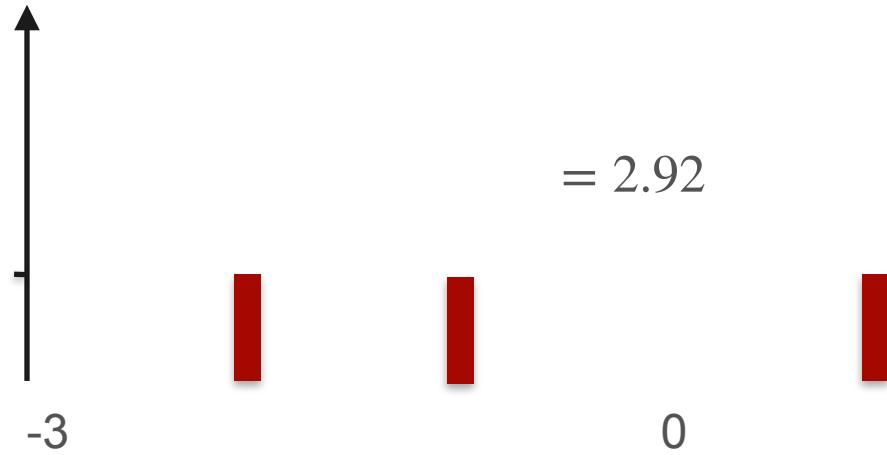
$$= 2.92$$



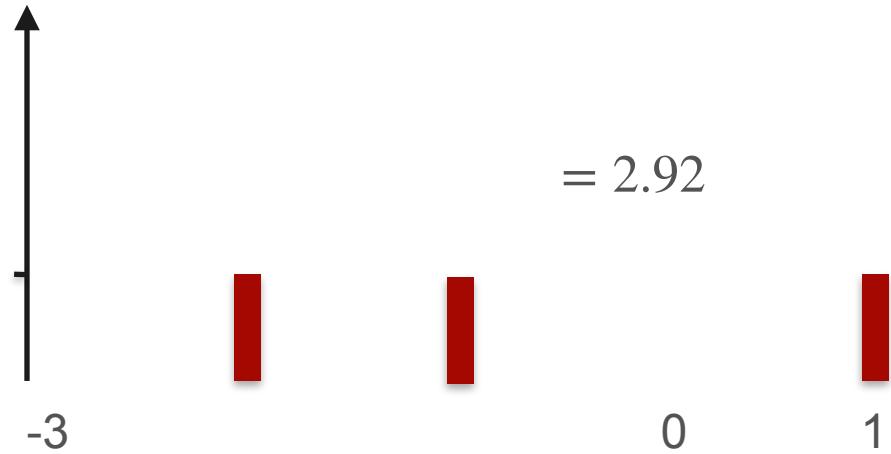
Properties of the Variance



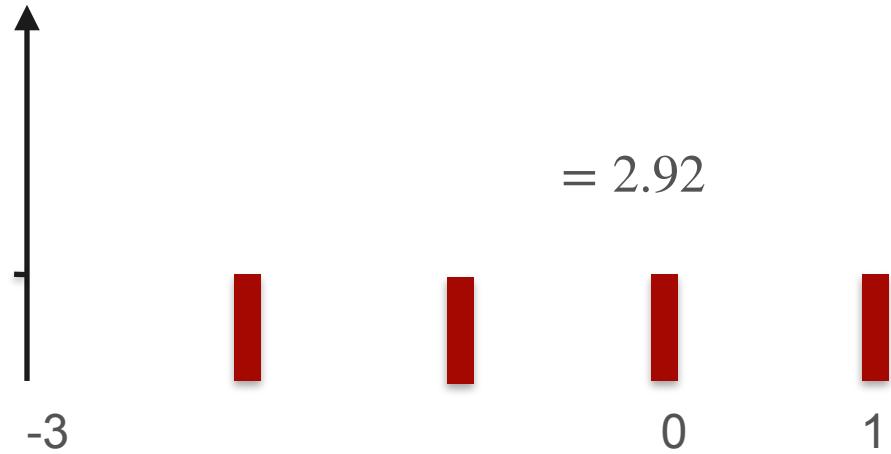
Properties of the Variance



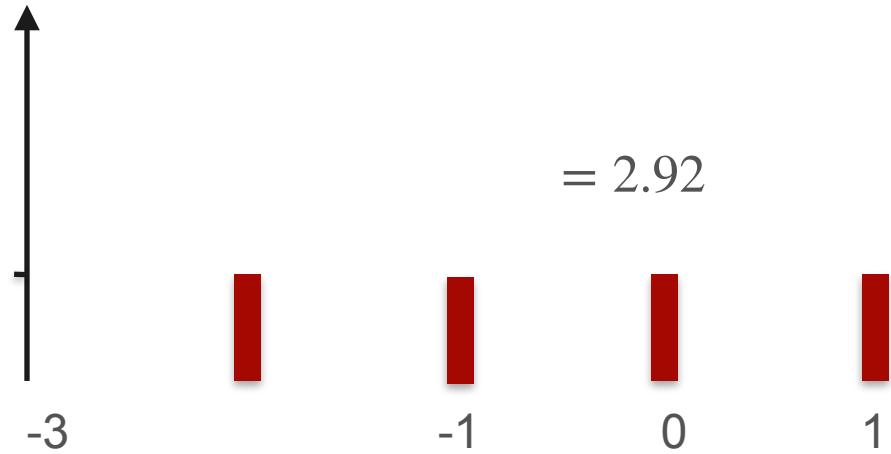
Properties of the Variance



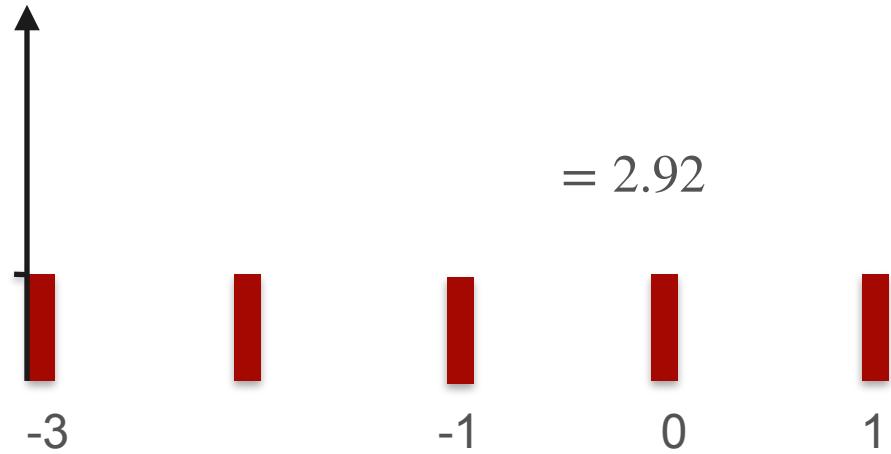
Properties of the Variance



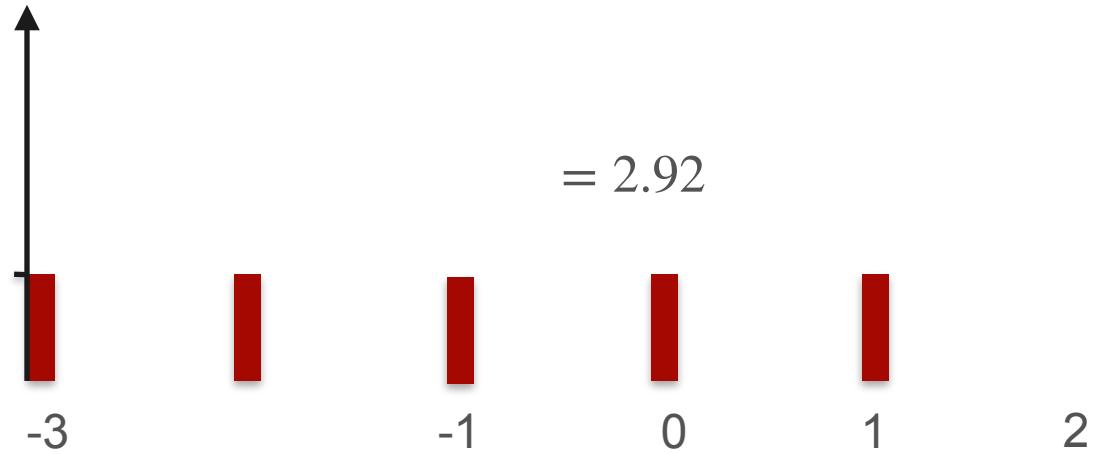
Properties of the Variance



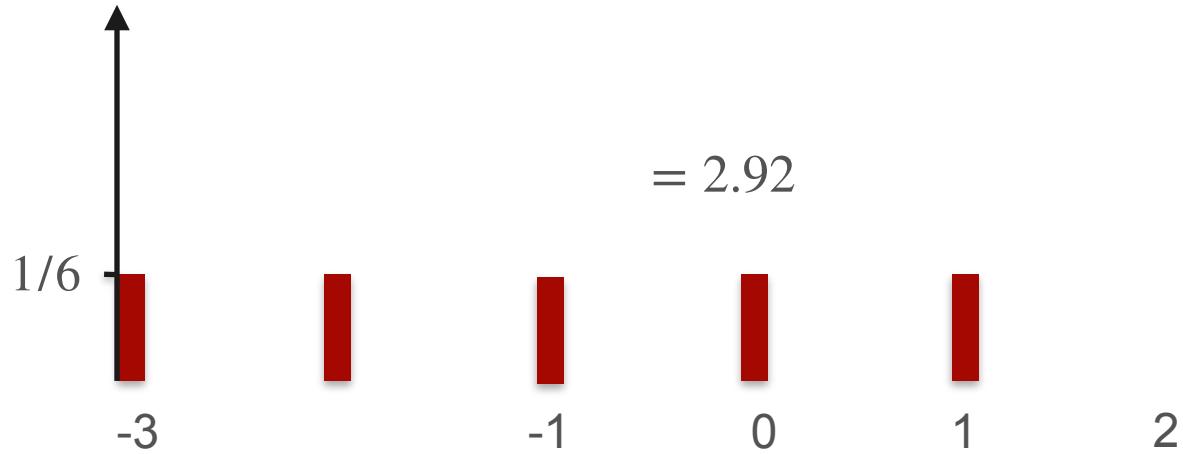
Properties of the Variance



Properties of the Variance

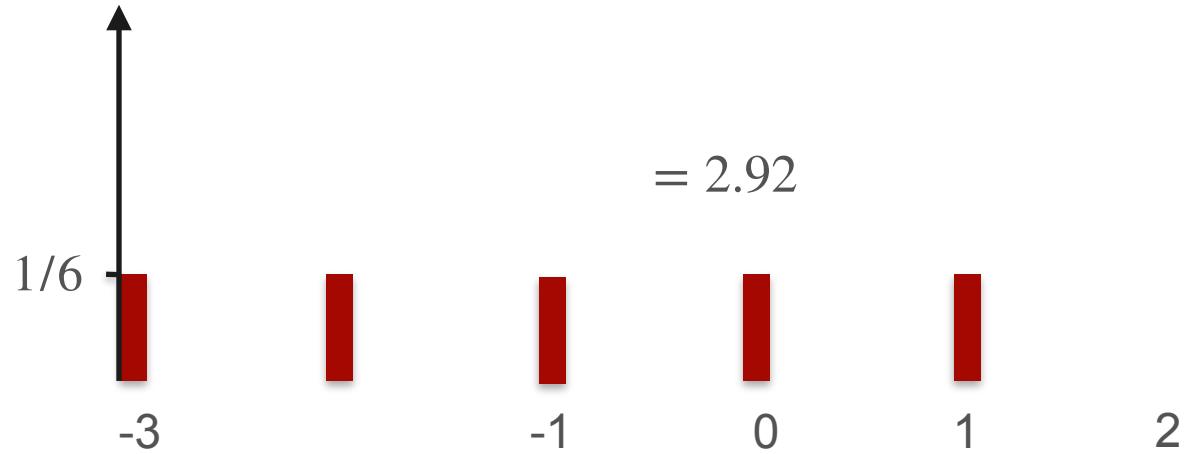


Properties of the Variance



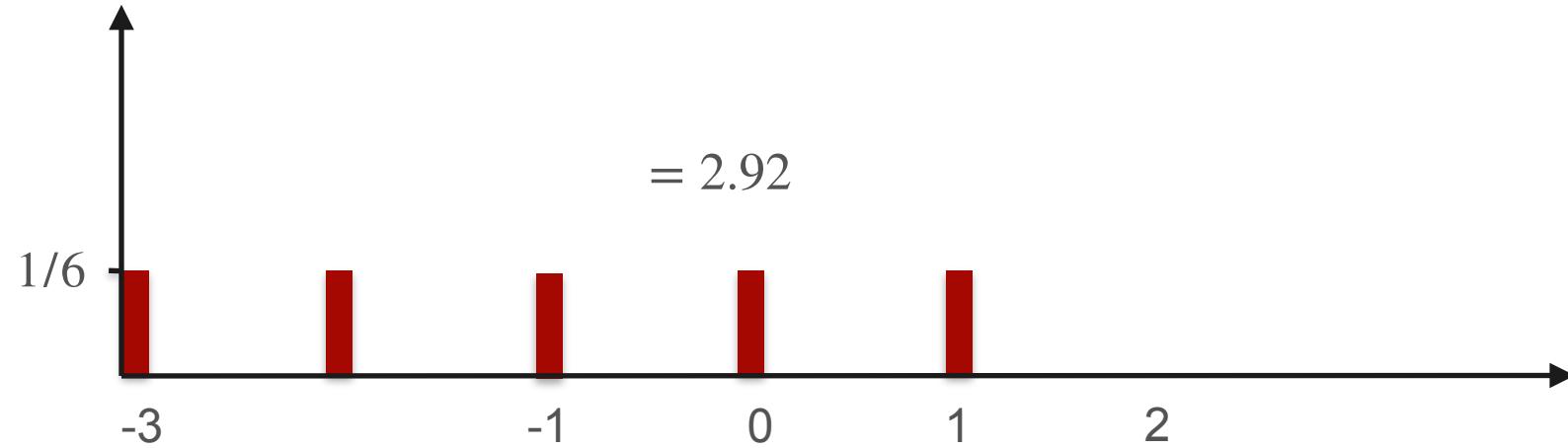
Properties of the Variance

Probability



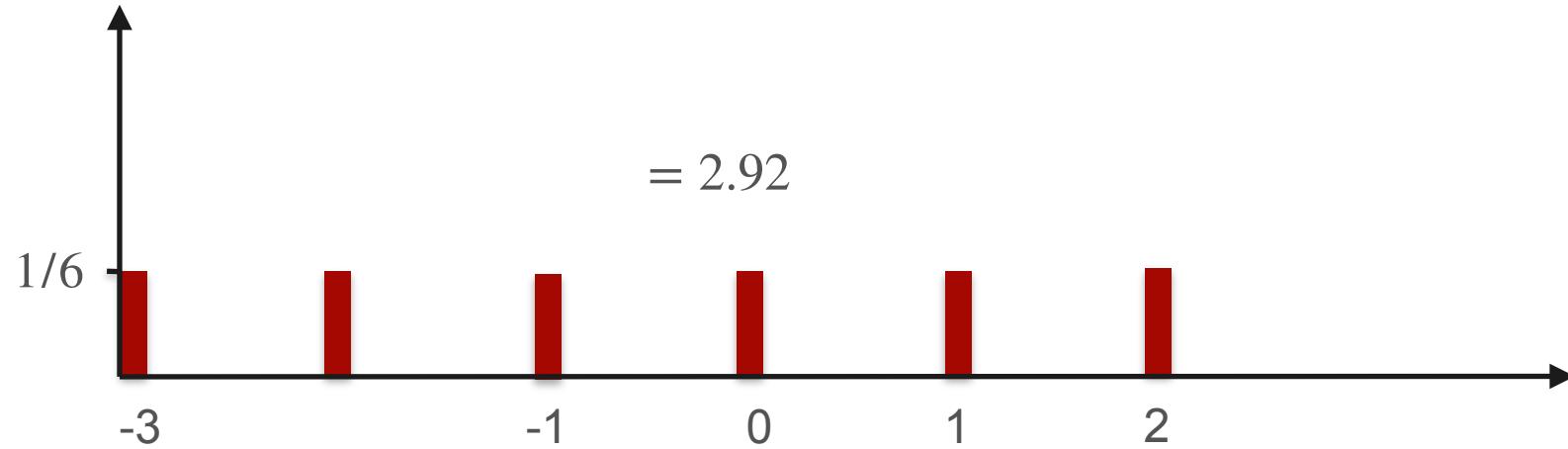
Properties of the Variance

Probability



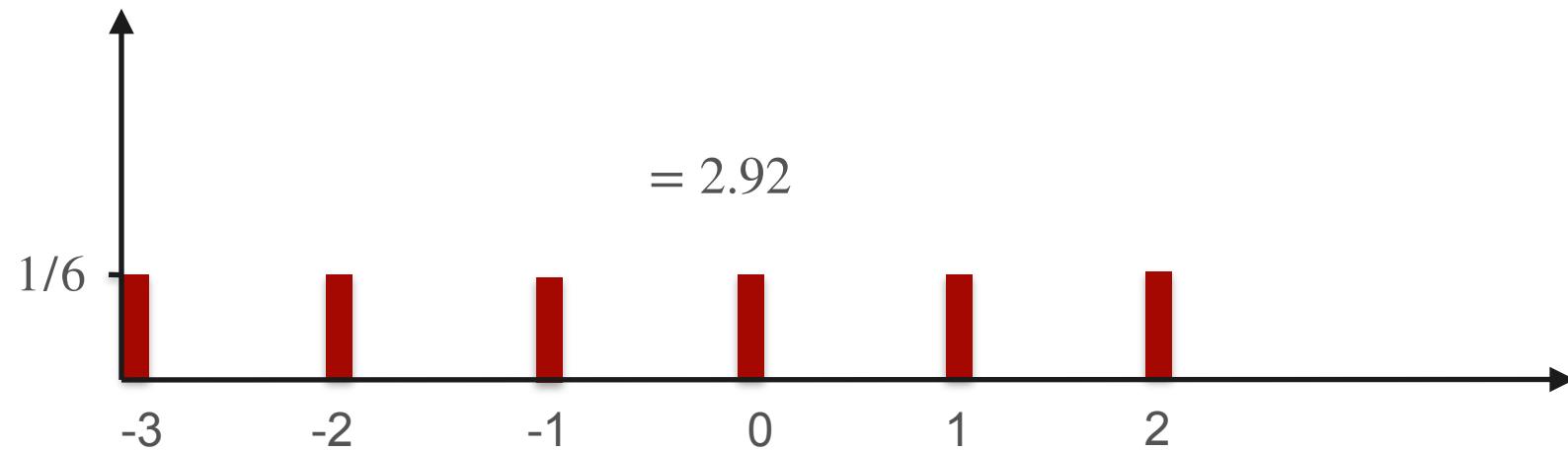
Properties of the Variance

Probability



Properties of the Variance

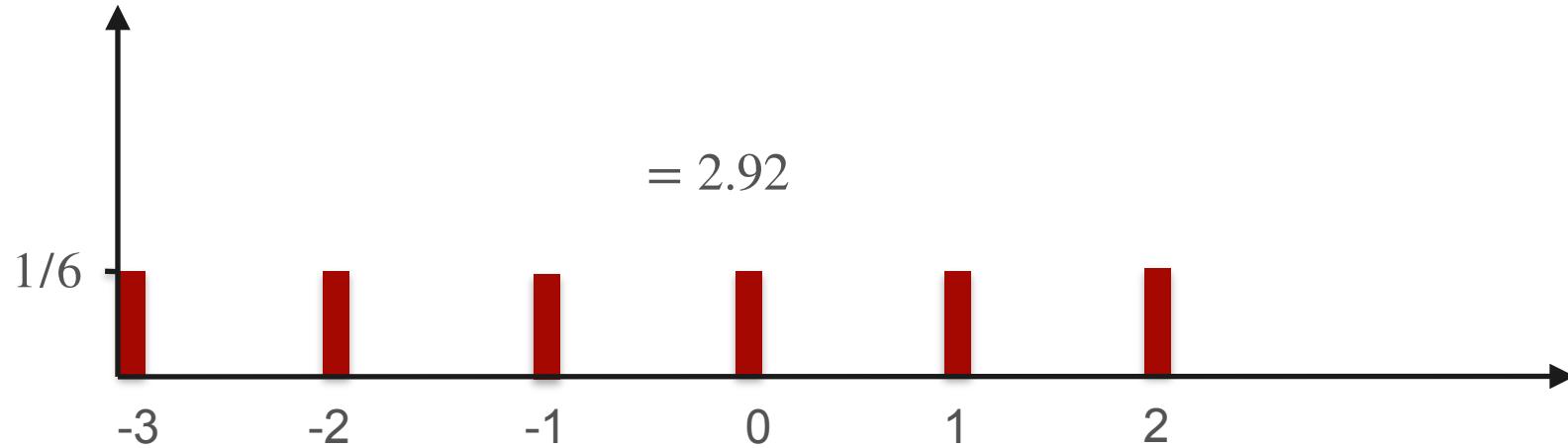
Probability



Properties of the Variance

$$\text{Var}(2X - 5) = \mathbb{E} \left[(2X - 5 - \mathbb{E}[2X - 5])^2 \right]$$

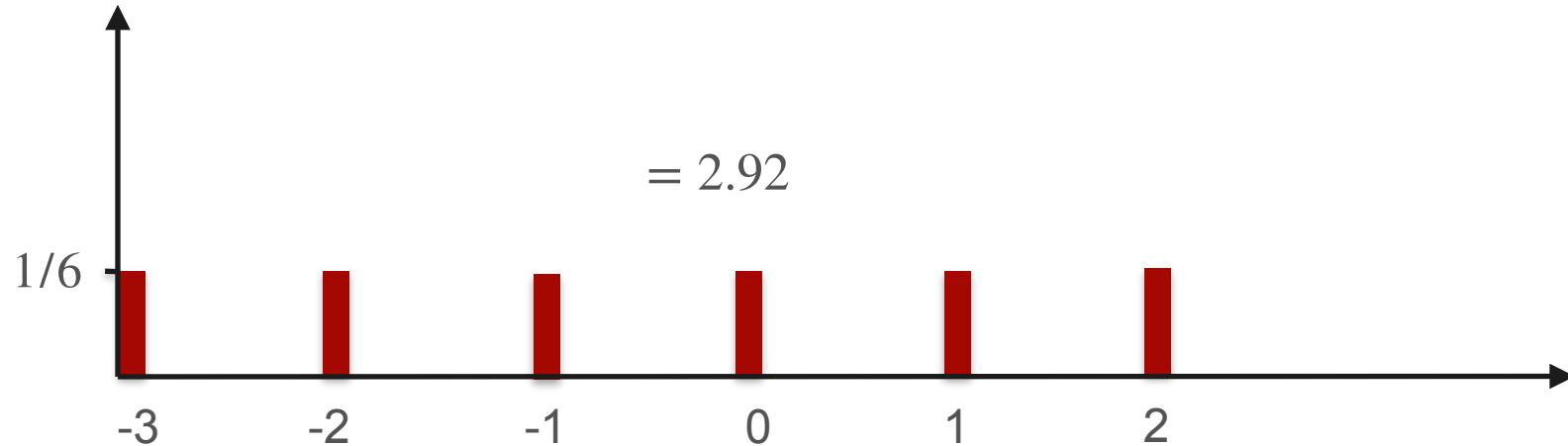
Probability



Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= \mathbb{E} \left[(2X - 5 - \mathbb{E}[2X - 5])^2 \right] \\ &= \mathbb{E} \left[(2X - 5)^2 \right] - \mathbb{E}[2X - 5]^2 \end{aligned}$$

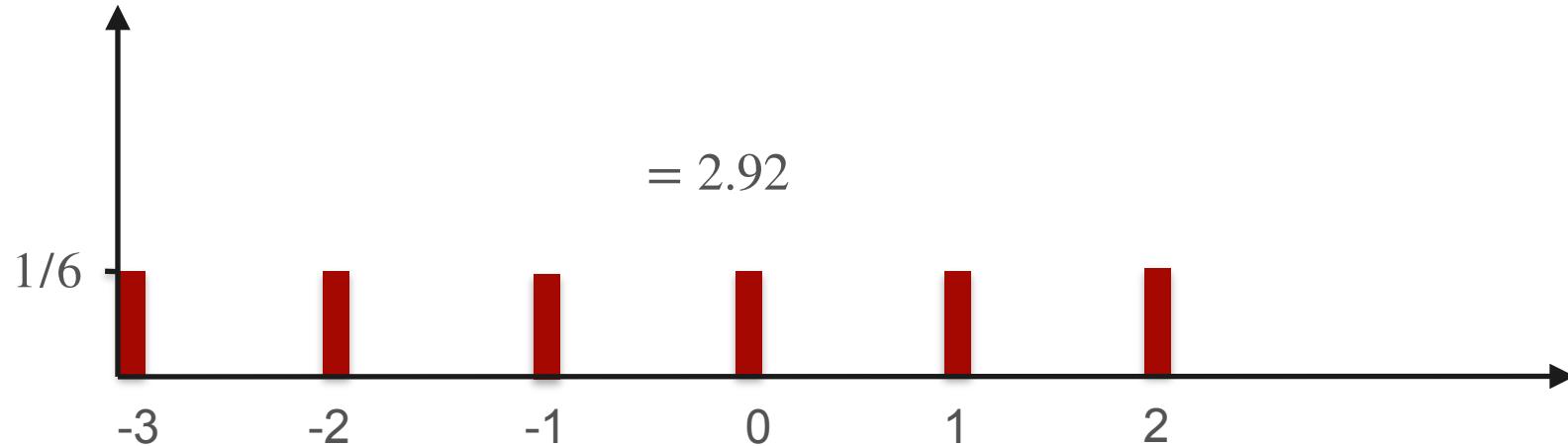
Probability



Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= \mathbb{E} \left[(2X - 5 - \mathbb{E}[2X - 5])^2 \right] \\ &= \mathbb{E} \left[(2X - 5)^2 \right] - \boxed{\mathbb{E}[2X - 5]^2} \end{aligned}$$

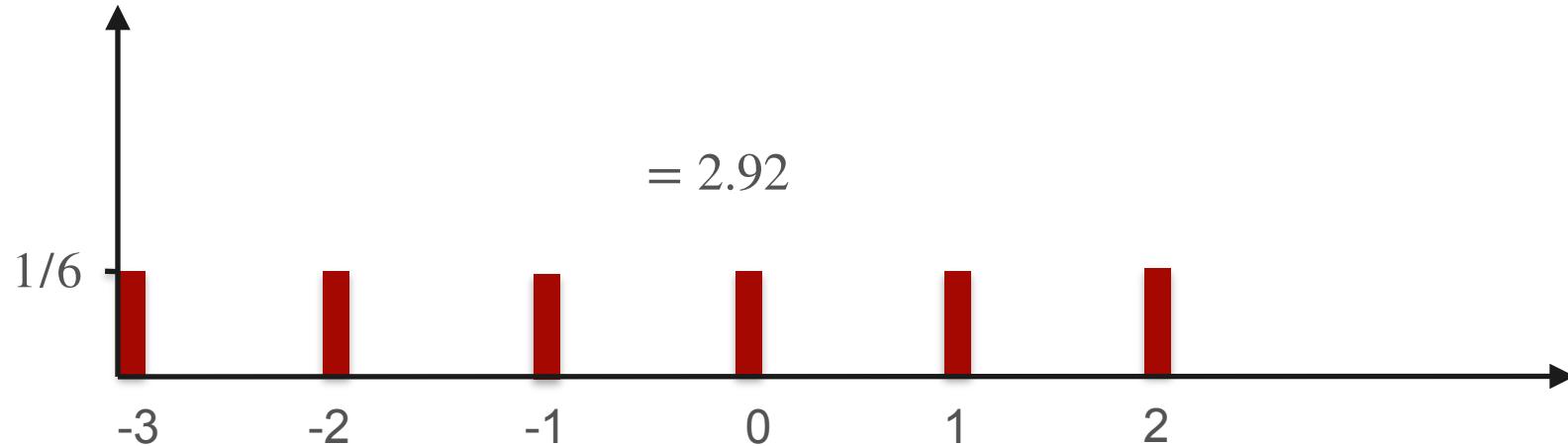
Probability



Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= \mathbb{E} \left[(2X - 5 - \mathbb{E}[2X - 5])^2 \right] \\ &= \mathbb{E} \left[(2X - 5)^2 \right] - \boxed{\mathbb{E}[2X - 5]^2} \end{aligned}$$

Probability



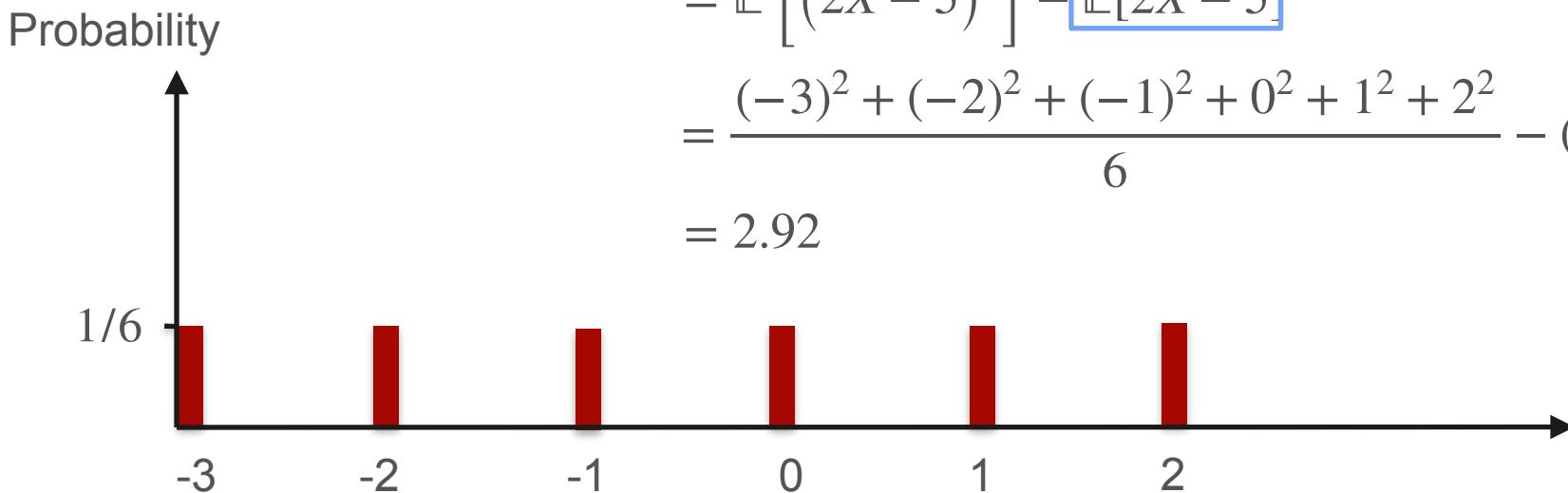
Properties of the Variance

$$Var(2X - 5) = \mathbb{E} \left[(2X - 5 - \mathbb{E}[2X - 5])^2 \right]$$

$$= \mathbb{E} \left[(2X - 5)^2 \right] - \boxed{\mathbb{E}[2X - 5]}^2$$

$$= \frac{(-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2}{6} - (-0.5)^2$$

$$= 2.92$$



Properties of the Variance

Properties of the Variance

$$\mathbb{E} [(2X - 5)^2]$$

Properties of the Variance

$$\mathbb{E} [(2X - 5)^2] = \mathbb{E} [2^2 X^2 + (-5)^2 - 2(-5)X]$$

Properties of the Variance

$$\begin{aligned}\mathbb{E}[(2X - 5)^2] &= \mathbb{E}[2^2X^2 + (-5)^2 - 2(-5)X] \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X]\end{aligned}$$

Properties of the Variance

$$\begin{aligned}\mathbb{E}[(2X - 5)^2] &= \mathbb{E}[2^2X^2 + (-5)^2 - 2(-5)X] \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X]\end{aligned}$$

$$Var(2X - 5) = \mathbb{E}[(2X - 5)^2] - \mathbb{E}[2X - 5]^2$$

Properties of the Variance

$$\begin{aligned}\mathbb{E}[(2X - 5)^2] &= \mathbb{E}[2^2X^2 + (-5)^2 - 2(-5)X] \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X]\end{aligned}$$

$$Var(2X - 5) = \mathbb{E}[(2X - 5)^2] - \boxed{\mathbb{E}[2X - 5]^2}$$

2 $\mathbb{E}[X]$ - 5



Properties of the Variance

$$\begin{aligned}\mathbb{E}[(2X - 5)^2] &= \mathbb{E}[2^2X^2 + (-5)^2 - 2(-5)X] \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X]\end{aligned}$$

$$\begin{aligned}Var(2X - 5) &= \mathbb{E}[(2X - 5)^2] - \boxed{\mathbb{E}[2X - 5]^2} \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2\mathbb{E}[X] - 5)^2\end{aligned}$$



Properties of the Variance

$$\begin{aligned}\mathbb{E}[(2X - 5)^2] &= \mathbb{E}[2^2X^2 + (-5)^2 - 2(-5)X] \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X]\end{aligned}$$

$$\begin{aligned}Var(2X - 5) &= \mathbb{E}[(2X - 5)^2] - \boxed{\mathbb{E}[2X - 5]^2} \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2\mathbb{E}[X] - 5)^2 \\ &= 2^2\mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2^2\mathbb{E}[X]^2 + (-5)^2 - 2(-5)\mathbb{E}[X])\end{aligned}$$

$2\mathbb{E}[X] - 5$


Properties of the Variance

$$Var(2X - 5) = \boxed{2^2\mathbb{E}[X^2]} + (-5)^2 - 2(-5)\mathbb{E}[X] - (\boxed{2^2\mathbb{E}[X]^2} + (-5)^2 - 2(-5)\mathbb{E}[X])$$

Properties of the Variance

$$\text{Var}(2X - 5) = \boxed{2^2\mathbb{E}[X^2]} + (-5)^2 - 2(-5)\mathbb{E}[X] - (\boxed{2^2\mathbb{E}[X]^2} + (-5)^2 - 2(-5)\mathbb{E}[X])$$

Properties of the Variance

$$\text{Var}(2X - 5) = \boxed{2^2 \mathbb{E}[X^2]} + (-5)^2 - 2(-5)\mathbb{E}[X] - (\boxed{2^2 \mathbb{E}[X]^2} + (-5)^2 - 2(-5)\mathbb{E}[X])$$

Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= 2^2 \mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2^2 \mathbb{E}[X]^2 + (-5)^2 - 2(-5)\mathbb{E}[X]) \\ &= 2^2 (\mathbb{E}[X^2] - \mathbb{E}[X]^2) \end{aligned}$$

Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= 2^2 \mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2^2 \mathbb{E}[X]^2 + (-5)^2 - 2(-5)\mathbb{E}[X]) \\ &= 2^2 \underbrace{(\mathbb{E}[X^2] - \mathbb{E}[X]^2)}_{\text{Var}(X)} \end{aligned}$$

Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= 2^2 \mathbb{E}[X^2] + (-5)^2 - 2(-5)\mathbb{E}[X] - (2^2 \mathbb{E}[X]^2 + (-5)^2 - 2(-5)\mathbb{E}[X]) \\ &= 2^2 \underbrace{(\mathbb{E}[X^2] - \mathbb{E}[X]^2)}_{\text{Var}(X)} = 2^2 \text{Var}(X) \end{aligned}$$

Properties of the Variance

$$\begin{aligned} \text{Var}(2X - 5) &= \cancel{2^2 \mathbb{E}[X^2]} + (-5)^2 - 2(-5)\mathbb{E}[X] - \cancel{(2^2 \mathbb{E}[X]^2 + (-5)^2 - 2(-5)\mathbb{E}[X])} \\ &= 2^2 \underbrace{(\mathbb{E}[X^2] - \mathbb{E}[X]^2)}_{\text{Var}(X)} = 2^2 \text{Var}(X) \end{aligned}$$

In general: $\text{Var}(aX + b) = a^2 \text{Var}(X)$



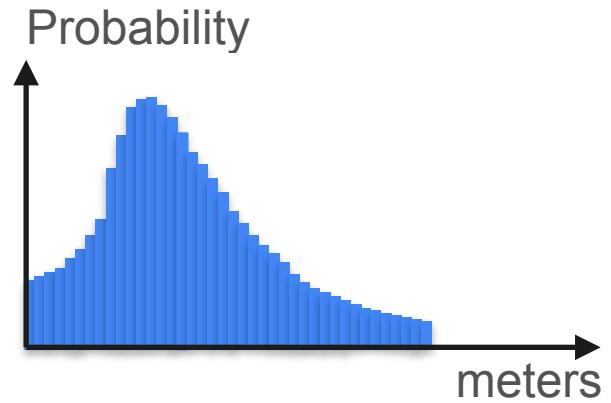
DeepLearning.AI

Describing Distributions

Standard deviation

Standard Deviation

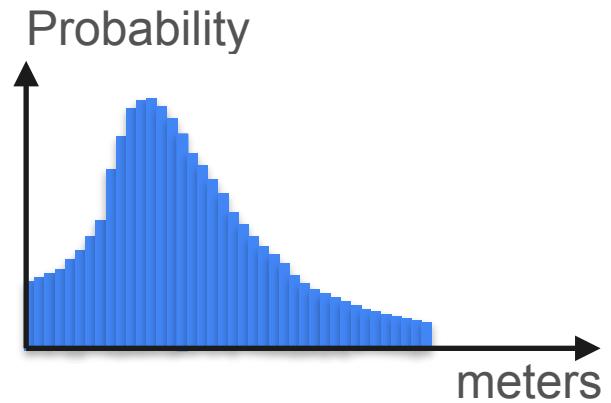
$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$



Standard Deviation

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

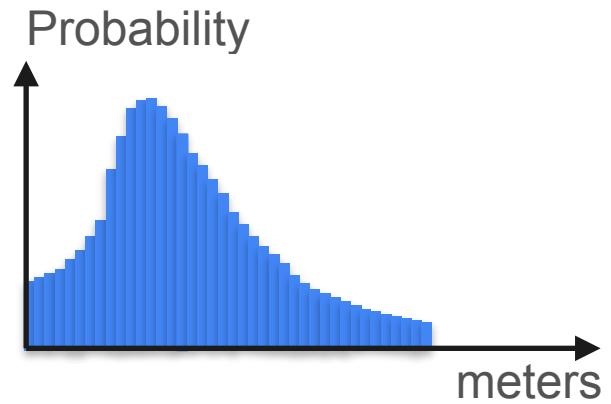


Standard Deviation

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.

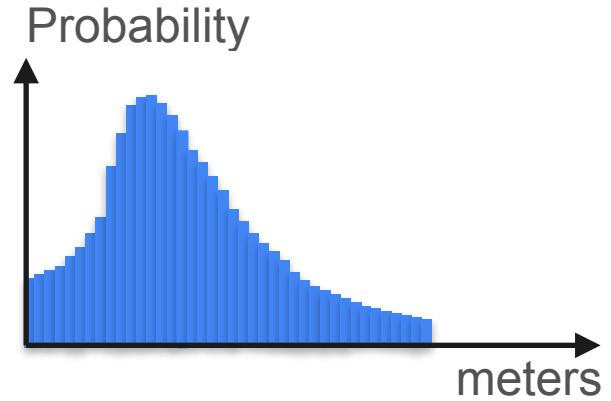


Standard Deviation

$$Var(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.

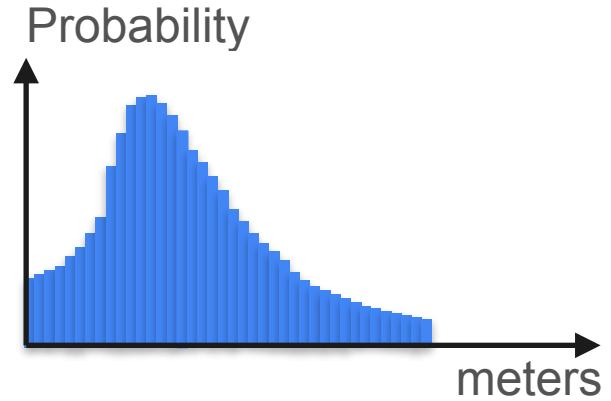


Standard Deviation

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.



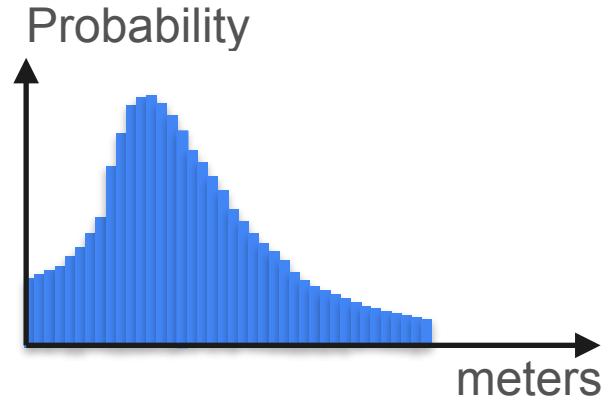
Standard Deviation

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.

Then $\text{Var}(X)$ is measured in meters².



Standard Deviation

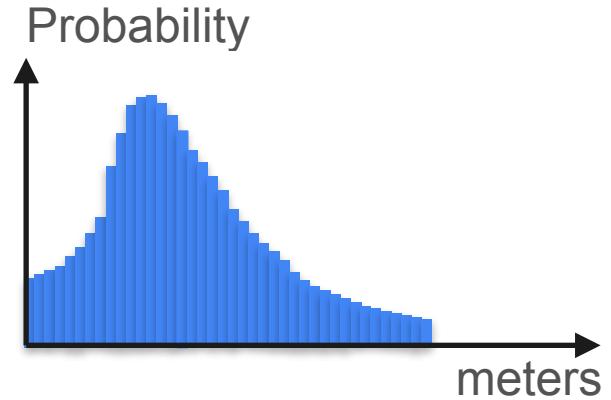
$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.

Then $\text{Var}(X)$ is measured in meters².

Then $\sqrt{\text{Var}(X)}$ is measured in meters.



Standard Deviation

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

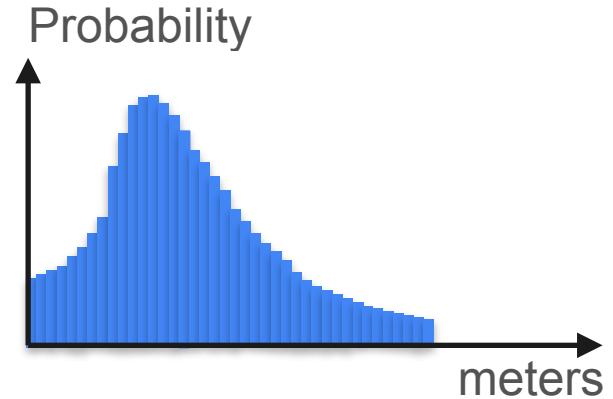
Say X is measured in meters.

Then $\mathbb{E}[X]$ is measured in meters.

Then $\text{Var}(X)$ is measured in meters².

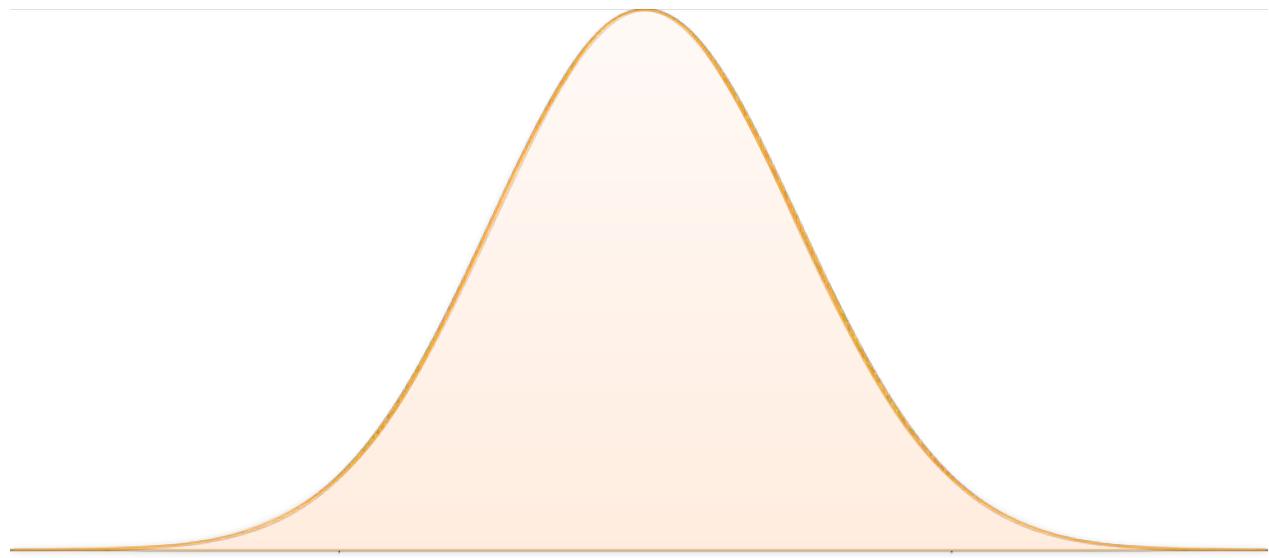
Then $\sqrt{\text{Var}(X)}$ is measured in meters.

Let's call $std(X) = \sqrt{\text{Var}(X)}$, the *standard deviation* of X



Normal Distribution: 68-95-99.7 Rule

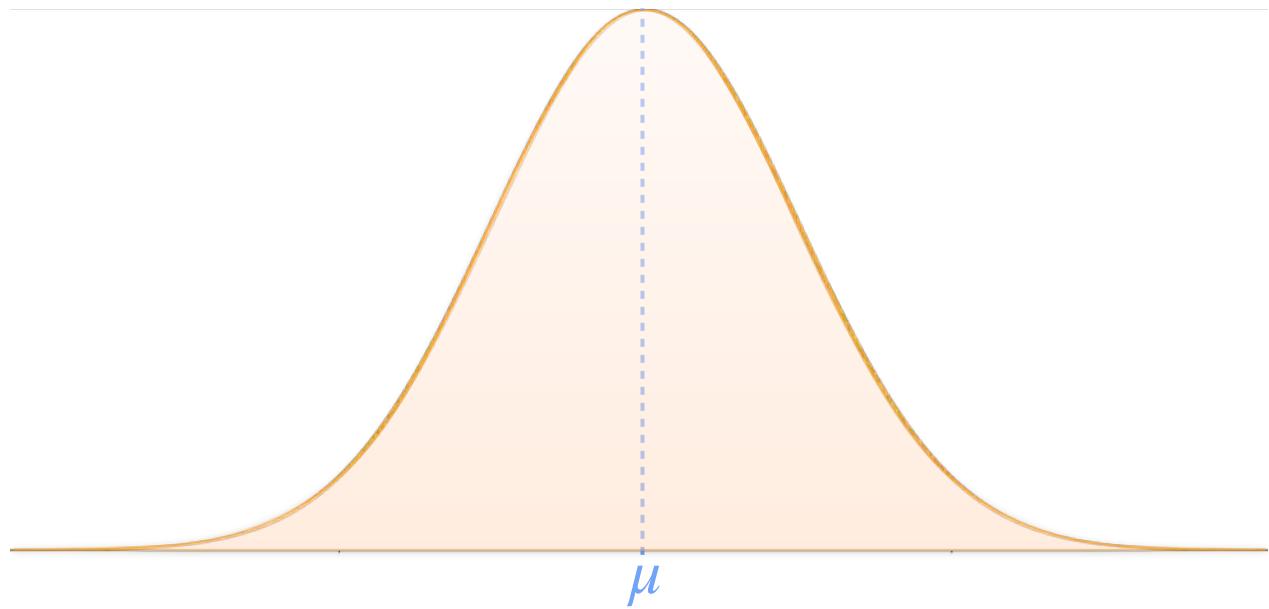
Normal Distribution: 68-95-99.7 Rule



Normal Distribution: 68-95-99.7 Rule

Parameters:

- μ : center of the bell

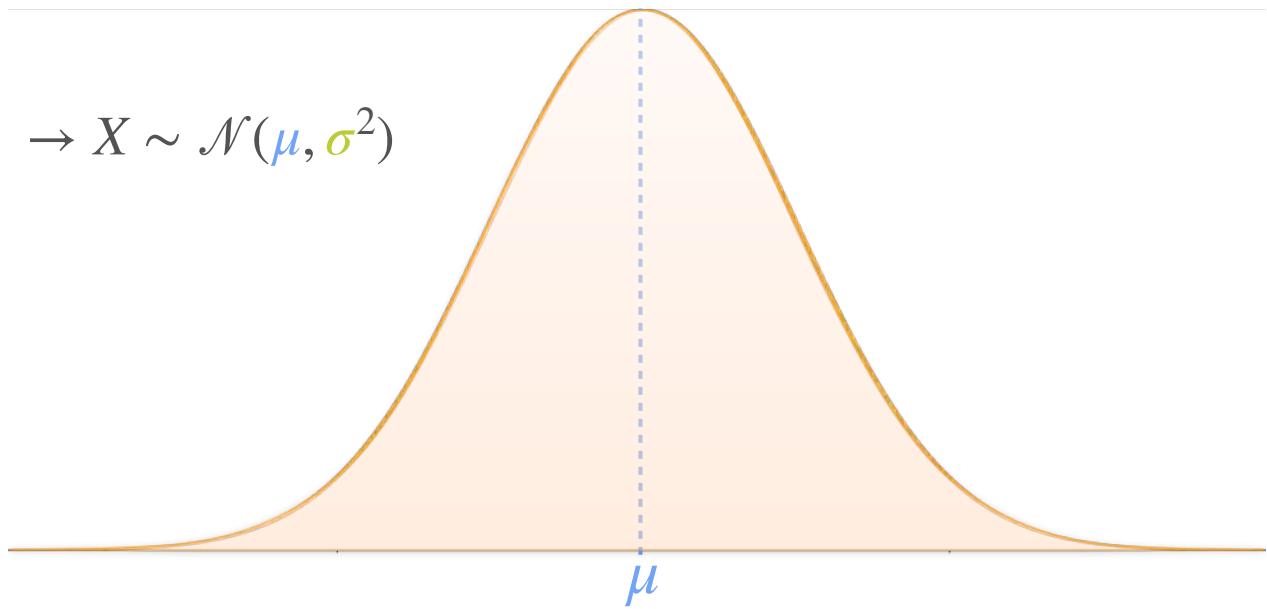


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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



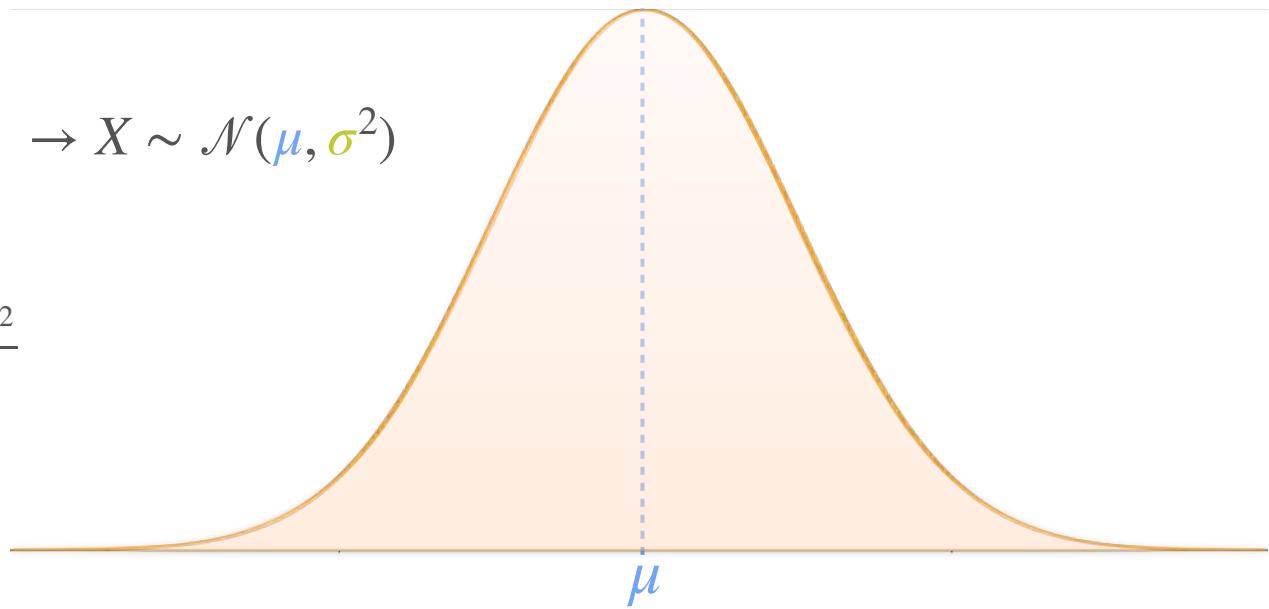
Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

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



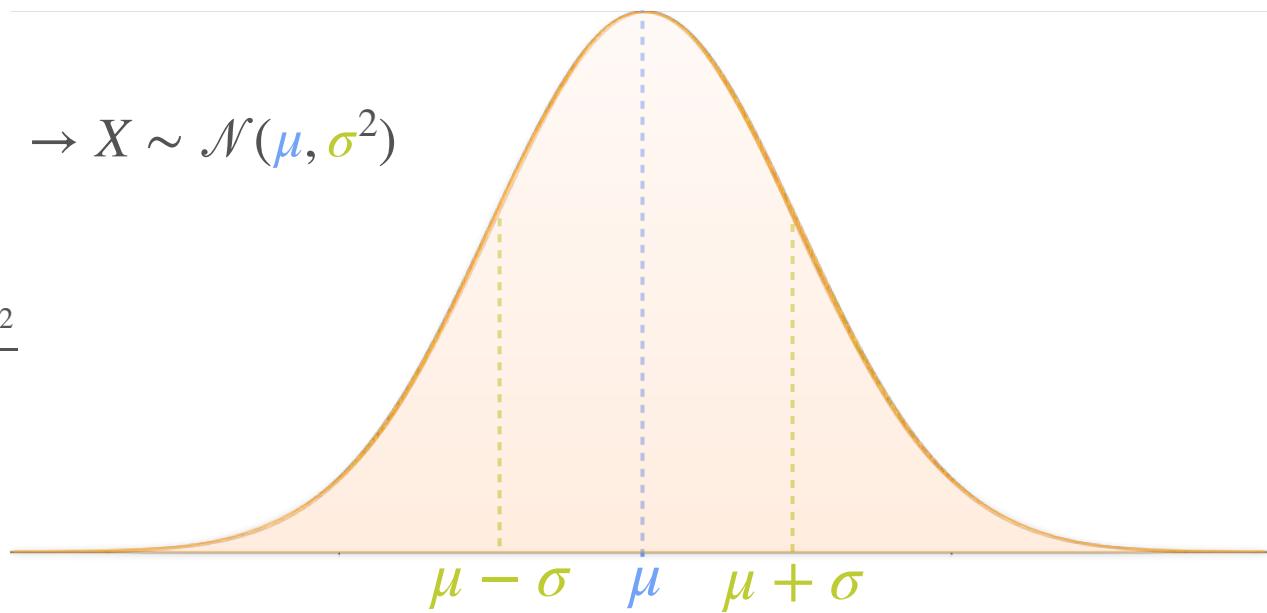
Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

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



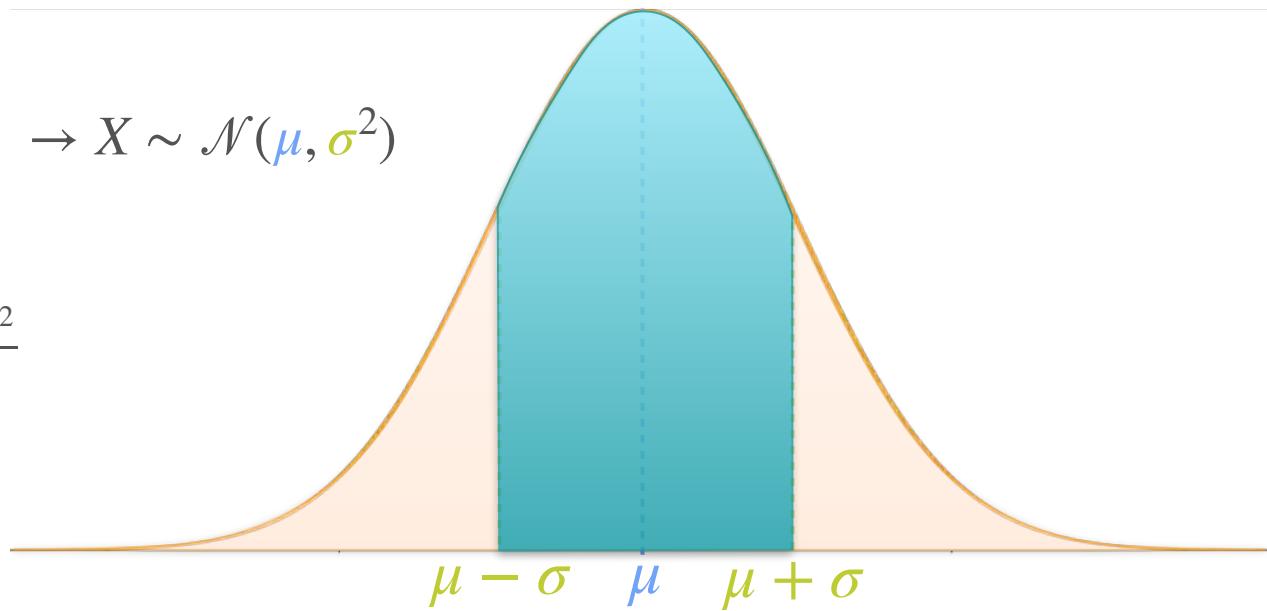
Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

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



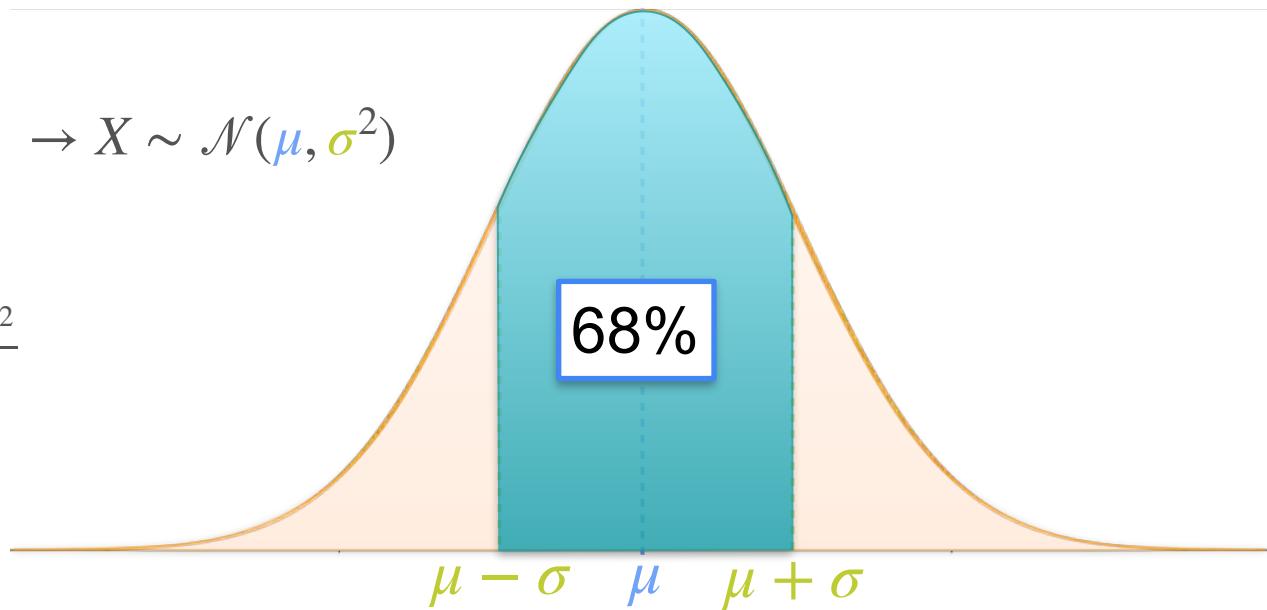
Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

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

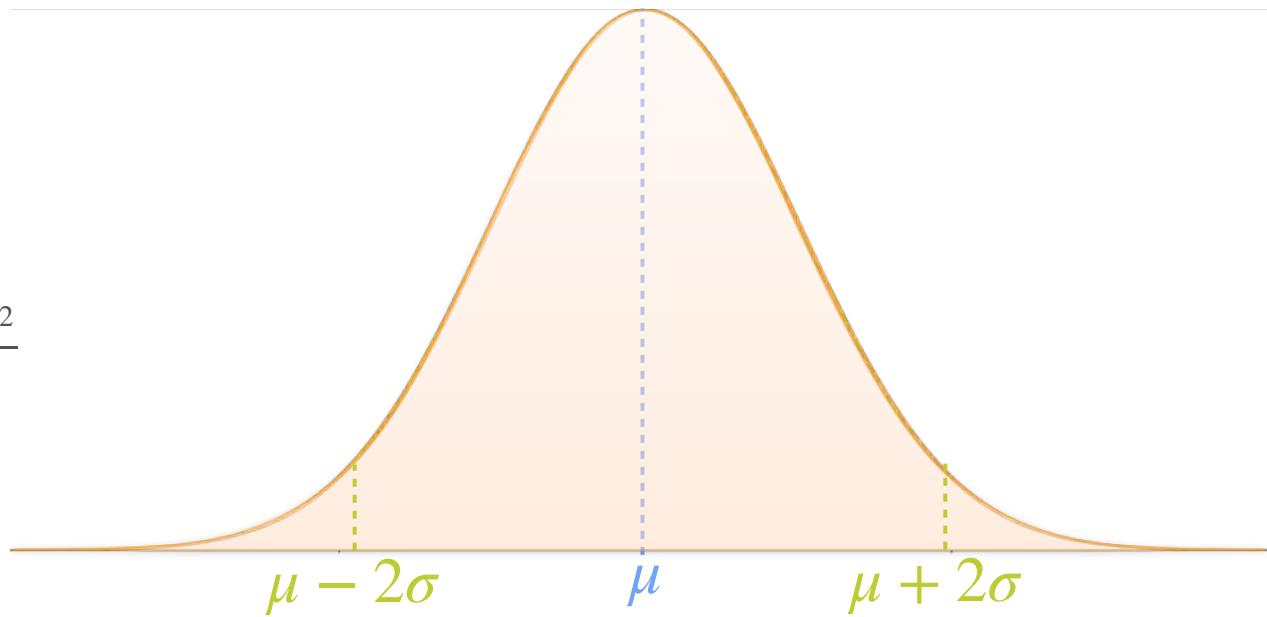


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

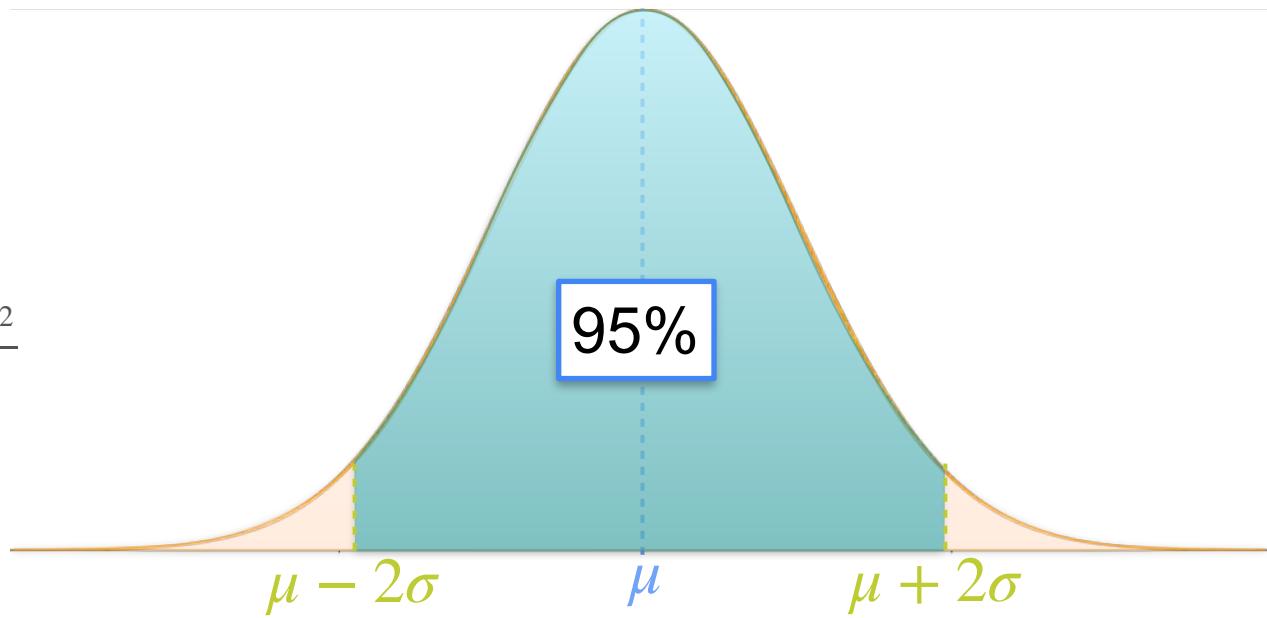


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

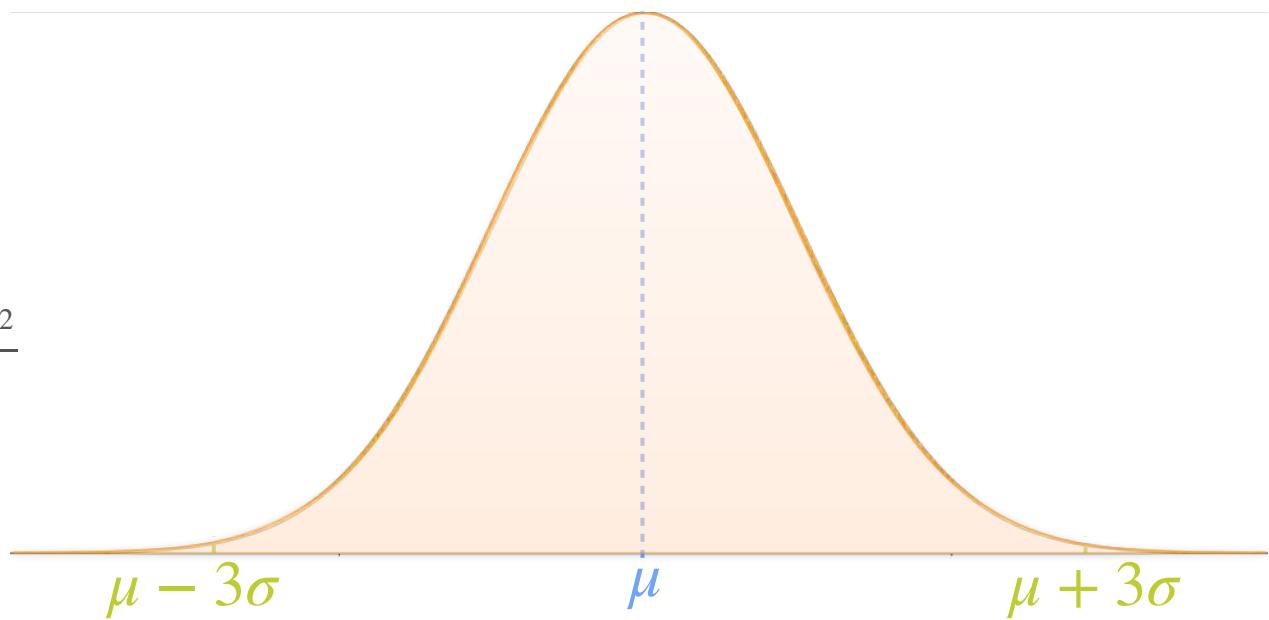


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

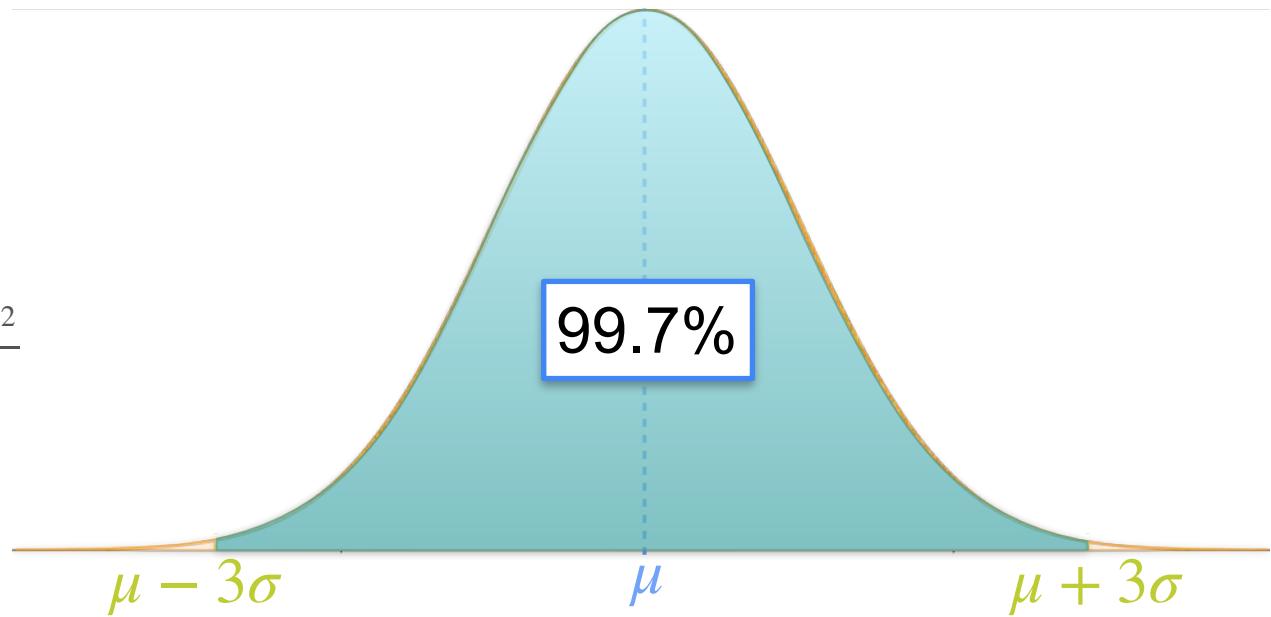


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

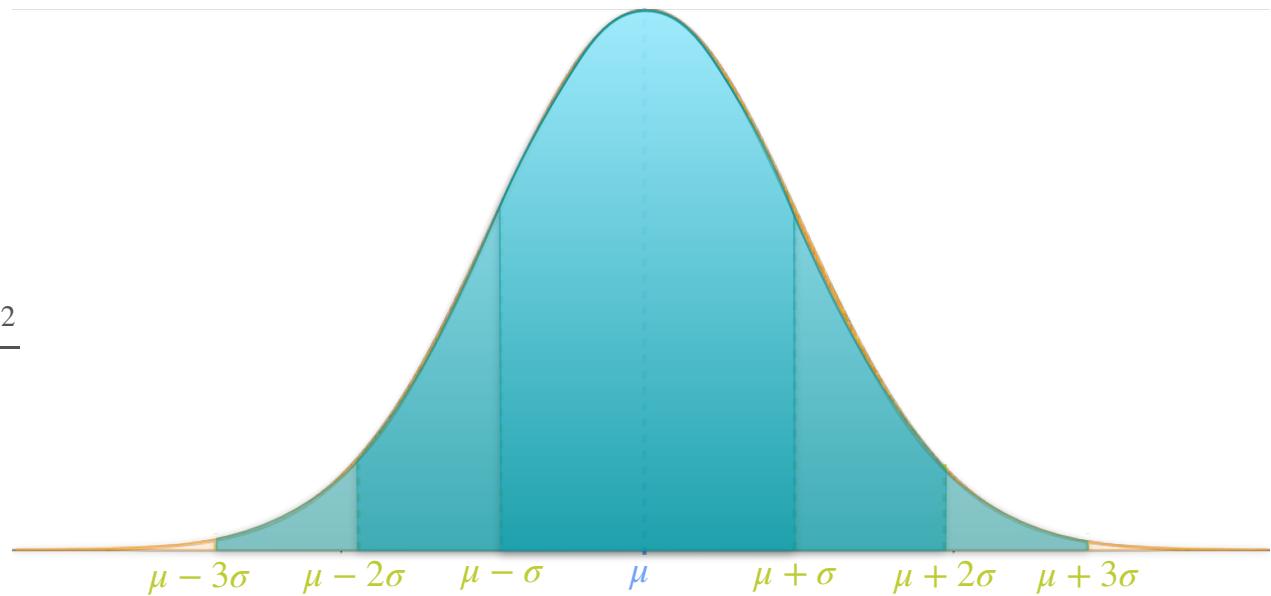


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

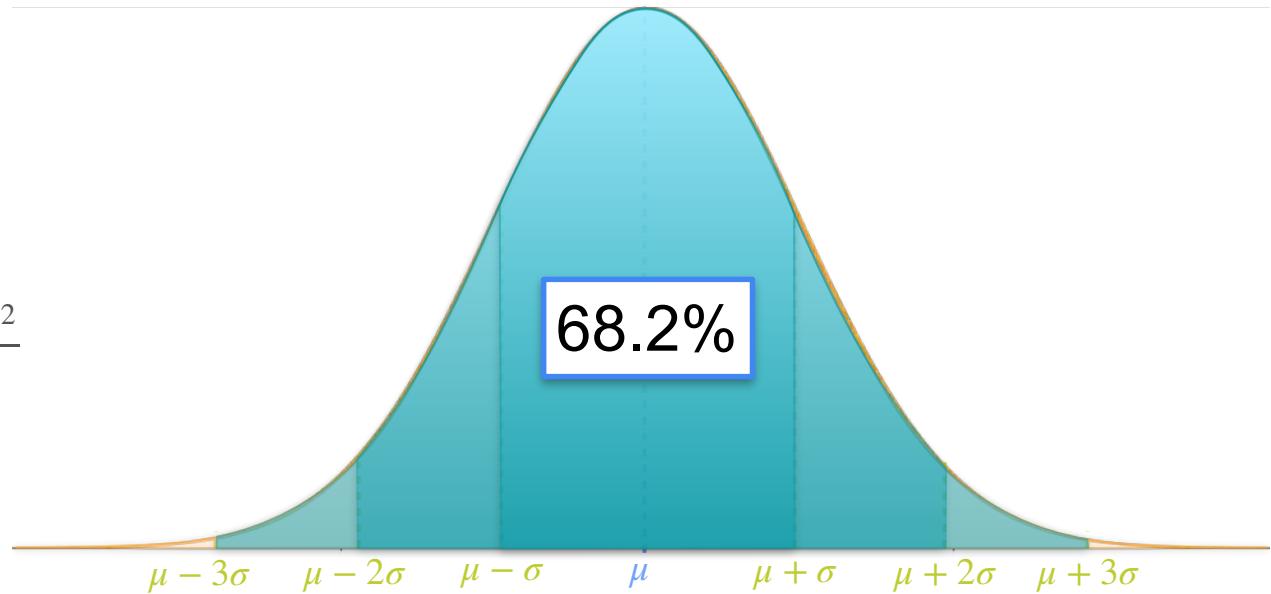


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

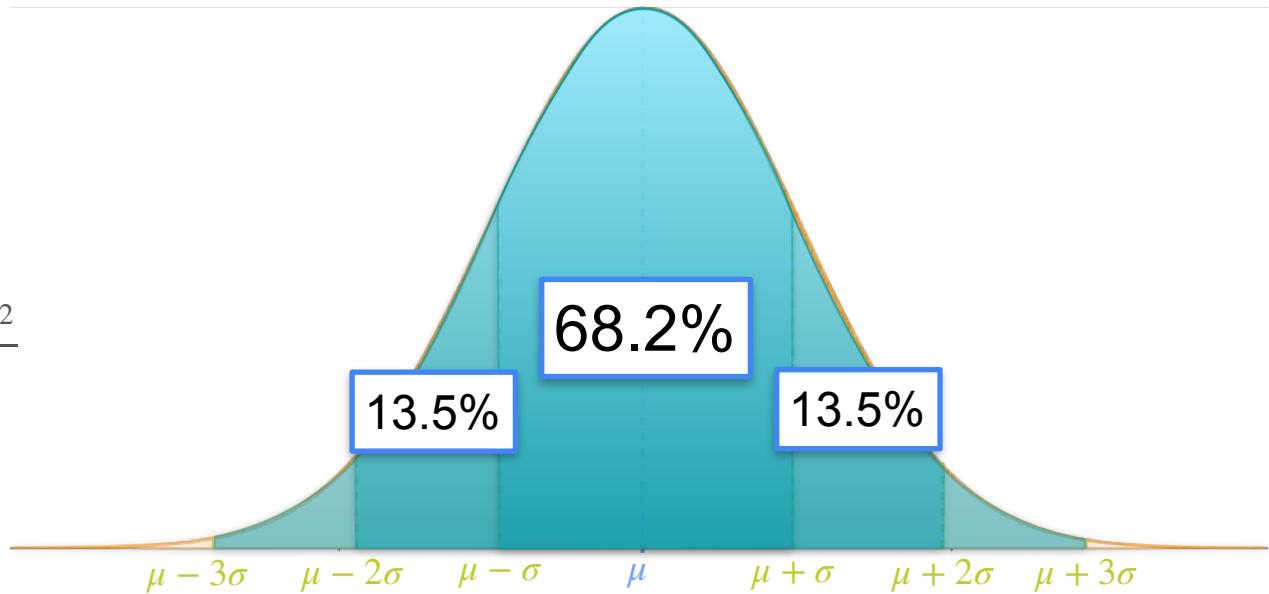


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

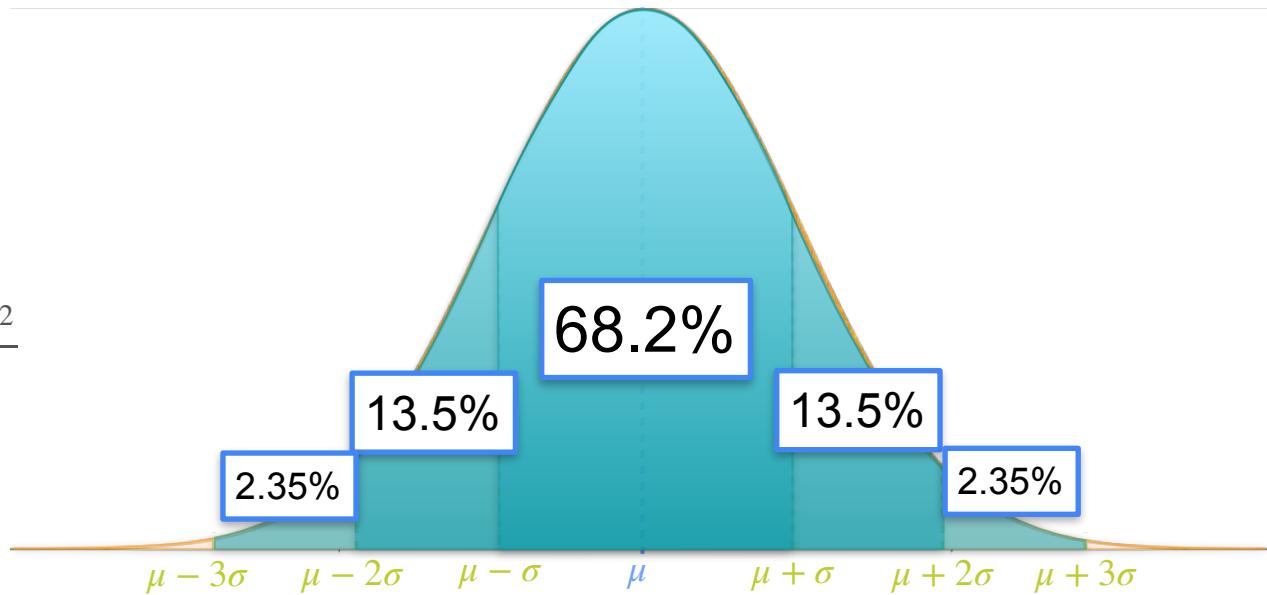


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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

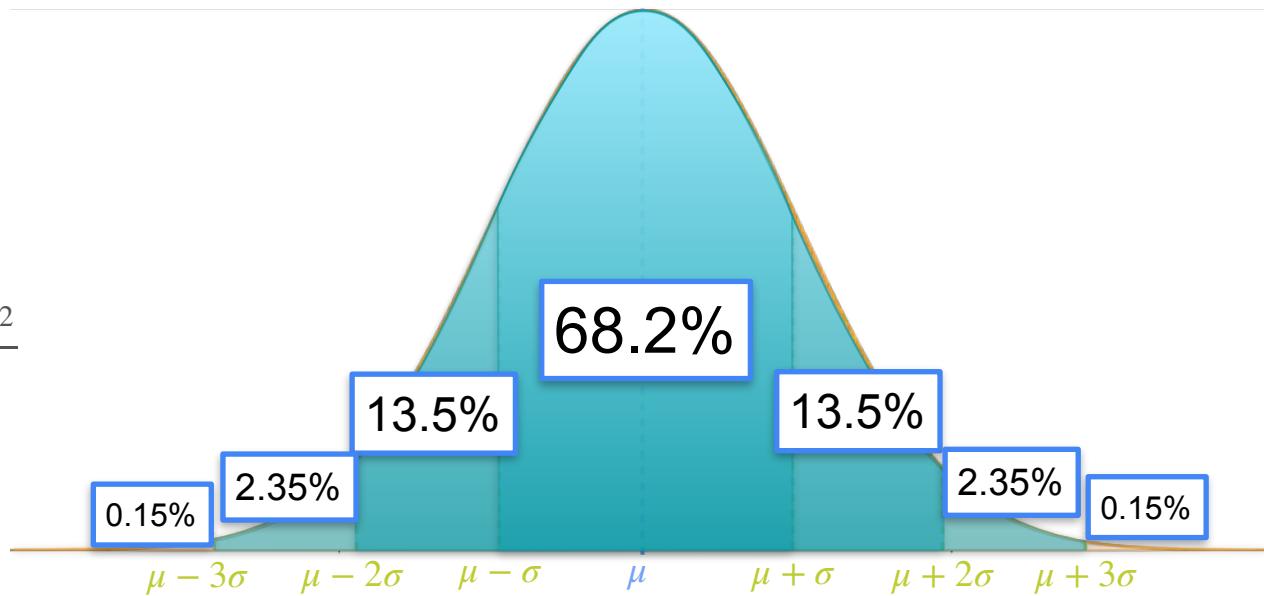


Normal Distribution: 68-95-99.7 Rule

Parameters:

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

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



Sum of Gaussians: an Example

Sum of Gaussians: an Example

Total response time of a computer system

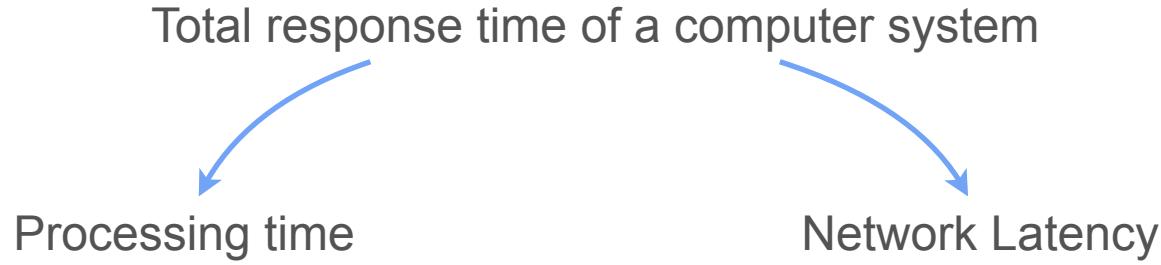
Sum of Gaussians: an Example

Total response time of a computer system



Processing time

Sum of Gaussians: an Example



Sum of Gaussians: an Example

R : Total response time of a computer system

T : Processing time

L : Network Latency



Sum of Gaussians: an Example

R : Total response time of a computer system

T : Processing time

L : Network Latency

$$R = T + L$$

Sum of Gaussians: an Example

R : Total response time of a computer system

T : Processing time

L : Network Latency

$$T \sim \mathcal{N}(10, 2^2)$$

$$R = T + L$$

Sum of Gaussians: an Example

R : Total response time of a computer system

T : Processing time

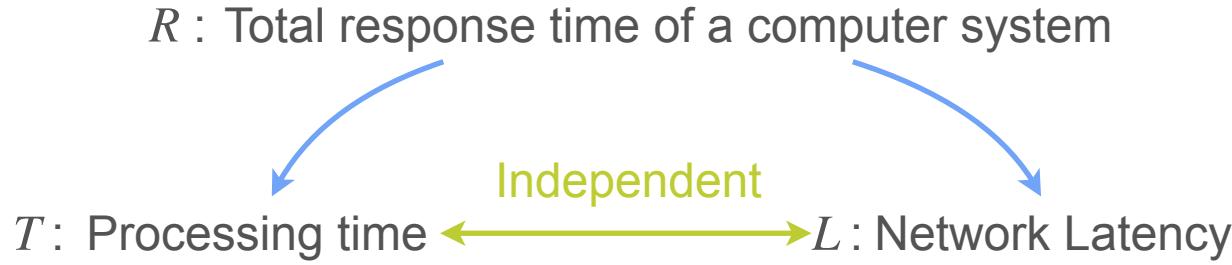
$$T \sim \mathcal{N}(10, 2^2)$$

L : Network Latency

$$L \sim \mathcal{N}(5, 1^2)$$

$$R = T + L$$

Sum of Gaussians: an Example



$$T \sim \mathcal{N}(10, 2^2)$$

$$L \sim \mathcal{N}(5, 1^2)$$

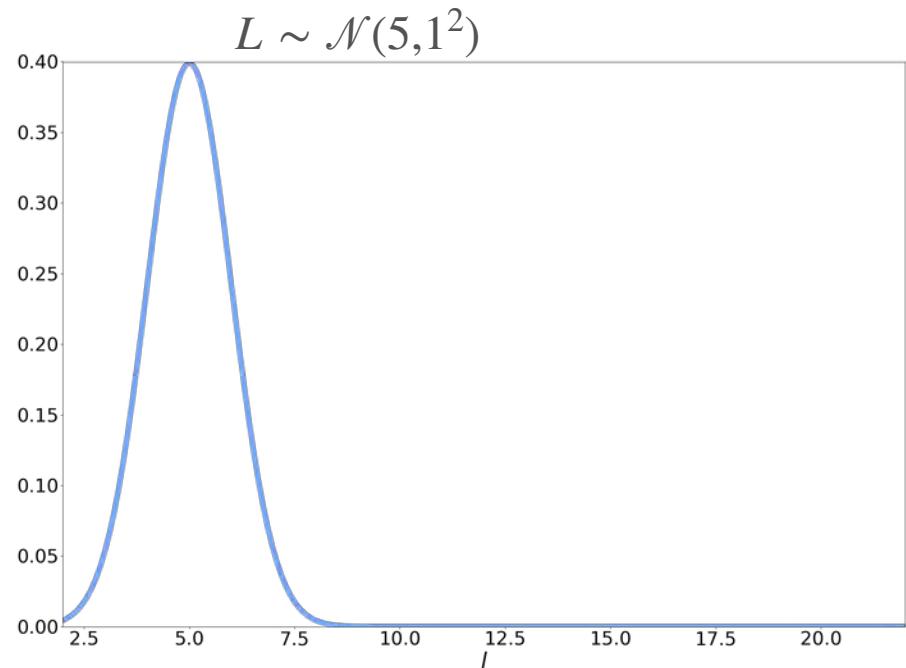
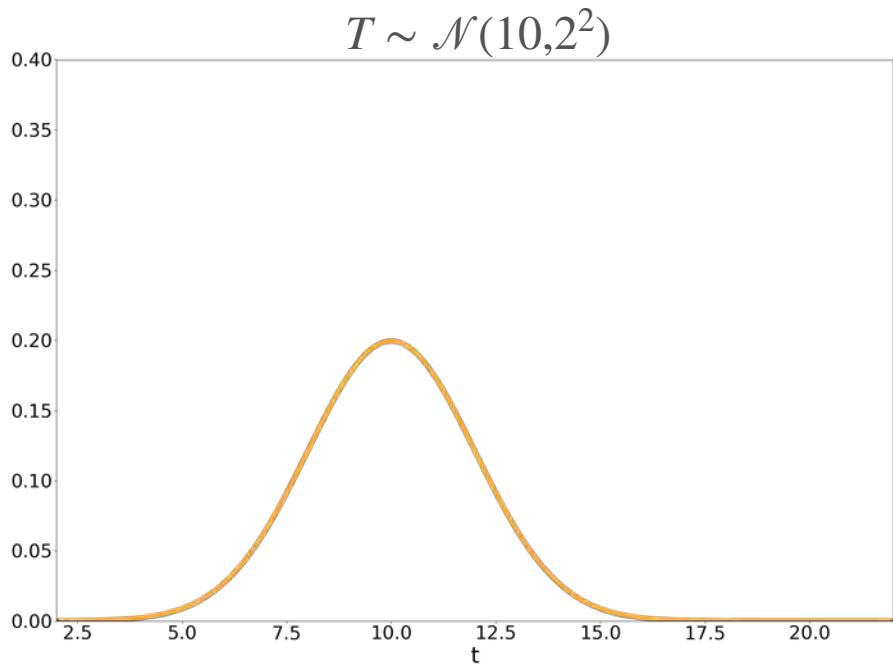
$$R = T + L$$

Sum Of Gaussians

$$T \sim \mathcal{N}(10, 2^2)$$

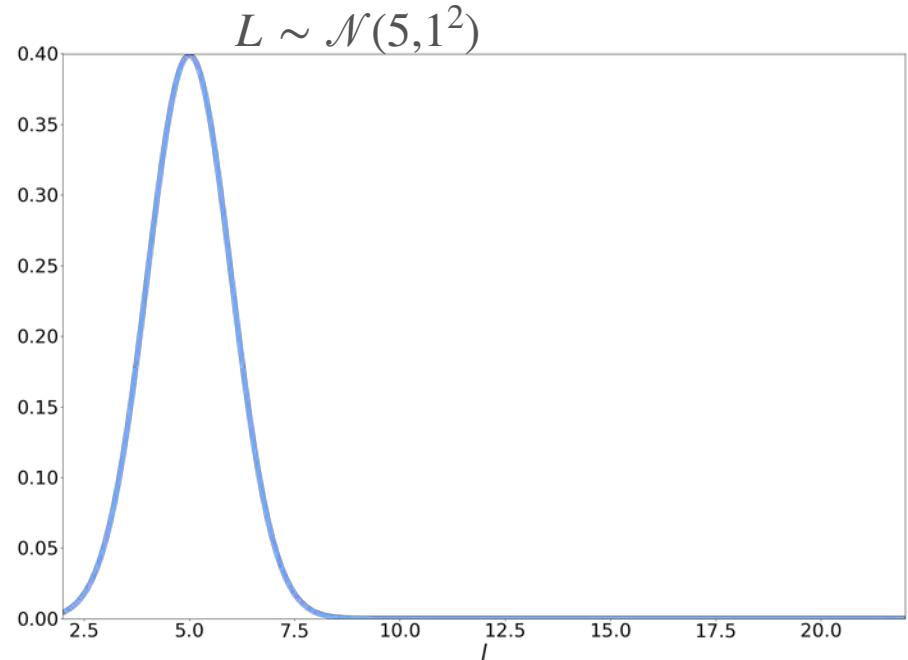
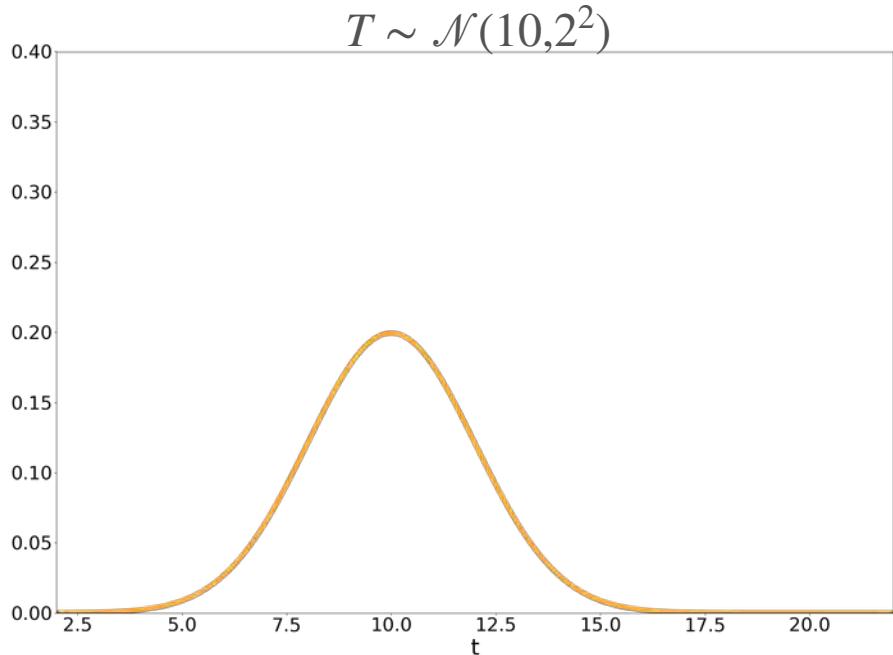
$$L \sim \mathcal{N}(5, 1^2)$$

Sum Of Gaussians



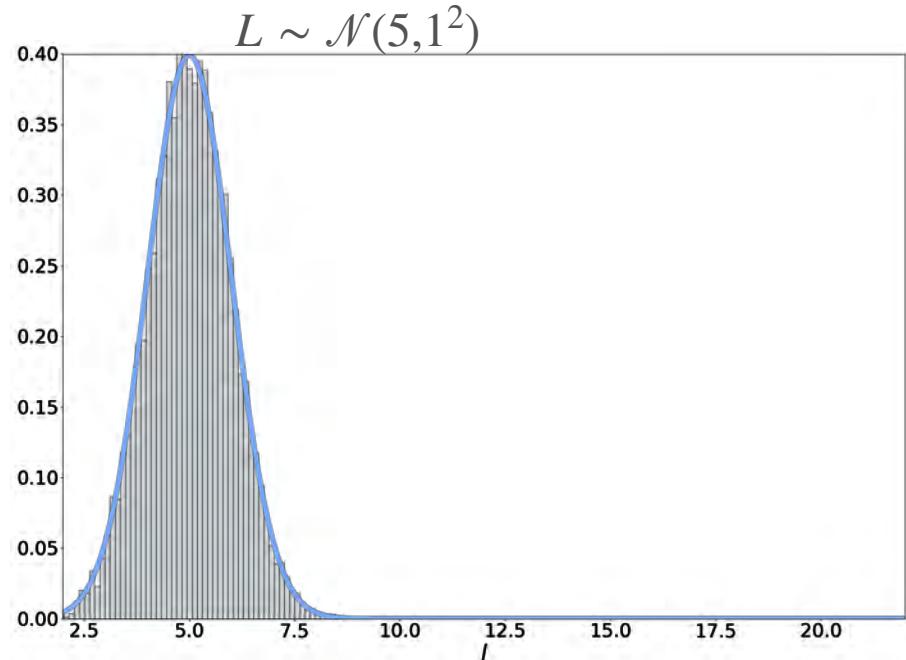
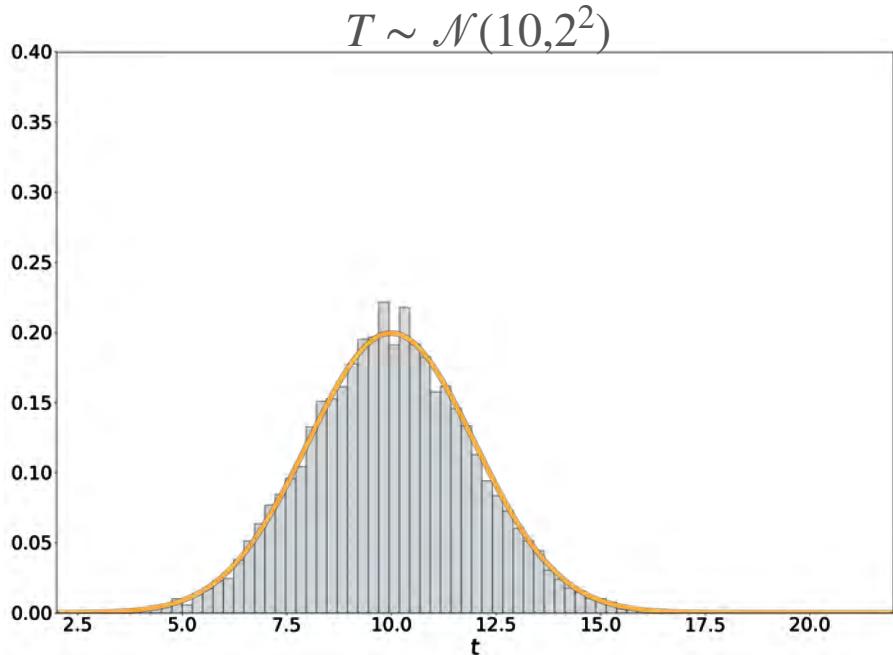
Sum Of Gaussians

Sample each variable 10000 times



Sum Of Gaussians

Sample each variable 10000 times

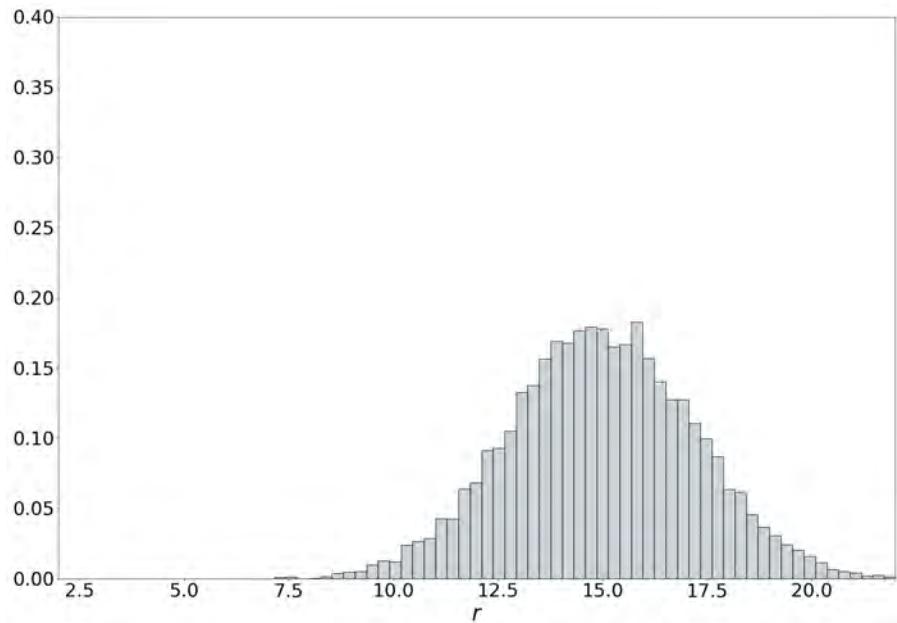


Sum Of Gaussians

$$R = T + L$$

Sum Of Gaussians

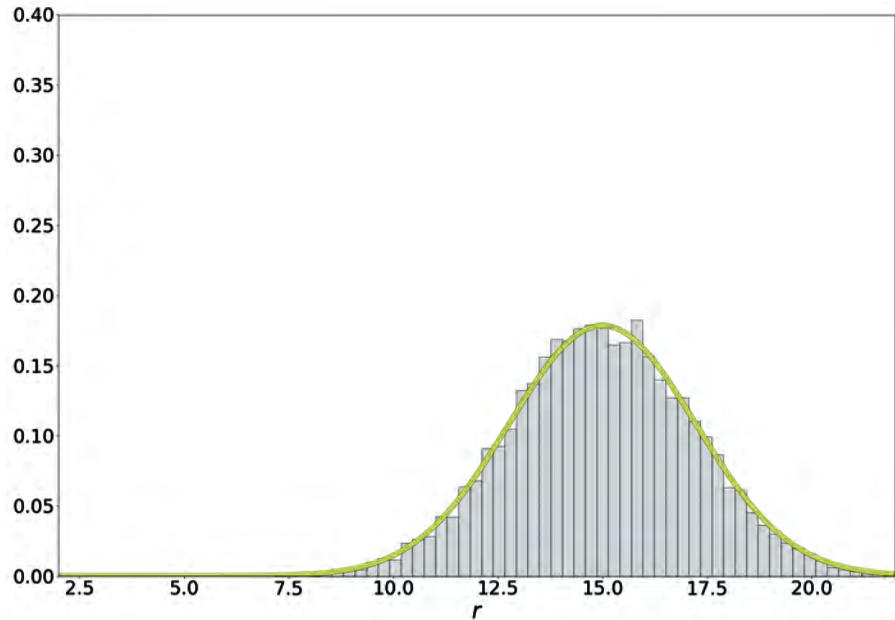
$$R = T + L$$



Sum Of Gaussians

$$R = T + L$$

R is still Gaussian!

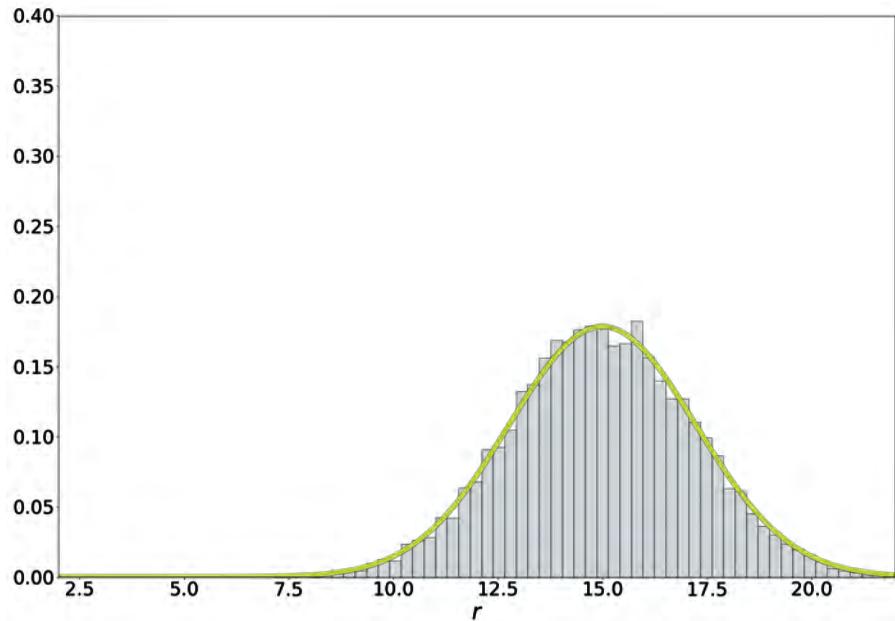


Sum Of Gaussians

$$R = T + L$$

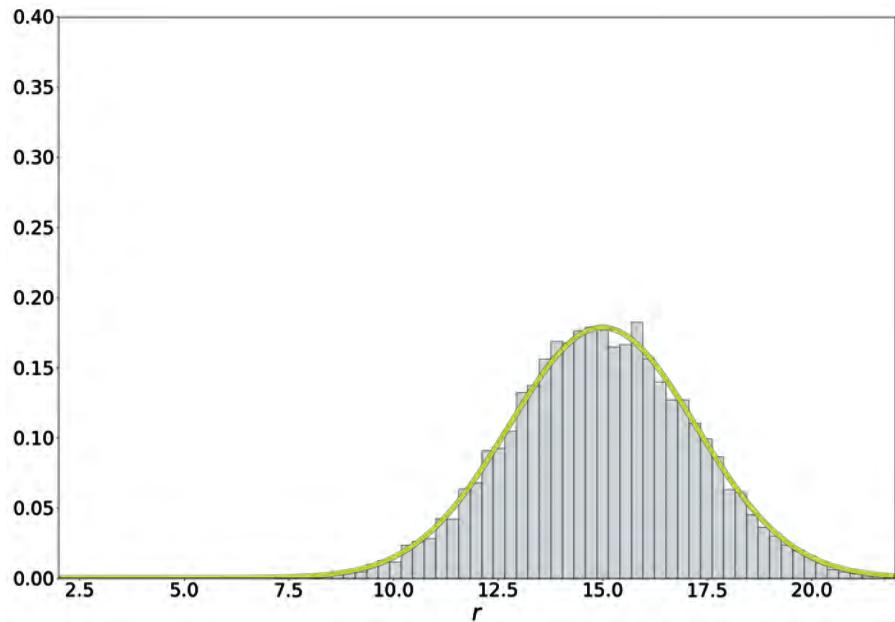
R is still Gaussian!

$$\mu_R = \mathbb{E}[R]$$



Sum Of Gaussians

$$R = T + L$$

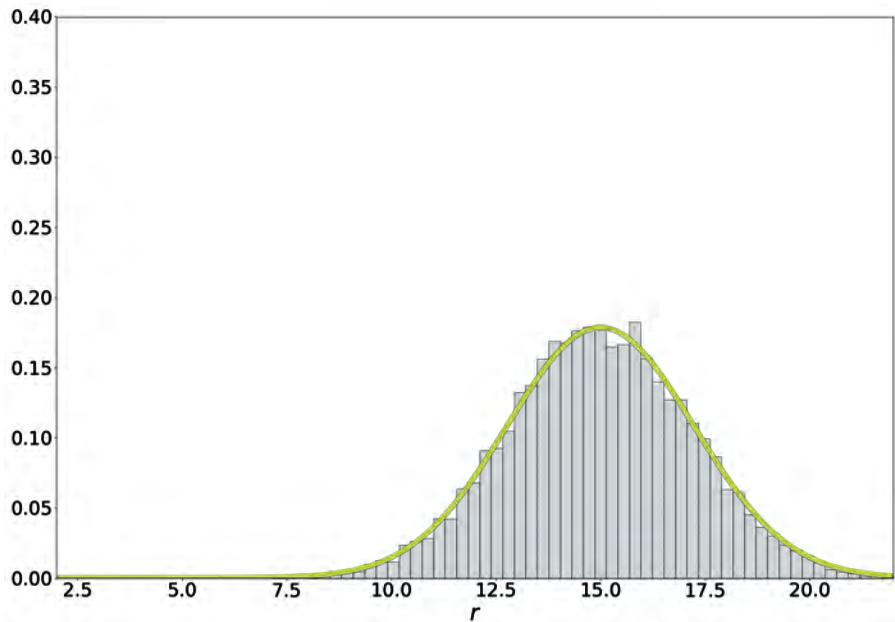


R is still Gaussian!

$$\mu_R = \mathbb{E}[R] = \mathbb{E}[T + L]$$

Sum Of Gaussians

$$R = T + L$$



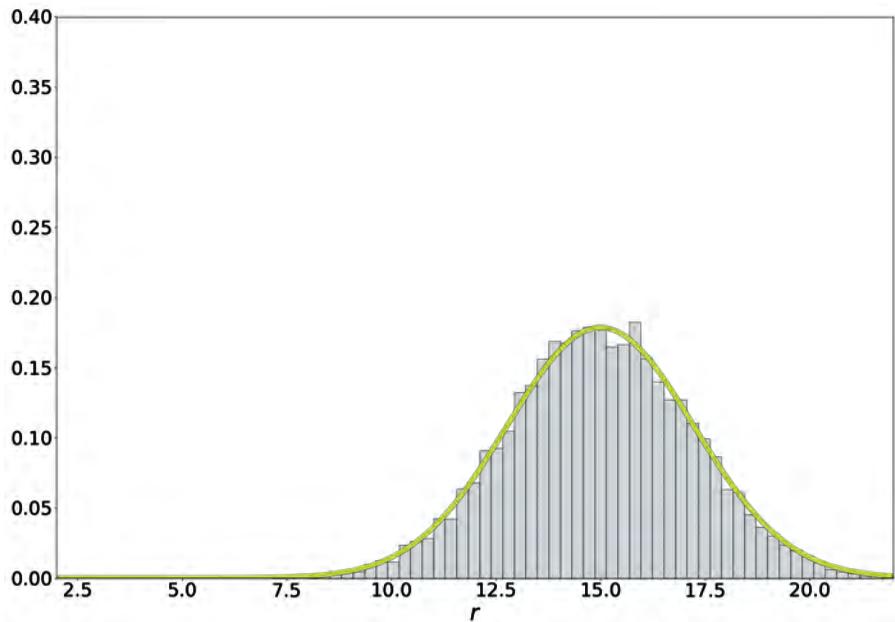
R is still Gaussian!

Expectation is linear

$$\mu_R = \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L]$$

Sum Of Gaussians

$$R = T + L$$



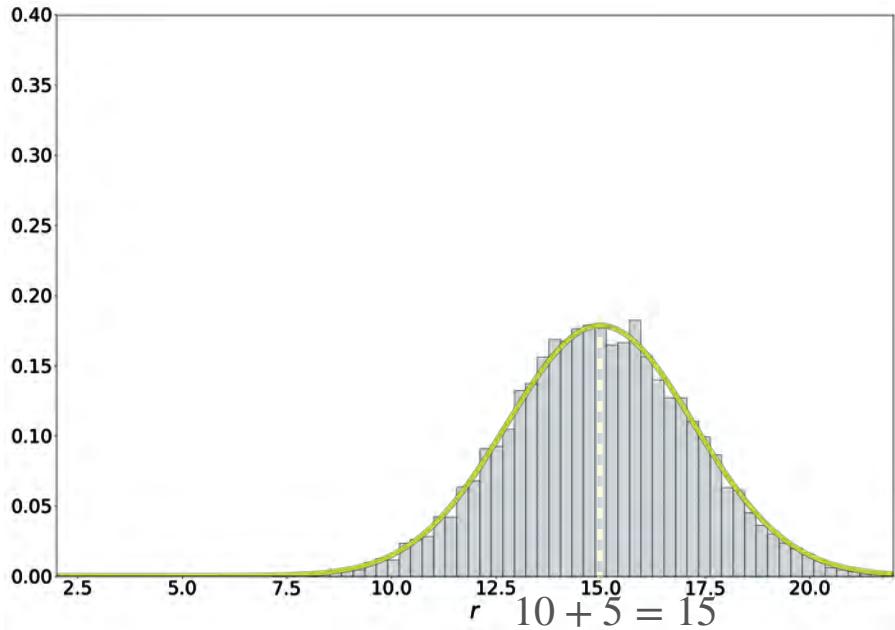
R is still Gaussian!

Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L\end{aligned}$$

Sum Of Gaussians

$$R = T + L$$



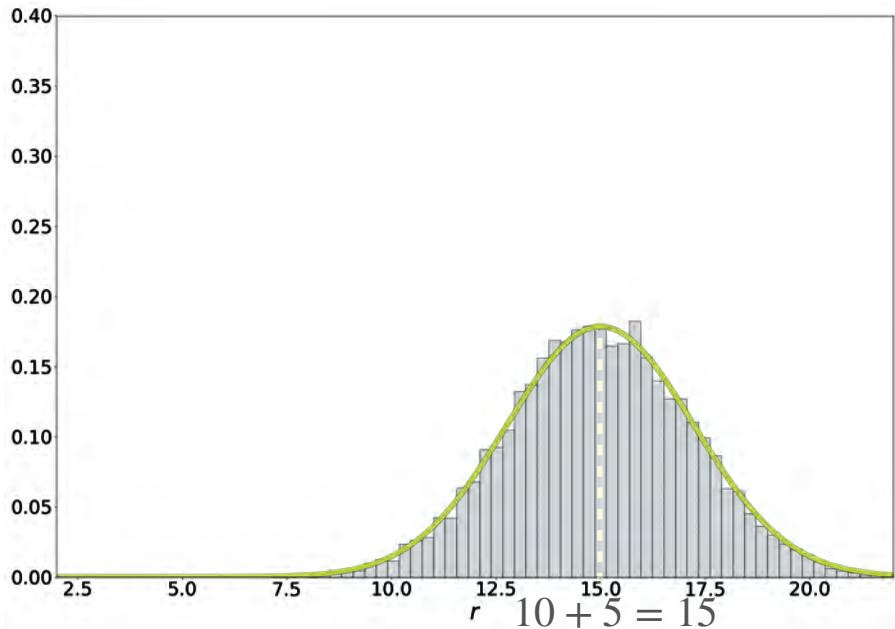
R is still Gaussian!

Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

Sum Of Gaussians

$$R = T + L$$



R is still Gaussian!

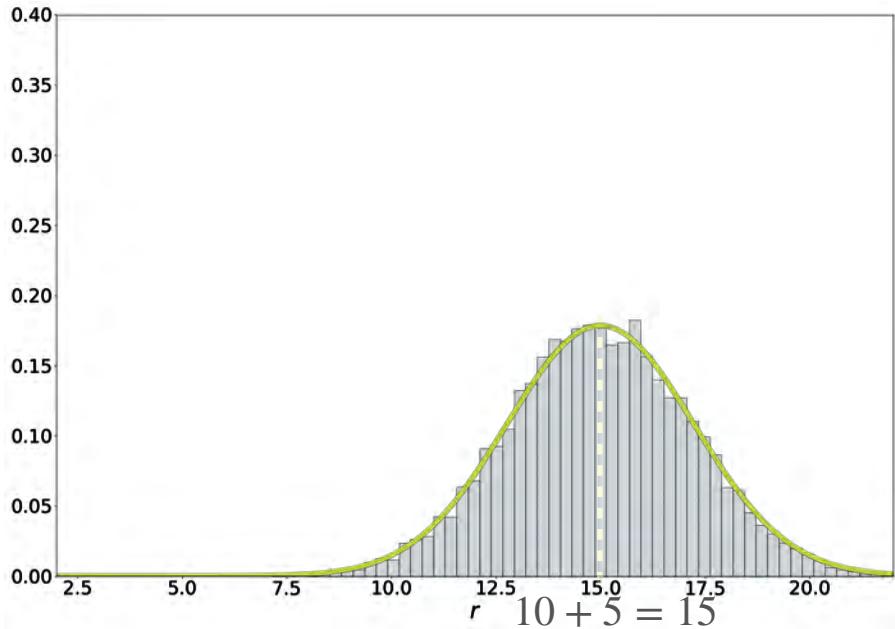
Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

$$\sigma_R^2 = Var(R)$$

Sum Of Gaussians

$$R = T + L$$



R is still Gaussian!

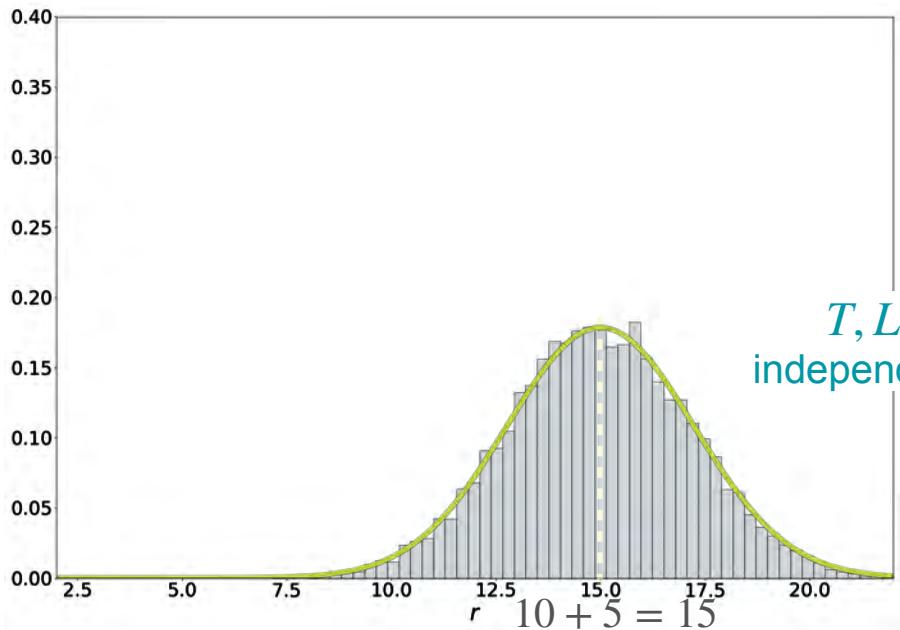
Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

$$\sigma_R^2 = Var(R) = Var(T + L)$$

Sum Of Gaussians

$$R = T + L$$



R is still Gaussian!

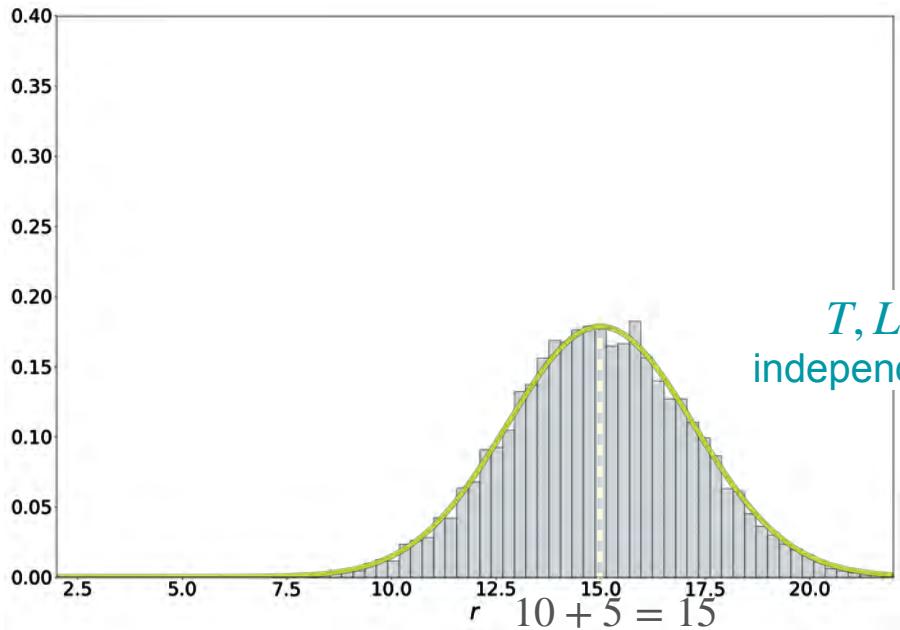
Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

$$\begin{aligned}\sigma_R^2 &= \text{Var}(R) = \text{Var}(T + L) \\ &= \text{Var}(T) + \text{Var}(L)\end{aligned}$$

Sum Of Gaussians

$$R = T + L$$



R is still Gaussian!

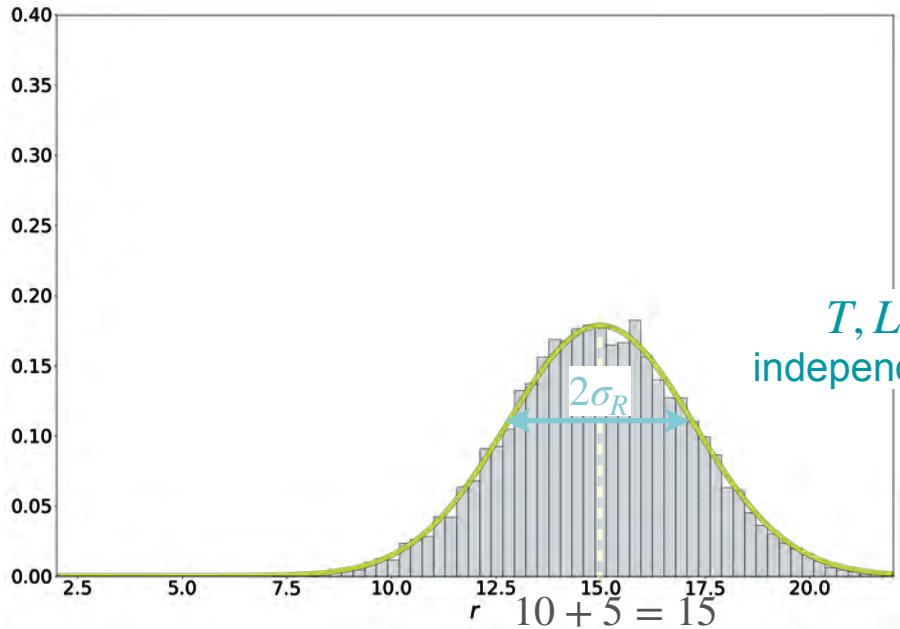
Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

$$\begin{aligned}\sigma_R^2 &= \text{Var}(R) = \text{Var}(T + L) \\ &= \text{Var}(T) + \text{Var}(L) = \sigma_T^2 + \sigma_L^2\end{aligned}$$

Sum Of Gaussians

$$R = T + L$$



R is still Gaussian!

Expectation is linear

$$\begin{aligned}\mu_R &= \mathbb{E}[R] = \mathbb{E}[T + L] = \mathbb{E}[T] + \mathbb{E}[L] \\ &= \mu_T + \mu_L = 10 + 5 = 15\end{aligned}$$

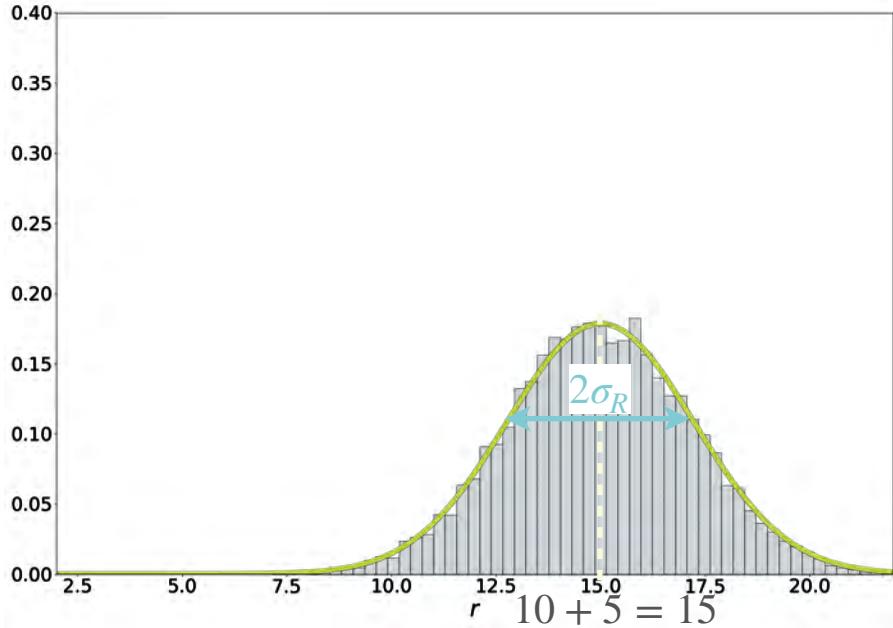
$$\begin{aligned}\sigma_R^2 &= \text{Var}(R) = \text{Var}(T + L) \\ &= \text{Var}(T) + \text{Var}(L) = \sigma_T^2 + \sigma_L^2\end{aligned}$$

$$= 4 + 1 = 5$$

Sum Of Gaussians

$$R = T + L$$

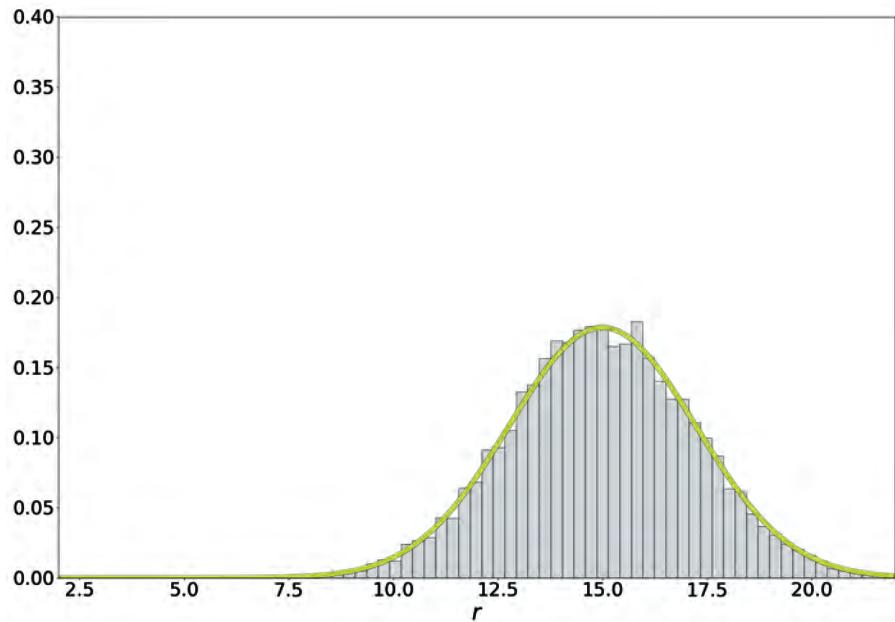
R is still Gaussian!



$$R = (T + L) \sim \mathcal{N} \left(10 + 5, \quad 4 + 1 \quad \right)$$

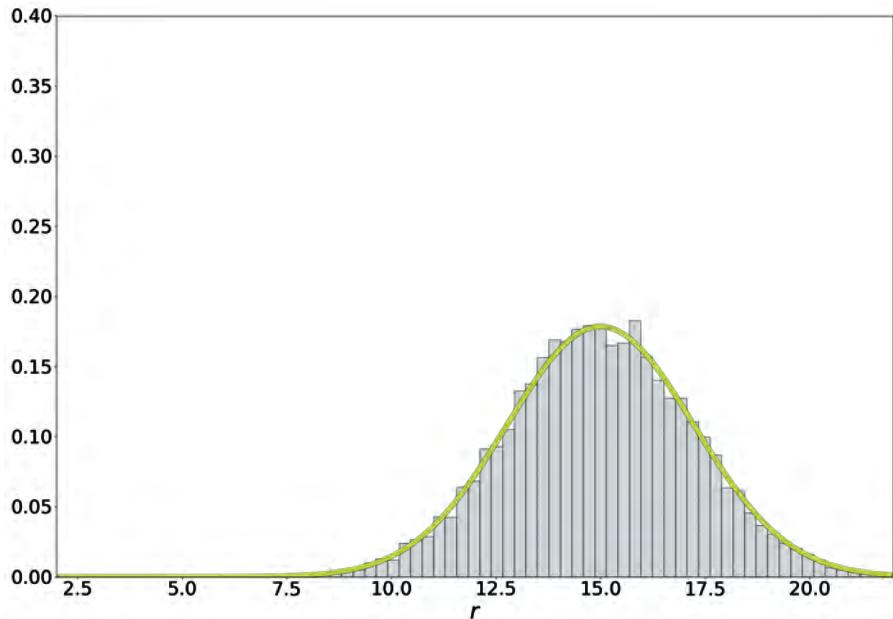
Sum Of Gaussians

$$R = T + L$$



Sum Of Gaussians

$$R = T + L$$

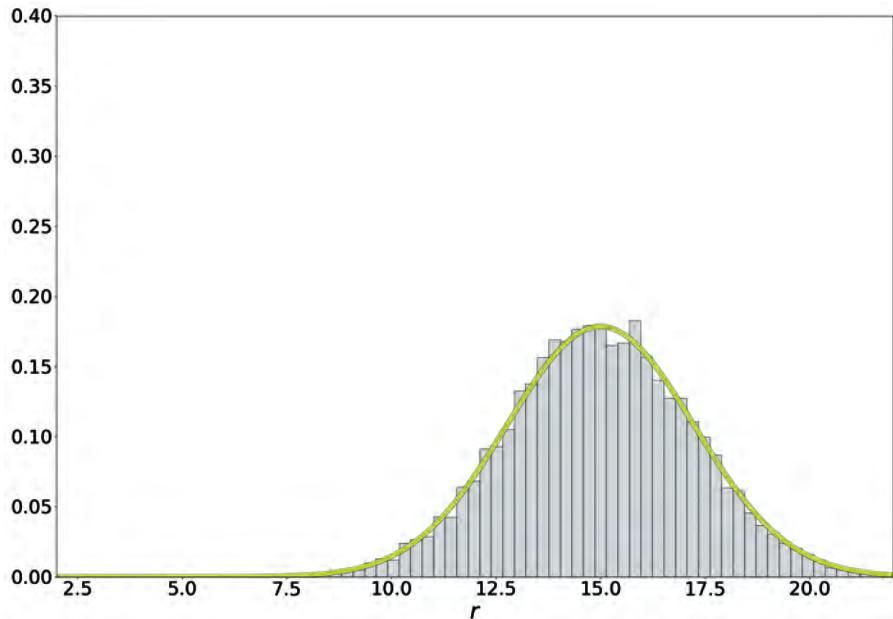


In general: $W = aX + bY$

Independent $\begin{cases} X \sim \mathcal{N}(\mu_X, \sigma_X^2) \\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \end{cases}$

Sum Of Gaussians

$$R = T + L$$



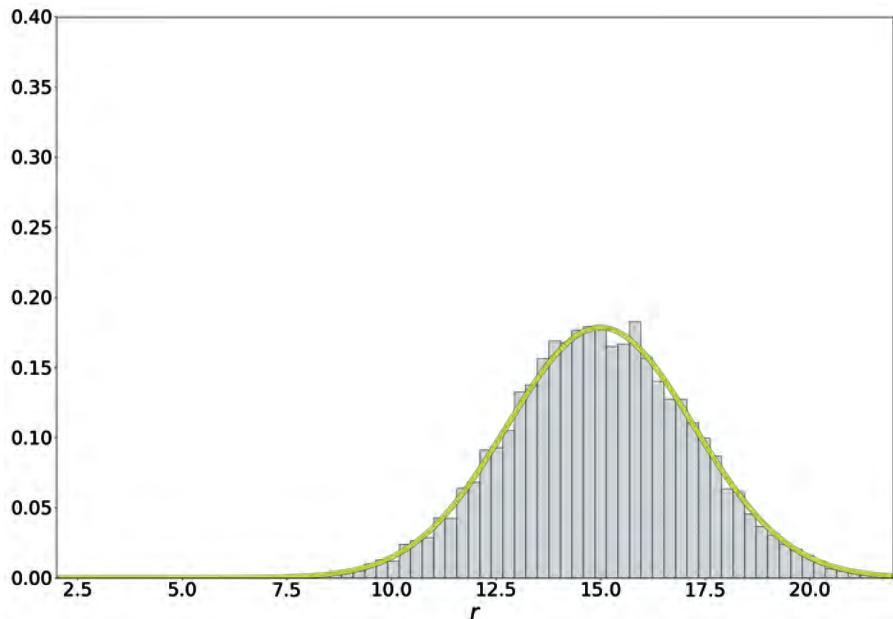
In general: $W = aX + bY$

Independent $\begin{cases} X \sim \mathcal{N}(\mu_X, \sigma_X^2) \\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \end{cases}$

$$\rightarrow W \sim \mathcal{N} \left(\quad , \quad \right)$$

Sum Of Gaussians

$$R = T + L$$



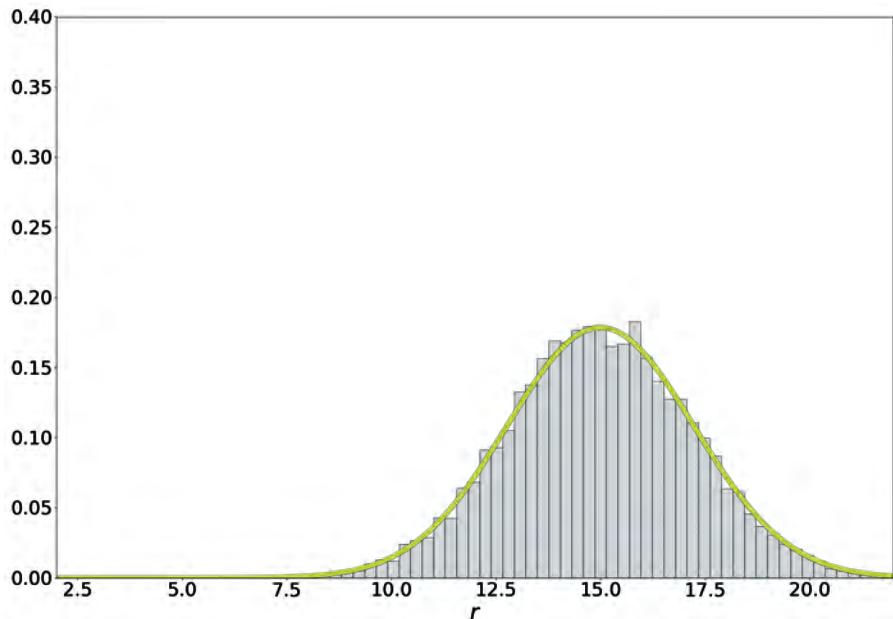
In general: $W = aX + bY$

Independent $\begin{cases} X \sim \mathcal{N}(\mu_X, \sigma_X^2) \\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \end{cases}$

$$\rightarrow W \sim \mathcal{N} \left(a\mu_x + b\mu_y, \quad \right)$$

Sum Of Gaussians

$$R = T + L$$



In general: $W = aX + bY$

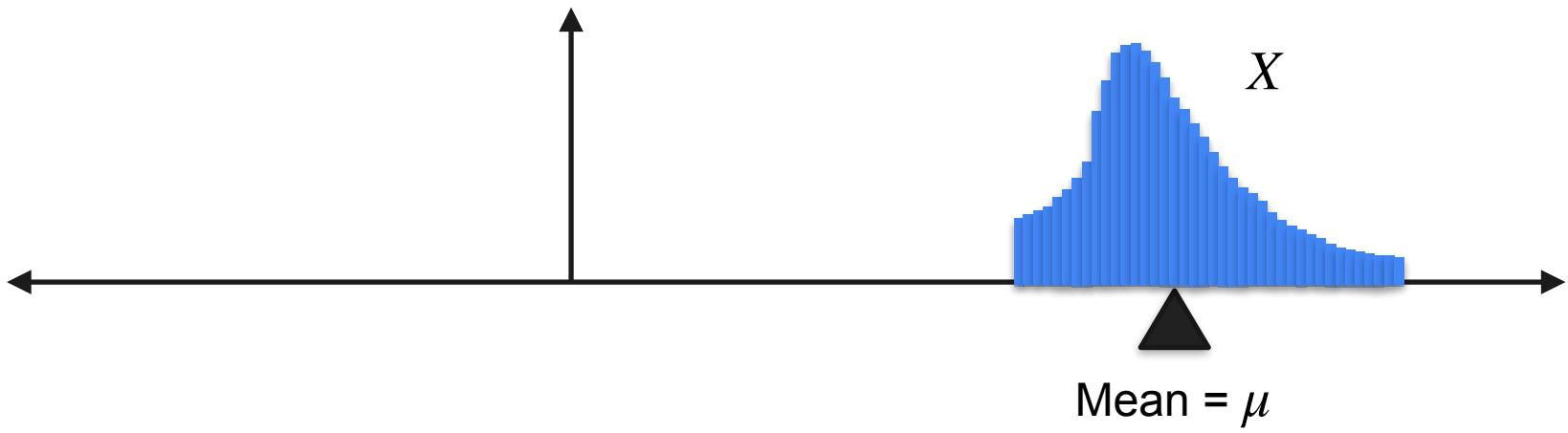
Independent $\begin{cases} X \sim \mathcal{N}(\mu_X, \sigma_X^2) \\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \end{cases}$

$$\rightarrow W \sim \mathcal{N} \left(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 \right)$$

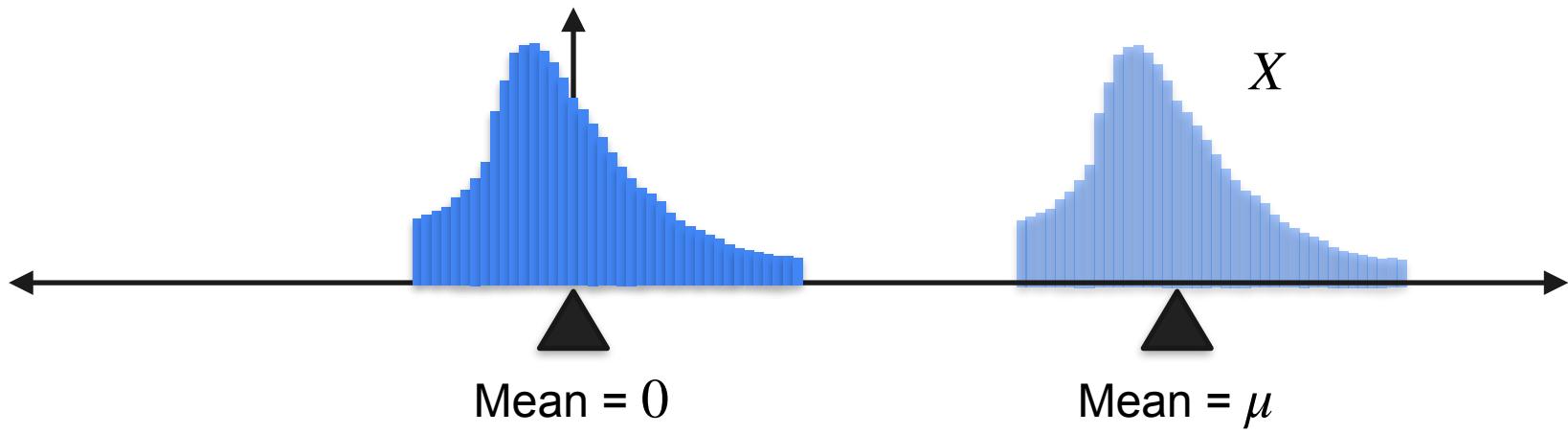
Everything Is Nicer When the Mean Is 0



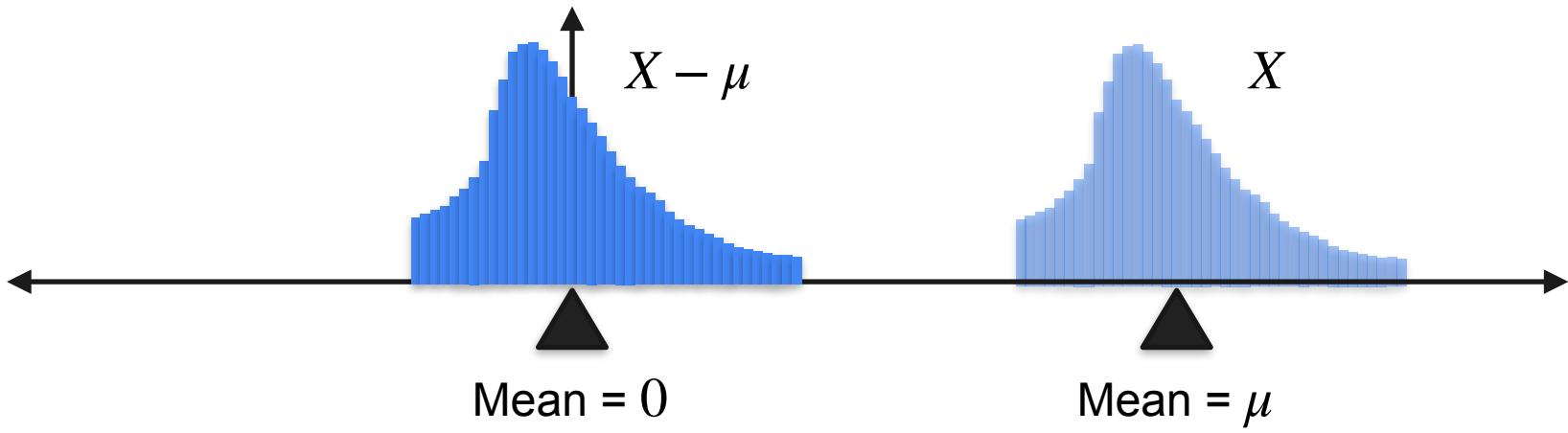
Everything Is Nicer When the Mean Is 0



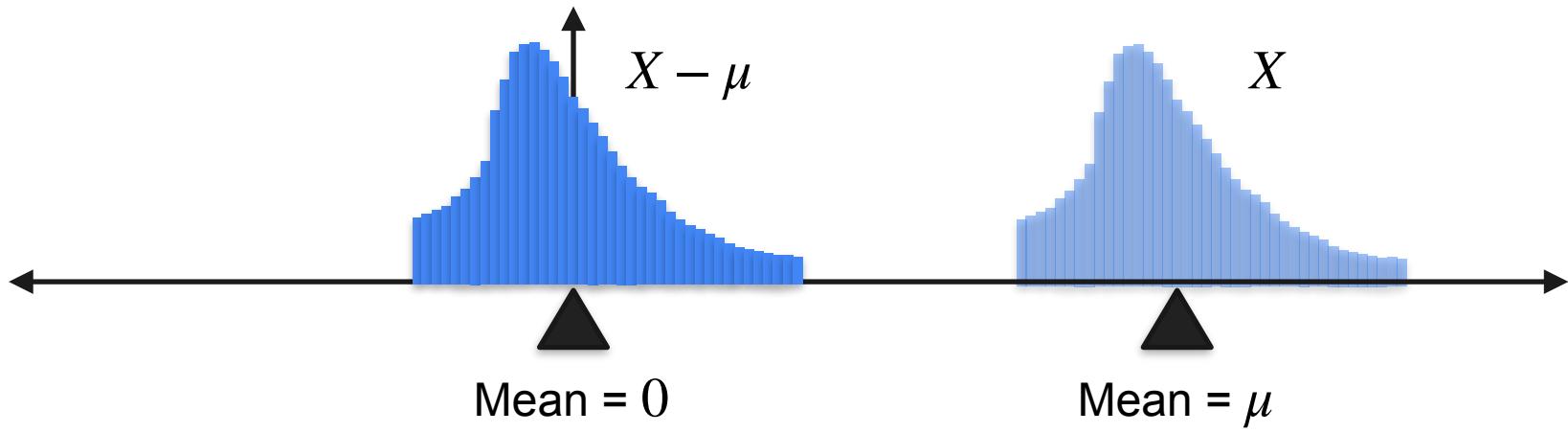
Everything Is Nicer When the Mean Is 0



Everything Is Nicer When the Mean Is 0



Everything Is Nicer When the Mean Is 0



$$X \rightarrow X - \mu$$

Everything Is Nicer When the Mean Is 0

Why?

Everything Is Nicer When the Mean Is 0

Why?

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Everything Is Nicer When the Mean Is 0

Why?

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathbb{E}[X - \mu]$$

Everything Is Nicer When the Mean Is 0

Why?

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mathbb{E}[\mu]$$

Everything Is Nicer When the Mean Is 0

Why?

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\begin{aligned}\mathbb{E}[X - \mu] &= \mathbb{E}[X] - \mathbb{E}[\mu] \\ &= \mathbb{E}[X] - \mu\end{aligned}$$

Everything Is Nicer When the Mean Is 0

Why?

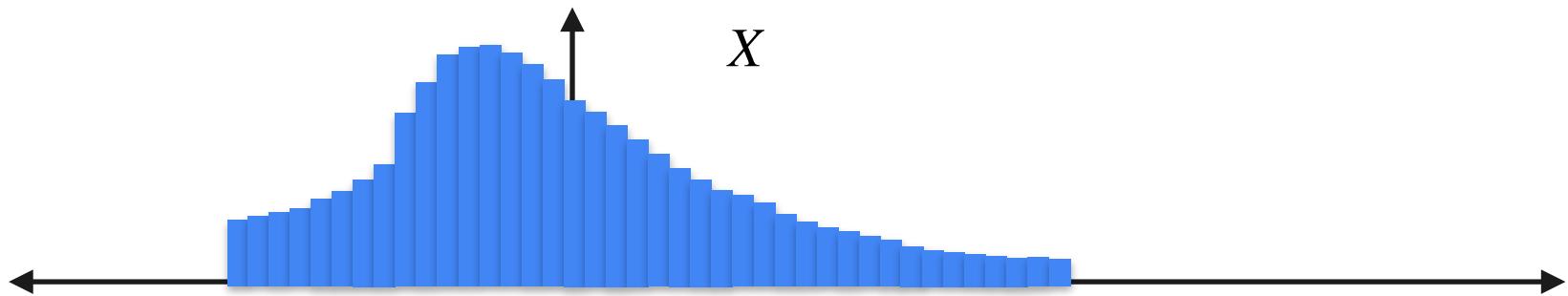
$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mathbb{E}[\mu]$$

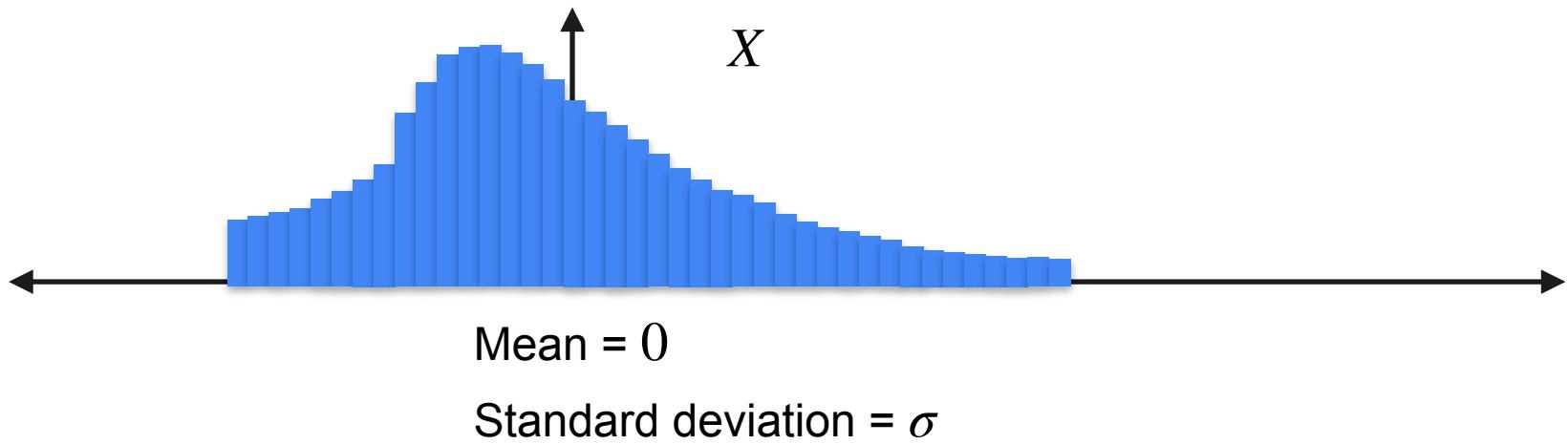
$$= \mathbb{E}[X] - \mu$$

$$= 0$$

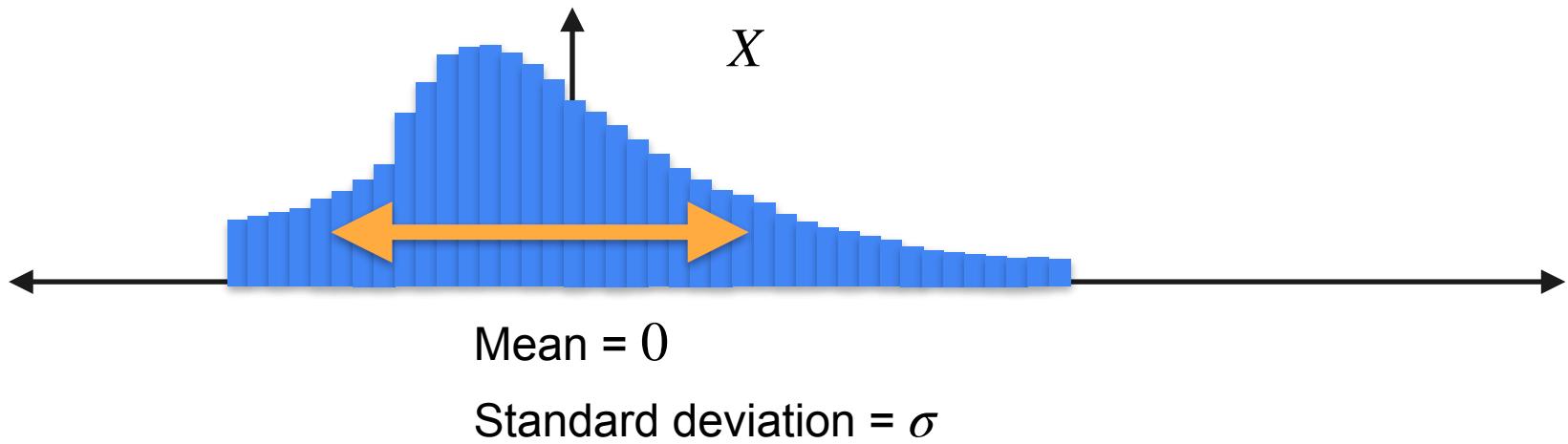
Everything Is Nicer When the Standard Deviation Is 1



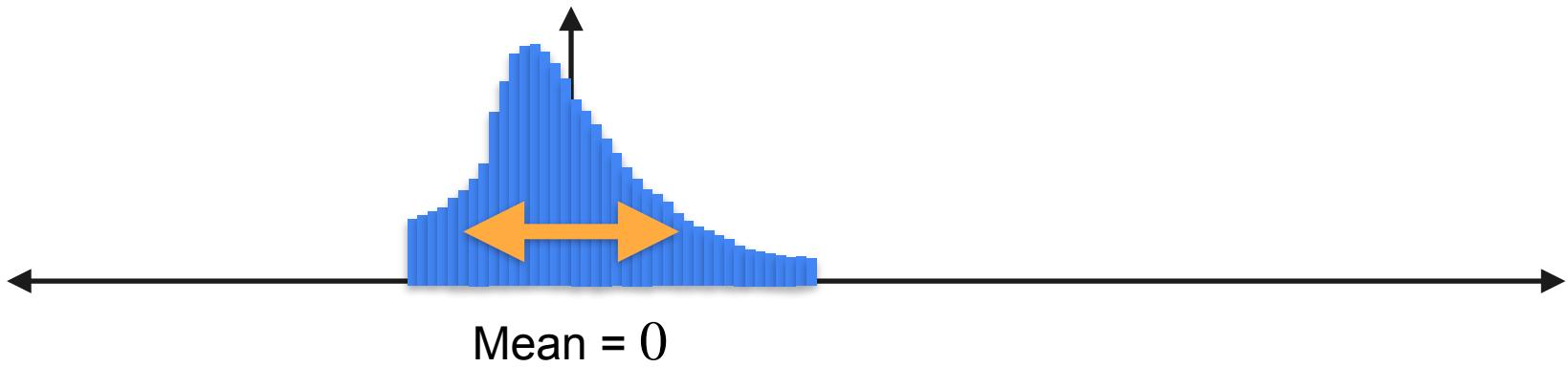
Everything Is Nicer When the Standard Deviation Is 1



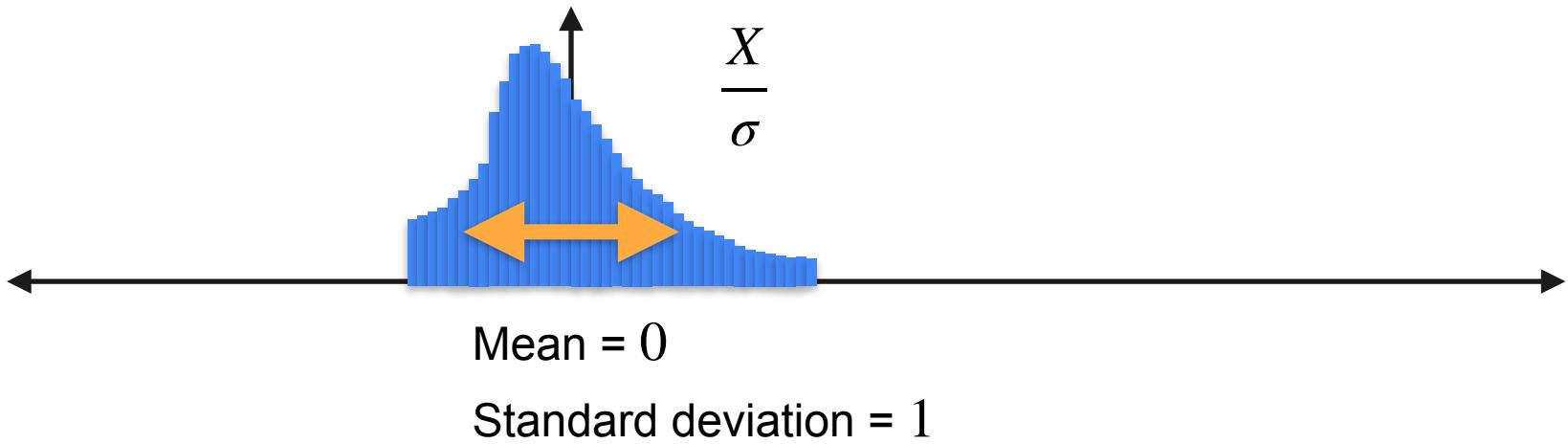
Everything Is Nicer When the Standard Deviation Is 1



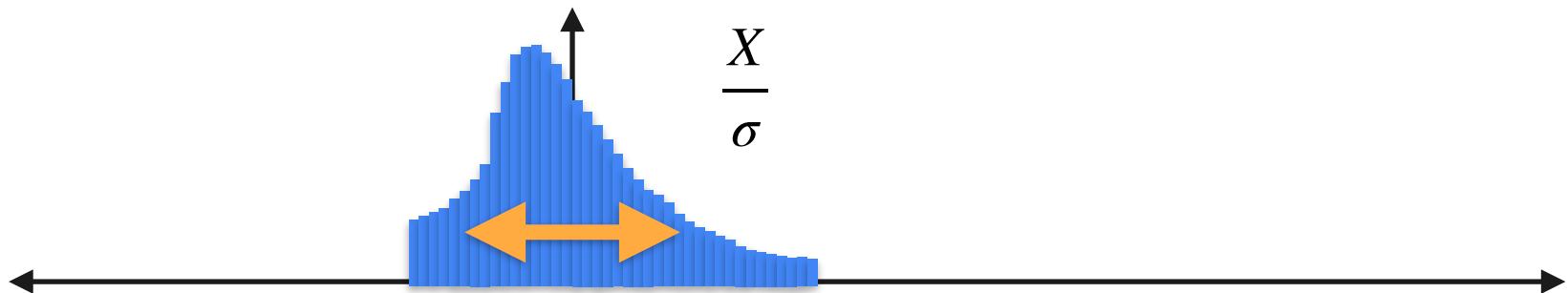
Everything Is Nicer When the Standard Deviation Is 1



Everything Is Nicer When the Standard Deviation Is 1



Everything Is Nicer When the Standard Deviation Is 1



Mean = 0

Standard deviation = 1

$$X \rightarrow \frac{X}{\sigma}$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

$$= c^2Var(X)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2}Var(X)$$

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

$$= c^2Var(X)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

$$= c^2Var(X)$$

$$Var\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2}Var(X)$$

$$std\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma}std(X)$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

$$= c^2Var(X)$$

$$Var\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2}Var(X)$$

$$std\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma}std(X)$$

$$= \frac{\sigma}{\sigma}$$

Everything Is Nicer When the Standard Deviation Is 1

Why?

$$Var(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$$

$$= \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2)$$

$$= c^2Var(X)$$

$$Var\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2}Var(X)$$

$$std\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma}std(X)$$

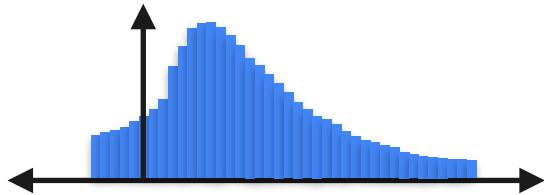
$$= \frac{\sigma}{\sigma}$$

$$= 1$$

Standardize a Distribution

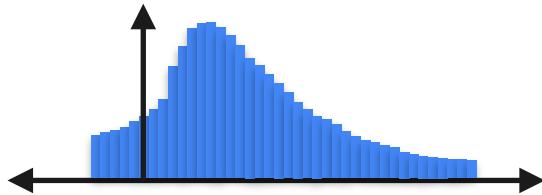
Standardize a Distribution

X

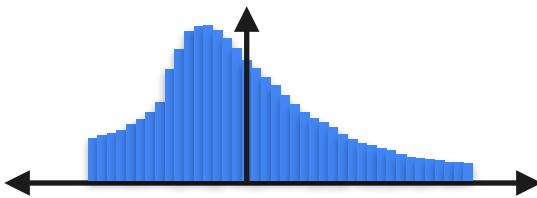


$$\begin{aligned}\text{Mean} &= \mu \\ \text{std} &= \sigma\end{aligned}$$

Standardize a Distribution

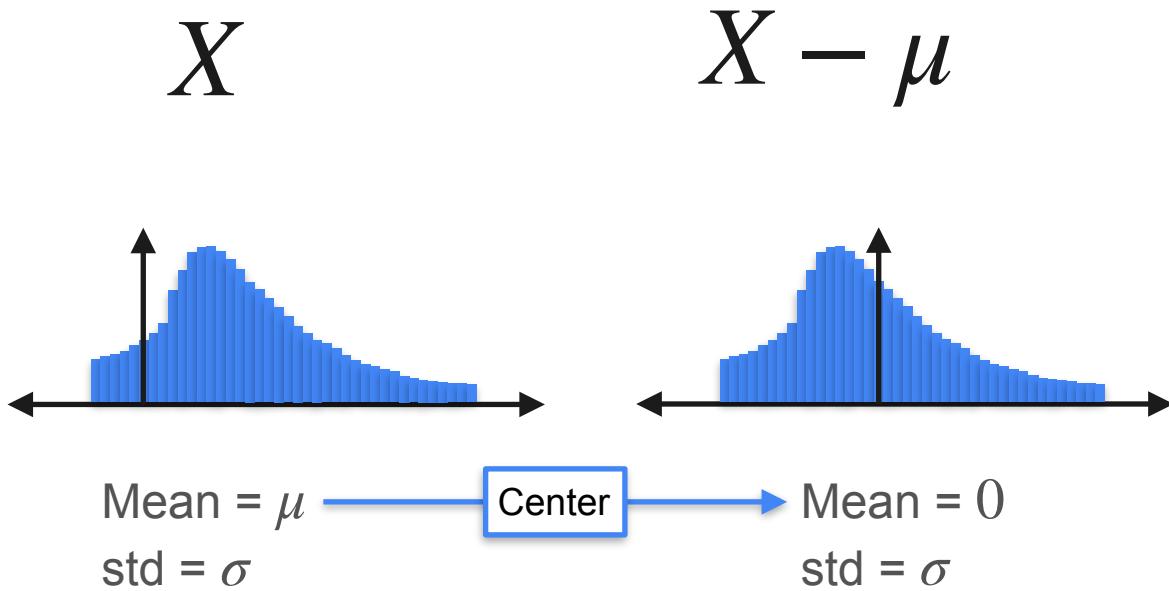
$$X$$
$$X - \mu$$


Mean = μ
std = σ

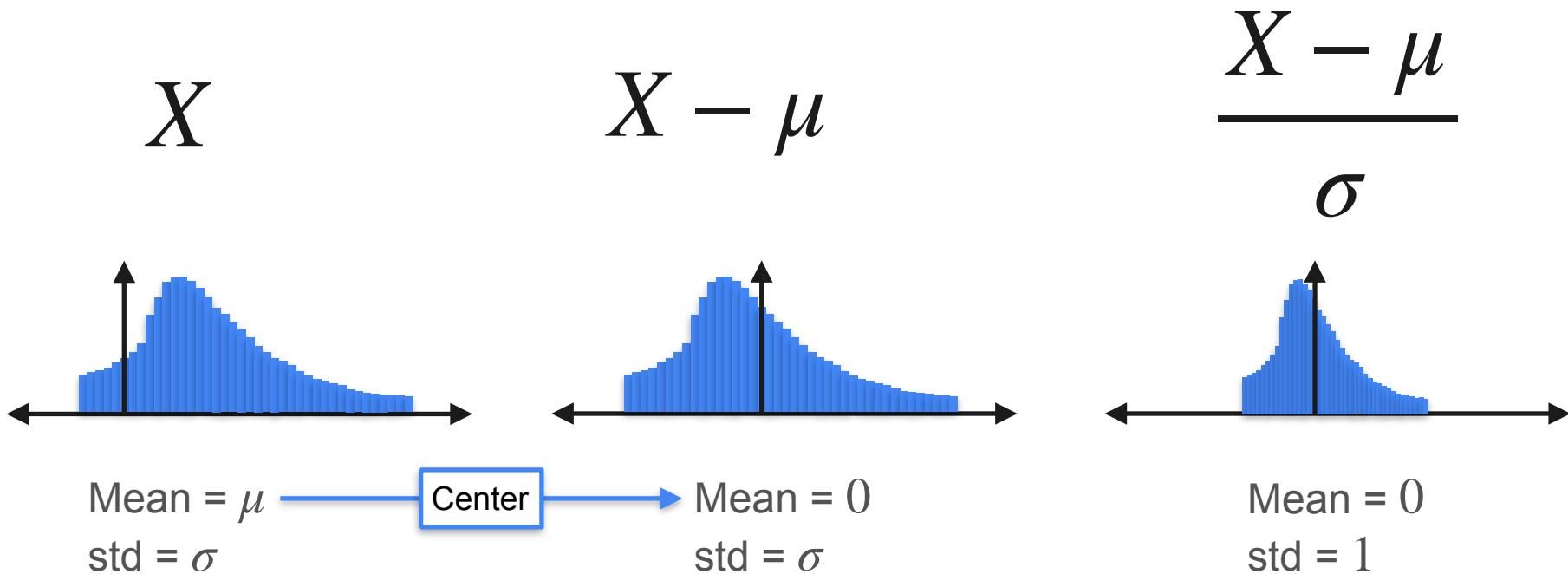


Mean = 0
std = σ

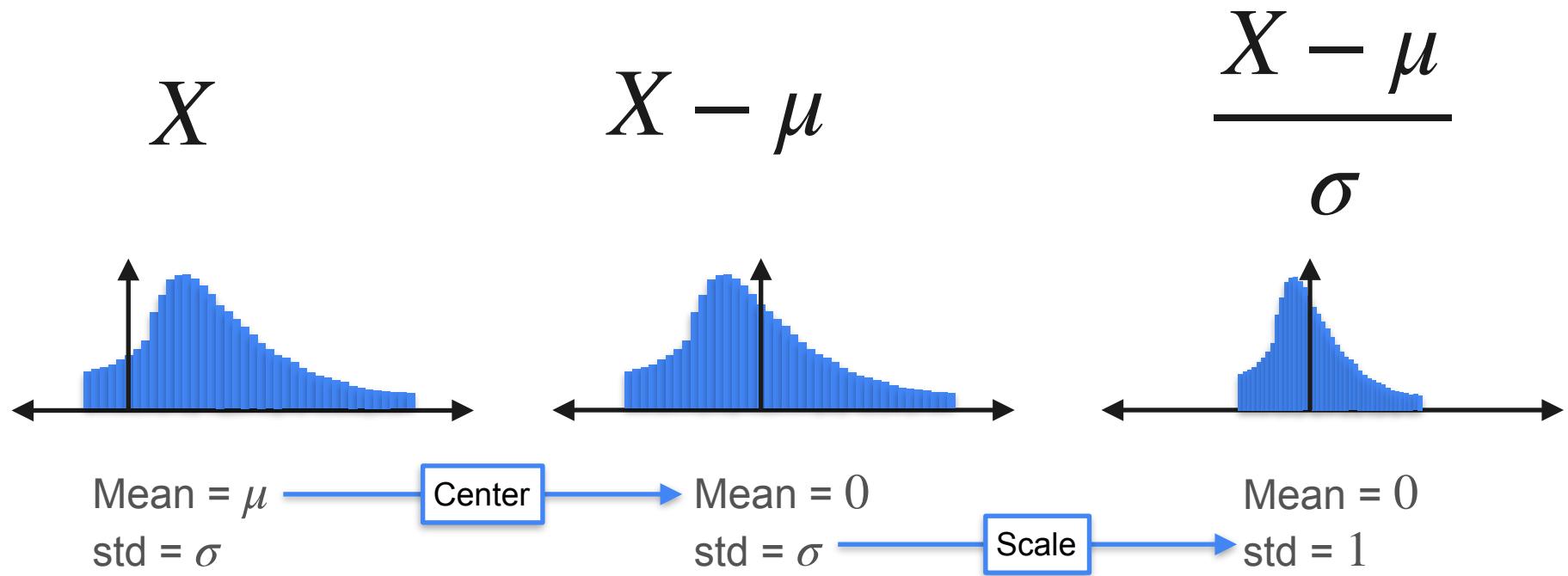
Standardize a Distribution



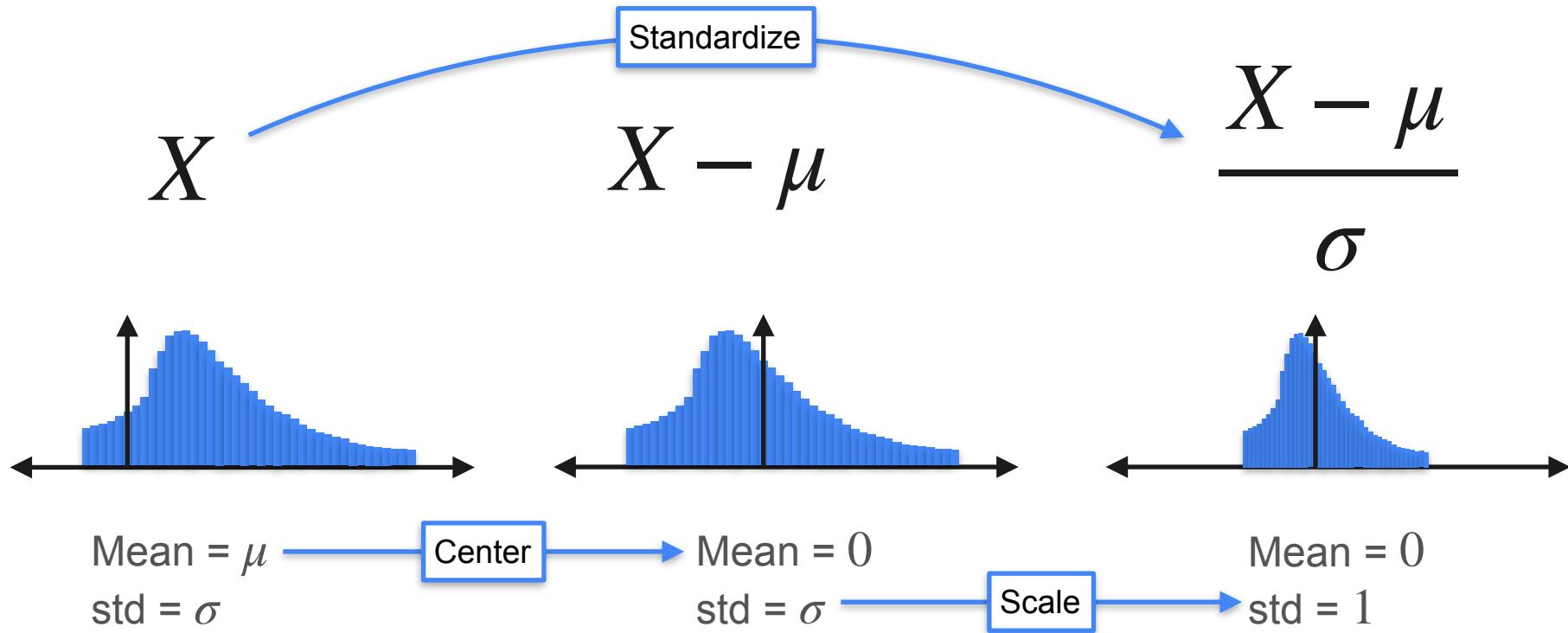
Standardize a Distribution



Standardize a Distribution



Standardize a Distribution





DeepLearning.AI

Describing Distributions

Skewness and Kurtosis

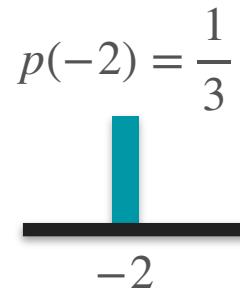
Moments of a Distribution

Random variable X



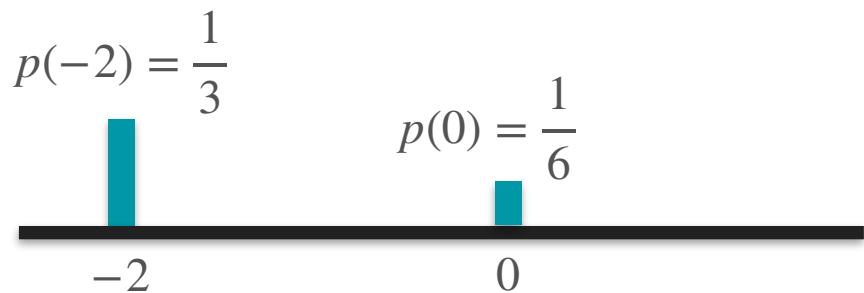
Moments of a Distribution

Random variable X

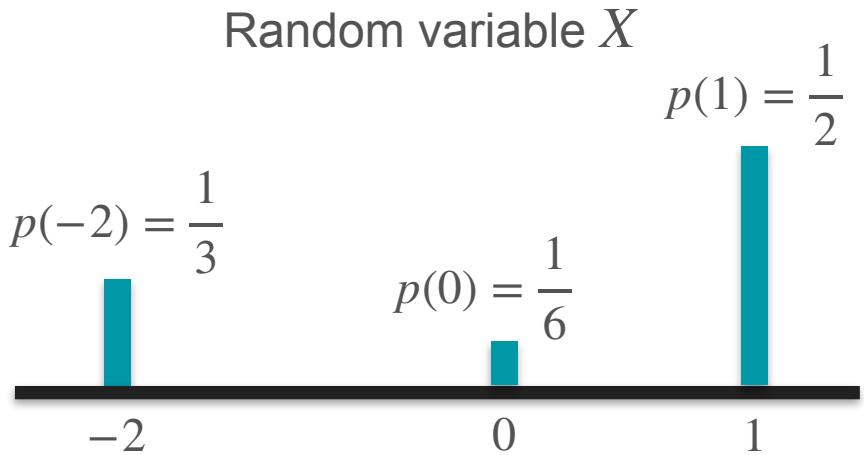


Moments of a Distribution

Random variable X



Moments of a Distribution



Moments of a Distribution

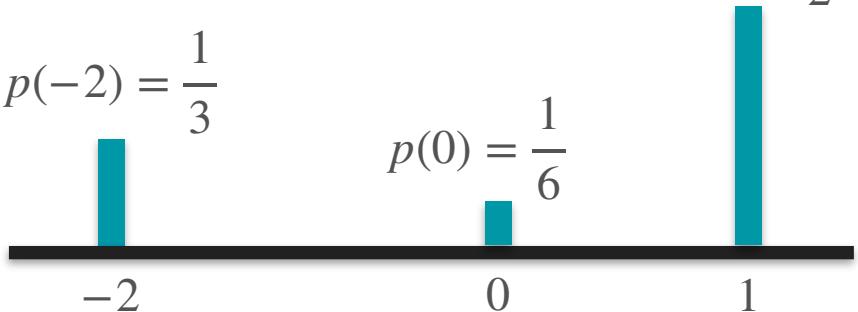
$$\mathbb{E}[X] = \frac{1}{3}(-2) + \frac{1}{6}(0) + \frac{1}{2}(1)$$

Random variable X

$$p(1) = \frac{1}{2}$$

$$p(-2) = \frac{1}{3}$$

$$p(0) = \frac{1}{6}$$



Moments of a Distribution

$$\mathbb{E}[X] = \frac{1}{3}(-2) + \frac{1}{6}(0) + \frac{1}{2}(1)$$

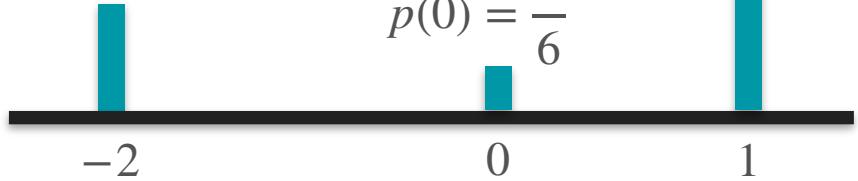
$$\mathbb{E}[X^2] = \frac{1}{3}(-2)^2 + \frac{1}{6}(0)^2 + \frac{1}{2}(1)^2$$

Random variable X

$$p(1) = \frac{1}{2}$$

$$p(-2) = \frac{1}{3}$$

$$p(0) = \frac{1}{6}$$

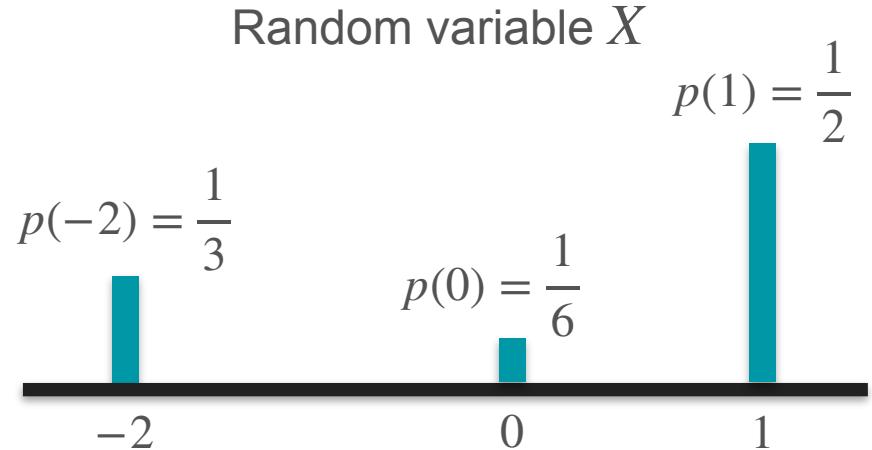


Moments of a Distribution

$$\mathbb{E}[X] = \frac{1}{3}(-2) + \frac{1}{6}(0) + \frac{1}{2}(1)$$

$$\mathbb{E}[X^2] = \frac{1}{3}(-2)^2 + \frac{1}{6}(0)^2 + \frac{1}{2}(1)^2$$

$$\mathbb{E}[X^3] = \frac{1}{3}(-2)^3 + \frac{1}{6}(0)^3 + \frac{1}{2}(1)^3$$



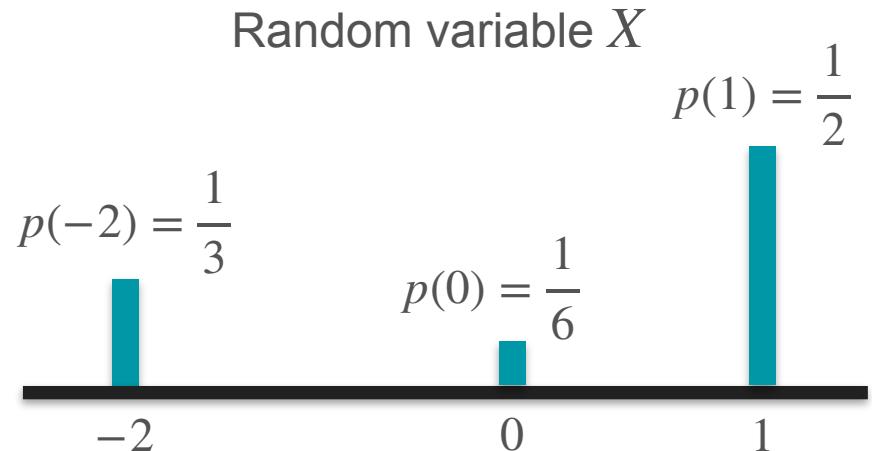
Moments of a Distribution

$$\mathbb{E}[X] = \frac{1}{3}(-2) + \frac{1}{6}(0) + \frac{1}{2}(1)$$

$$\mathbb{E}[X^2] = \frac{1}{3}(-2)^2 + \frac{1}{6}(0)^2 + \frac{1}{2}(1)^2$$

$$\mathbb{E}[X^3] = \frac{1}{3}(-2)^3 + \frac{1}{6}(0)^3 + \frac{1}{2}(1)^3$$

...



Moments of a Distribution

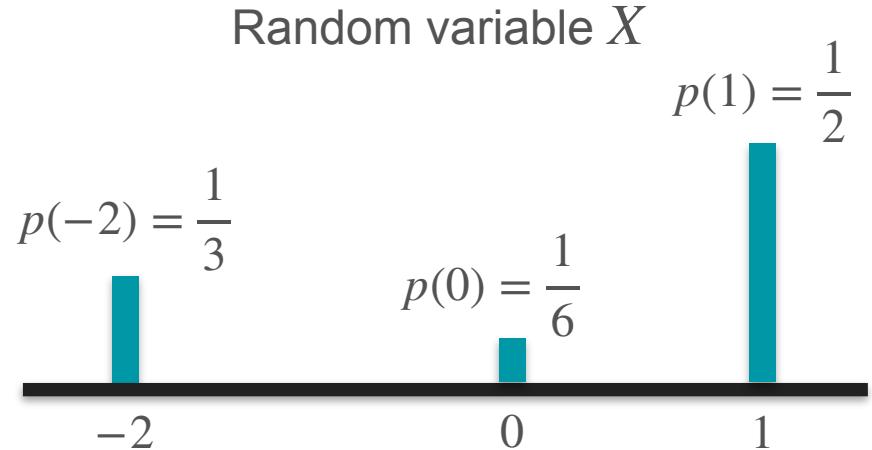
$$\mathbb{E}[X] = \frac{1}{3}(-2) + \frac{1}{6}(0) + \frac{1}{2}(1)$$

$$\mathbb{E}[X^2] = \frac{1}{3}(-2)^2 + \frac{1}{6}(0)^2 + \frac{1}{2}(1)^2$$

$$\mathbb{E}[X^3] = \frac{1}{3}(-2)^3 + \frac{1}{6}(0)^3 + \frac{1}{2}(1)^3$$

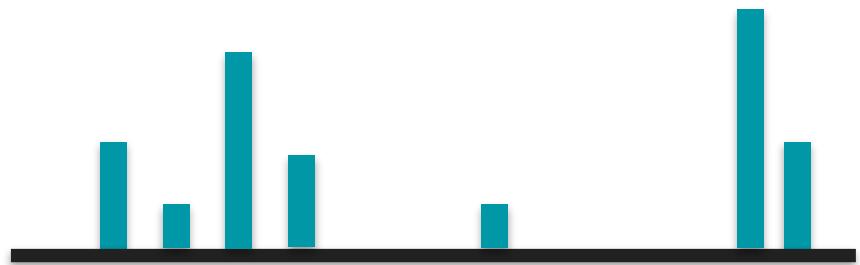
...

$$\mathbb{E}[X^k] = \frac{1}{3}(-2)^k + \frac{1}{6}(0)^k + \frac{1}{2}(1)^k$$



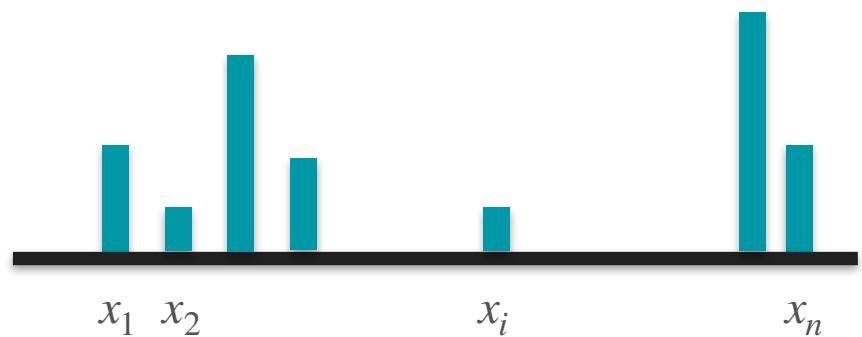
Moments of a Distribution

Random variable X



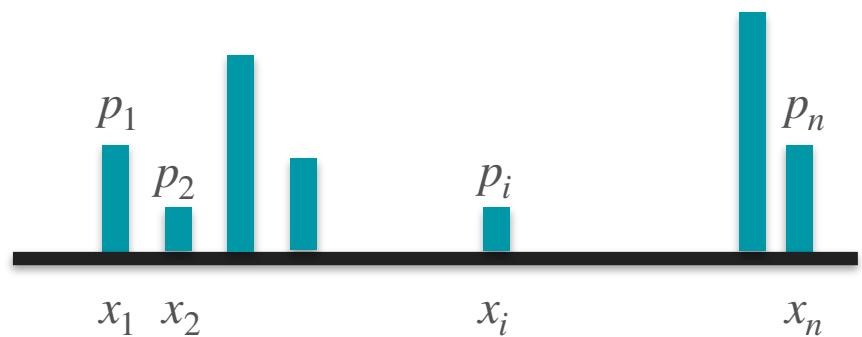
Moments of a Distribution

Random variable X



Moments of a Distribution

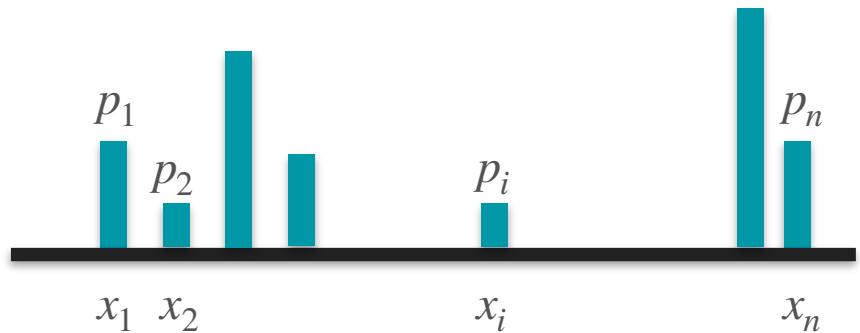
Random variable X



Moments of a Distribution

$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

Random variable X

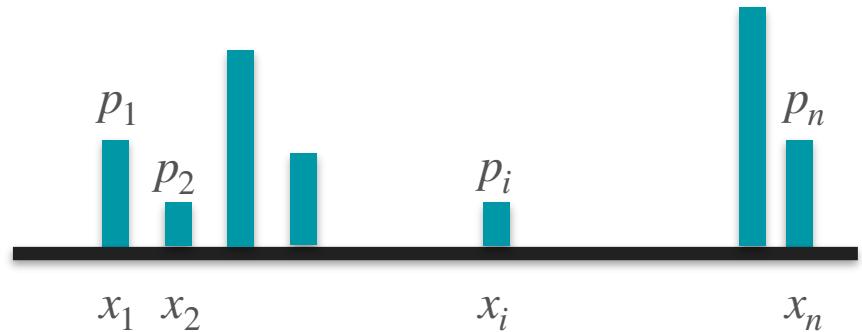


Moments of a Distribution

$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

$$\mathbb{E}[X^2] = p_1x_1^2 + p_2x_2^2 + \cdots + p_nx_n^2$$

Random variable X



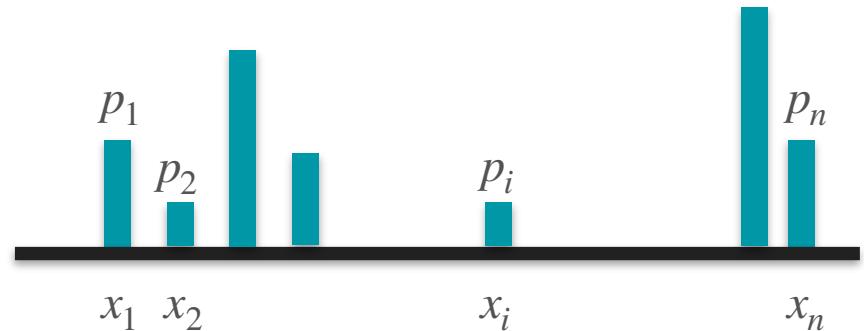
Moments of a Distribution

$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

$$\mathbb{E}[X^2] = p_1x_1^2 + p_2x_2^2 + \cdots + p_nx_n^2$$

$$\mathbb{E}[X^3] = p_1x_1^3 + p_2x_2^3 + \cdots + p_nx_n^3$$

Random variable X



Moments of a Distribution

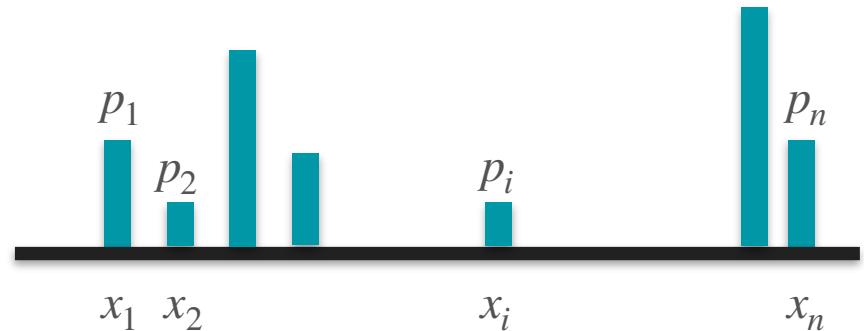
$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

$$\mathbb{E}[X^2] = p_1x_1^2 + p_2x_2^2 + \cdots + p_nx_n^2$$

$$\mathbb{E}[X^3] = p_1x_1^3 + p_2x_2^3 + \cdots + p_nx_n^3$$

$$\mathbb{E}[X^4] = p_1x_1^4 + p_2x_2^4 + \cdots + p_nx_n^4$$

Random variable X



Moments of a Distribution

$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

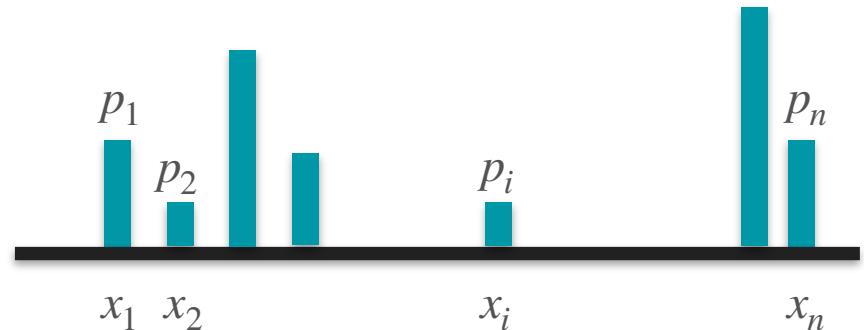
$$\mathbb{E}[X^2] = p_1x_1^2 + p_2x_2^2 + \cdots + p_nx_n^2$$

$$\mathbb{E}[X^3] = p_1x_1^3 + p_2x_2^3 + \cdots + p_nx_n^3$$

$$\mathbb{E}[X^4] = p_1x_1^4 + p_2x_2^4 + \cdots + p_nx_n^4$$

...

Random variable X



Moments of a Distribution

$$\mathbb{E}[X] = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

$$\mathbb{E}[X^2] = p_1x_1^2 + p_2x_2^2 + \cdots + p_nx_n^2$$

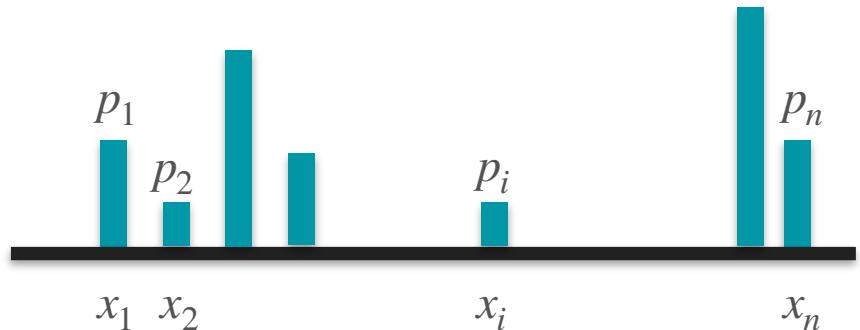
$$\mathbb{E}[X^3] = p_1x_1^3 + p_2x_2^3 + \cdots + p_nx_n^3$$

$$\mathbb{E}[X^4] = p_1x_1^4 + p_2x_2^4 + \cdots + p_nx_n^4$$

...

$$\mathbb{E}[X^k] = p_1x_1^k + p_2x_2^k + \cdots + p_nx_n^k$$

Random variable X

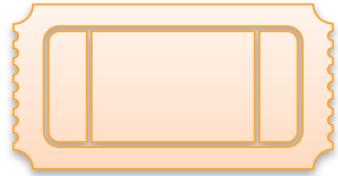


Video 7b:

- Skewness

Lottery vs Insurance

Lottery vs Insurance



Lottery

Lottery vs Insurance



Lottery



Car insurance

Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Car insurance

Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Car insurance

You **win** \$99 with 1% probability

You **lose** \$1 with 99% probability

Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance

You **win** \$99 with 1% probability

You **lose** \$1 with 99% probability

Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery

You **win** \$99 with 1% probability

You **lose** \$1 with 99% probability



Cost: \$1
Crash Reparation: \$100

Car insurance

You **win** \$1 with 99% probability

You **lose** \$99 with 1% probability

Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance



Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance



Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance



Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance



Lottery vs Insurance



Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance





Ticket: \$1
Jackpot: \$100

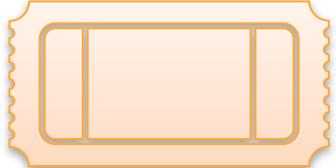
Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance





Ticket: \$1
Jackpot: \$100

Lottery

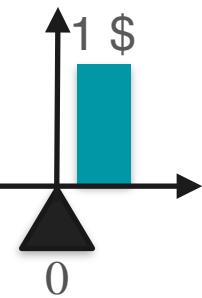


Cost: \$1
Crash Reparation: \$100

Car insurance



-99 \$





Ticket: \$1
Jackpot: \$100

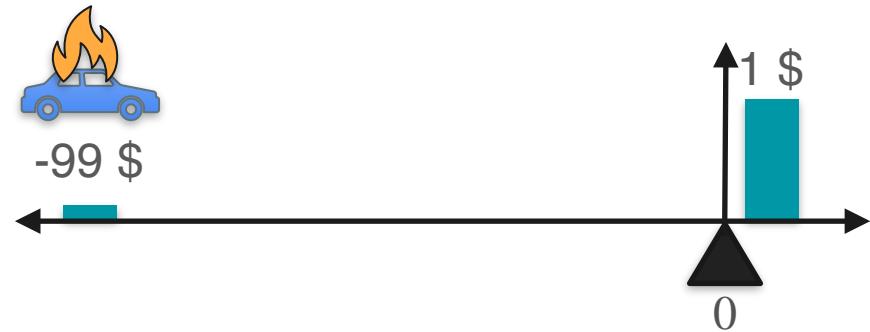
Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$





Ticket: \$1
Jackpot: \$100

Lottery



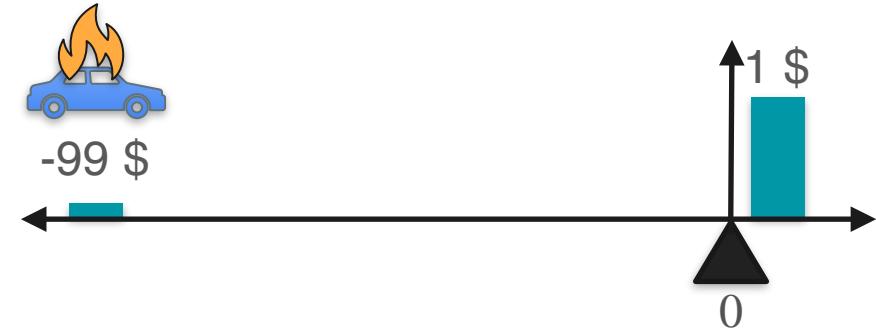
Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

Lose 1
With probability 0.99

Win 99
With probability 0.01





Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

Lose 1
With probability 0.99

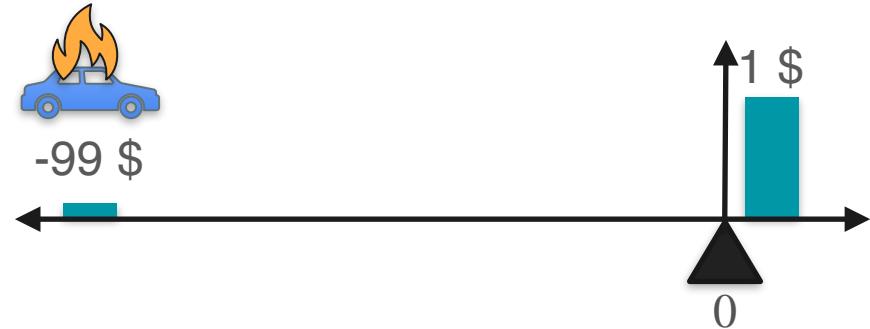
Win 99
With probability 0.01



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$





Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

Lose 1
With probability 0.99

Win 99
With probability 0.01



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$

Lose 99
With probability 0.01

Win 1
With probability 0.99





Ticket: \$1
Jackpot: \$100

Lottery



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$





Ticket: \$1
Jackpot: \$100

Lottery

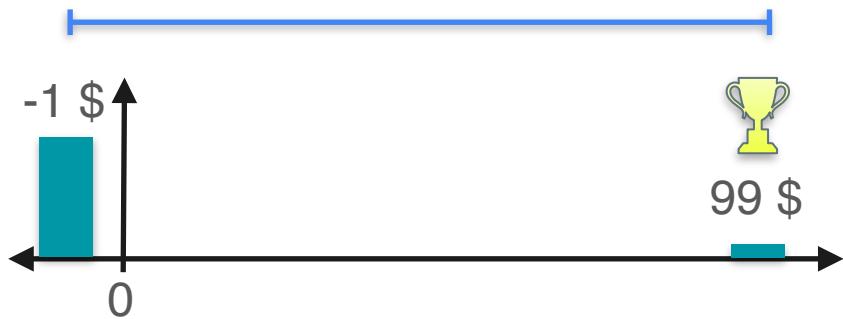
$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$





Ticket: \$1
Jackpot: \$100

Lottery

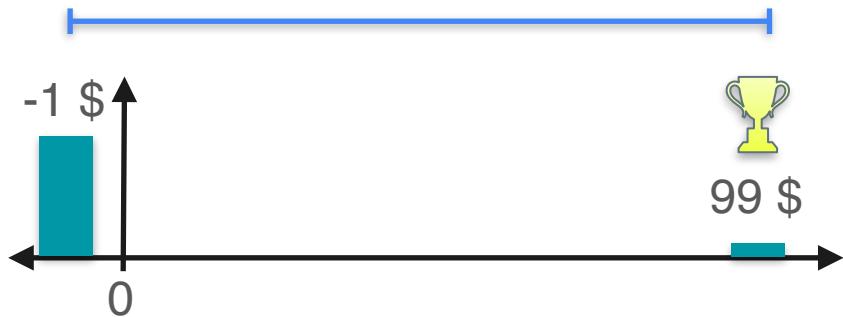
$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$



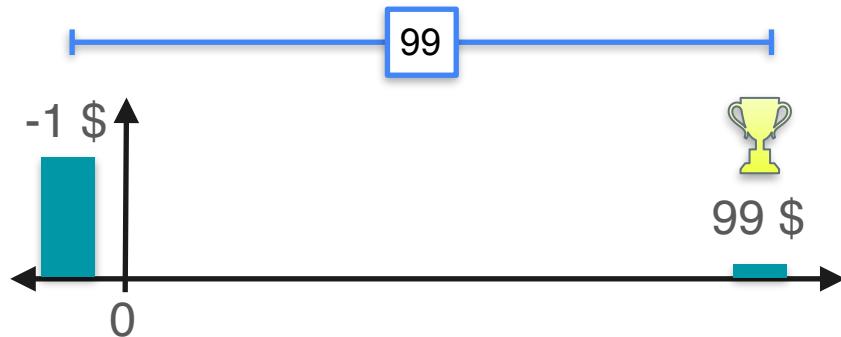


Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

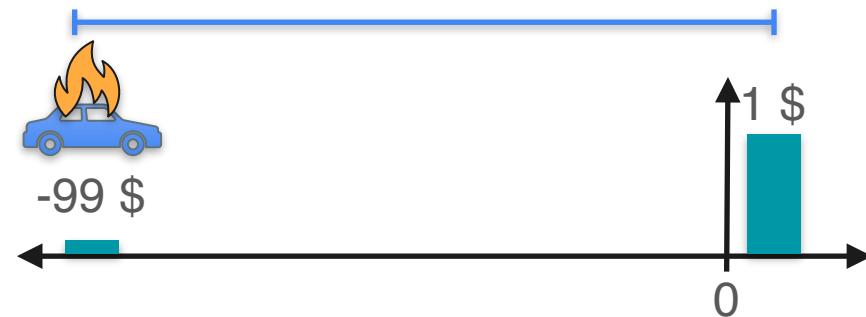
$$Var(X_1) = (-1)^2 \cdot 0.99 + (99)^2 \cdot 0.01 = 99$$



Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$



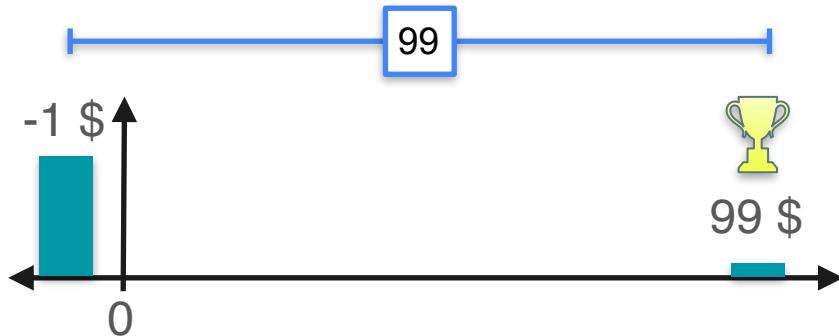


Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = -1 \cdot 0.99 + 99 \cdot 0.01 = 0$$

$$Var(X_1) = (-1)^2 \cdot 0.99 + (99)^2 \cdot 0.01 = 99$$

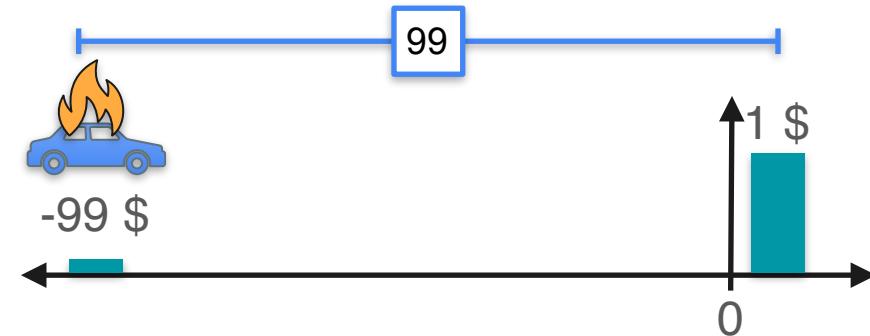


Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = -99 \cdot 0.01 + 1 \cdot 0.99 = 0$$

$$Var(X_2) = (-99)^2 \cdot 0.01 + (1)^2 \cdot 0.99 = 99$$



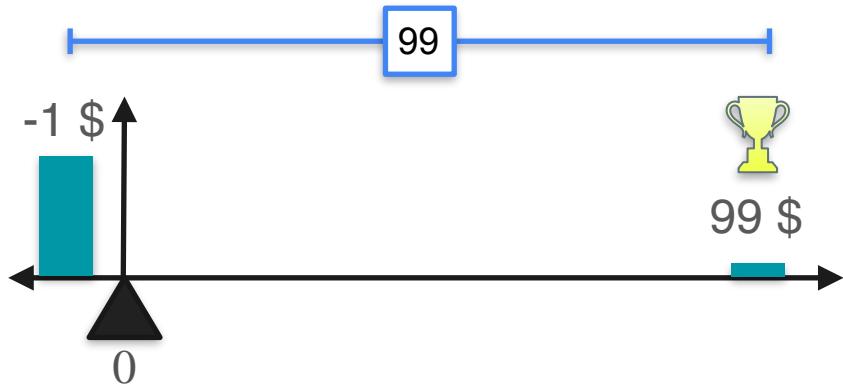


Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$

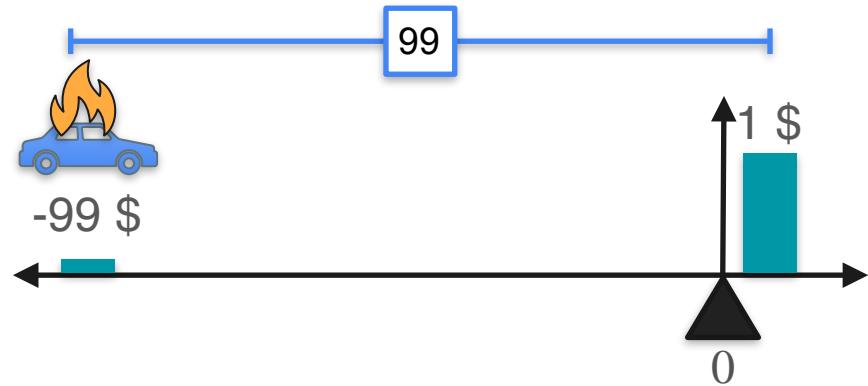


Cost: \$1
Crash Reparation: \$100

Car insurance

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$





Ticket: \$1
Jackpot: \$100

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$



Cost: \$1
Crash Reparation: \$100

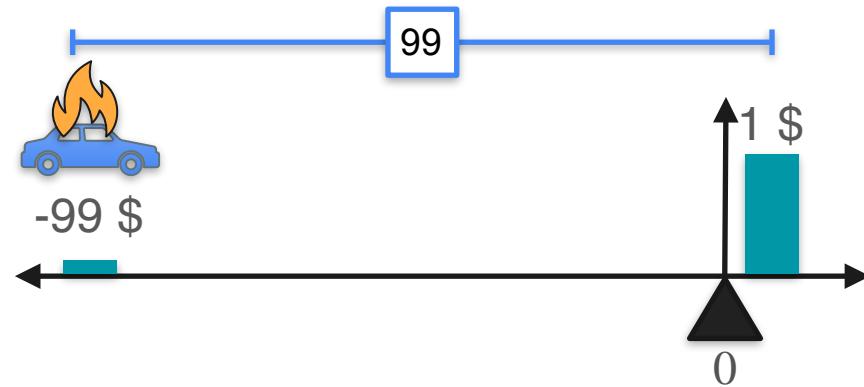
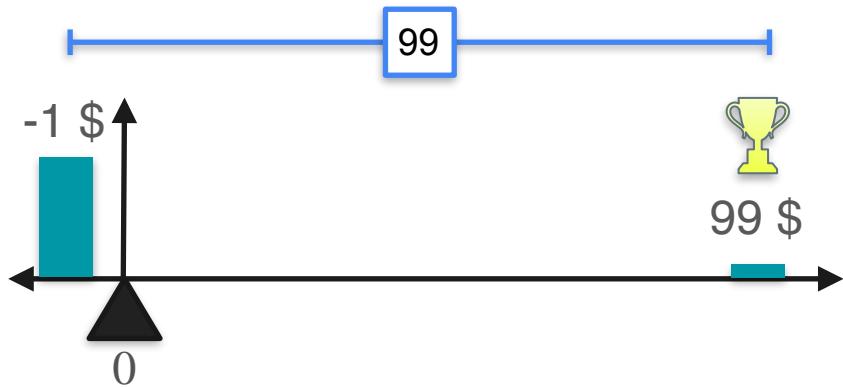
Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$

$$\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = 99$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$

$$\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 = 99$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$

$$\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = 99$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$

$$\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 = 99$$





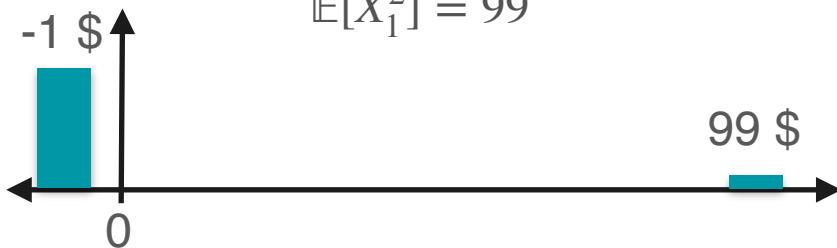
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$Var(X_1) = 99$$

$$\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = 99$$



Cost: \$1
Crash Reparation: \$99

Car insurance

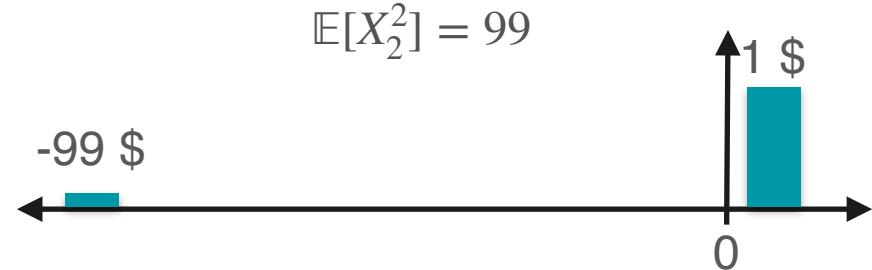
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$Var(X_2) = 99$$

$$\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 = 99$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$





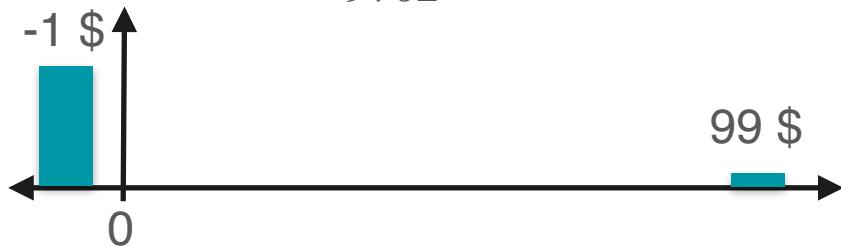
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_1^3] &= (-1)^3 \cdot 0.99 + (99)^3 \cdot 0.01 \\ &= 9702\end{aligned}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$





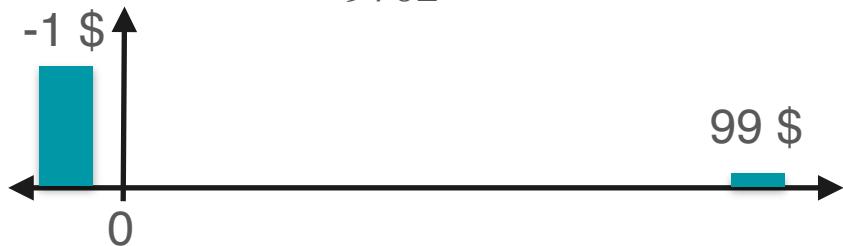
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_1^3] &= (-1)^3 \cdot 0.99 + (99)^3 \cdot 0.01 \\ &= 9702\end{aligned}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

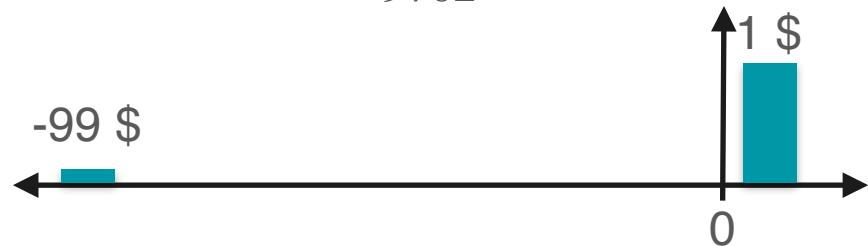
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_2^3] &= (1)^3 \cdot 0.99 + (-99)^3 \cdot 0.01 \\ &= -9702\end{aligned}$$





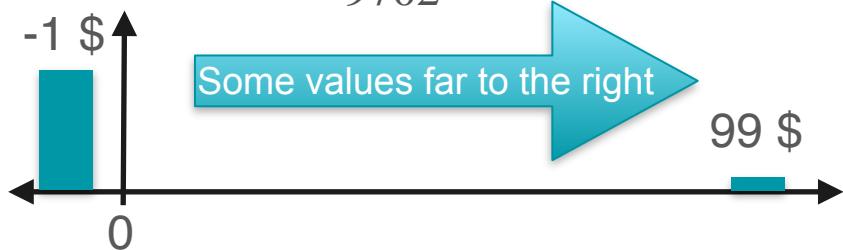
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_1^3] &= (-1)^3 \cdot 0.99 + (99)^3 \cdot 0.01 \\ &= 9702\end{aligned}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

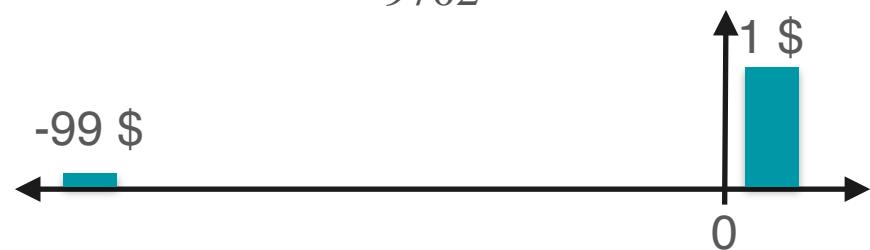
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_2^3] &= (1)^3 \cdot 0.99 + (-99)^3 \cdot 0.01 \\ &= -9702\end{aligned}$$





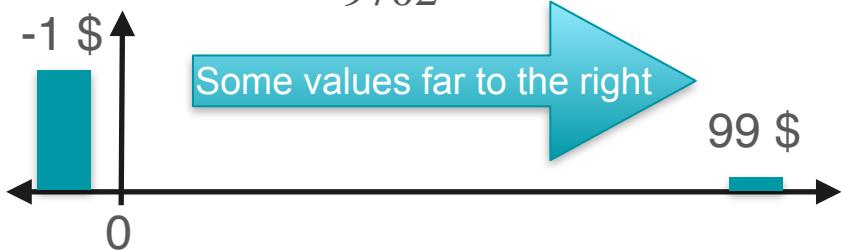
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_1^3] &= (-1)^3 \cdot 0.99 + (99)^3 \cdot 0.01 \\ &= 9702\end{aligned}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

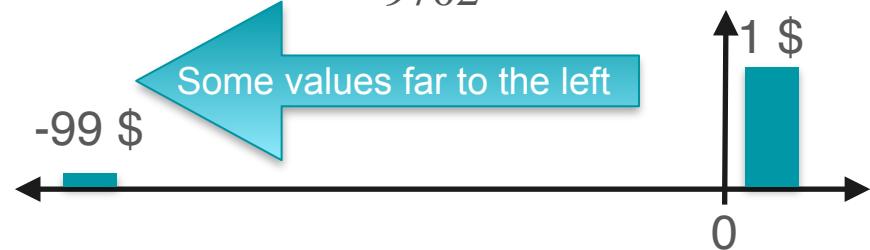
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\begin{aligned}\mathbb{E}[X_2^3] &= (1)^3 \cdot 0.99 + (-99)^3 \cdot 0.01 \\ &= -9702\end{aligned}$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\mathbb{E}[X_1^3] = \text{Large positive value}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\mathbb{E}[X_2^3] = \text{Large negative value}$$





Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\mathbb{E}[X_1^3] = \text{Large positive value}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

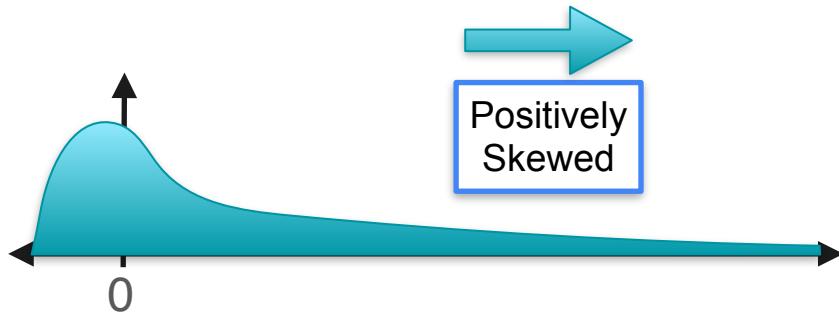
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\mathbb{E}[X_2^3] = \text{Large negative value}$$





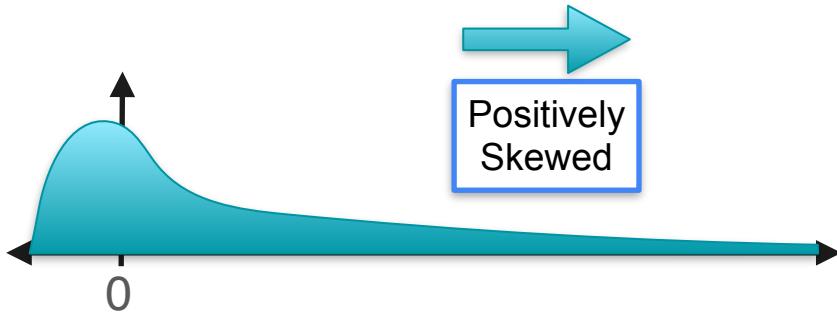
Ticket: \$1
Jackpot: \$99

Lottery

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_1^2] = 99$$

$$\mathbb{E}[X_1^3] = \text{Large positive value}$$



Cost: \$1
Crash Reparation: \$99

Car insurance

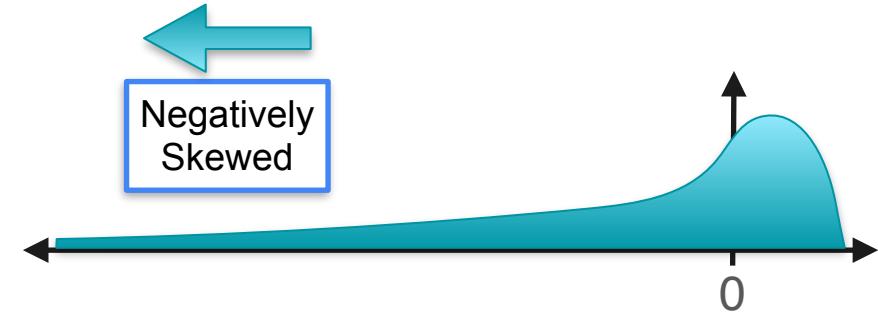
Same expectation
Same variance

How to tell them apart?

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_2^2] = 99$$

$$\mathbb{E}[X_2^3] = \text{Large negative value}$$



Skewness

$$\mathbb{E}[X^3]$$

Skewness

$$\mathbb{E}[X^3]$$

Almost...

Skewness

$$\mathbb{E}[X^3]$$

Almost...

Need to standardize...

Skewness

Skewness

$$\text{Skewness} = \mathbb{E} \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right]$$

Skewness

Skewness



Positively
Skewed



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] > 0$$

Skewness



Positively
Skewed



Not
Skewed



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] > 0$$

$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] = 0$$

Skewness



Positively
Skewed



Not
Skewed



Negatively
Skewed



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] > 0$$

$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] = 0$$

$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] < 0$$

Video 7c:

- Kurtosis

Kurtosis: Example

Game 1

Game 2

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

Game 2

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

probability $\frac{100}{202}$: You win 10 cents

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

probability $\frac{100}{202}$: You win 10 cents

probability $\frac{100}{202}$: You lose 10 cents

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

probability $\frac{100}{202}$: You win 10 cents

probability $\frac{100}{202}$: You lose 10 cents

probability $\frac{1}{202}$: You win 10 dollars

Kurtosis: Example

Game 1

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

probability $\frac{100}{202}$: You win 10 cents

probability $\frac{100}{202}$: You lose 10 cents

probability $\frac{1}{202}$: You win 10 dollars

probability $\frac{1}{202}$: You lose 10 dollars

Kurtosis: Example

Game 1

Which one
is riskier?

probability $\frac{1}{2}$: You win 1 dollar

probability $\frac{1}{2}$: You lose 1 dollar

Game 2

probability $\frac{100}{202}$: You win 10 cents

probability $\frac{100}{202}$: You lose 10 cents

probability $\frac{1}{202}$: You win 10 dollars

probability $\frac{1}{202}$: You lose 10 dollars

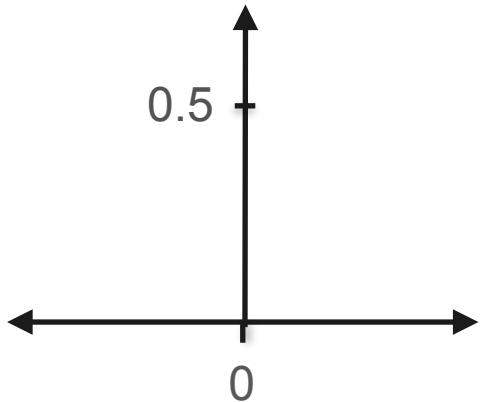
Kurtosis: Example

Game 1

Game 2

Kurtosis: Example

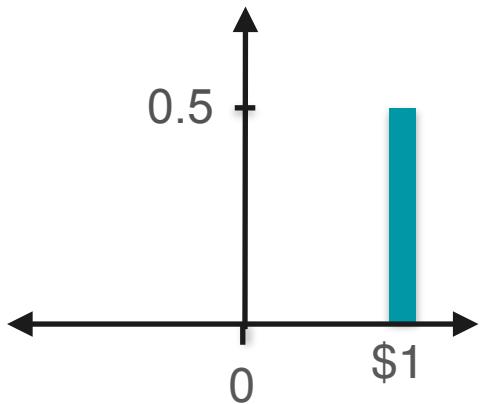
Game 1



Game 2

Kurtosis: Example

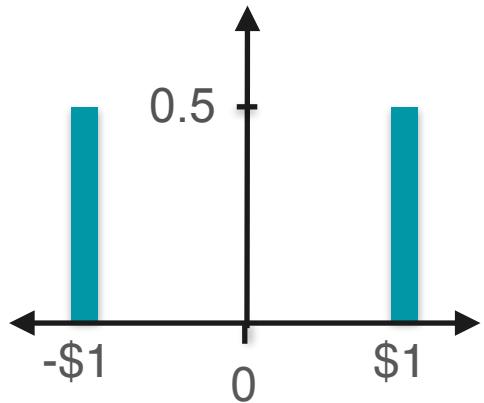
Game 1



Game 2

Kurtosis: Example

Game 1



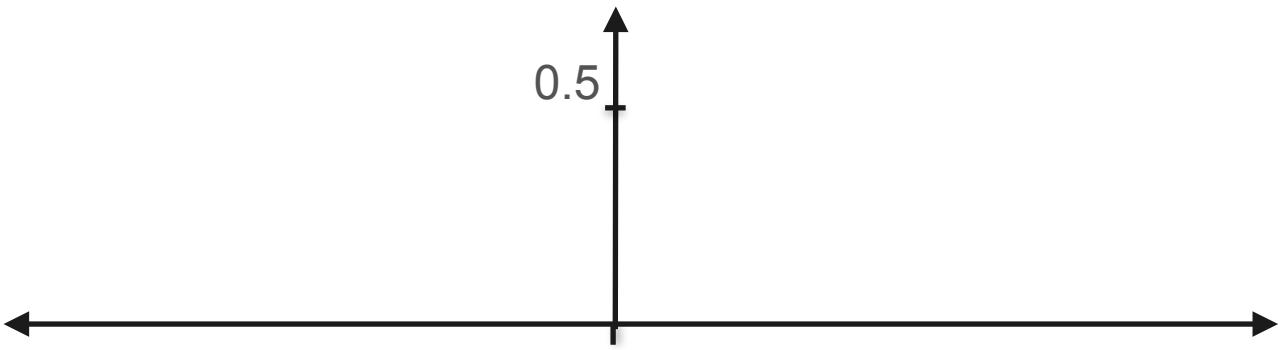
Game 2

Kurtosis: Example

Game 1



Game 2

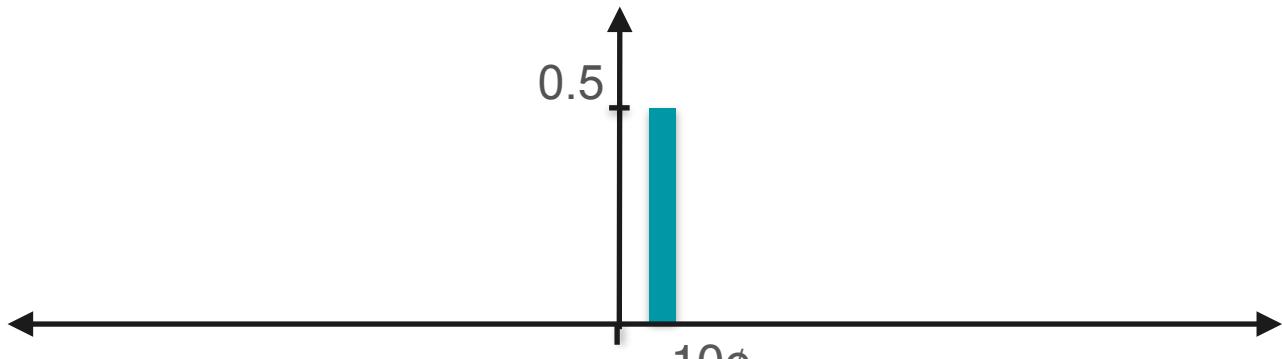


Kurtosis: Example

Game 1



Game 2

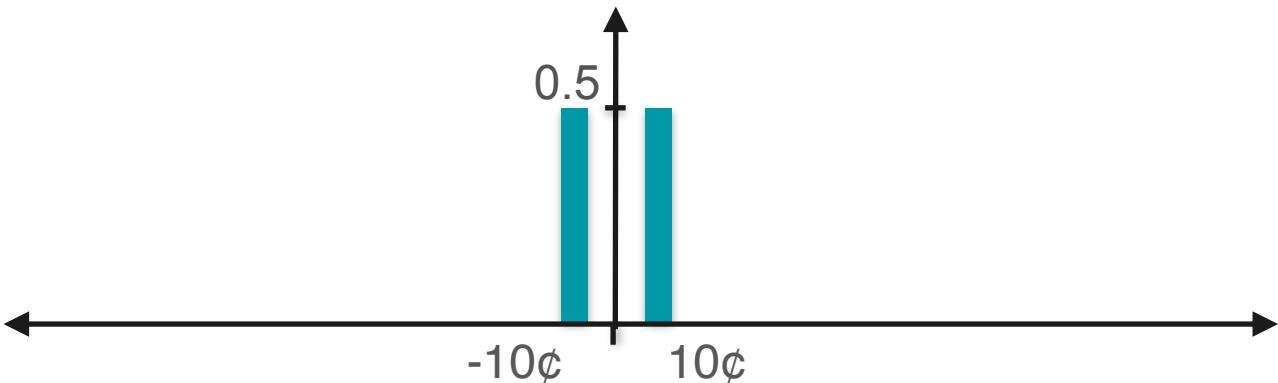


Kurtosis: Example

Game 1

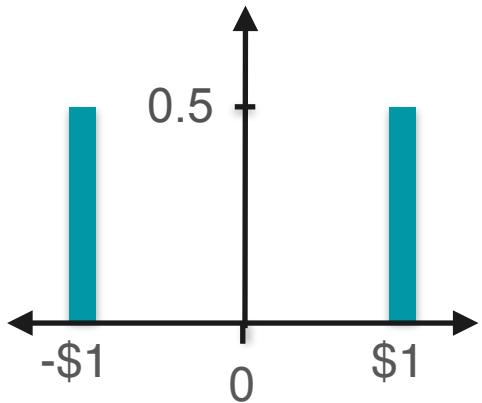


Game 2

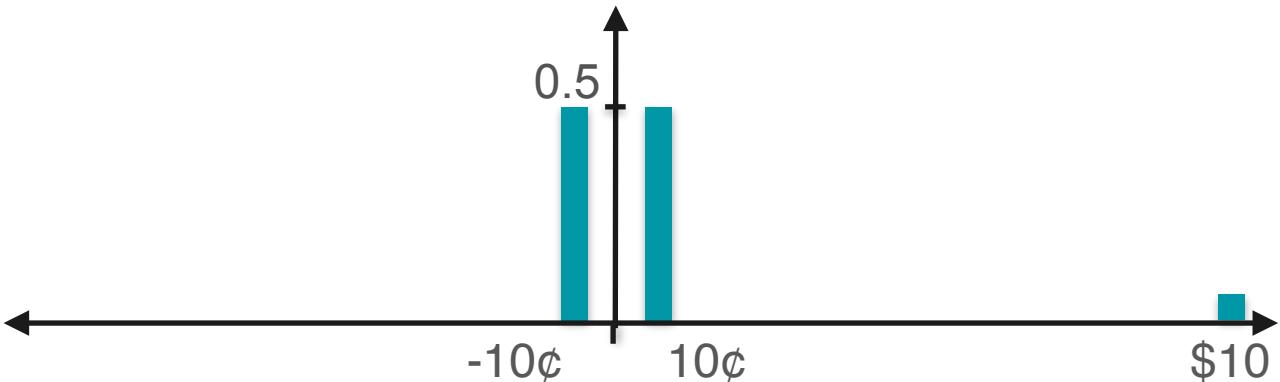


Kurtosis: Example

Game 1

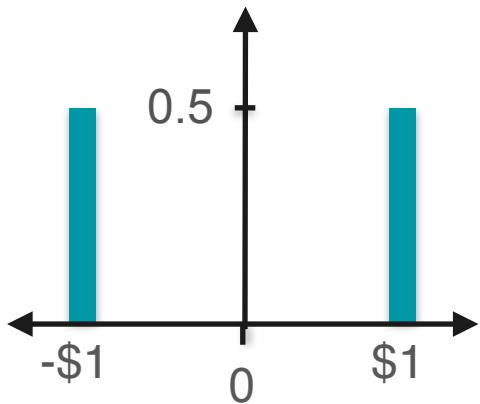


Game 2

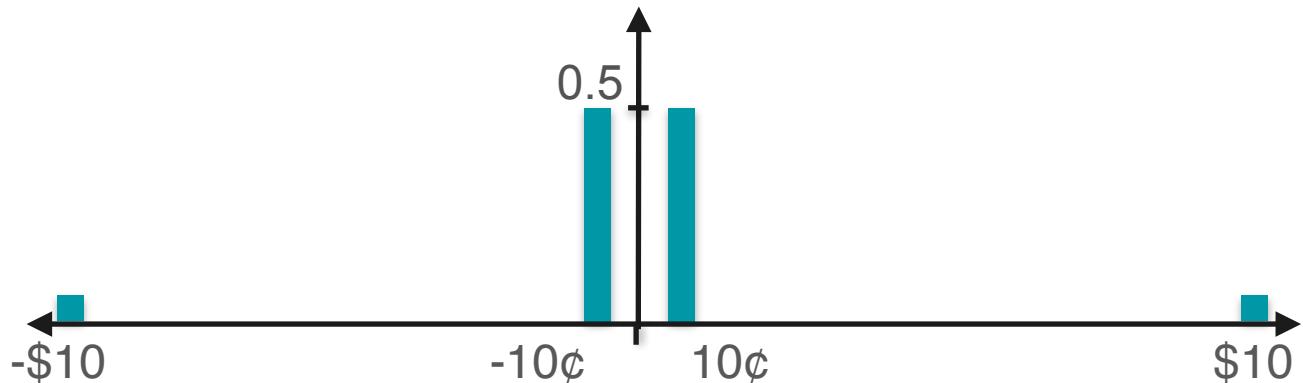


Kurtosis: Example

Game 1

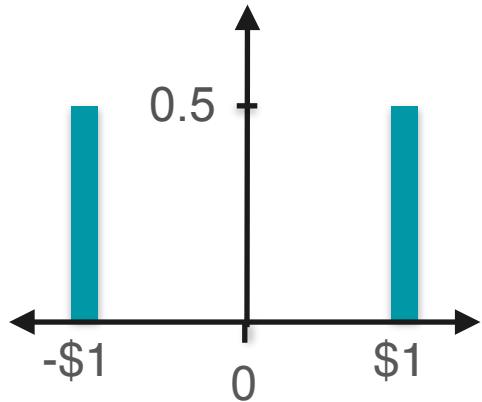


Game 2



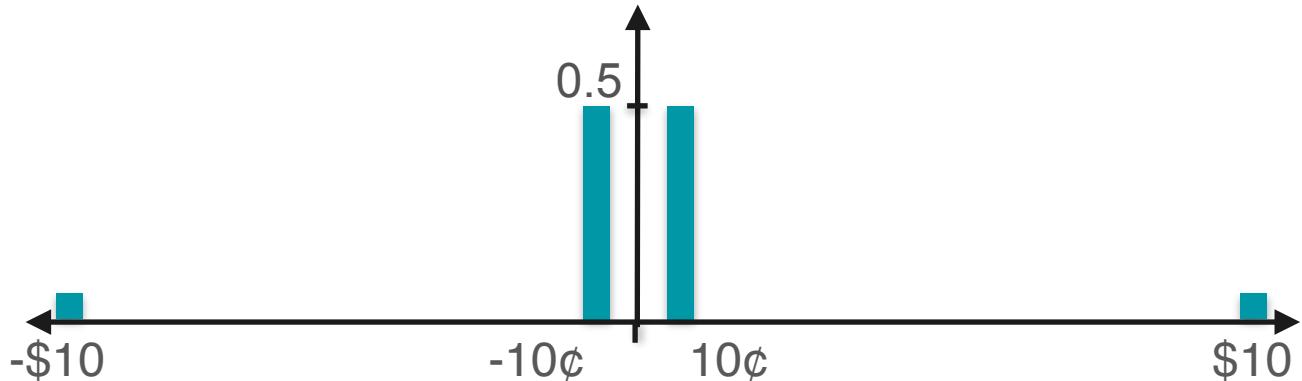
Kurtosis: Example

Game 1



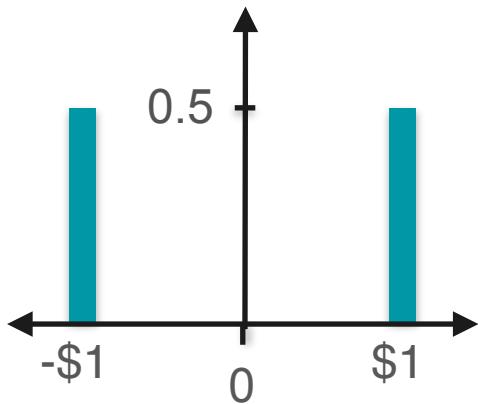
Expected value?

Game 2



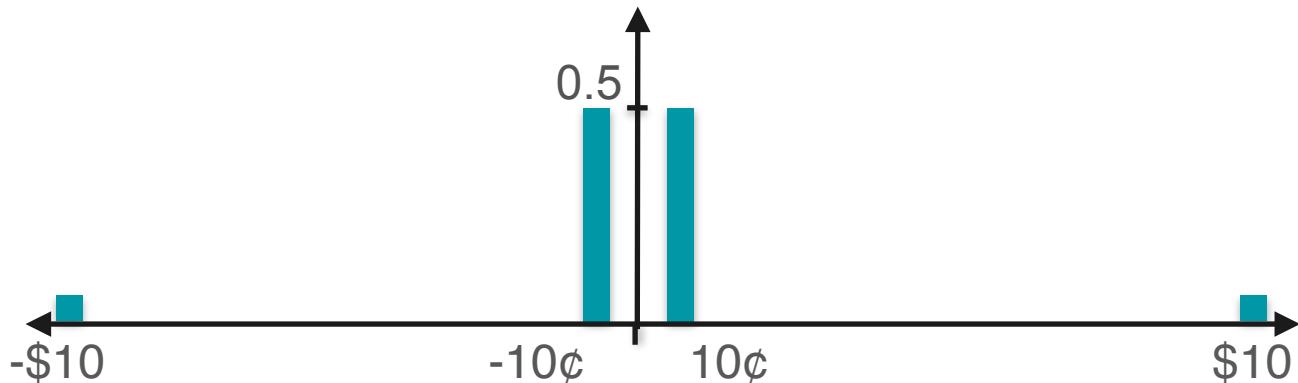
Kurtosis: Example

Game 1



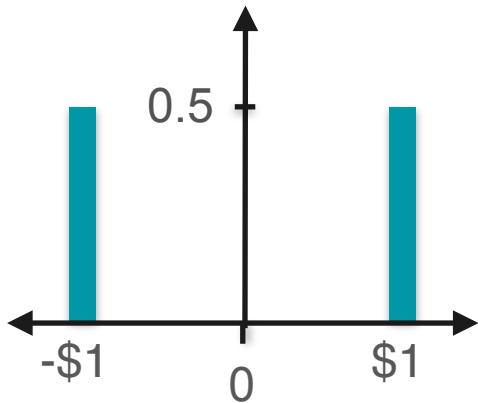
Expected value?
Standard deviation?

Game 2



Kurtosis: Example

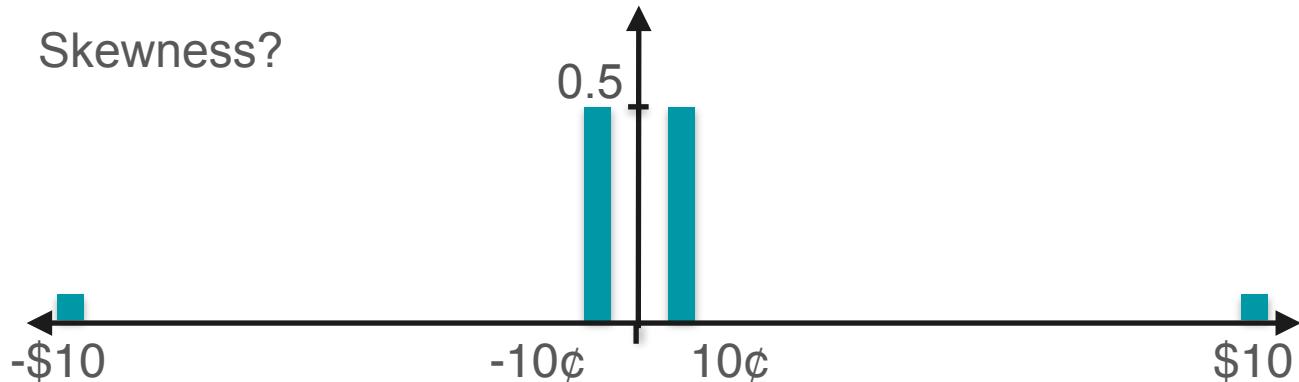
Game 1



Expected value?
Standard deviation?

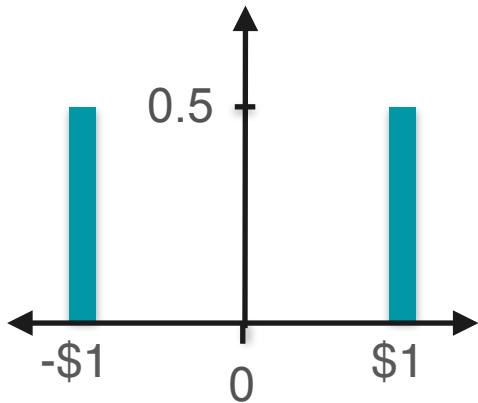
Skewness?

Game 2

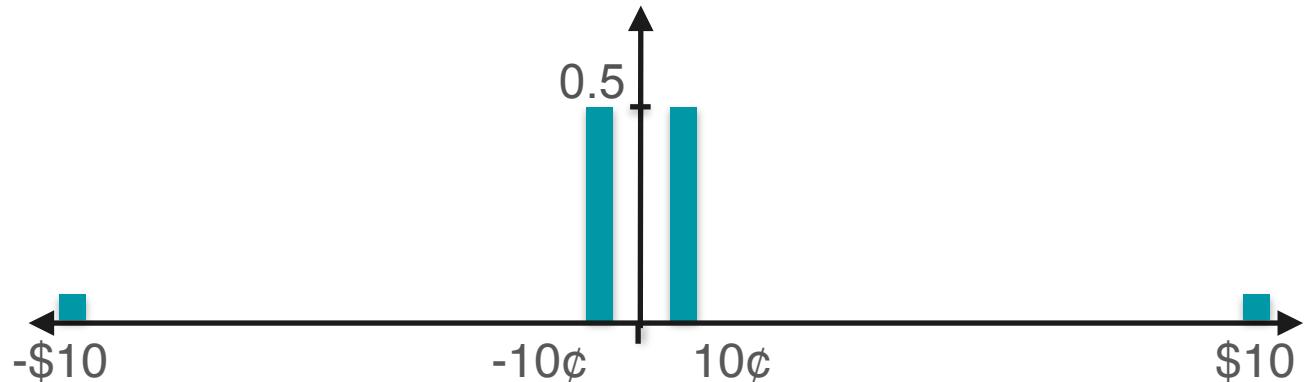


Kurtosis: Example Expected Value

Game 1

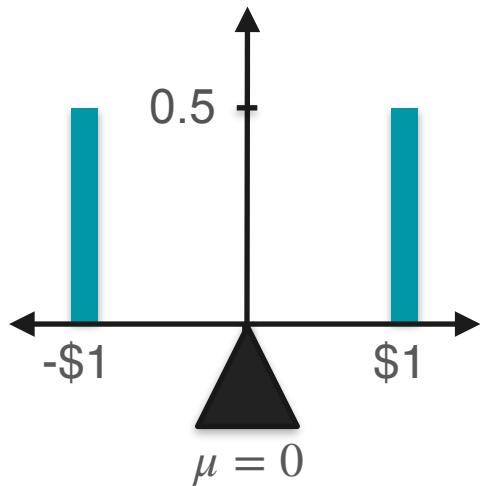


Game 2

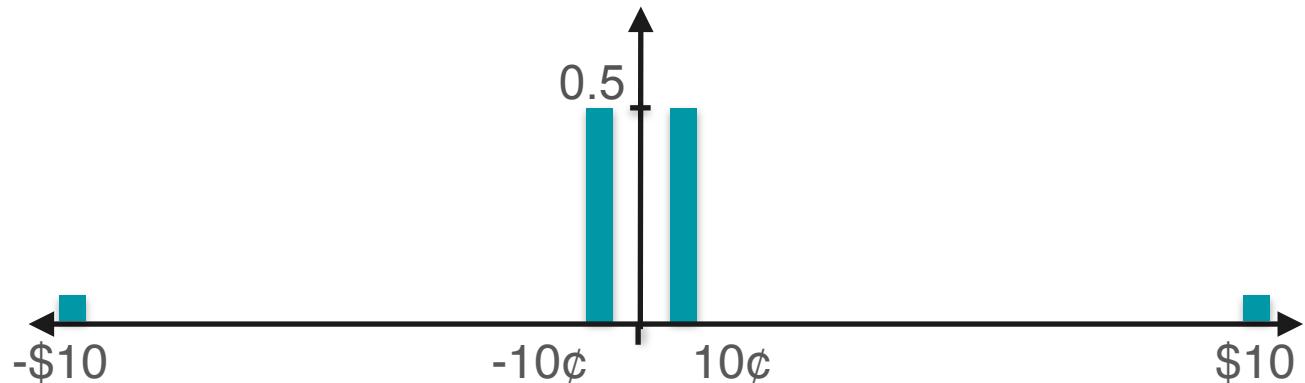


Kurtosis: Example Expected Value

Game 1

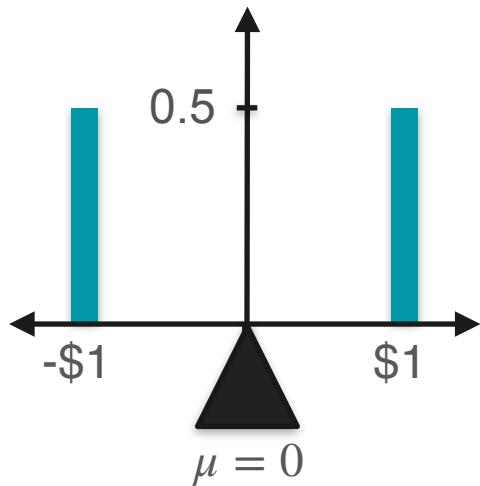


Game 2

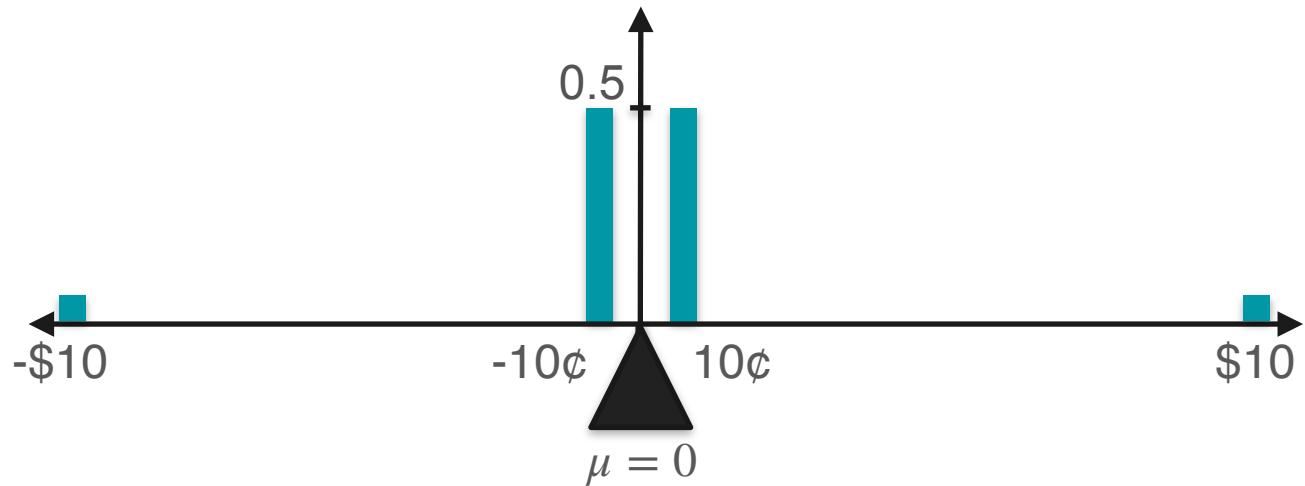


Kurtosis: Example Expected Value

Game 1



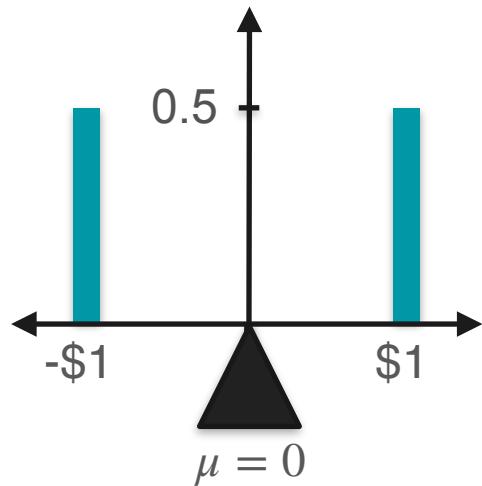
Game 2



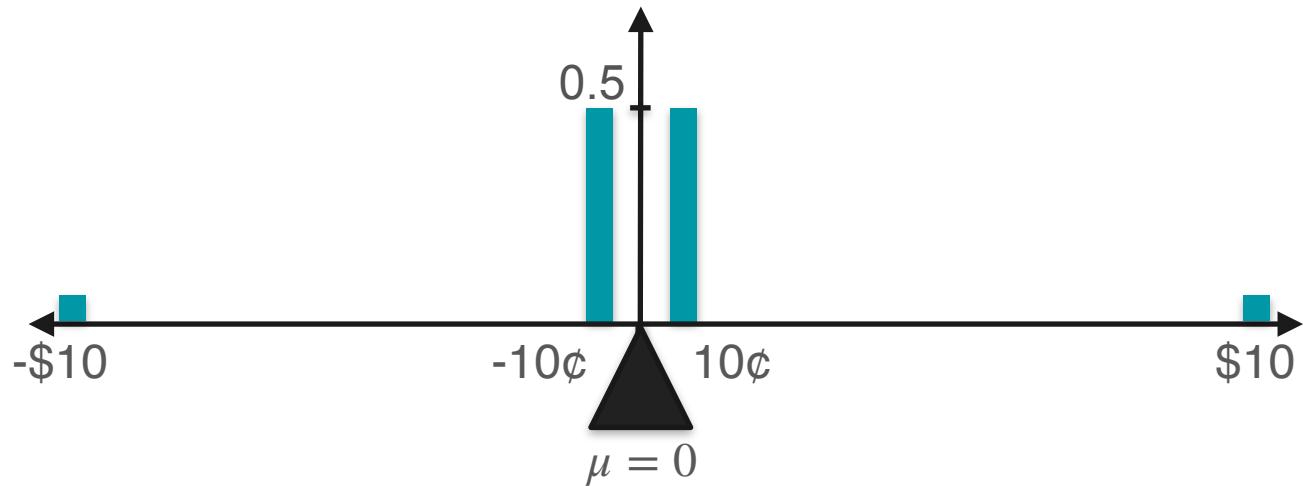
Kurtosis: Example Expected Value

Game 1

$$\mathbb{E}[X_1] = 0$$



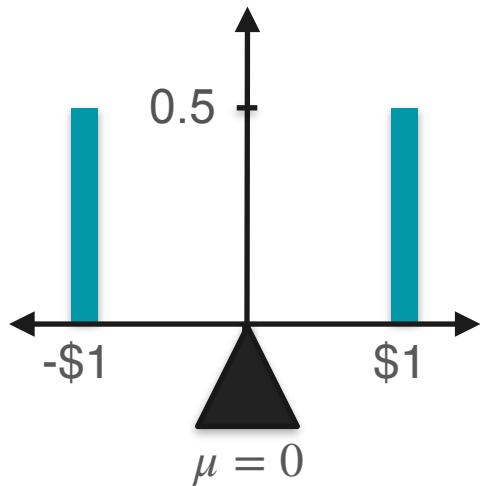
Game 2



Kurtosis: Example Expected Value

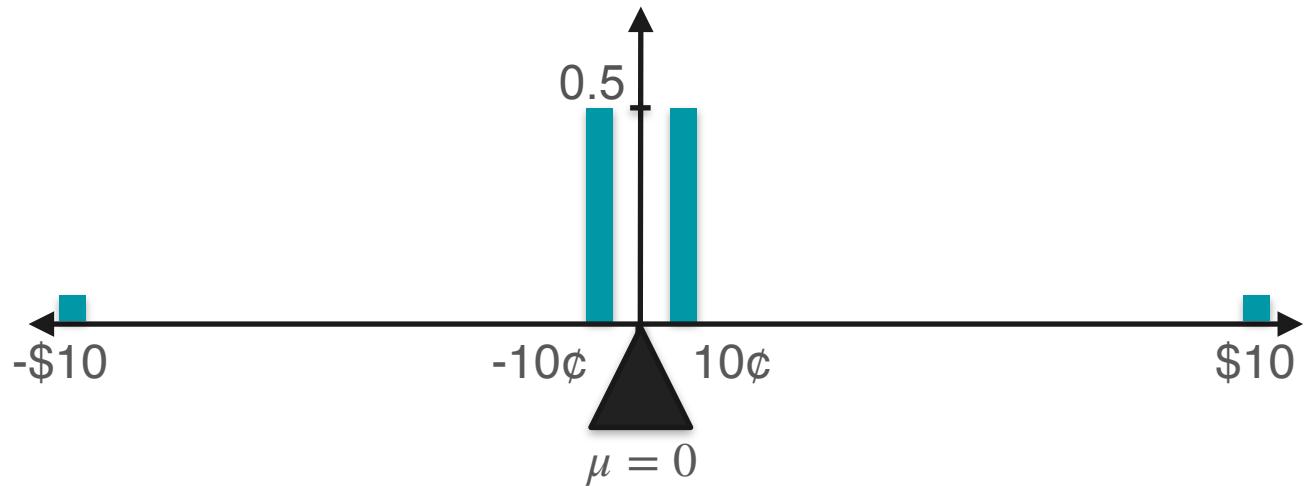
Game 1

$$\mathbb{E}[X_1] = 0$$

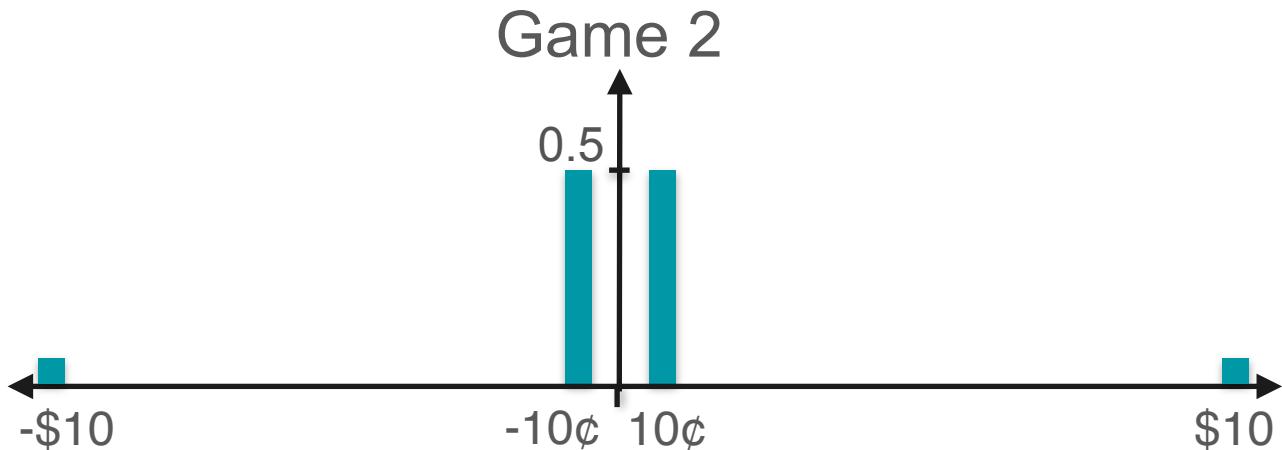
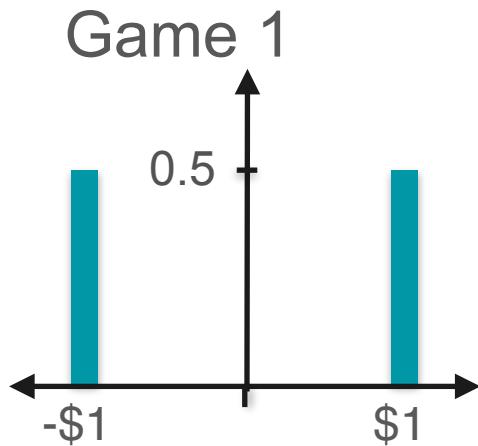


Game 2

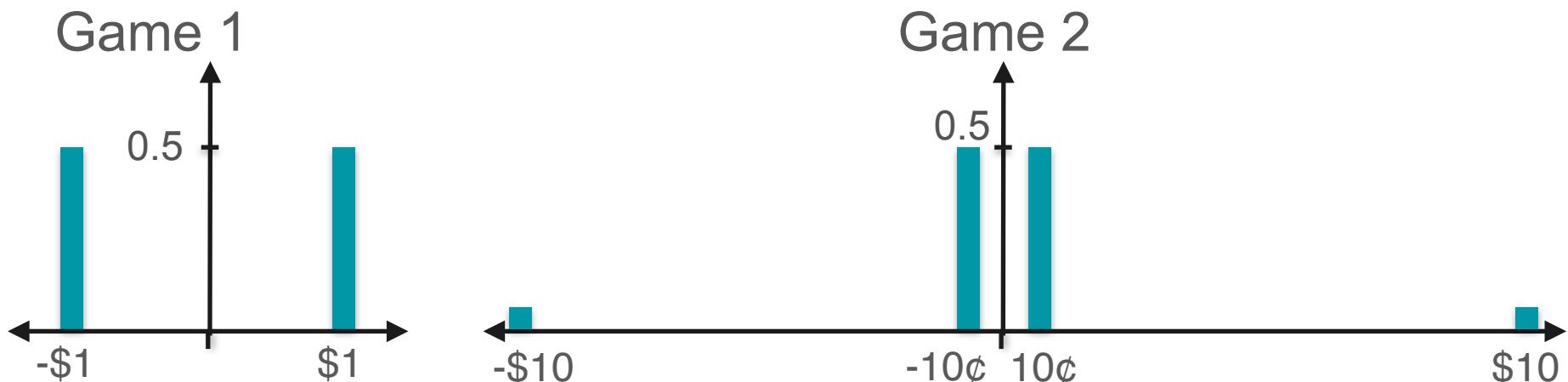
$$\mathbb{E}[X_2] = 0$$



Kurtosis: Example Variance

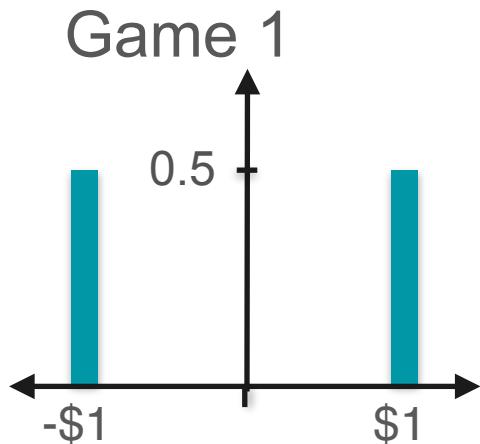


Kurtosis: Example Variance

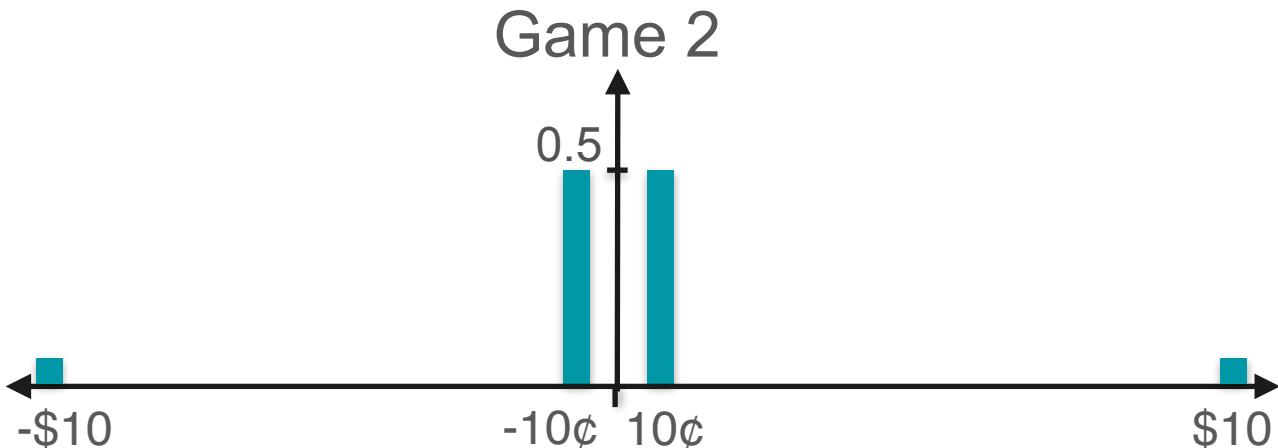


$$\mathbb{E}[X_1^2]$$

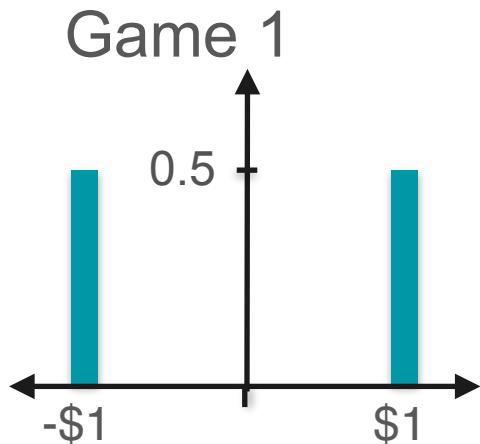
Kurtosis: Example Variance



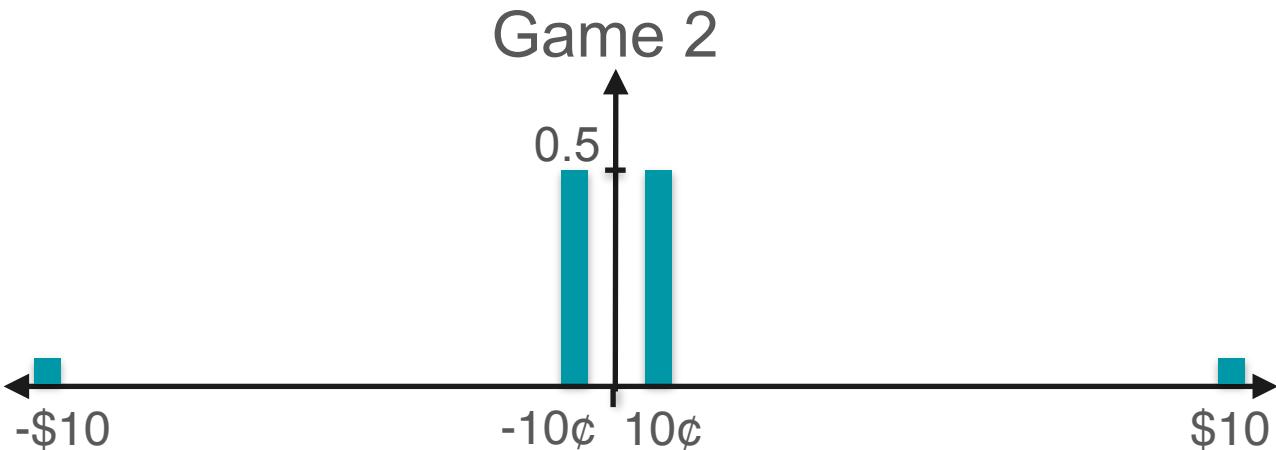
$$\mathbb{E}[X_1^2] = \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2$$



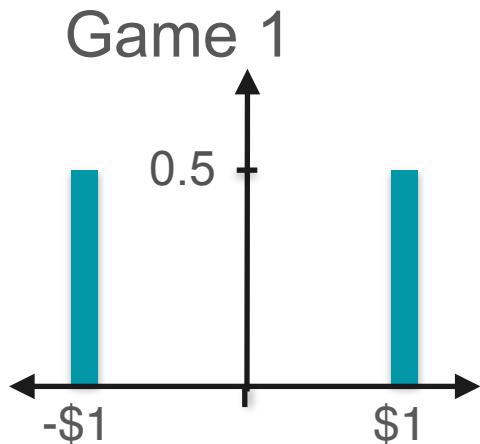
Kurtosis: Example Variance



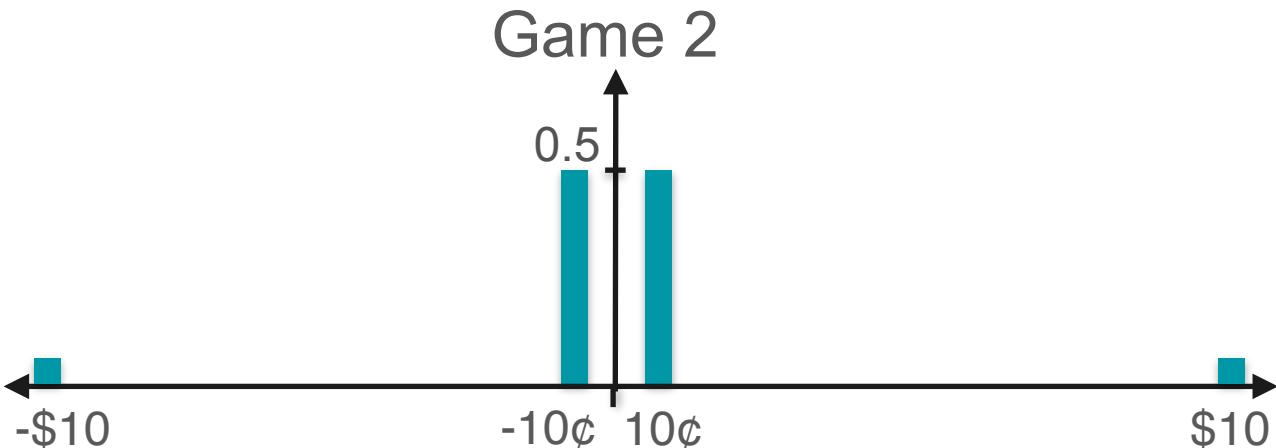
$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$



Kurtosis: Example Variance

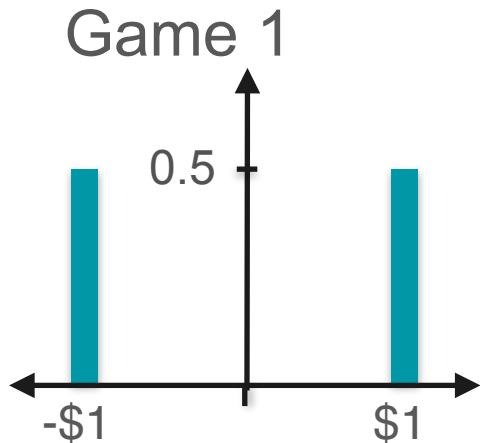


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

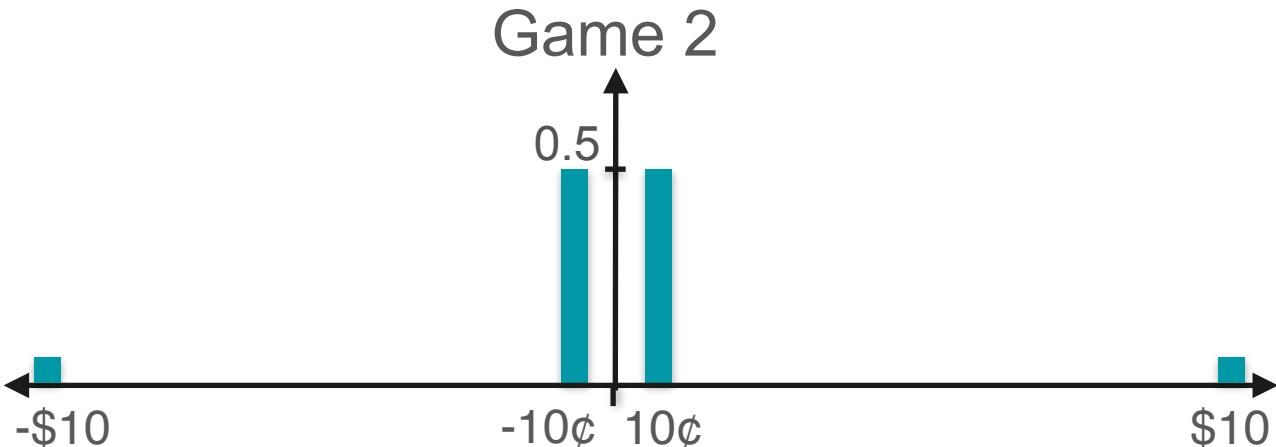


$$\mathbb{E}[X_2^2]$$

Kurtosis: Example Variance

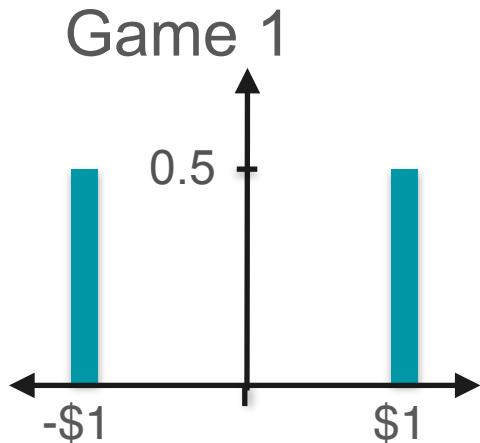


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

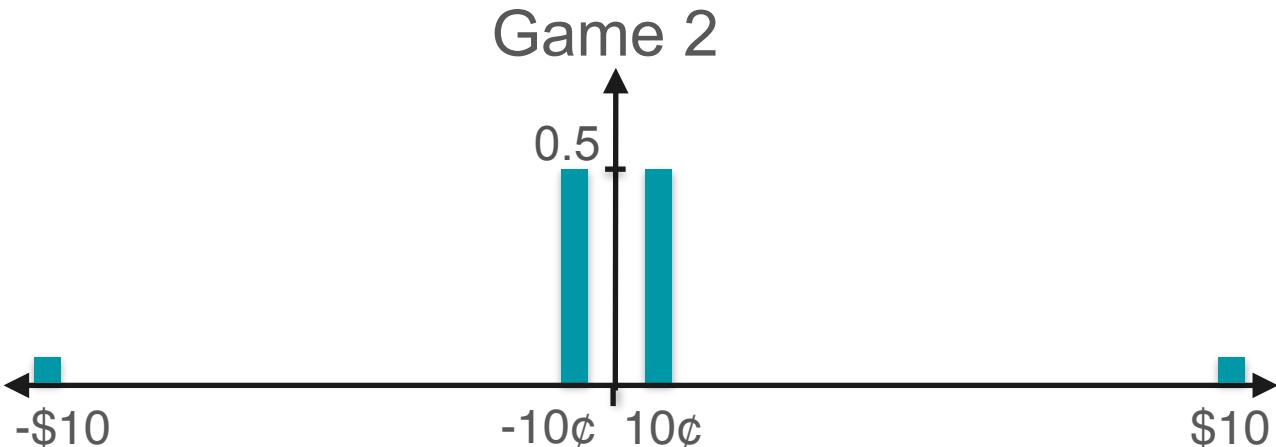


$$\mathbb{E}[X_2^2] = \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2$$

Kurtosis: Example Variance

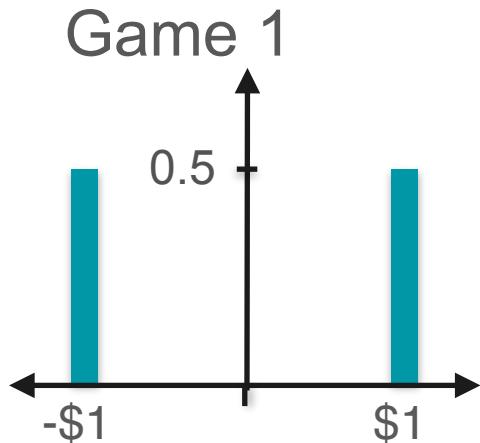


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

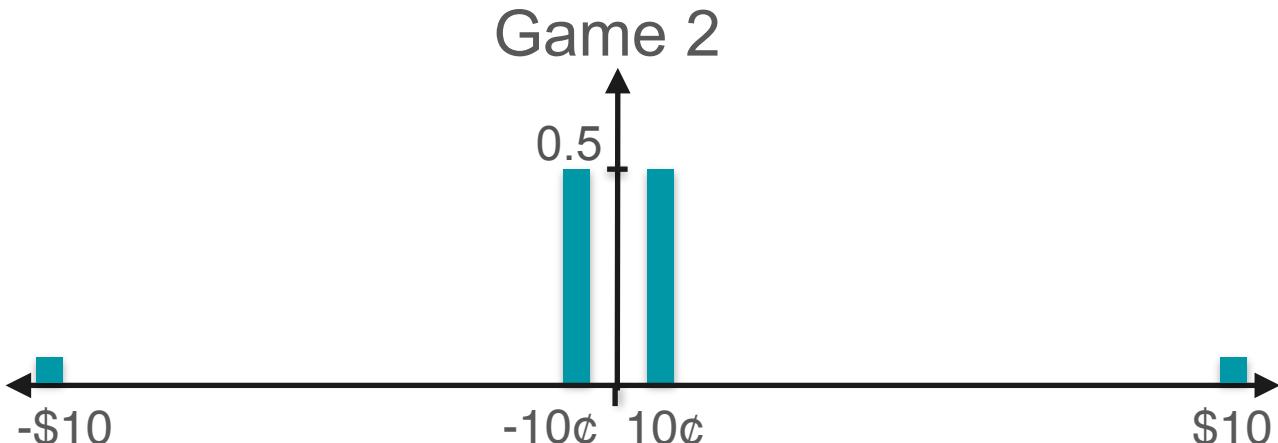


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100\end{aligned}$$

Kurtosis: Example Variance

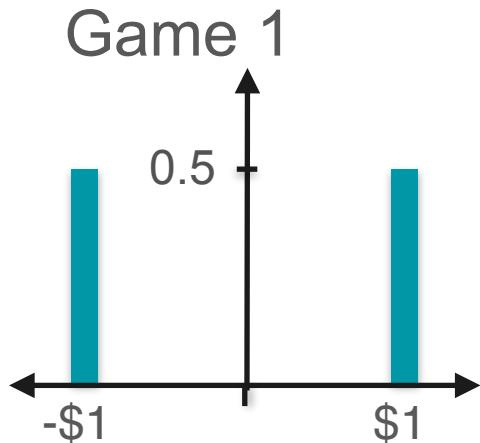


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

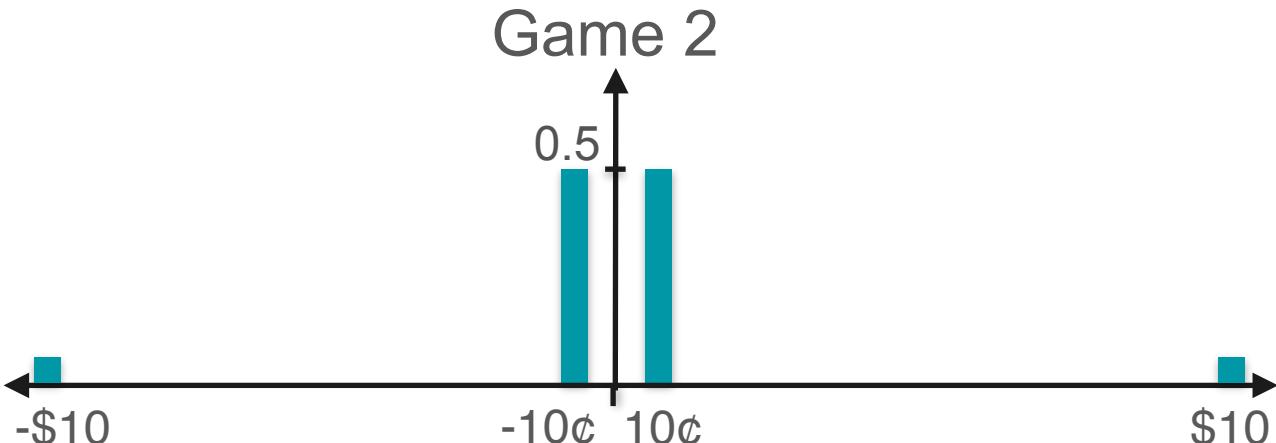


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202}\end{aligned}$$

Kurtosis: Example Variance

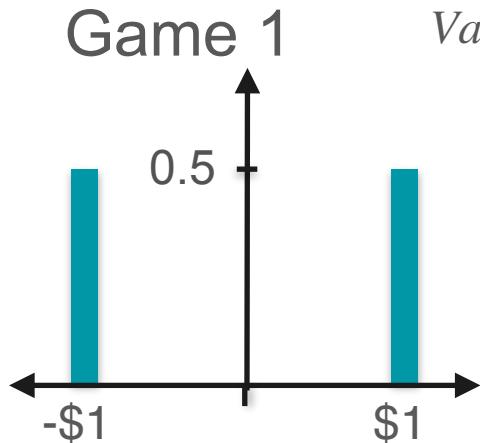


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

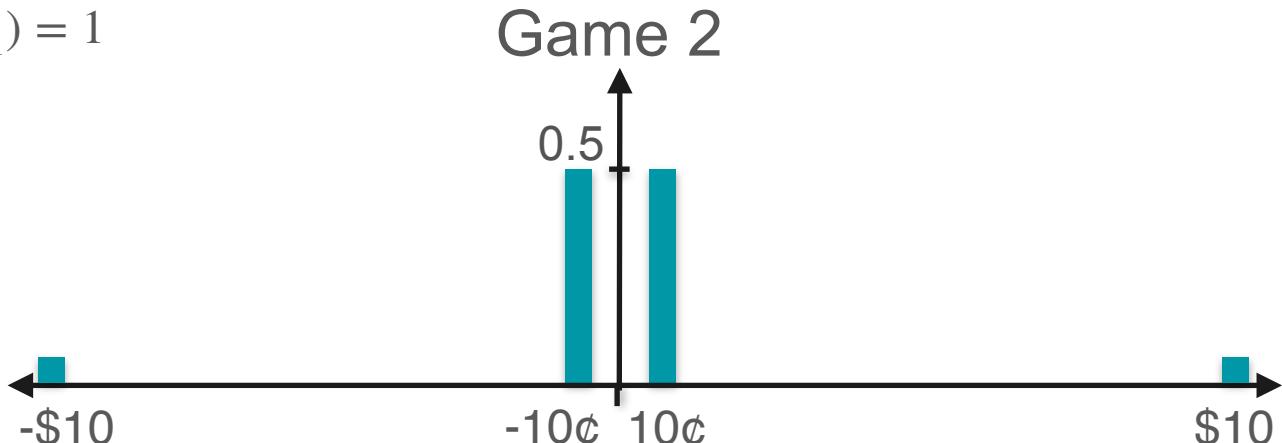


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202} = 1\end{aligned}$$

Kurtosis: Example Variance

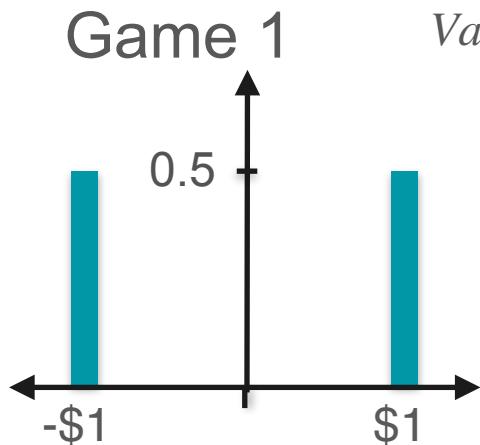


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

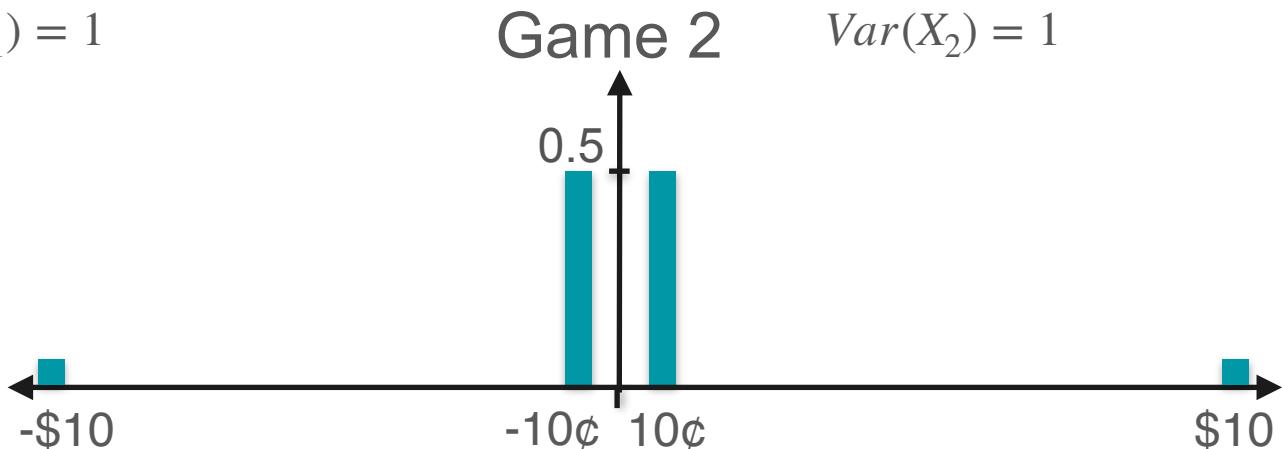


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202} = 1\end{aligned}$$

Kurtosis: Example Variance

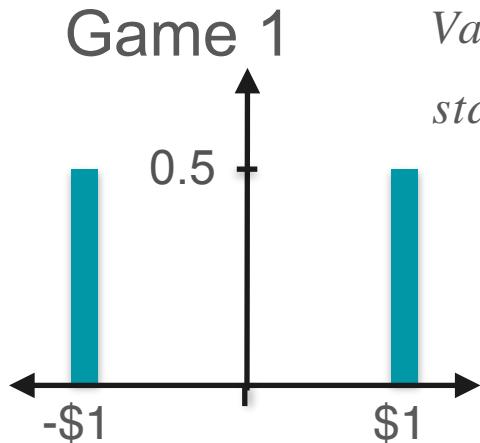


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

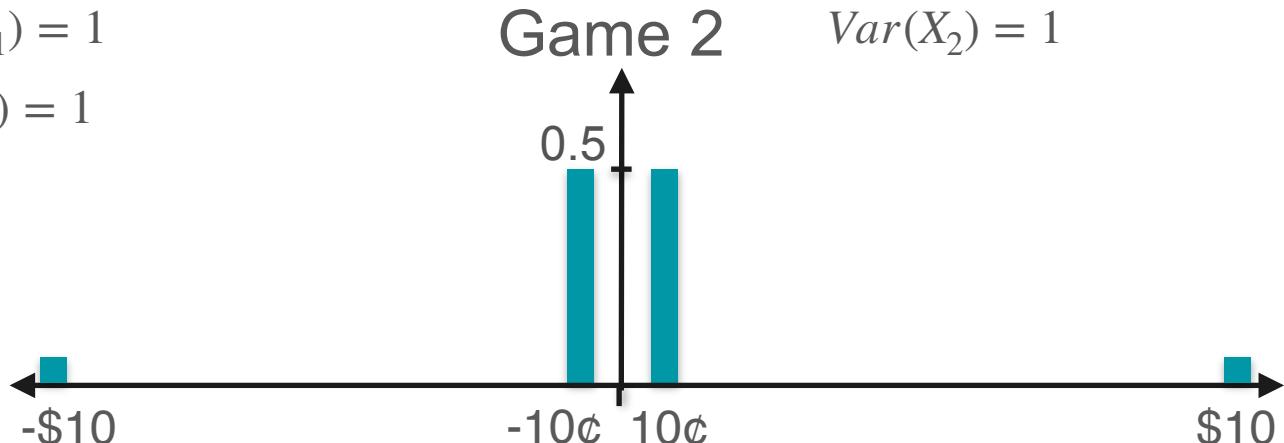


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202} = 1\end{aligned}$$

Kurtosis: Example Variance

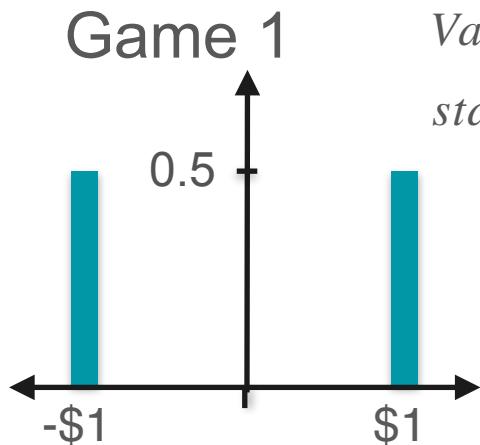


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$

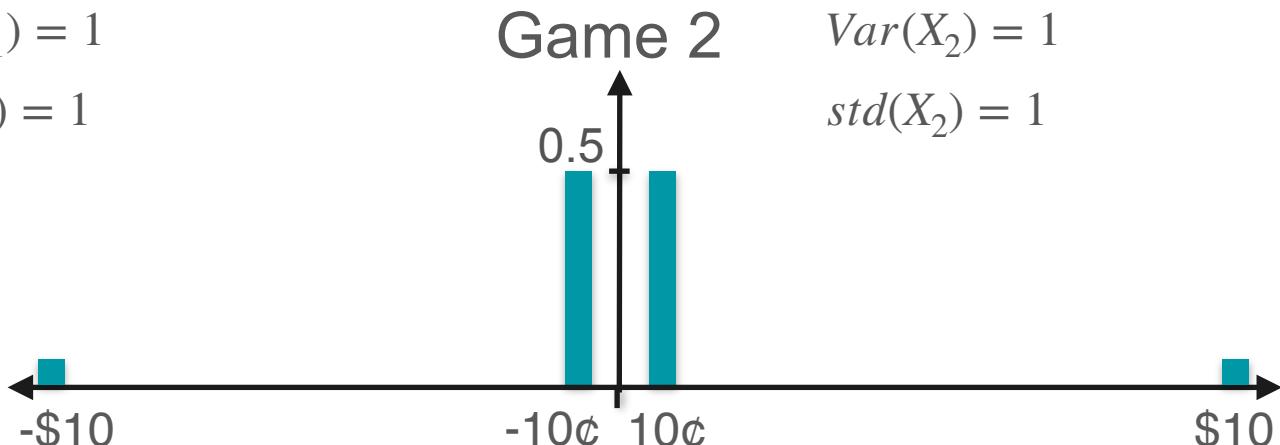


$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202} = 1\end{aligned}$$

Kurtosis: Example Variance

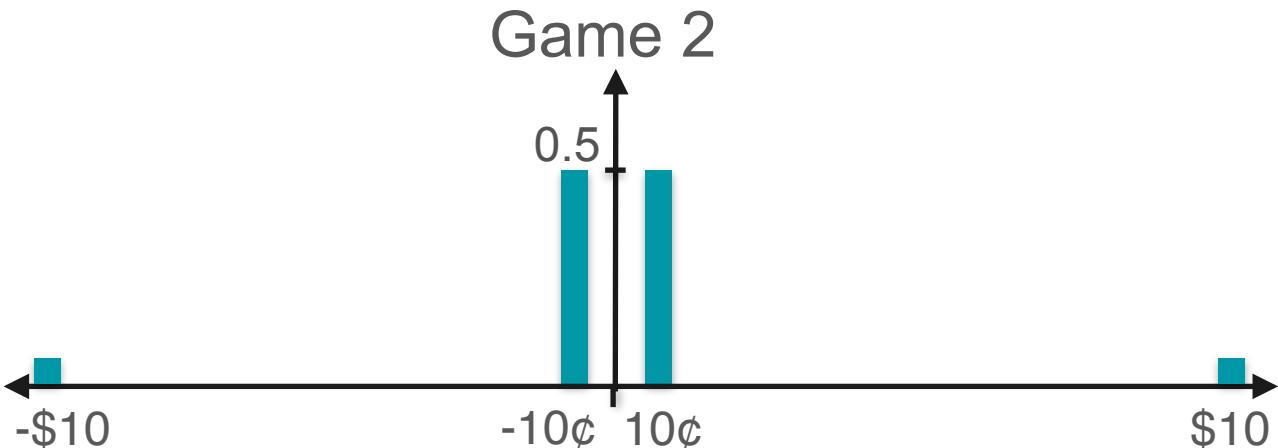
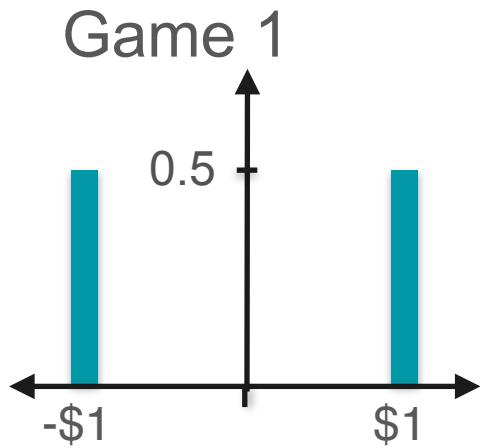


$$\begin{aligned}\mathbb{E}[X_1^2] &= \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 \\ &= 1\end{aligned}$$



$$\begin{aligned}\mathbb{E}[X_2^2] &= \frac{100}{202}(-0.1)^2 + \frac{100}{202}(0.1)^2 + \frac{1}{202}(-10)^2 + \frac{1}{202}(10)^2 \\ &= \frac{100}{202} \cdot \frac{1}{100} + \frac{100}{202} \cdot \frac{1}{100} + \frac{1}{202} \cdot 100 + \frac{1}{202} \cdot 100 \\ &= \frac{1}{202} + \frac{1}{202} + \frac{100}{202} + \frac{100}{202} = 1\end{aligned}$$

Kurtosis: Example Skewness



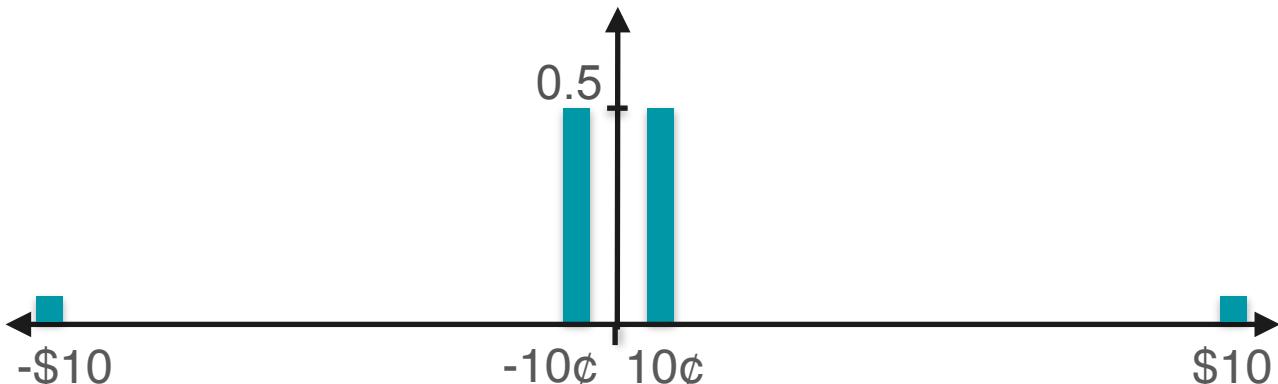
Kurtosis: Example Skewness

Game 1

$$Skew(X_1) = 0$$



Game 2



Kurtosis: Example Skewness

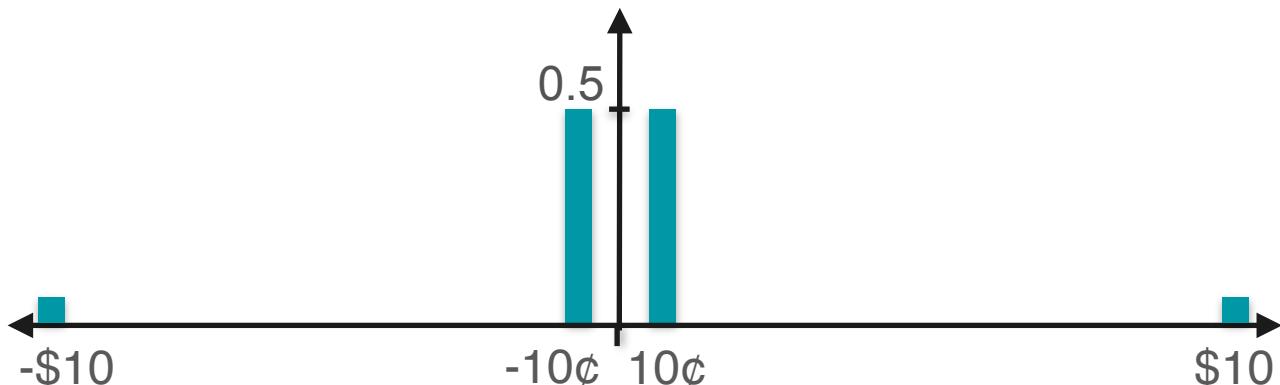
Game 1

$$Skew(X_1) = 0$$



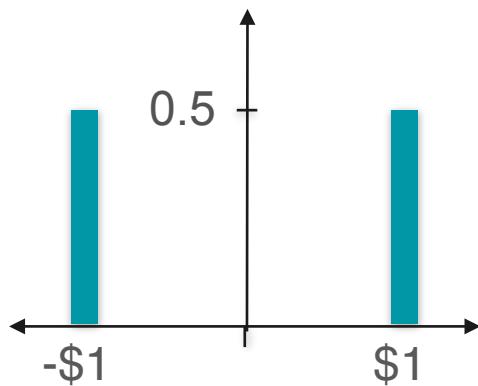
Game 2

$$Skew(X_2) = 0$$

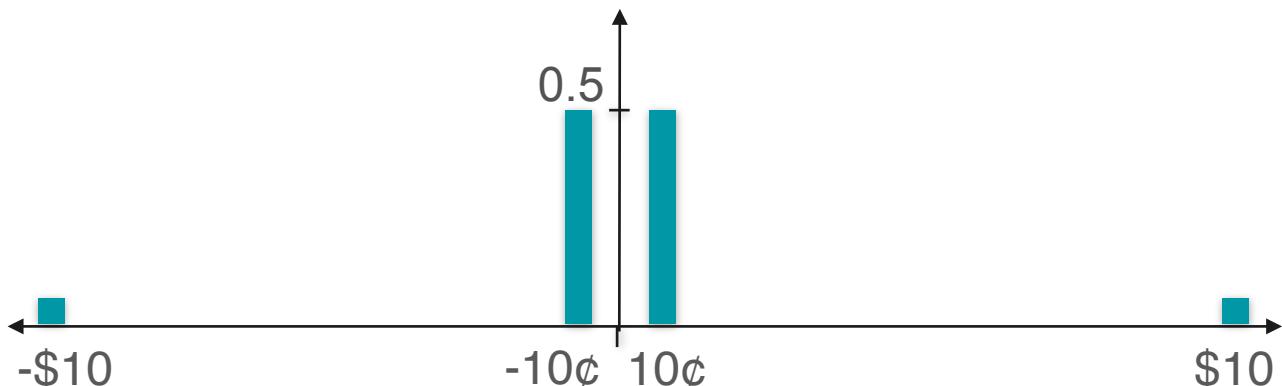


Kurtosis

Game 1



Game 2

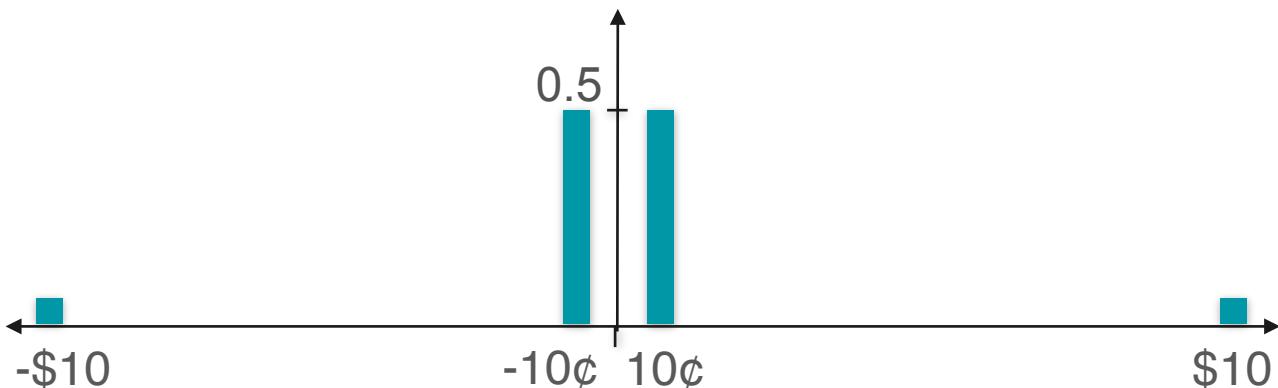


Kurtosis

Game 1



Game 2



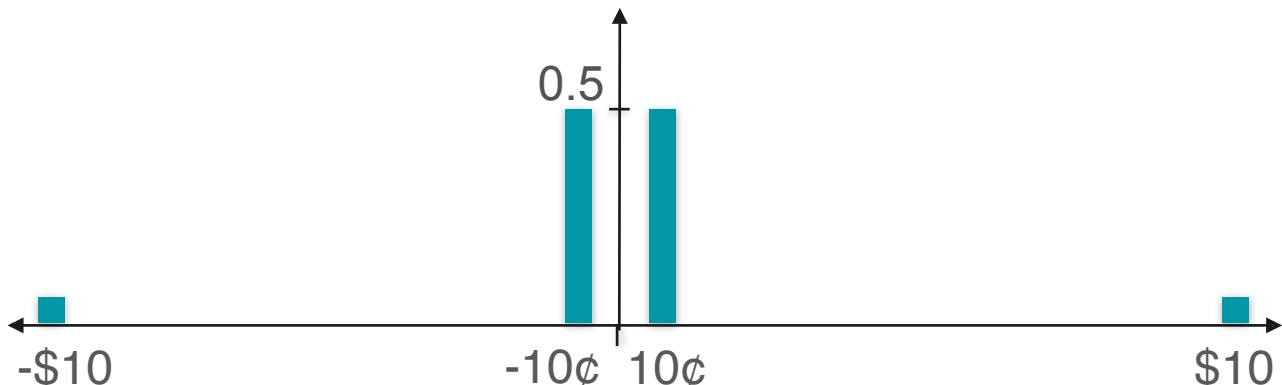
$$Skew(X_1) = 0$$

Kurtosis

Game 1



Game 2



$$Skew(X_1) = 0$$

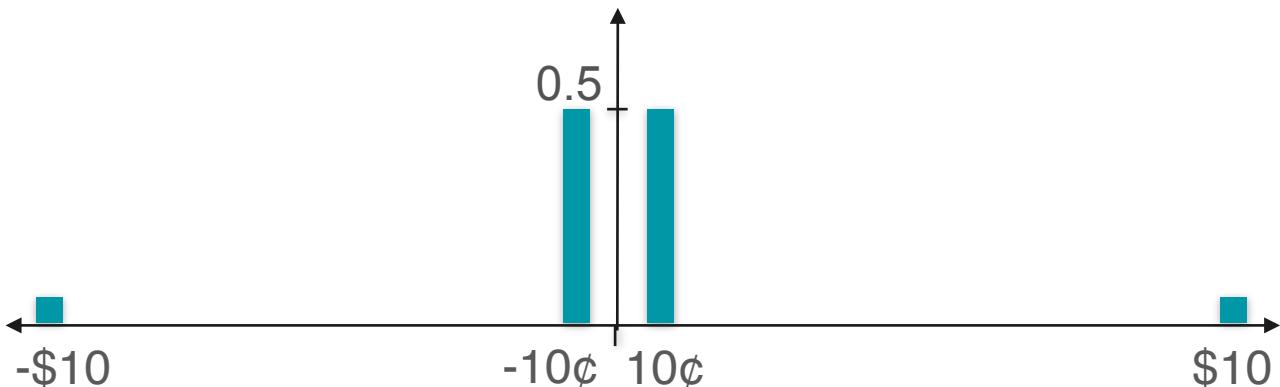
$$Skew(X_2) = 0$$

Kurtosis

Game 1



Game 2



$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$

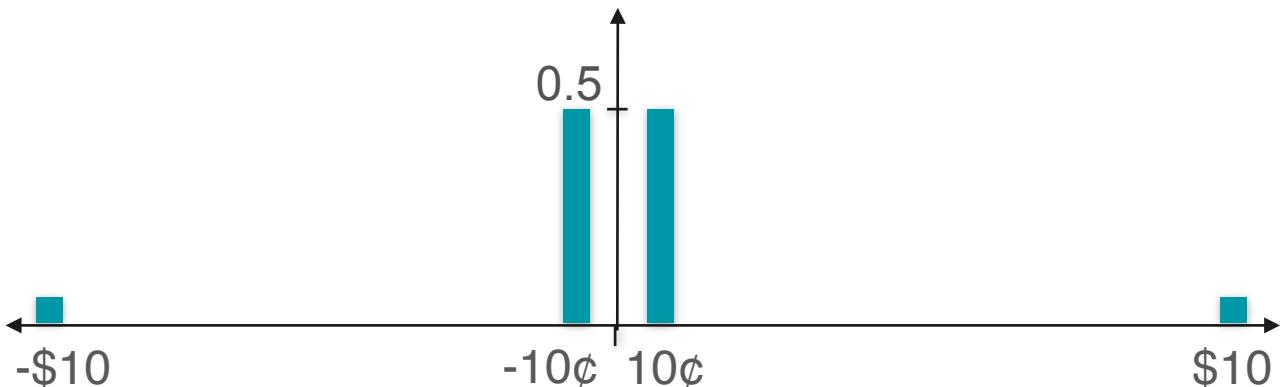
$$Skew(X_2) = 0$$

Kurtosis

Game 1



Game 2



$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$

$$Var(X_2) = 1$$

$$Skew(X_2) = 0$$

Kurtosis

Game 1

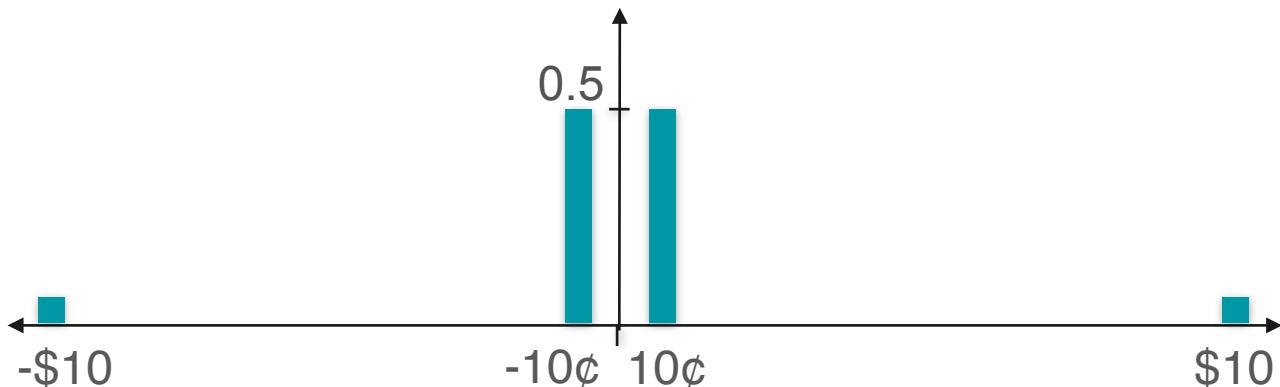


$$E[X_1] = 0$$

$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$

Game 2



$$Var(X_2) = 1$$

$$Skew(X_2) = 0$$

Kurtosis

Game 1

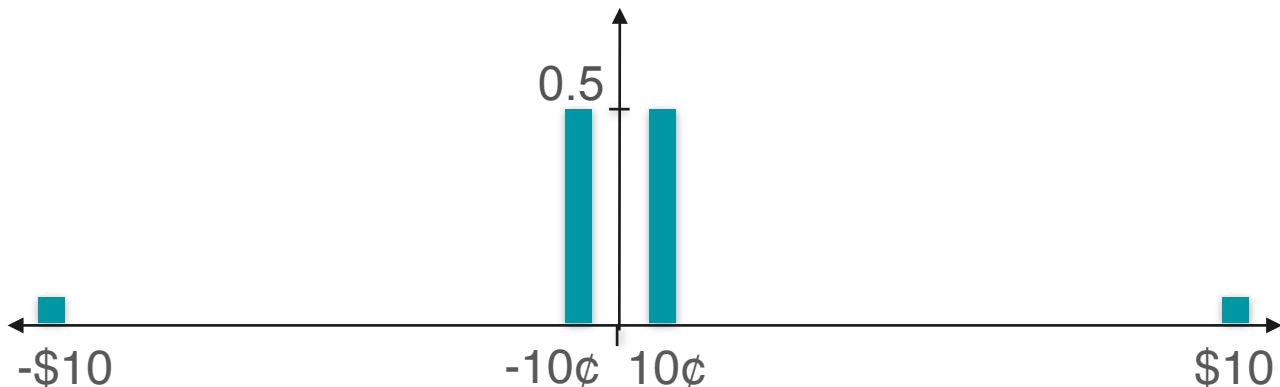


$$E[X_1] = 0$$

$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$

Game 2

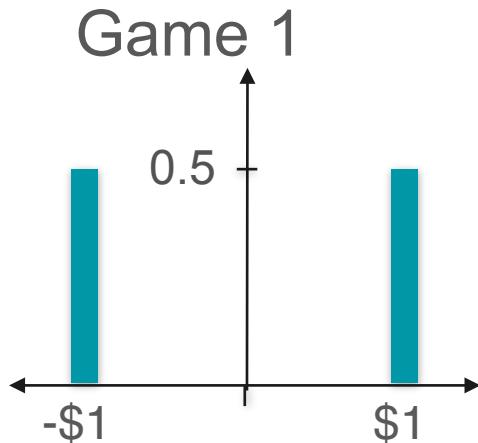


$$E[X_2] = 0$$

$$Var(X_2) = 1$$

$$Skew(X_2) = 0$$

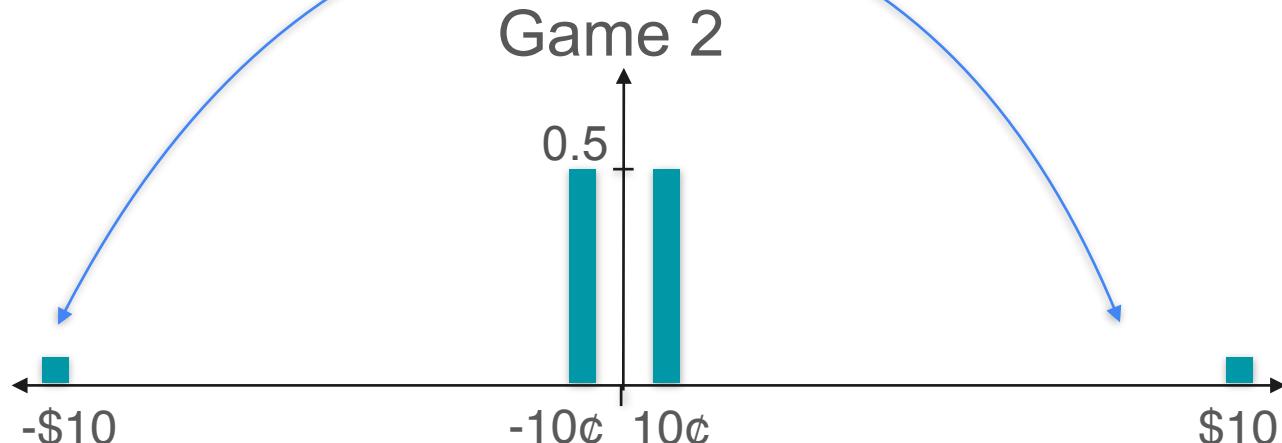
Kurtosis



$$E[X_1] = 0$$

$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$



$$E[X_2] = 0$$

$$Var(X_2) = 1$$

$$Skew(X_2) = 0$$

Has values way
farther from 0

Kurtosis

Game 1



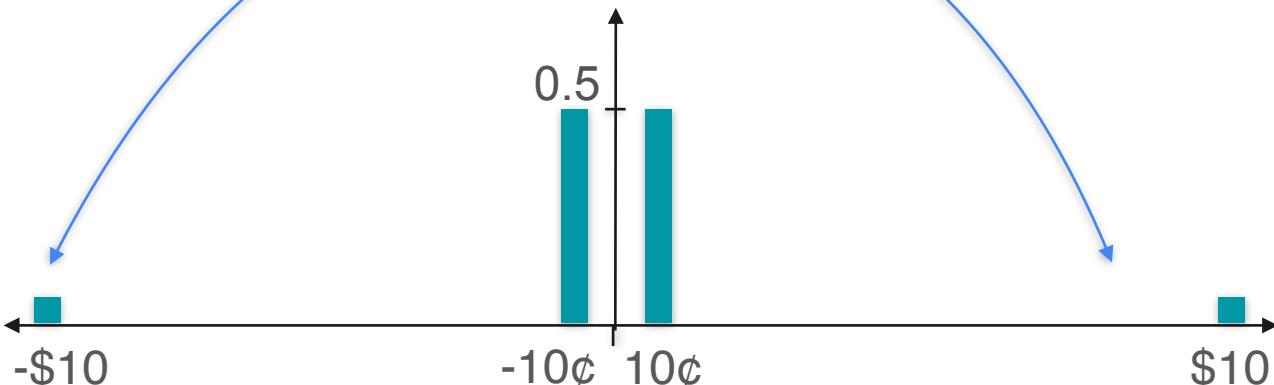
$$E[X_1] = 0$$

$$E[X_1] = 0$$

$$Var(X_1) = 1$$

$$Skew(X_1) = 0$$

Game 2



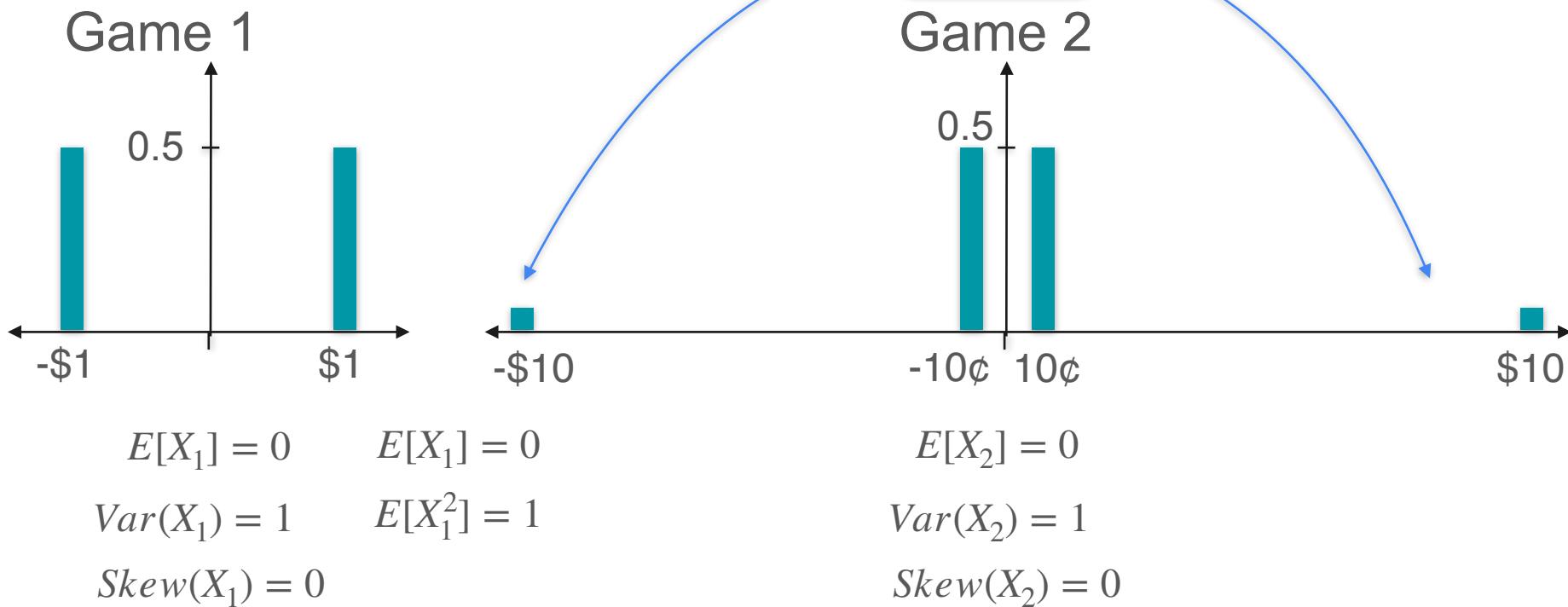
$$E[X_2] = 0$$

$$Var(X_2) = 1$$

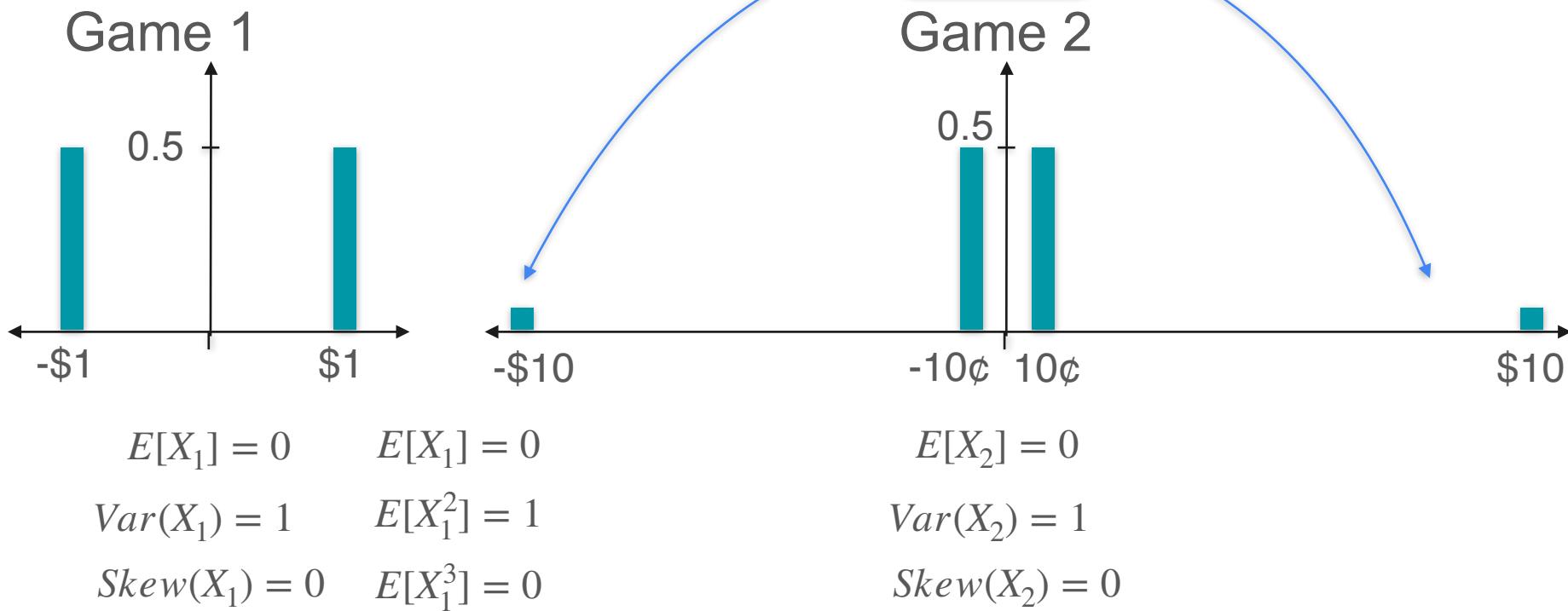
$$Skew(X_2) = 0$$

Has values way
farther from 0

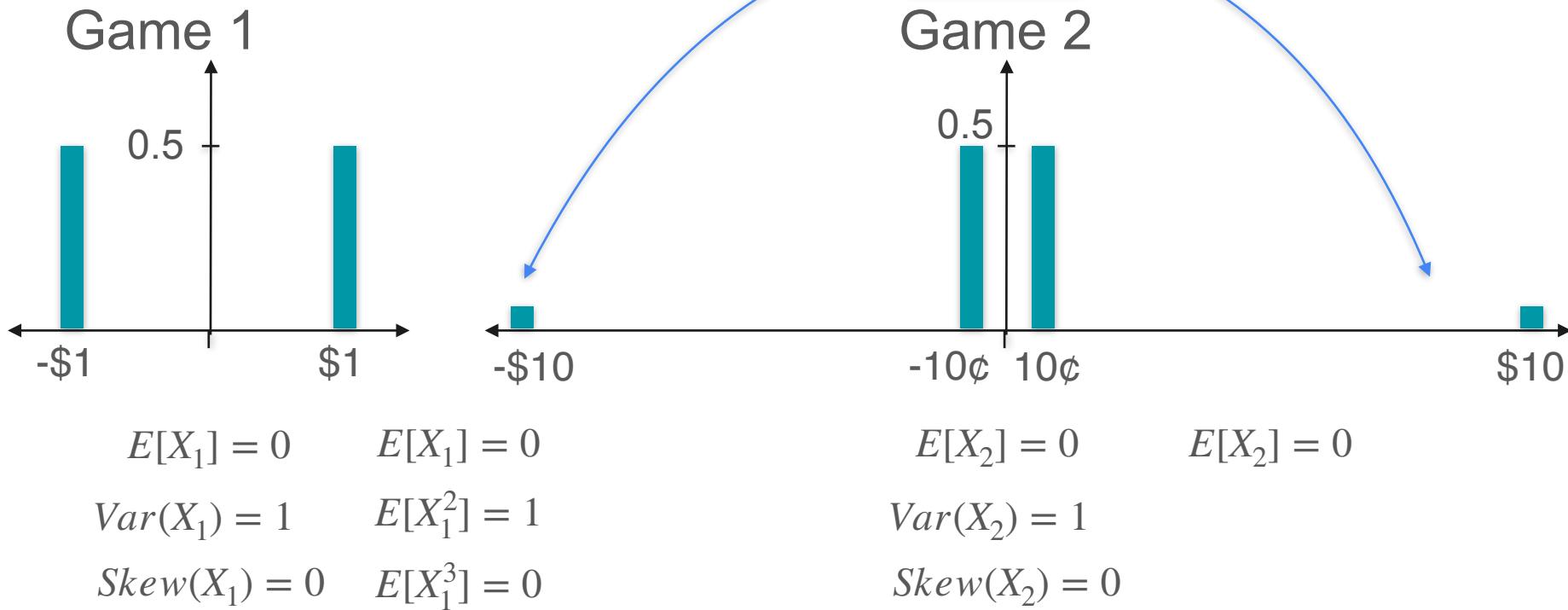
Kurtosis



Kurtosis



Kurtosis



Kurtosis

Game 1



$$E[X_1] = 0$$

$$E[X_1] = 0$$

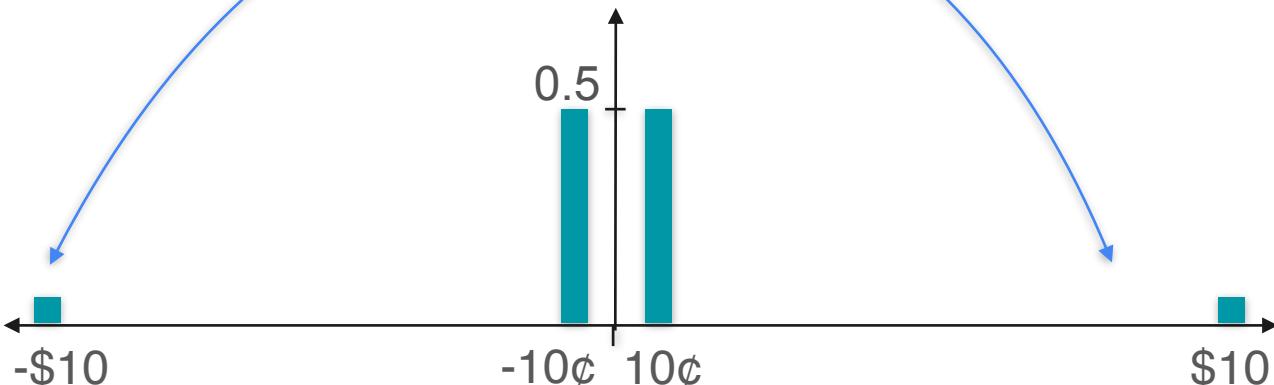
$$Var(X_1) = 1$$

$$E[X_1^2] = 1$$

$$Skew(X_1) = 0$$

$$E[X_1^3] = 0$$

Game 2



$$E[X_2] = 0$$

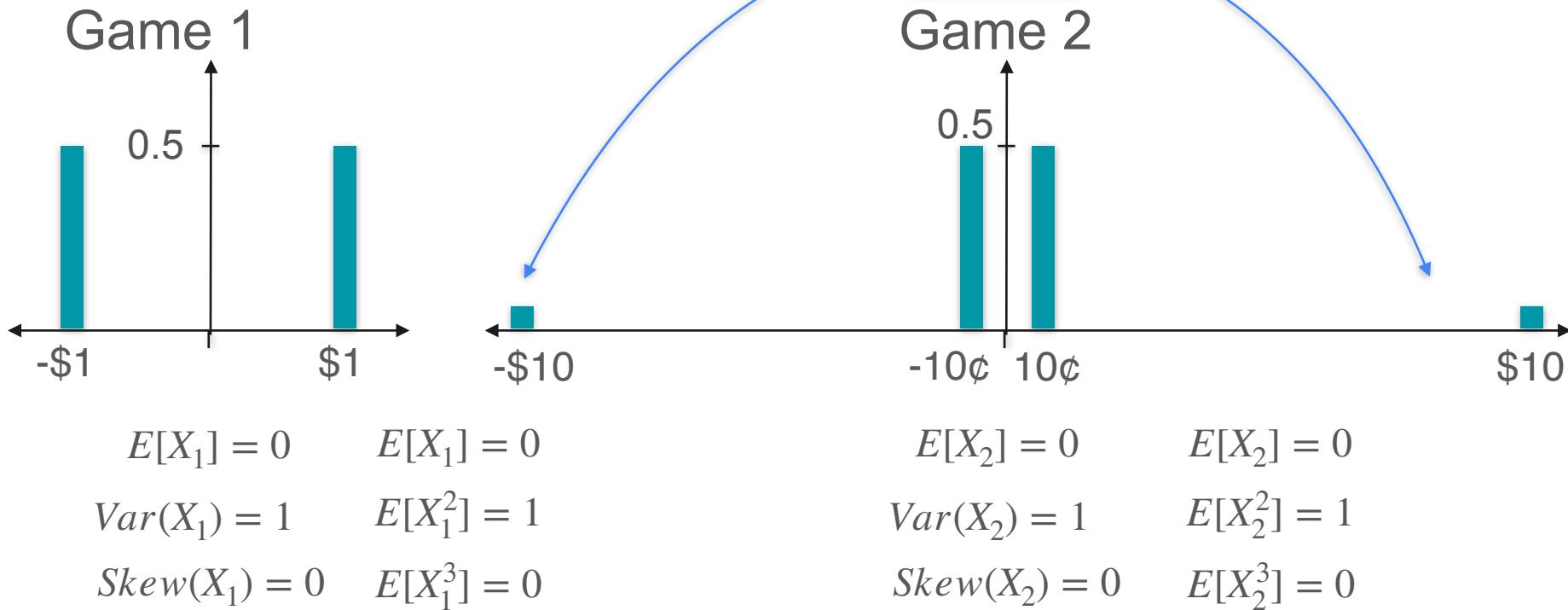
$$E[X_2] = 0$$

$$Var(X_2) = 1$$

$$E[X_2^2] = 1$$

$$Skew(X_2) = 0$$

Kurtosis



Kurtosis

Game 1



$$E[X_1] = 0$$

$$E[X_1] = 0$$

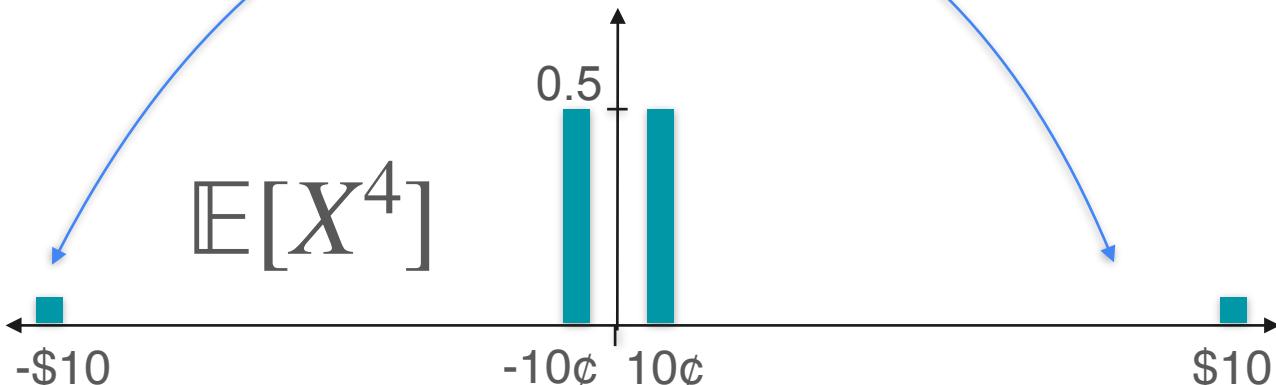
$$Var(X_1) = 1$$

$$E[X_1^2] = 1$$

$$Skew(X_1) = 0$$

$$E[X_1^3] = 0$$

Game 2



$$E[X_2] = 0$$

$$E[X_2] = 0$$

$$Var(X_2) = 1$$

$$E[X_2^2] = 1$$

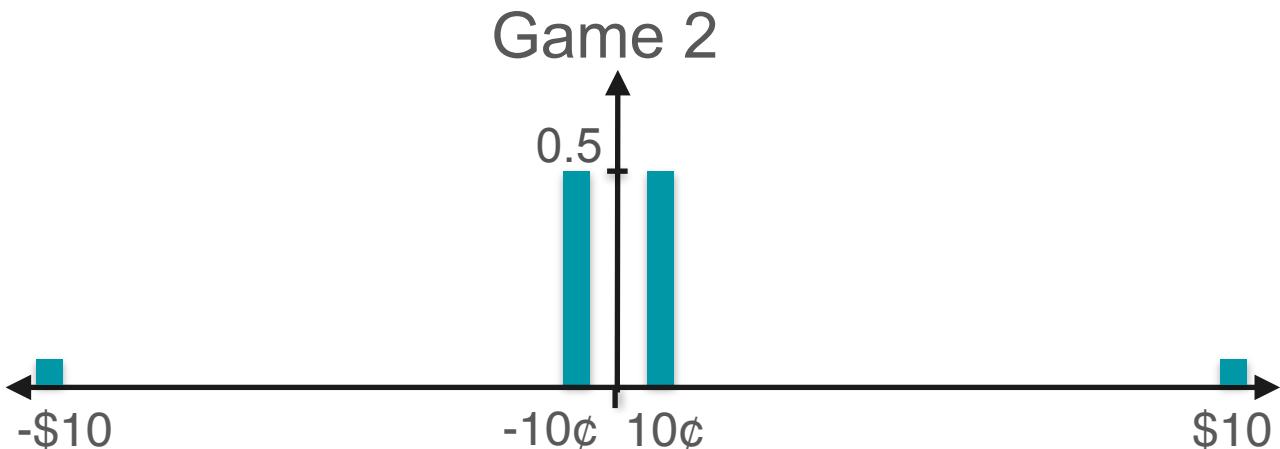
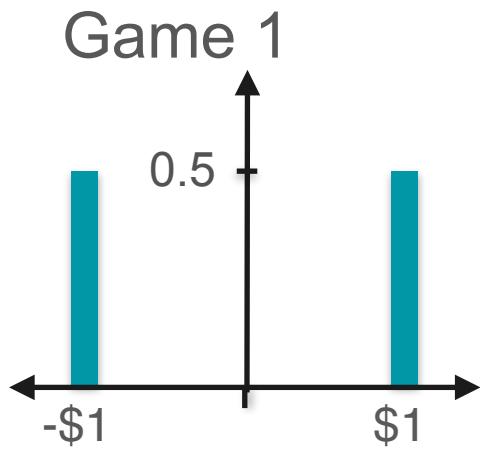
$$Skew(X_2) = 0$$

$$E[X_2^3] = 0$$

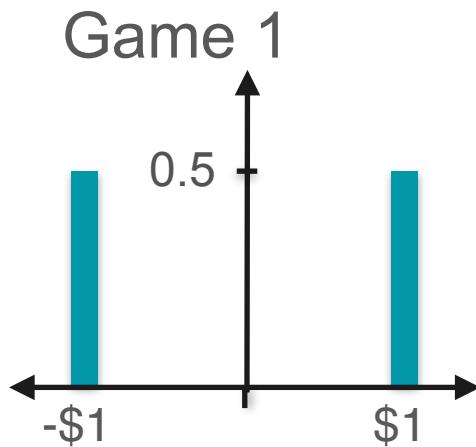
Has values way
farther from 0

$$\mathbb{E}[X^4]$$

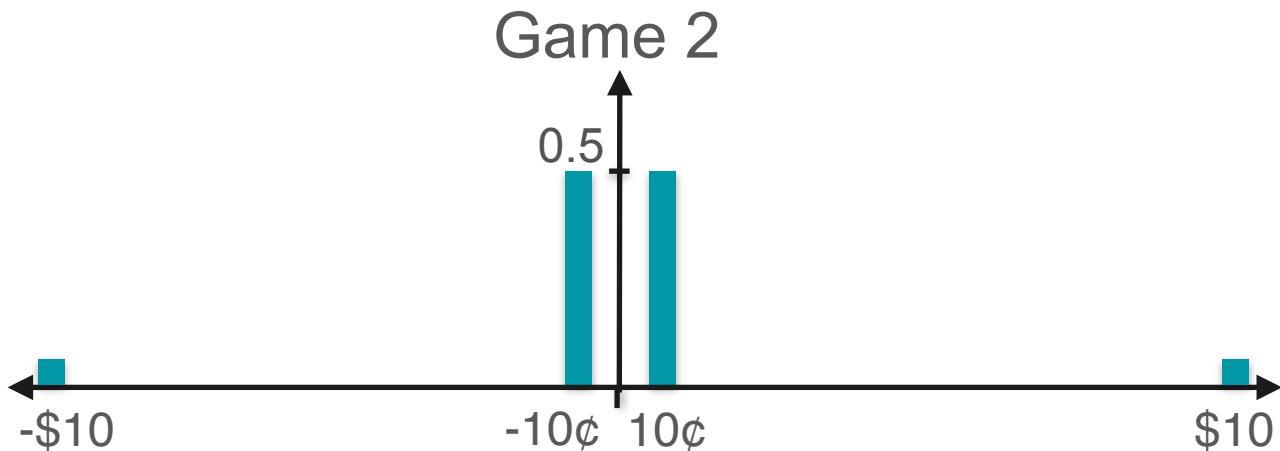
Kurtosis



Kurtosis



$$\mathbb{E}[X_1^4]$$



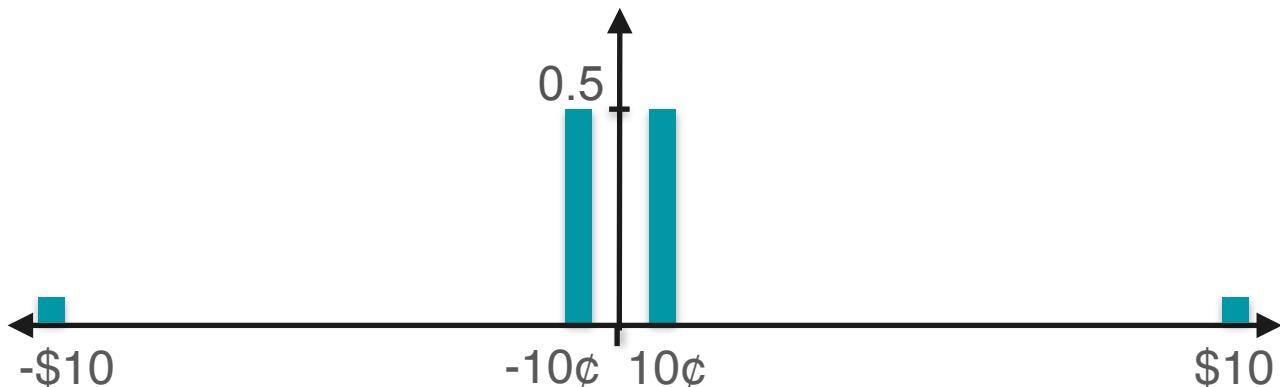
Kurtosis

Game 1



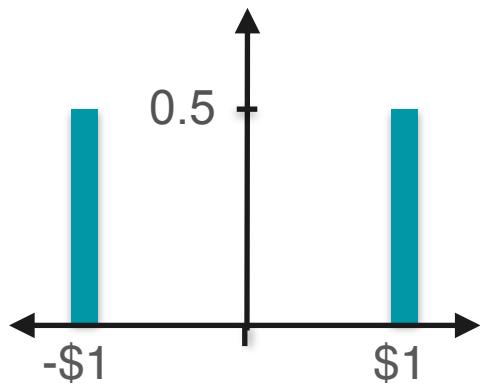
$$\mathbb{E}[X_1^4] = \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4$$

Game 2



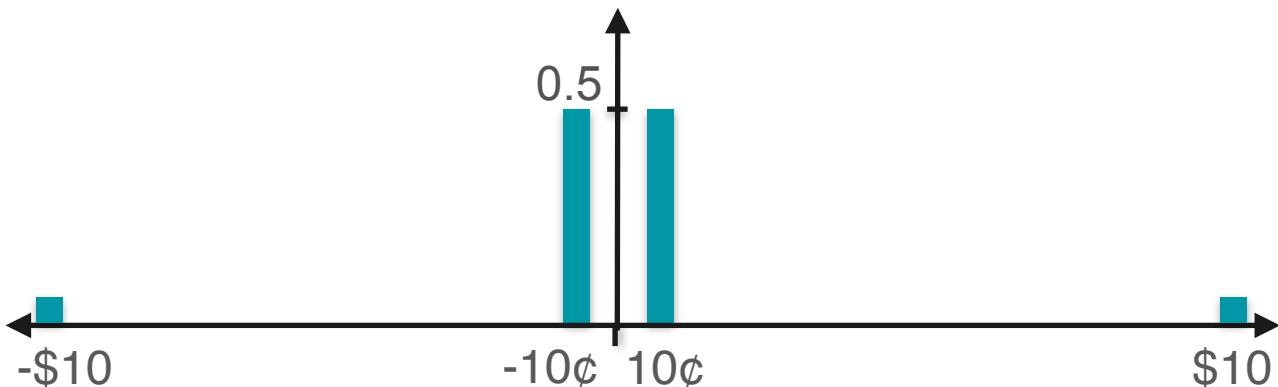
Kurtosis

Game 1



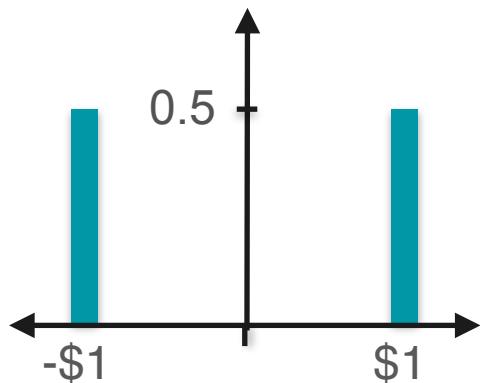
$$\begin{aligned}\mathbb{E}[X_1^4] &= \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4 \\ &= 1\end{aligned}$$

Game 2



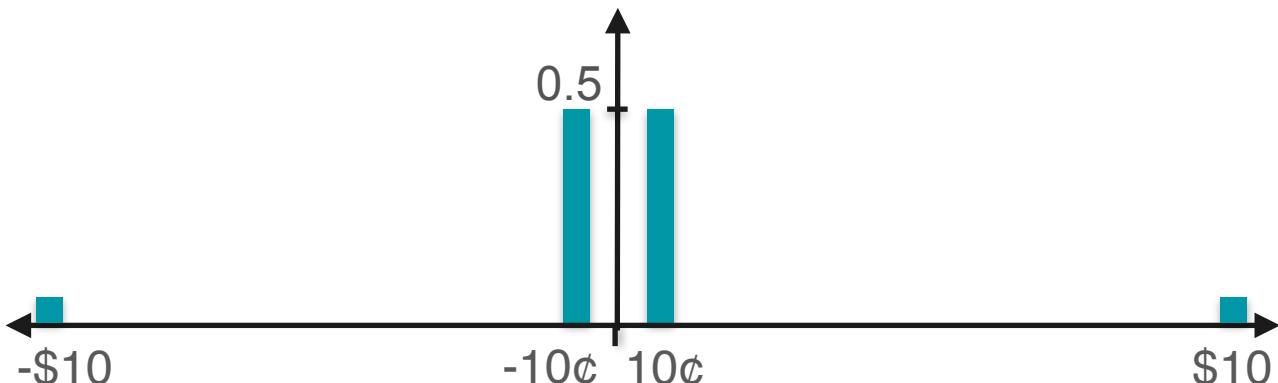
Kurtosis

Game 1



$$\begin{aligned}\mathbb{E}[X_1^4] &= \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4 \\ &= 1\end{aligned}$$

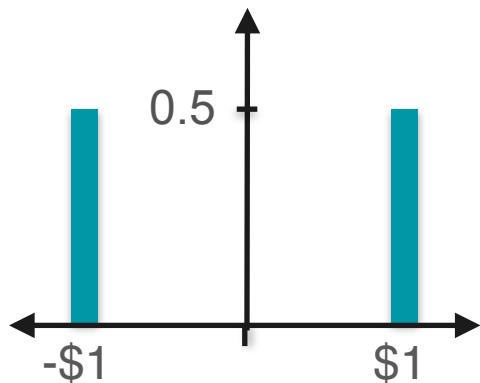
Game 2



$$\mathbb{E}[X_2^4]$$

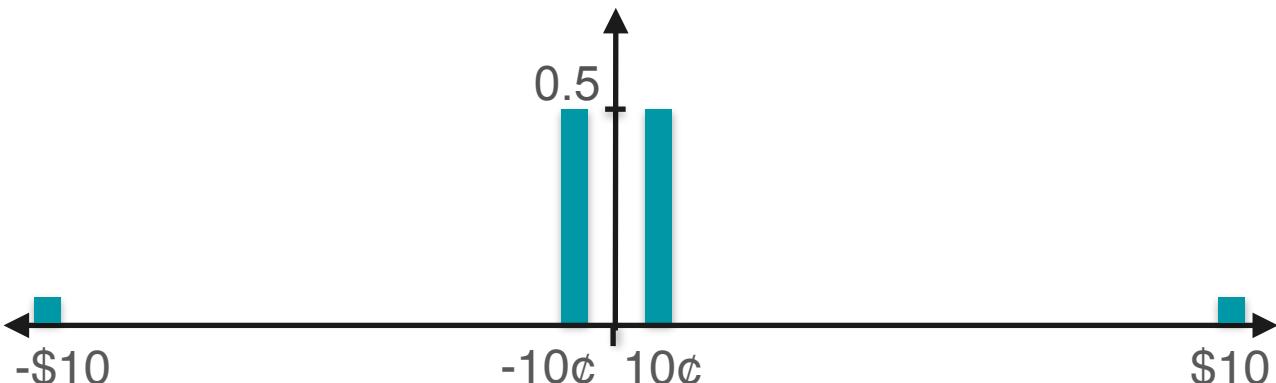
Kurtosis

Game 1



$$\begin{aligned}\mathbb{E}[X_1^4] &= \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4 \\ &= 1\end{aligned}$$

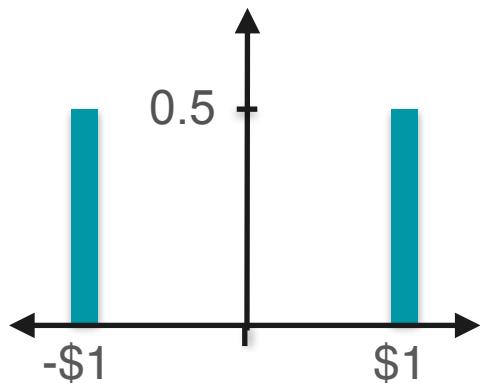
Game 2



$$\mathbb{E}[X_2^4] = \frac{100}{202}(-0.1)^4 + \frac{100}{202}(0.1)^4 + \frac{1}{202}(-10)^4 + \frac{1}{202}(10)^4$$

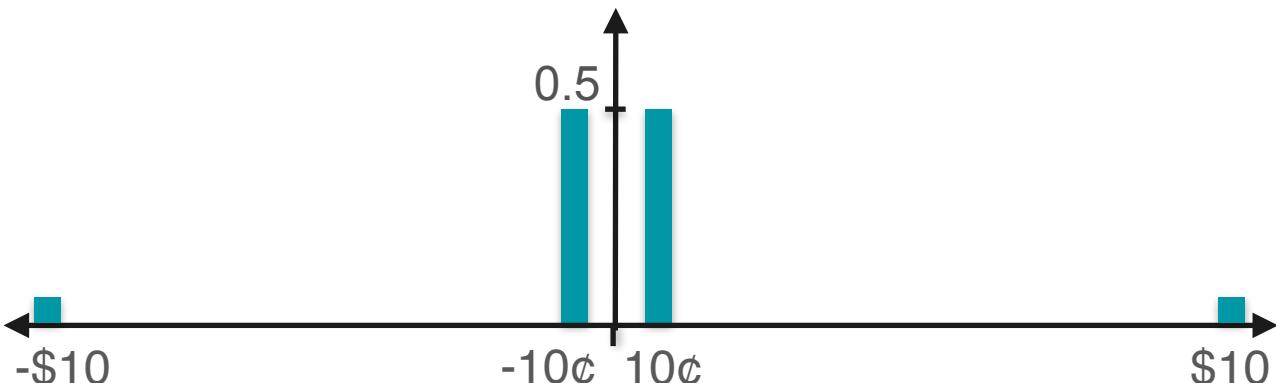
Kurtosis

Game 1



$$\begin{aligned}\mathbb{E}[X_1^4] &= \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4 \\ &= 1\end{aligned}$$

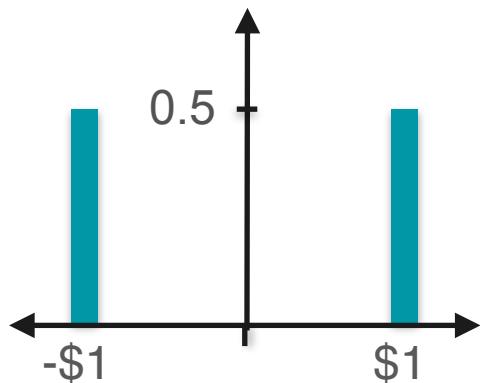
Game 2



$$\begin{aligned}\mathbb{E}[X_2^4] &= \frac{100}{202}(-0.1)^4 + \frac{100}{202}(0.1)^4 + \frac{1}{202}(-10)^4 + \frac{1}{202}(10)^4 \\ &= 99.01\end{aligned}$$

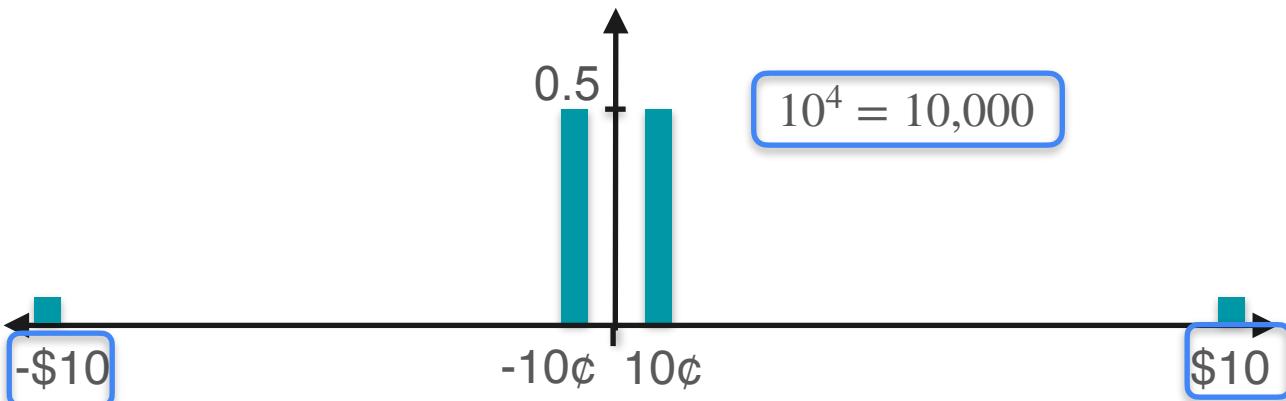
Kurtosis

Game 1



$$\begin{aligned}\mathbb{E}[X_1^4] &= \frac{1}{2}(-1)^4 + \frac{1}{2}(1)^4 \\ &= 1\end{aligned}$$

Game 2



$$\begin{aligned}\mathbb{E}[X_2^4] &= \frac{100}{202}(-0.1)^4 + \frac{100}{202}(0.1)^4 + \frac{1}{202}(-10)^4 + \frac{1}{202}(10)^4 \\ &= 99.01\end{aligned}$$

Kurtosis

$$\mathbb{E}[X^4]$$

Kurtosis

$$\mathbb{E}[X^4]$$

Almost...

Kurtosis

$$\mathbb{E}[X^4]$$

Almost...

Need to standardize...

Kurtosis

Kurtosis

$$\text{Kurtosis} = E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right]$$

Kurtosis: High and Low

Kurtosis: High and Low



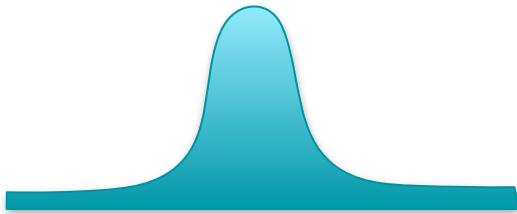
$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{small}$$

$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{large}$$

Kurtosis: High and Low



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{small}$$

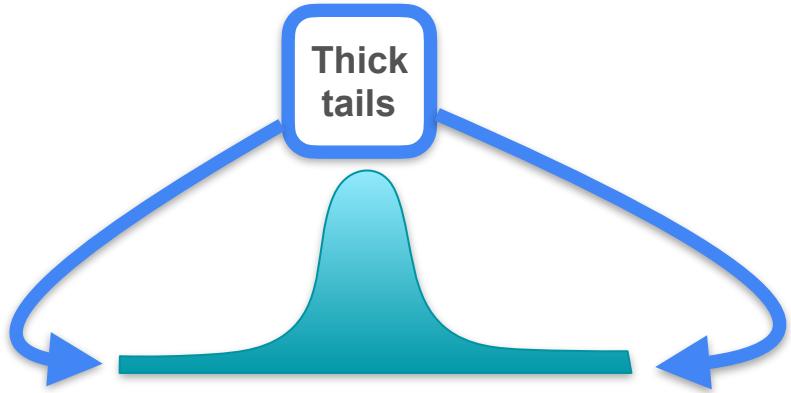


$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{large}$$

Kurtosis: High and Low

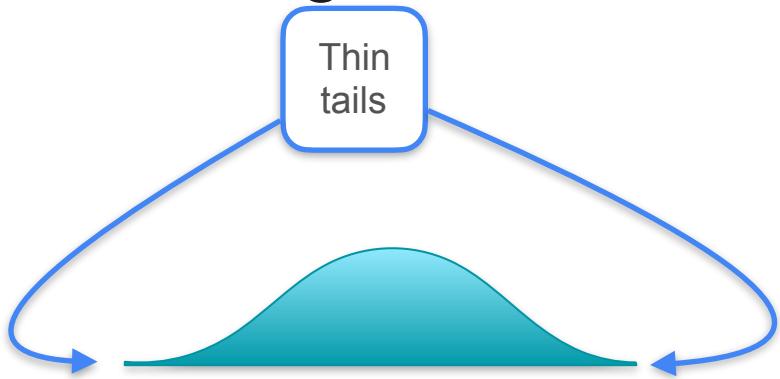


$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{small}$$

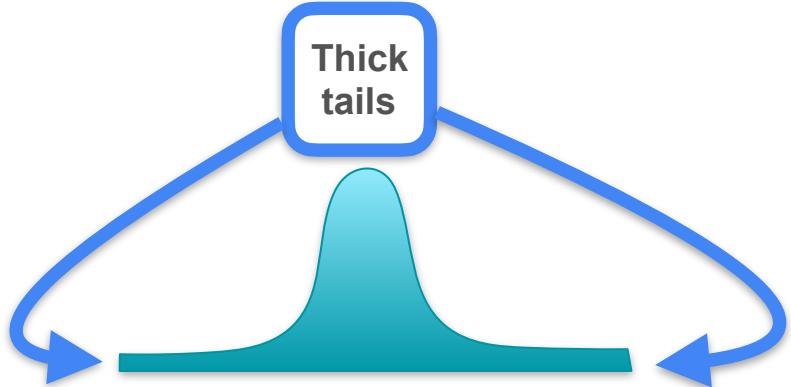


$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{large}$$

Kurtosis: High and Low

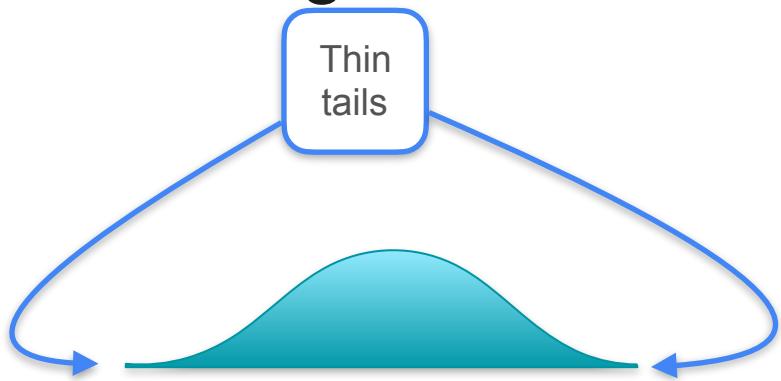


$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{small}$$

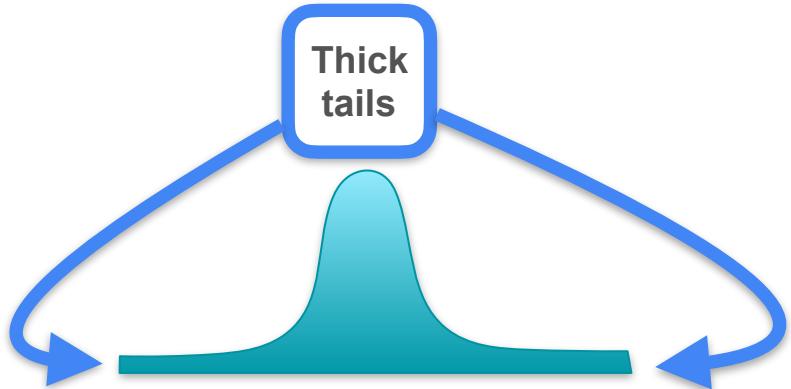


$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{large}$$

Kurtosis: High and Low



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{small}$$



$$E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \text{large}$$

Even if they have the same variance!



DeepLearning.AI

Describing Distributions

Quantiles and box-plots

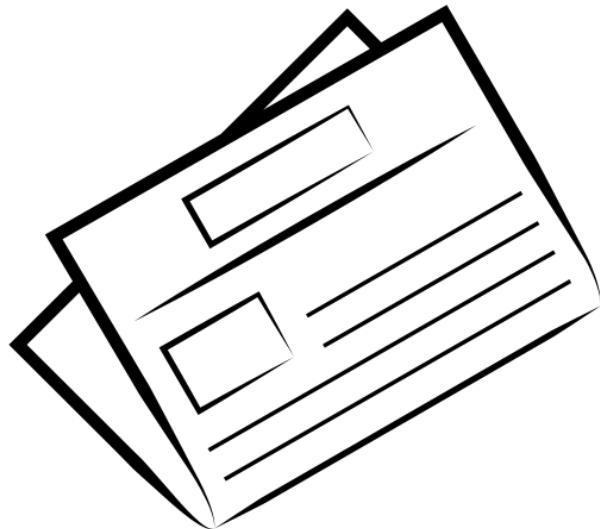
Quantiles: Example

Quantiles: Example



Newspaper advertisement

Quantiles: Example



Newspaper advertisement

Newspaper Advertisement (X)			
18.3	18.4	23.2	51.2
35.2	29.7	75	8.7
65.9	14.2	54.7	25.9

Quantiles: Example

Newspaper Advertisement (X)			
18.3	18.4	23.2	51.2
35.2	29.7	75	8.7
65.9	14.2	54.7	25.9

Quantiles: Example

What is the median here?

Newspaper Advertisement (X)			
18.3	18.4	23.2	51.2
35.2	29.7	75	8.7
65.9	14.2	54.7	25.9

Quantiles: Example

What is the median here?

The point that splits your data in half

Newspaper Advertisement (X)			
18.3	18.4	23.2	51.2
35.2	29.7	75	8.7
65.9	14.2	54.7	25.9

Quantiles: Example

What is the median here?

The point that splits your data in half

Newspaper Advertisement (X)			
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Quantiles: Example

What is the median here?

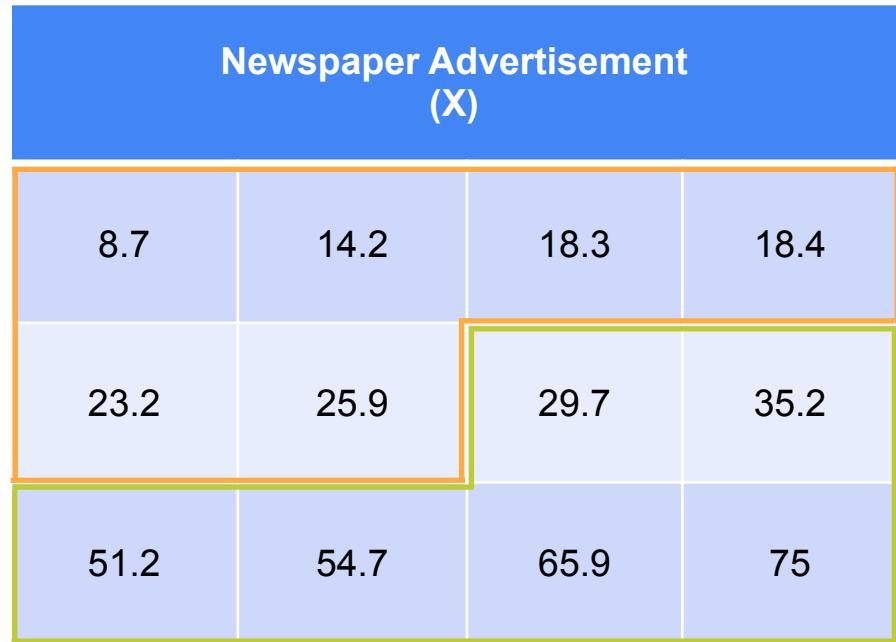
The point that splits your data in half

Newspaper Advertisement (X)			
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Quantiles: Example

What is the median here?

The point that splits your data in half



Quantiles: Example

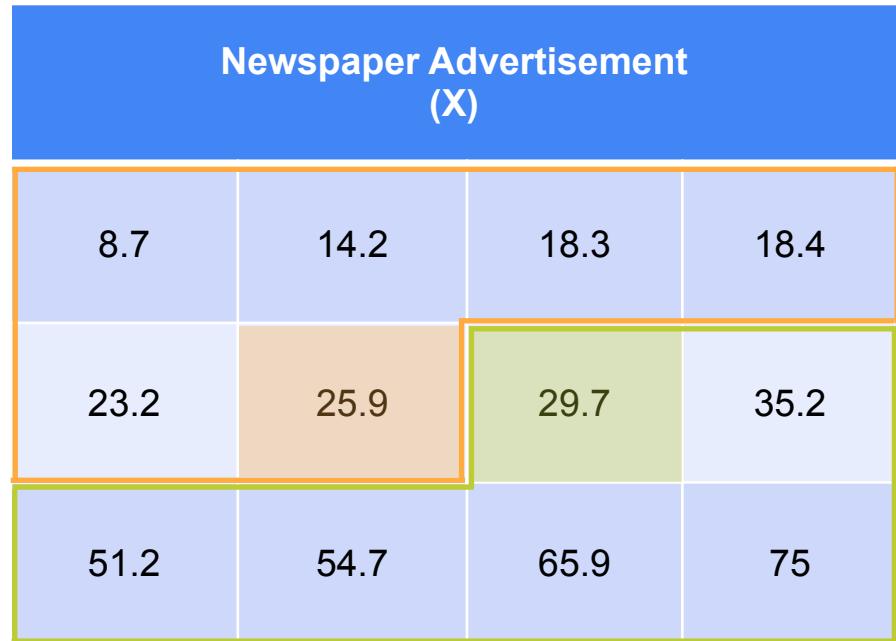
What is the median here?

The point that splits your data in half

$$\text{Median} = \frac{25.9 + 29.7}{2} = 27.8$$

50% quantile

Second quartile

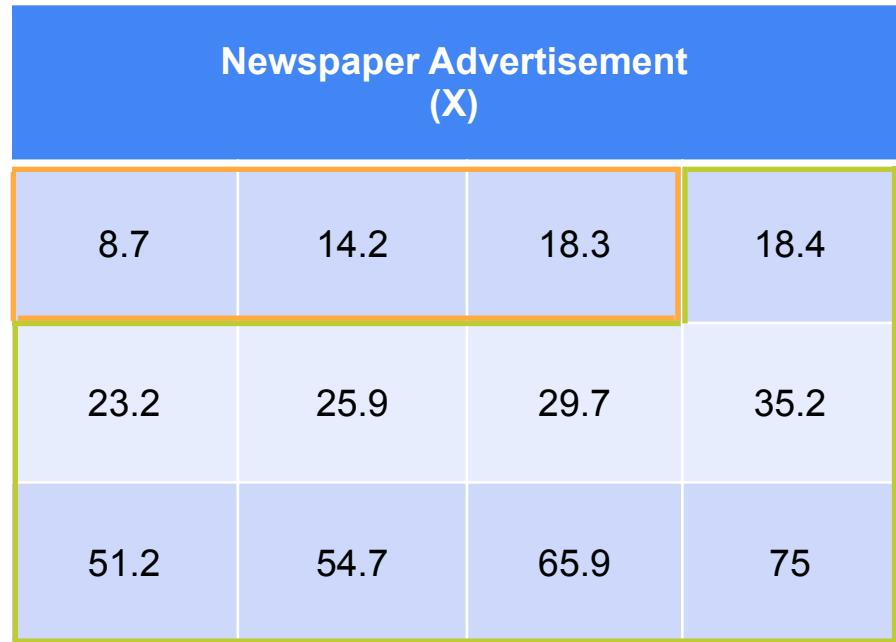


Quantiles: Example

Newspaper Advertisement (X)			
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Quantiles: Example

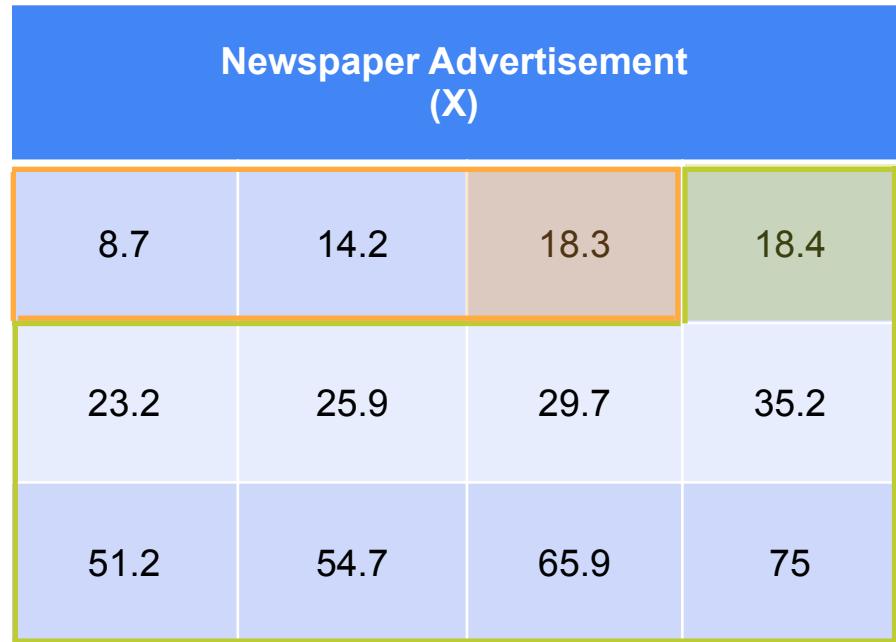
What about the point that leaves 1/4 of your data to the left and 3/4 to the right?



Quantiles: Example

What about the point that leaves 1/4 of your data to the left and 3/4 to the right?

$$\frac{18.3 + 18.4}{2} = 18.35$$



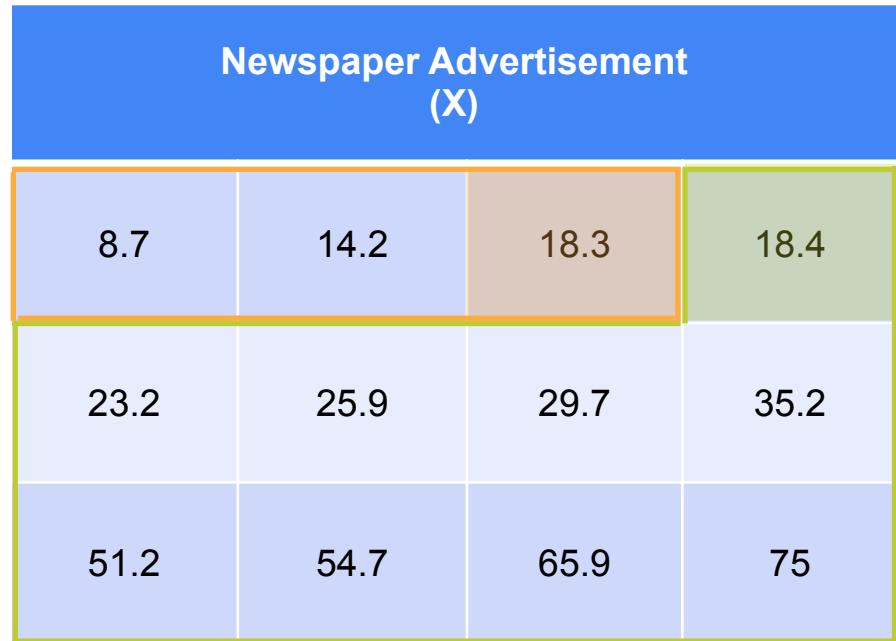
Quantiles: Example

What about the point that leaves 1/4 of your data to the left and 3/4 to the right?

$$\frac{18.3 + 18.4}{2} = 18.35$$

25% quantile

First quartile



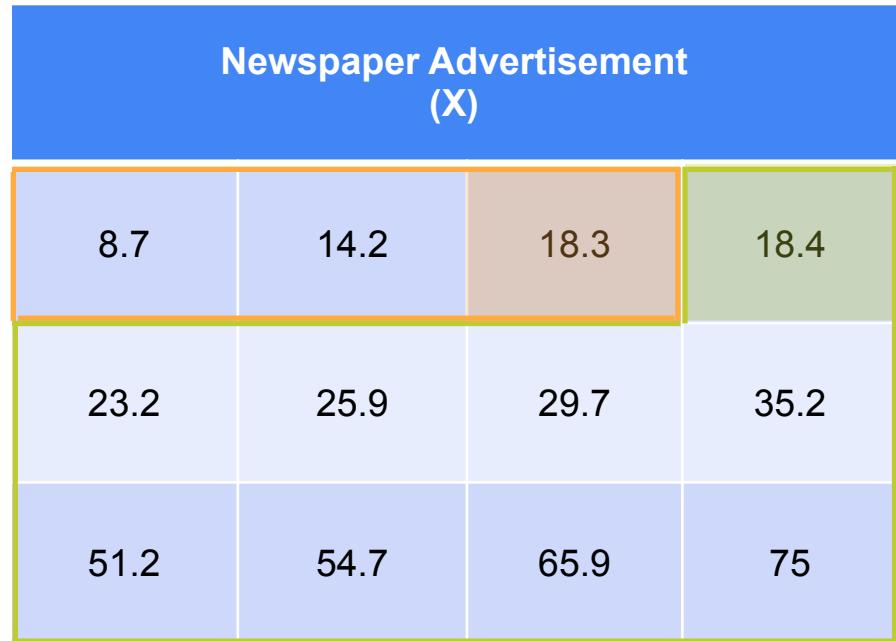
Quantiles: Example

What about the point that leaves 1/4 of your data to the left and 3/4 to the right?

$$q_{0.25} = Q1 = \frac{18.3 + 18.4}{2} = 18.35$$

25% quantile

First quartile



Quantiles

Quantiles

In general:

The **k%** quantile ($q_{k/100}$) is the value that leaves k% of your data to the left and (100-k)% of your data to the right

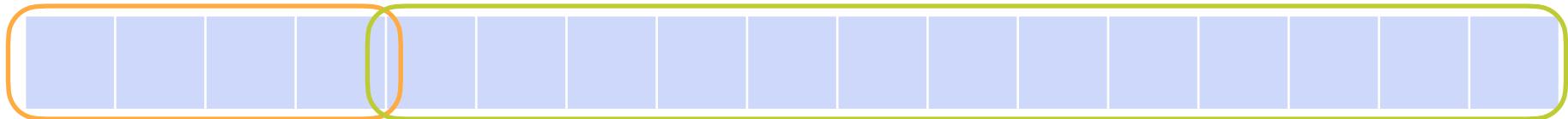
Quantiles

In general:

The **k%** quantile ($q_{k/100}$) is the value that leaves k% of your data to the left and $(100 - k)\%$ of your data to the right

$k\%$

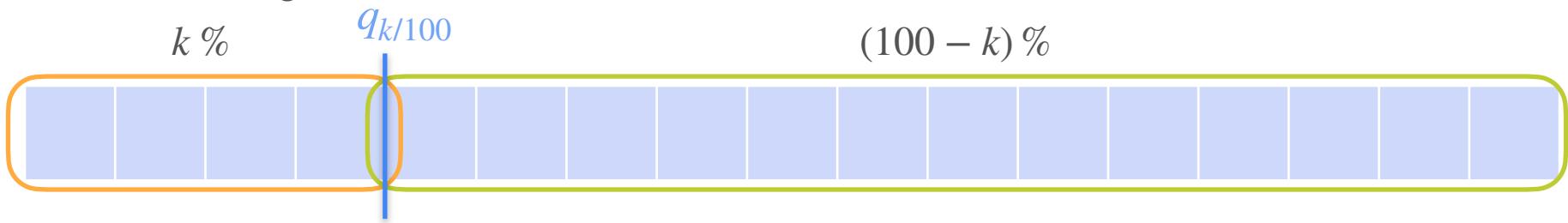
$(100 - k)\%$



Quantiles

In general:

The **k%** quantile ($q_{k/100}$) is the value that leaves k% of your data to the left and $(100 - k)\%$ of your data to the right



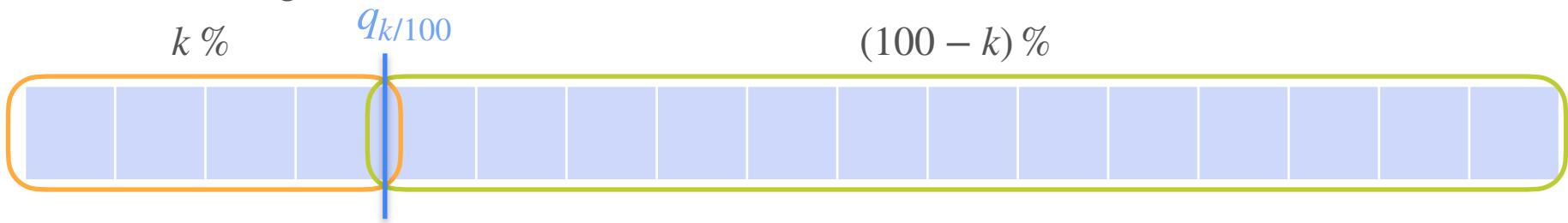
Quantiles

In general:

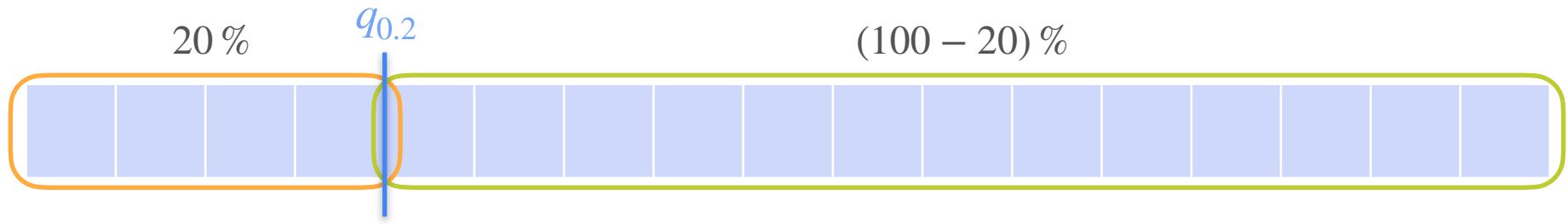
The **k%** quantile ($q_{k/100}$) is the value that leaves k% of your data to the left and $(100-k)\%$ of your data to the right

Some common quantiles:

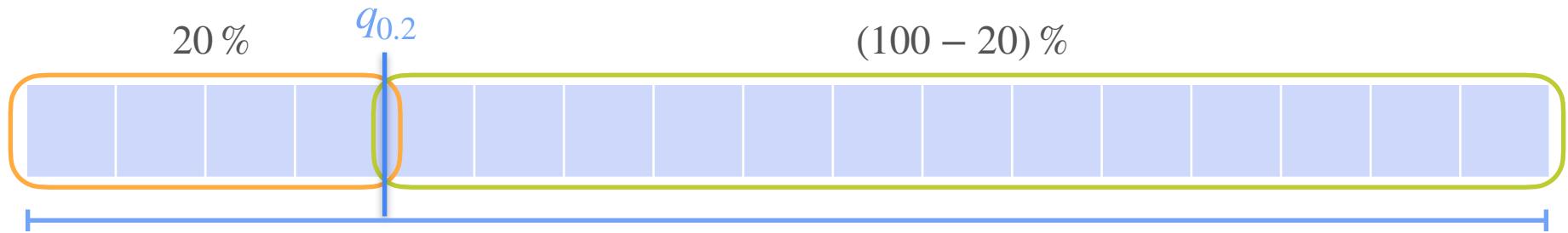
- 25% quantile (first quartile - Q1)
- 50% quantile (median - Q2)
- 75% quantile (third quartile - Q3)



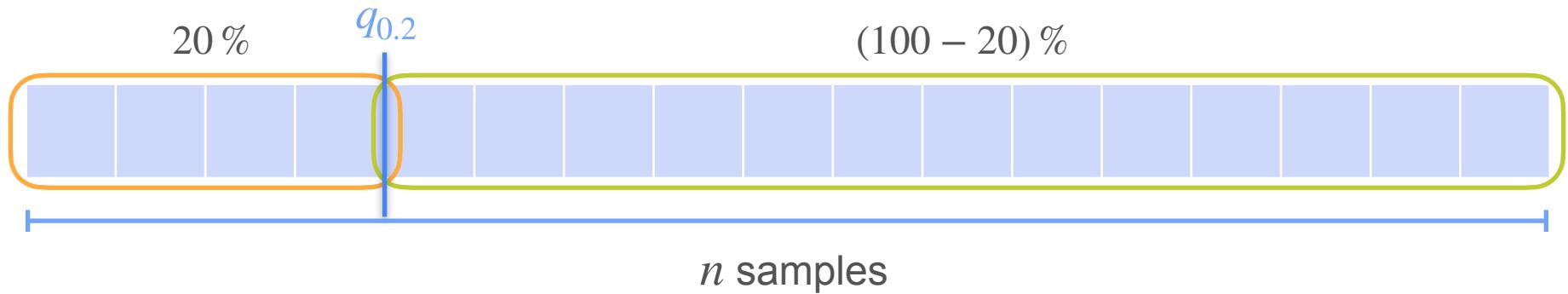
Quantiles



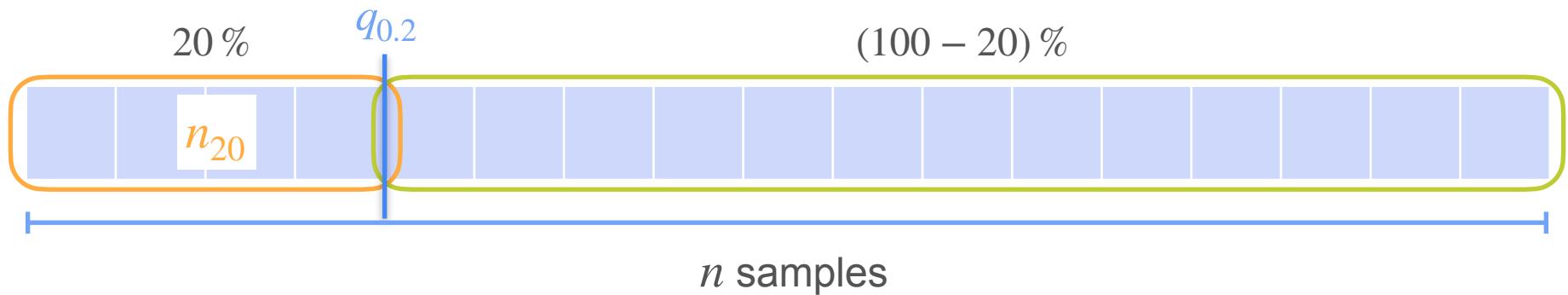
Quantiles



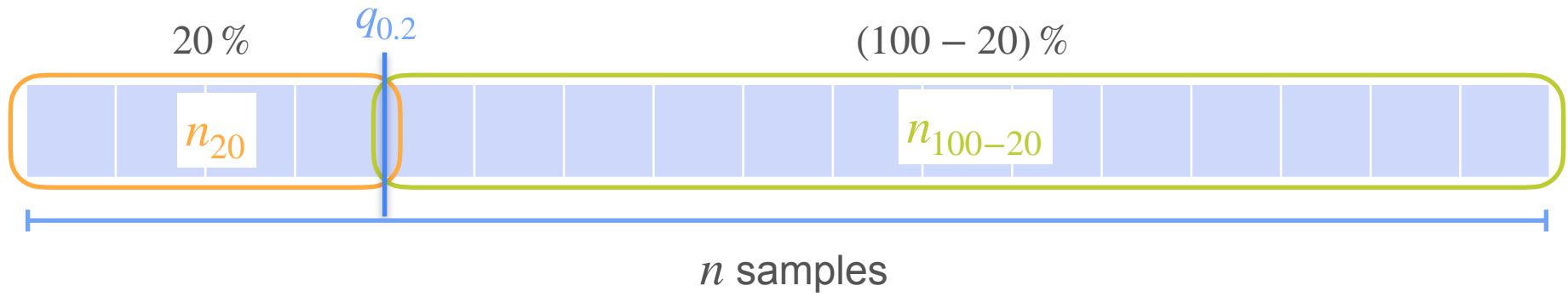
Quantiles



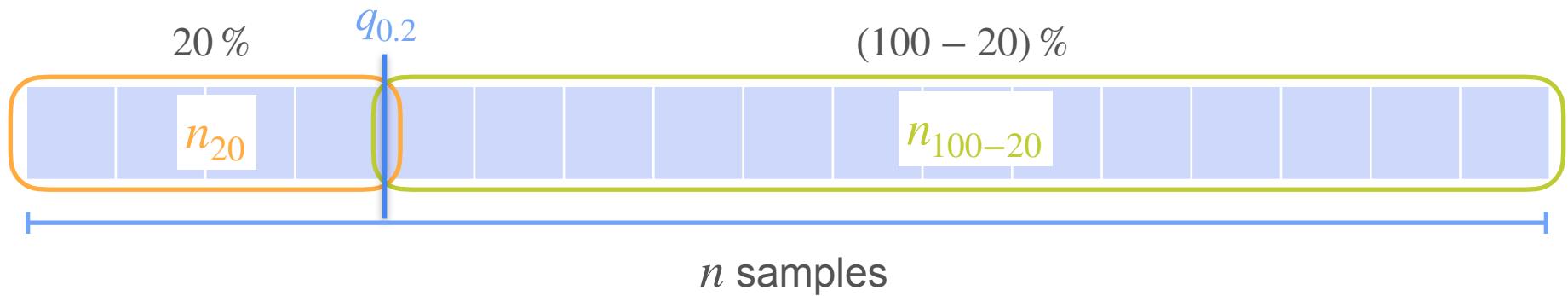
Quantiles



Quantiles

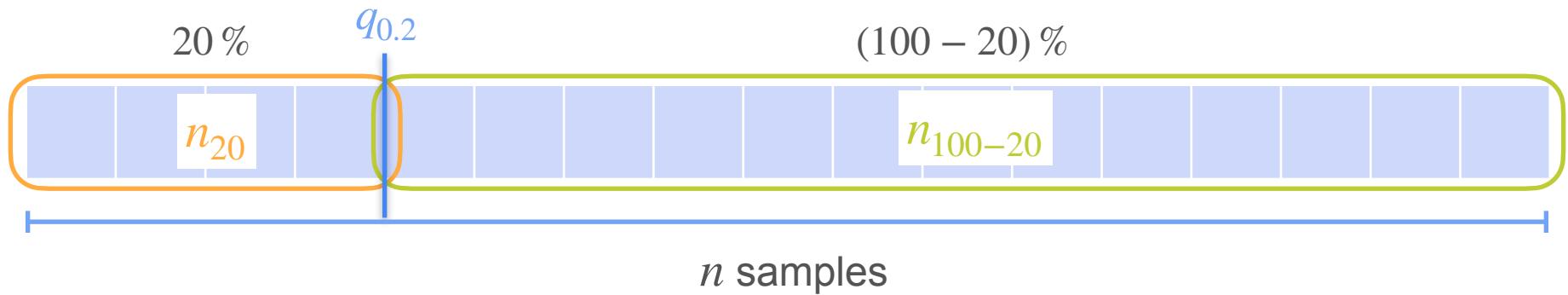


Quantiles



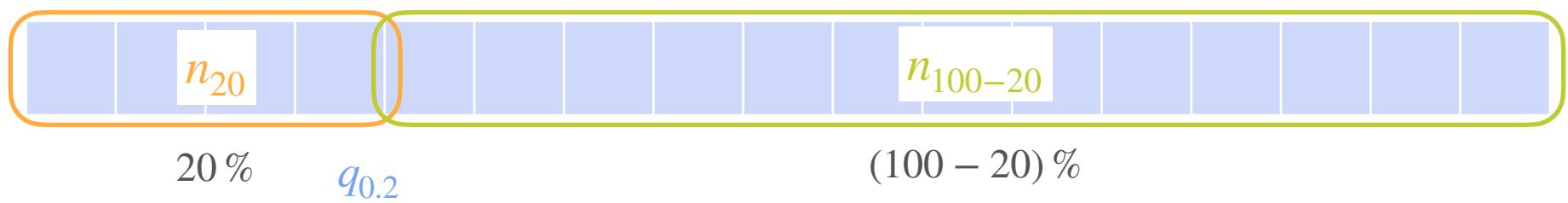
$$\frac{20}{100} = \frac{n_{20}}{n}$$

Quantiles

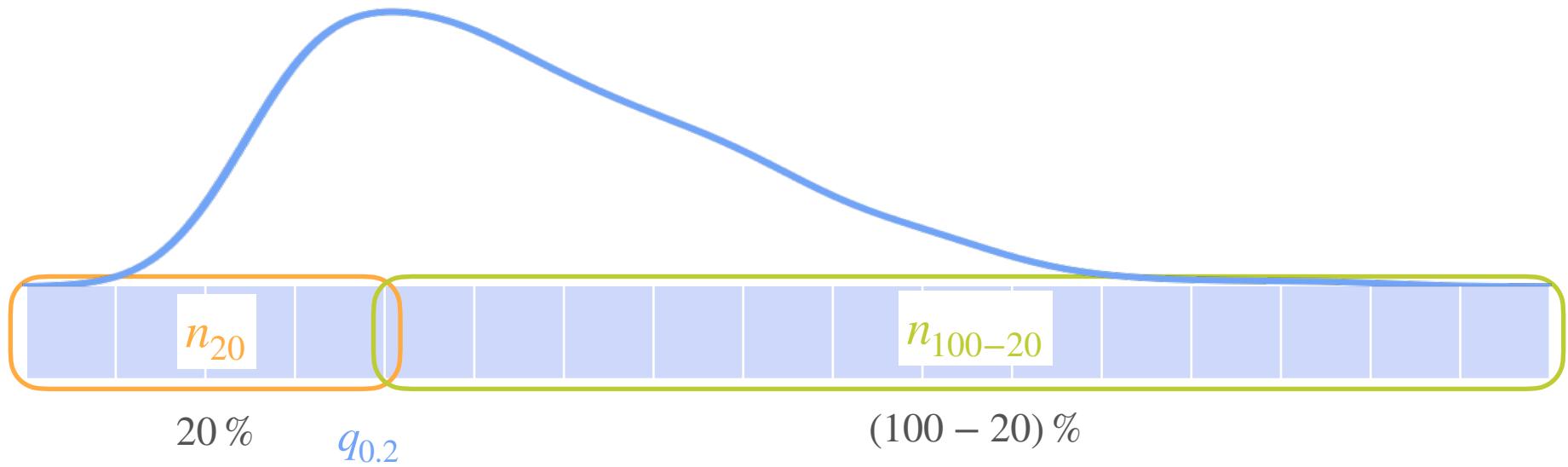


$$\frac{20}{100} = \frac{n_{20}}{n} \approx \mathbf{P}(X \leq q_{0.2})$$

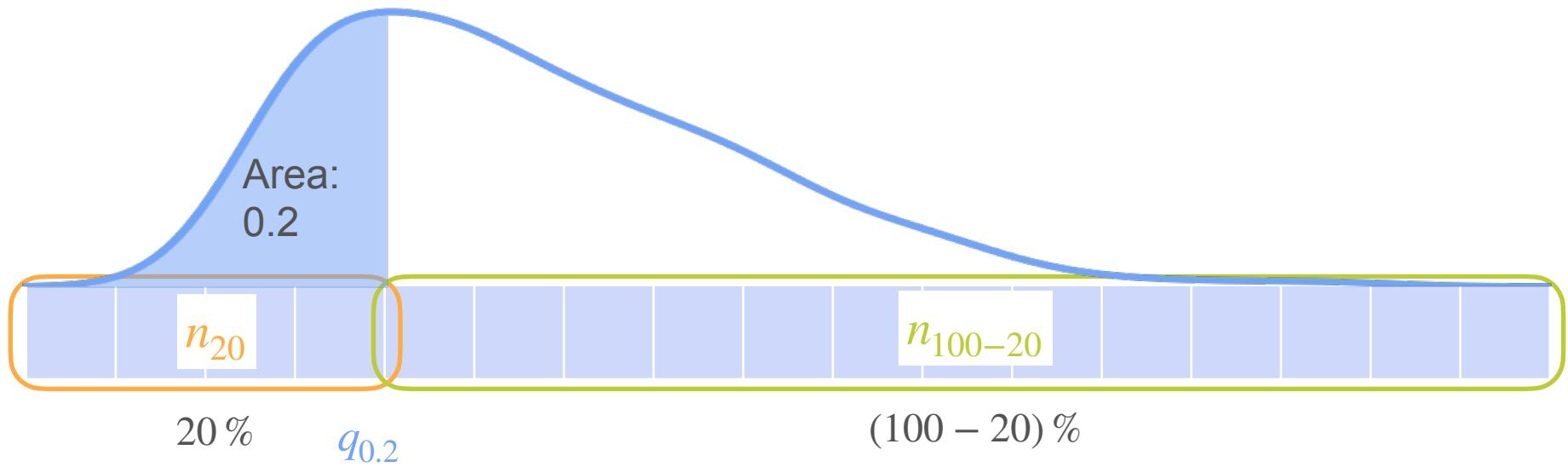
Quantiles



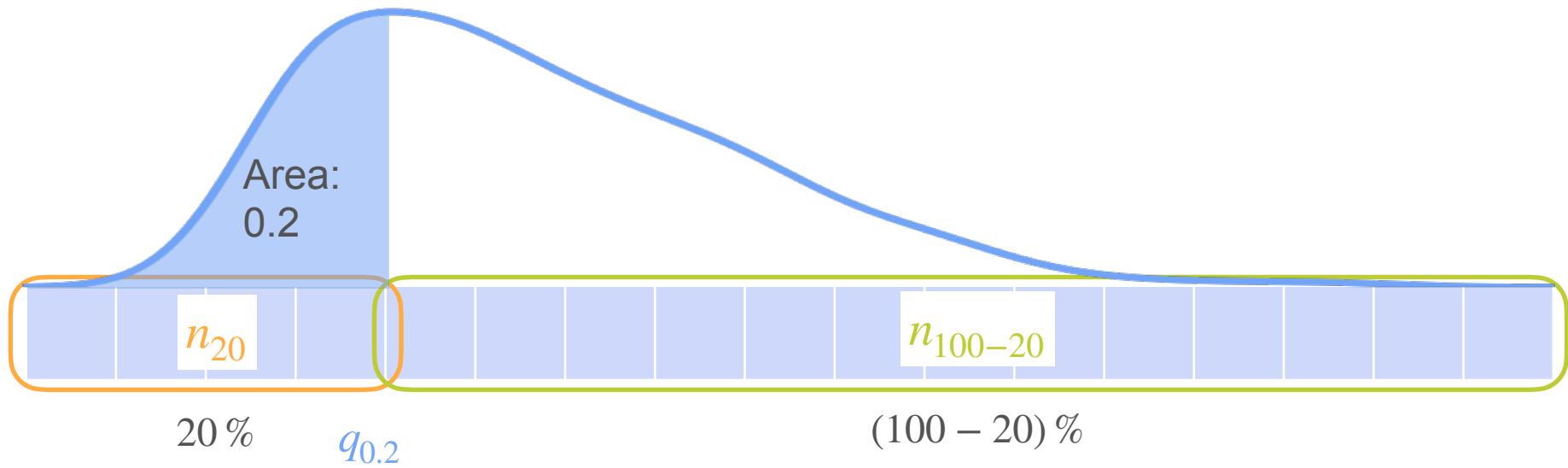
Quantiles



Quantiles



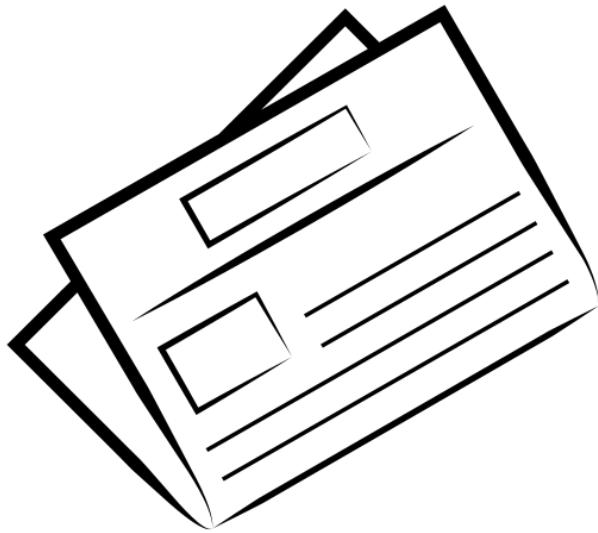
Quantiles



k% quantile ($q_{k/100}$) is the value such that $\mathbf{P}(X \leq q_{k/100}) = \frac{k}{100}$

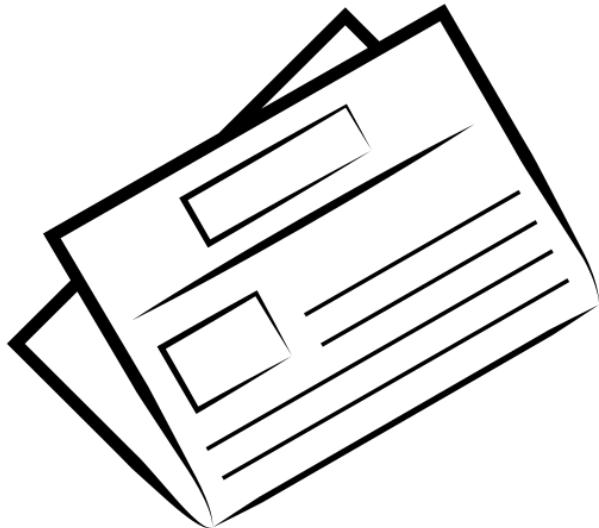
Box-Plots

Box-Plots



Newspaper advertisement

Box-Plots



Newspaper advertisement

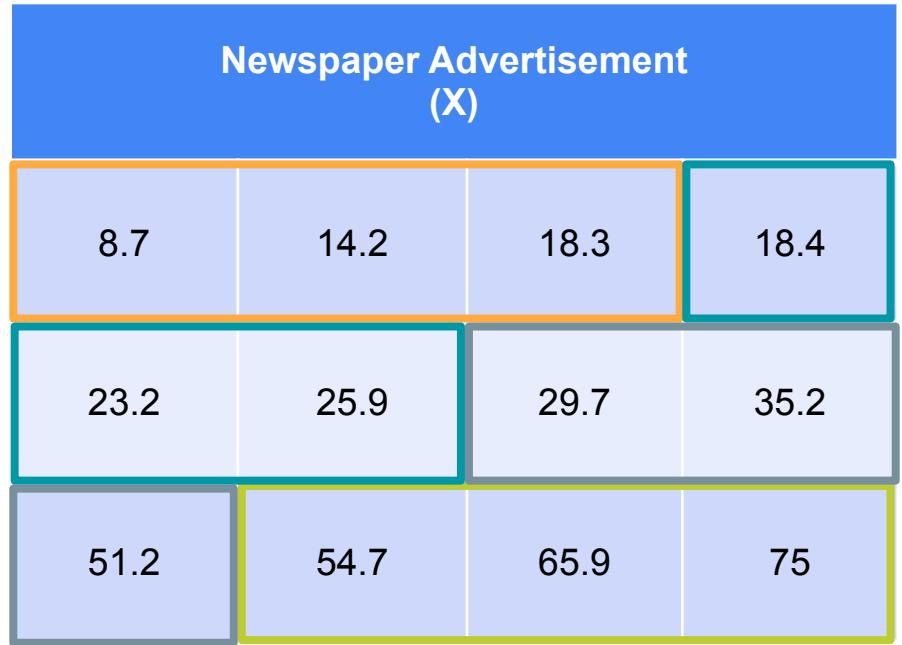
Newspaper Advertisement (X)			
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Box-Plots

Newspaper Advertisement
(X)

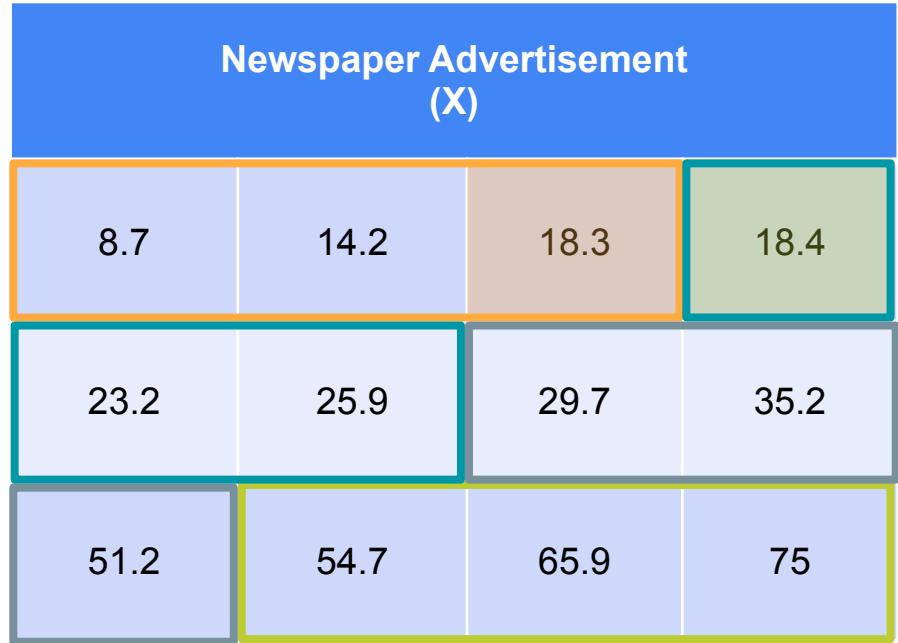
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Box-Plots



Box-Plots

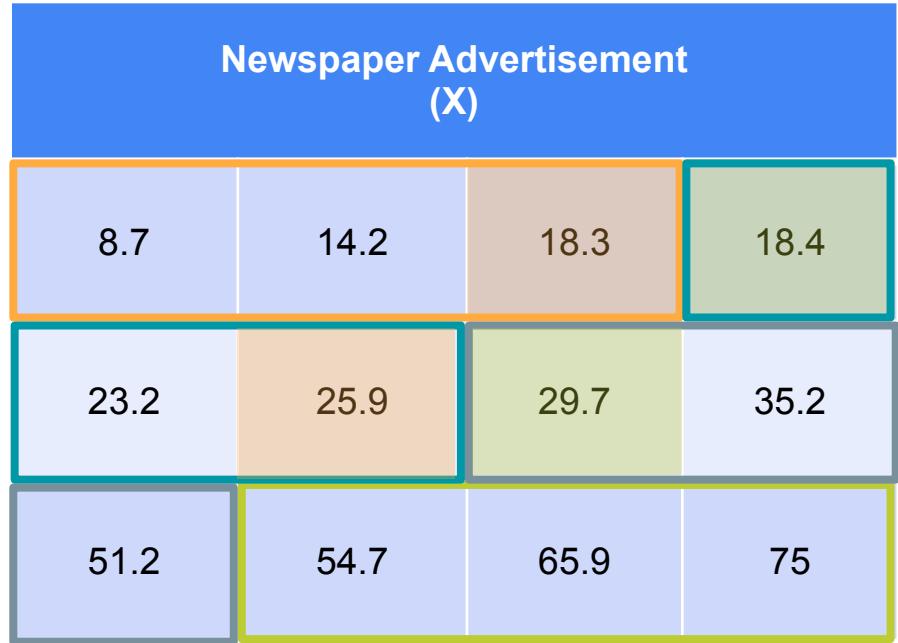
$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

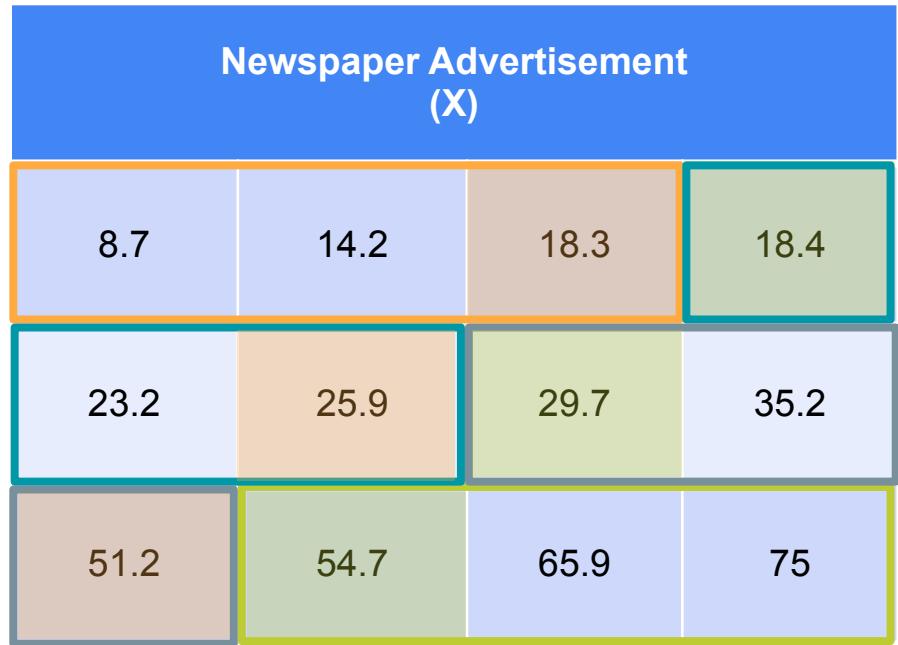


Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

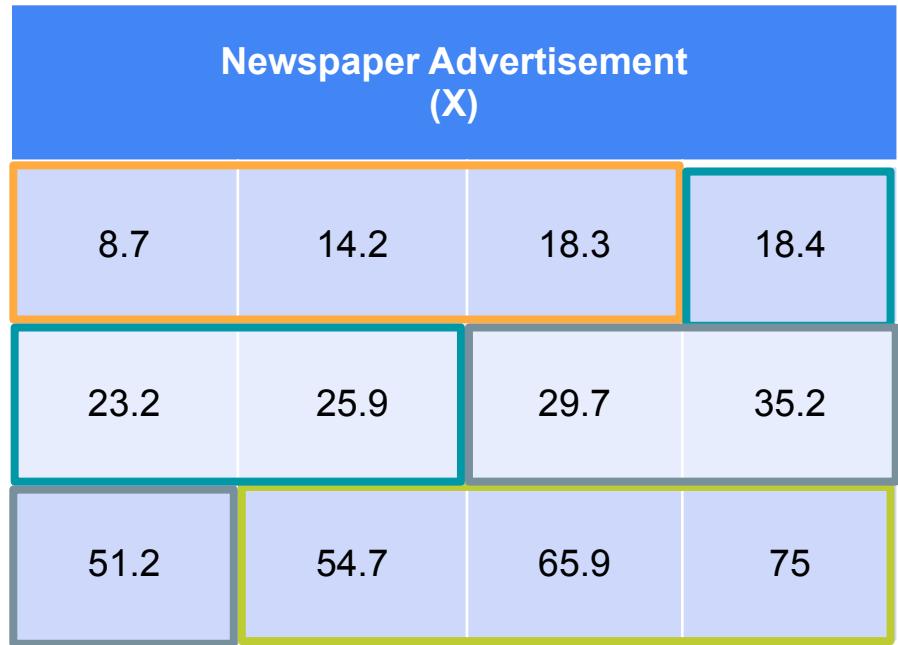


Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$



Box-Plots

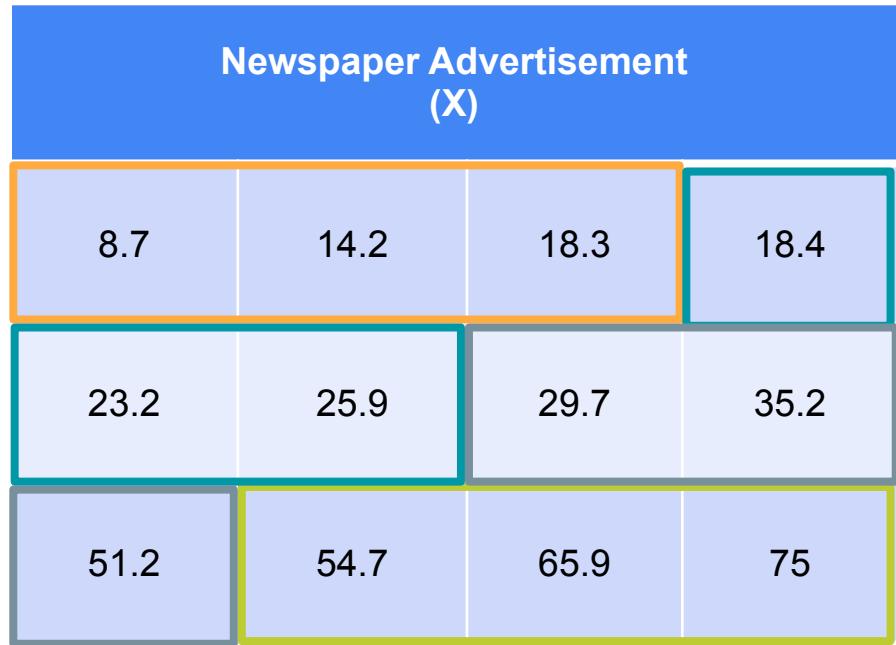
$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1$$



Box-Plots

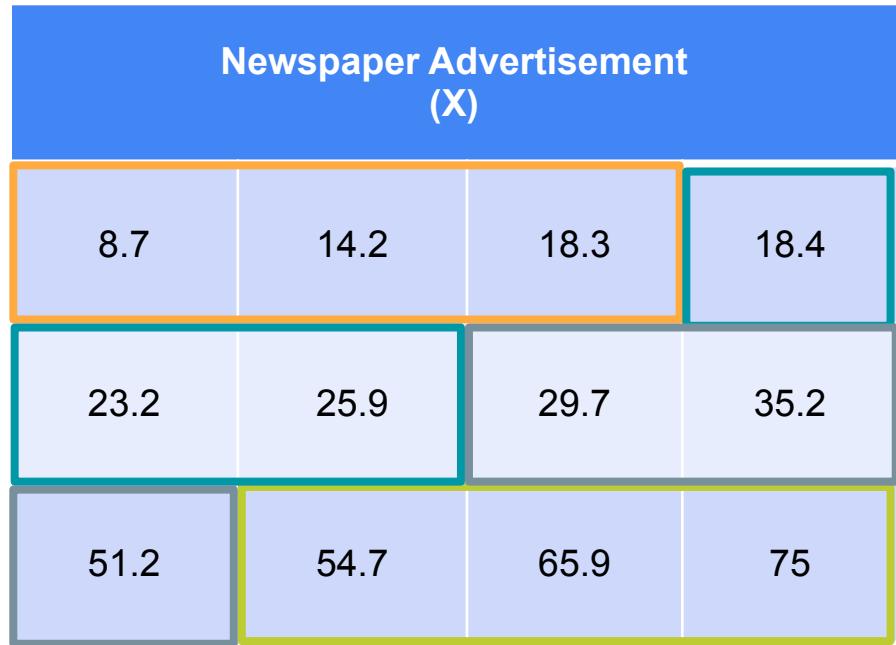
$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

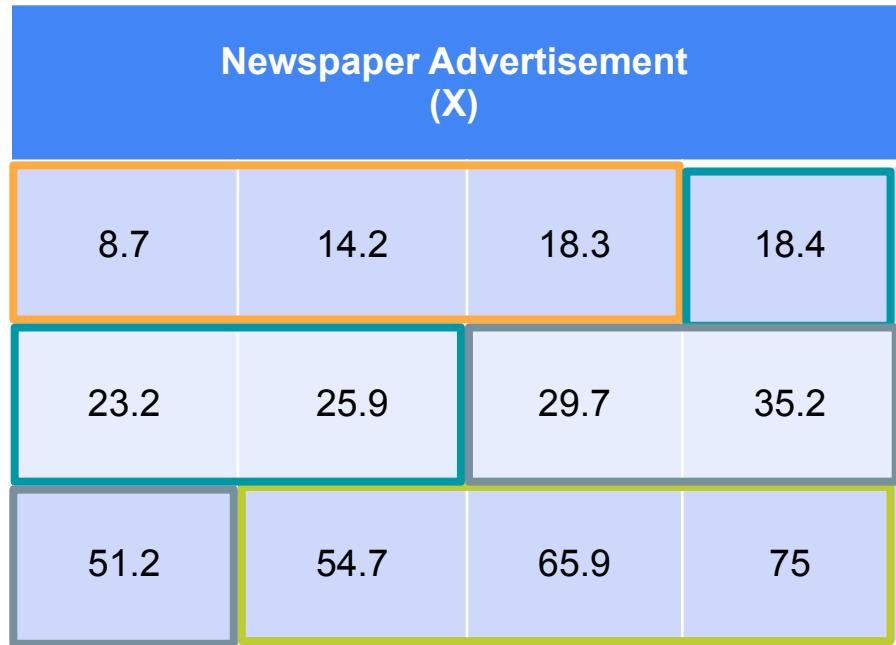
$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$\min = 8.7$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

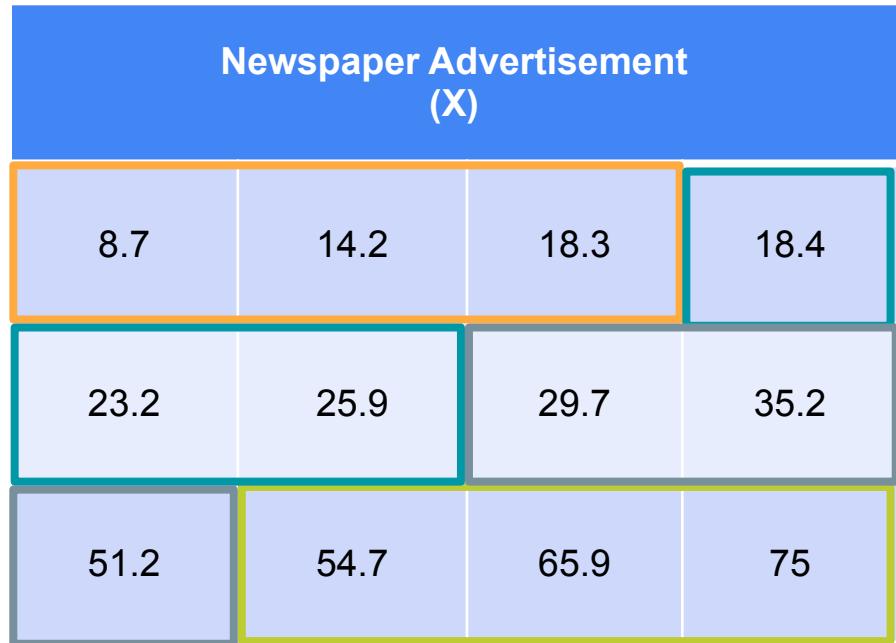
$$Q2 = q_{0.50} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$\min = 8.7 \quad \max = 75$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

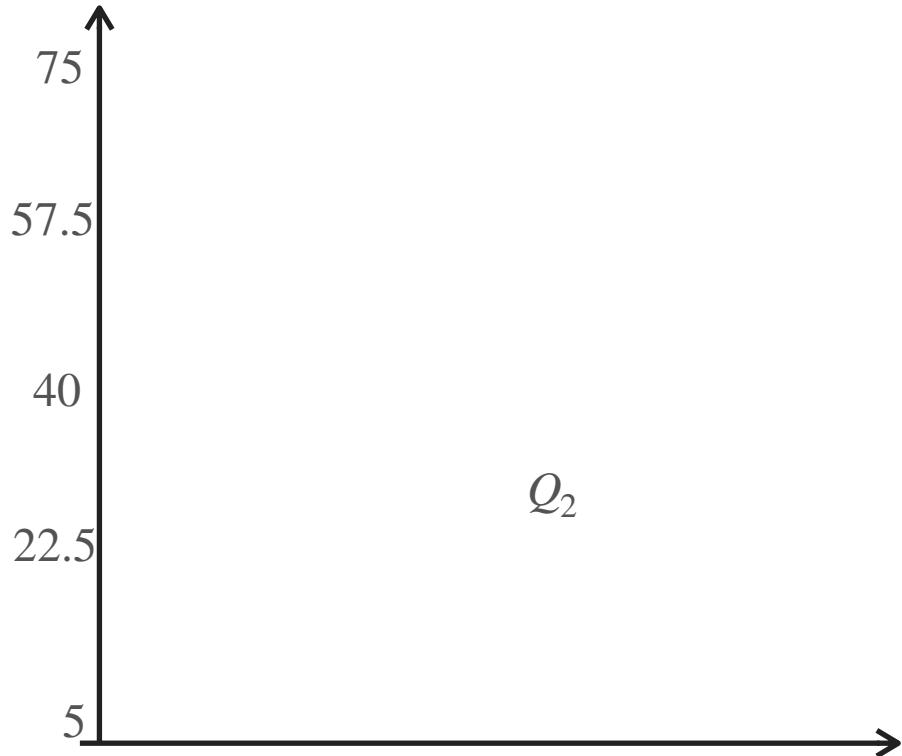
$$Q2 = q_{0.5} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$x_{\min} = 8.7 \quad x_{\max} = 75$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

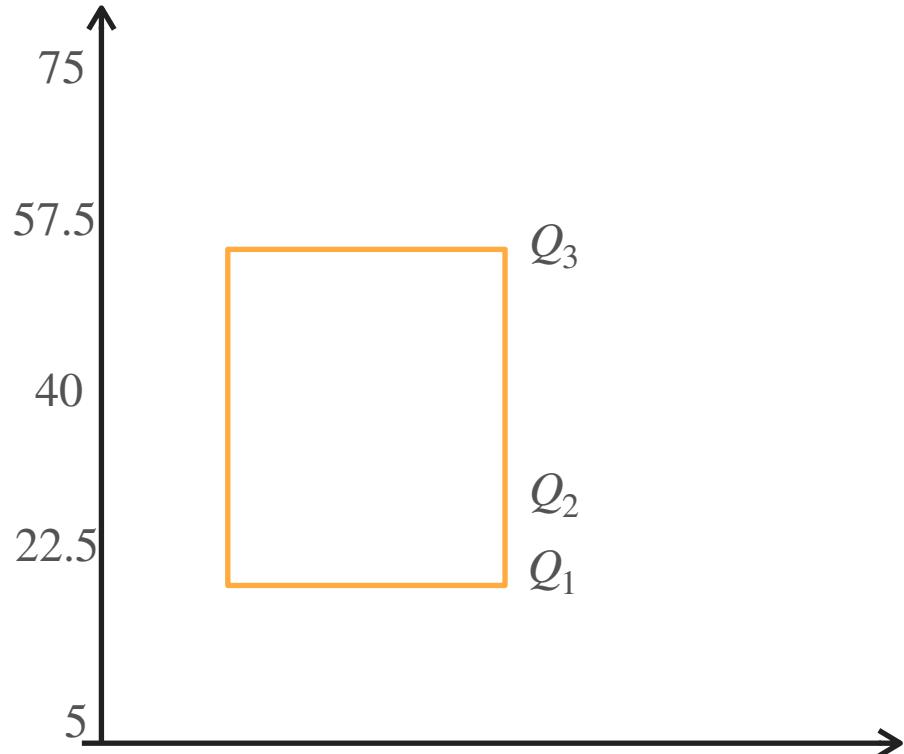
$$Q2 = q_{0.5} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$x_{\min} = 8.7 \quad x_{\max} = 75$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

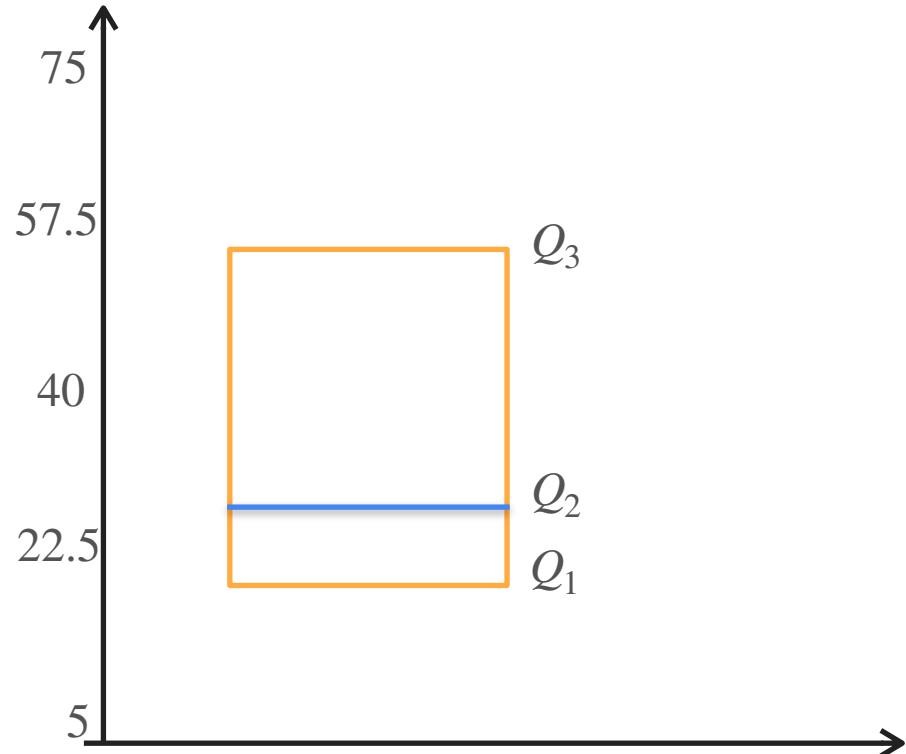
$$Q2 = q_{0.5} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$x_{\min} = 8.7 \quad x_{\max} = 75$$



Box-Plots

$$Q1 = q_{0.25} = \frac{18.3 + 18.4}{2} = 18.35$$

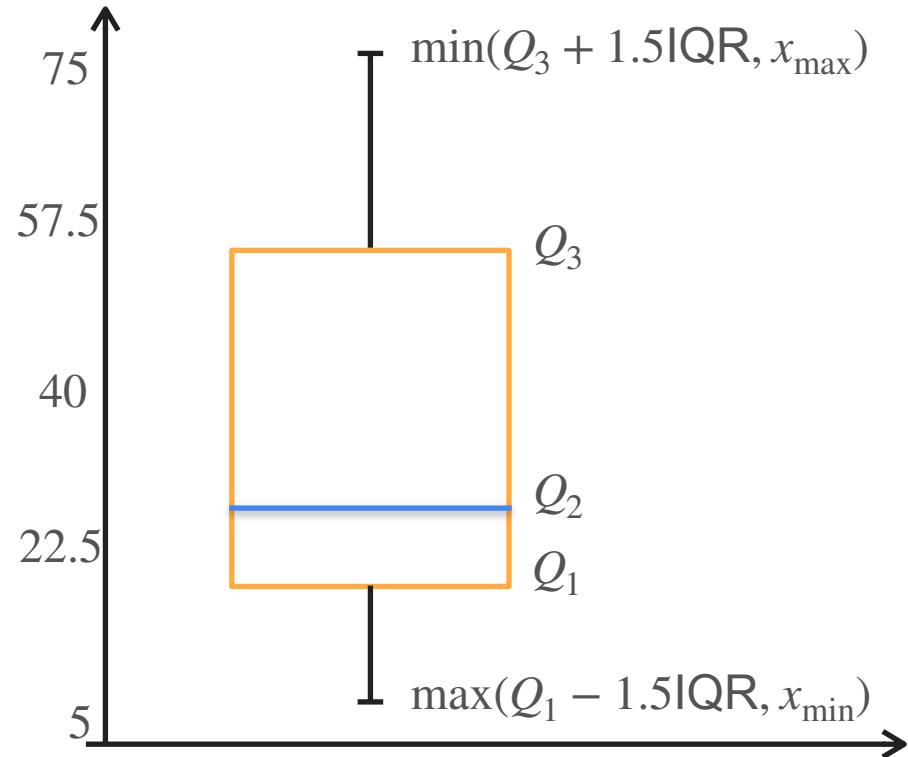
$$Q2 = q_{0.5} = \frac{25.9 + 29.7}{2} = 27.8$$

$$Q3 = q_{0.75} = \frac{51.2 + 54.7}{2} = 52.95$$

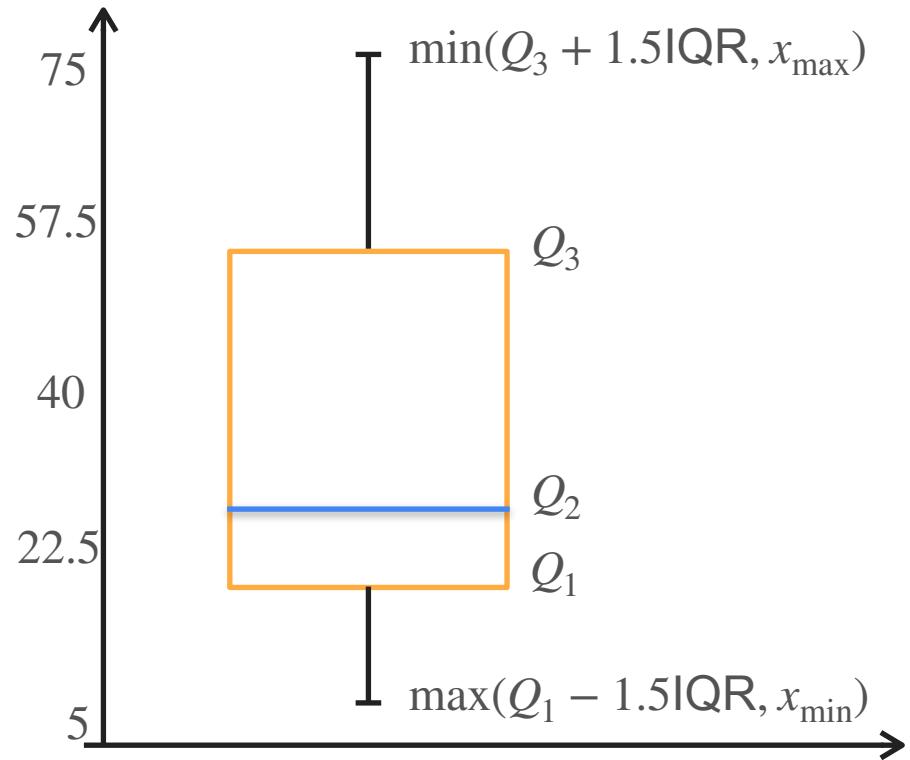
Interquartile range (IQR)

$$IQR = Q3 - Q1 = 52.95 - 18.35 = 34.6$$

$$x_{\min} = 8.7 \quad x_{\max} = 75$$



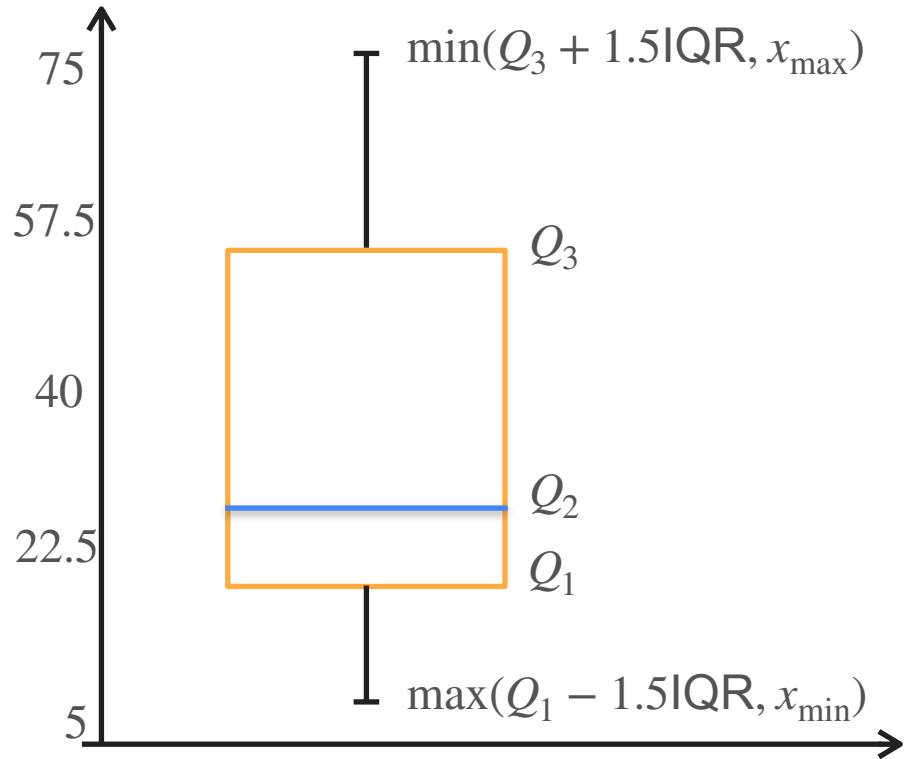
Box-Plots



Box-Plots

What can you tell from this plot?

- Data is skewed
- No outliers (whiskers were cut at max and min value)
- Analyze dispersion



Box-Plots

Box-Plots

Let's see how the box-plot looks for the whole dataset

Box-Plots

Let's see how the box-plot looks for the whole dataset

$$Q_1 = 12.75 \quad Q_2 = 25.75 \quad Q_3 = 45.1$$

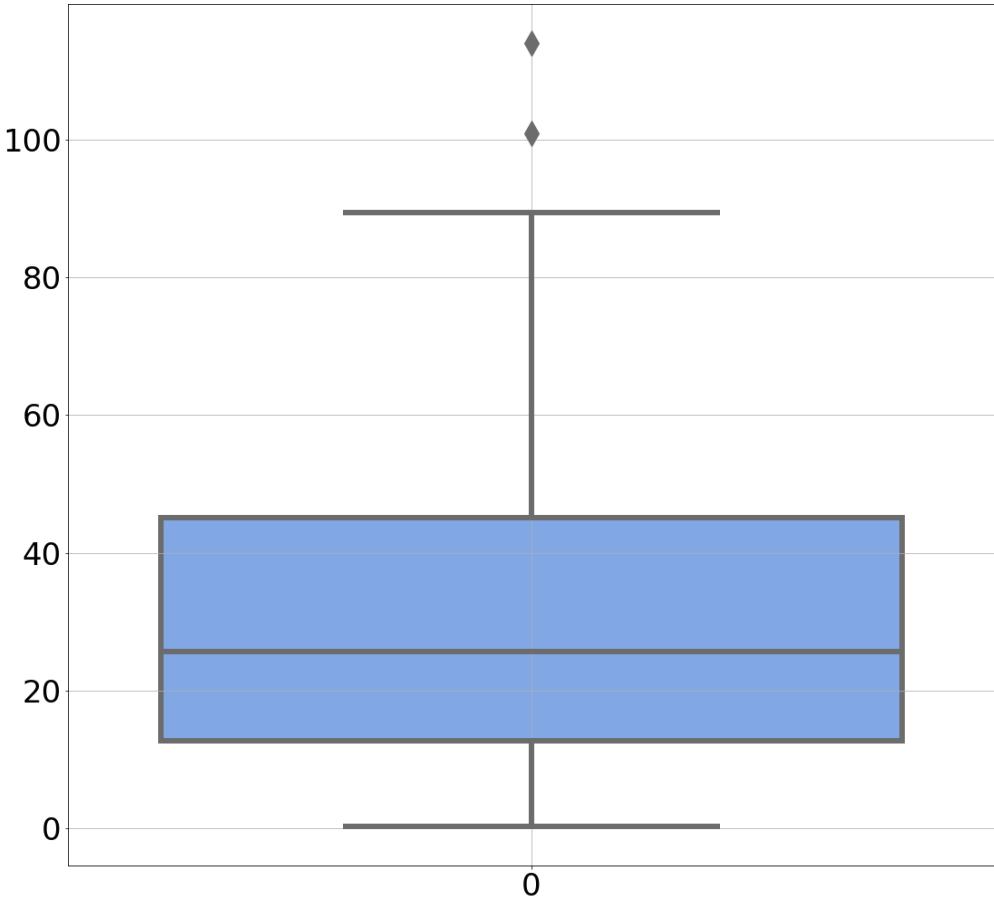
$$\text{IQR} = 32.35 \quad x_{\min} = 0.3 \quad x_{\max} = 114$$

Box-Plots

Let's see how the box-plot looks for the whole dataset

$$Q_1 = 12.75 \quad Q_2 = 25.75 \quad Q_3 = 45.1$$

$$\text{IQR} = 32.35 \quad x_{\min} = 0.3 \quad x_{\max} = 114$$

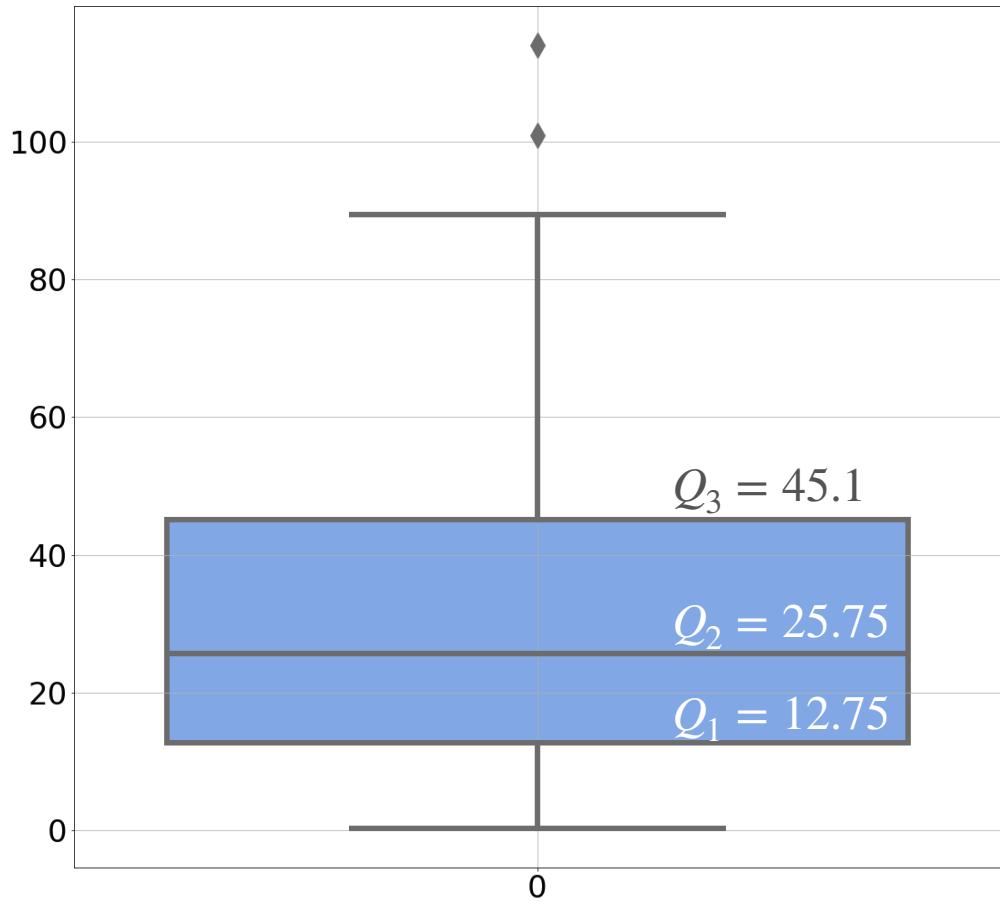


Box-Plots

Let's see how the box-plot looks for the whole dataset

$$Q_1 = 12.75 \quad Q_2 = 25.75 \quad Q_3 = 45.1$$

$$\text{IQR} = 32.35 \quad x_{\min} = 0.3 \quad x_{\max} = 114$$

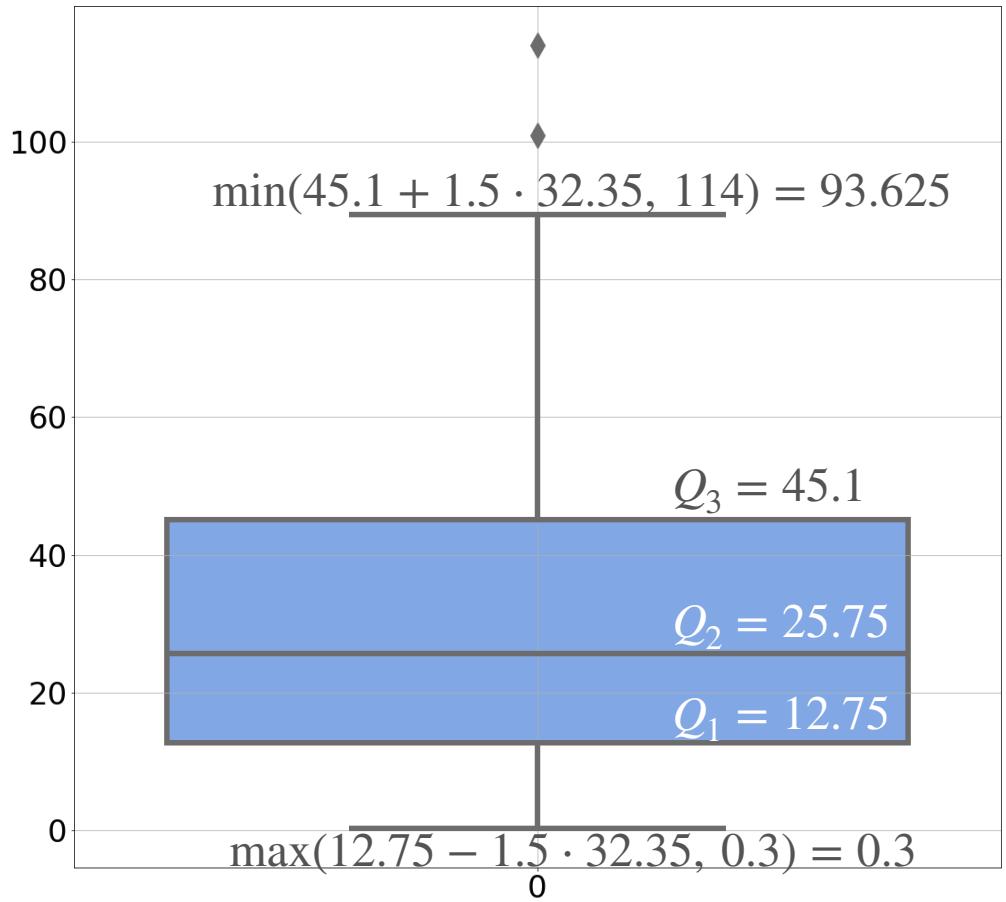


Box-Plots

Let's see how the box-plot looks for the whole dataset

$$Q_1 = 12.75 \quad Q_2 = 25.75 \quad Q_3 = 45.1$$

$$\text{IQR} = 32.35 \quad x_{\min} = 0.3 \quad x_{\max} = 114$$



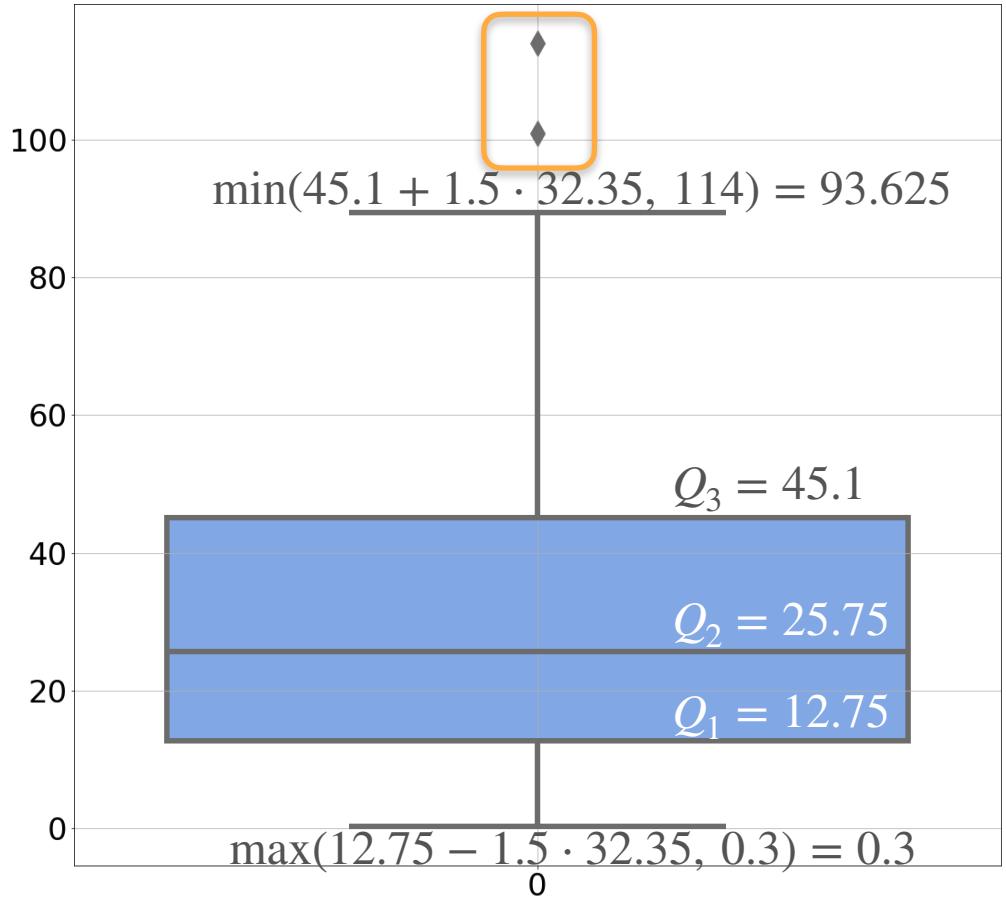
Box-Plots

Let's see how the box-plot looks for the whole dataset

$$Q_1 = 12.75 \quad Q_2 = 25.75 \quad Q_3 = 45.1$$

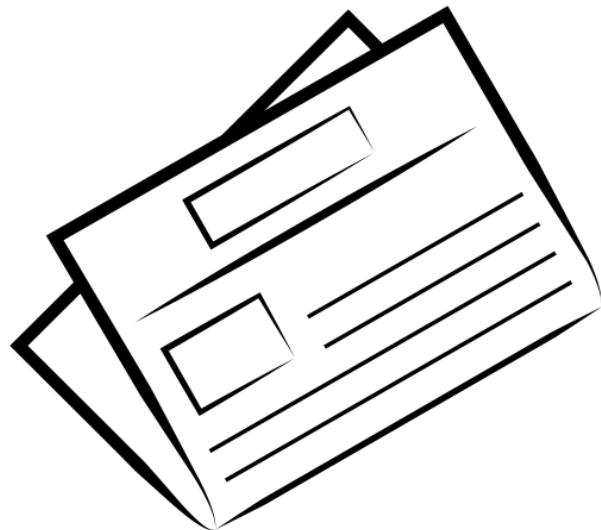
$$\text{IQR} = 32.35 \quad x_{\min} = 0.3 \quad x_{\max} = 114$$

Now you can see two outliers



Density Estimation

Density Estimation

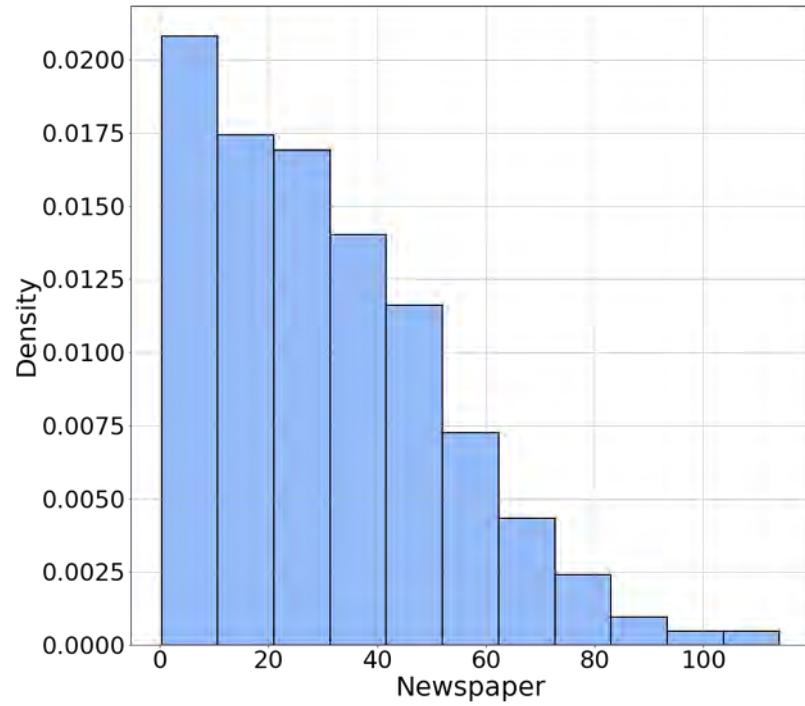


Newspaper advertisement

Newspaper Advertisement (X)			
8.7	14.2	18.3	18.4
23.2	25.9	29.7	35.2
51.2	54.7	65.9	75

Histograms

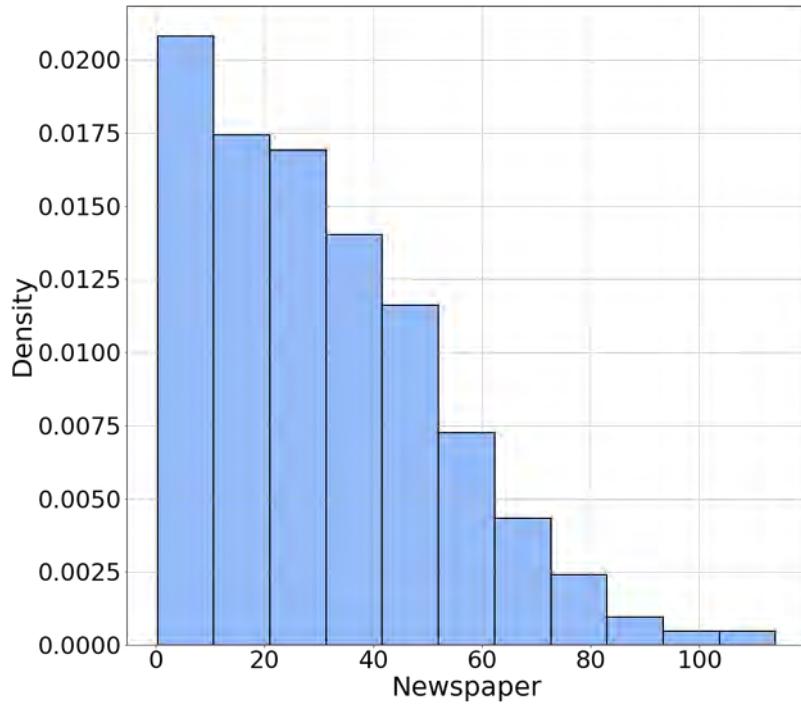
Histograms



Histograms

It represents a density function

- It is positive
- Area under the curve is 1



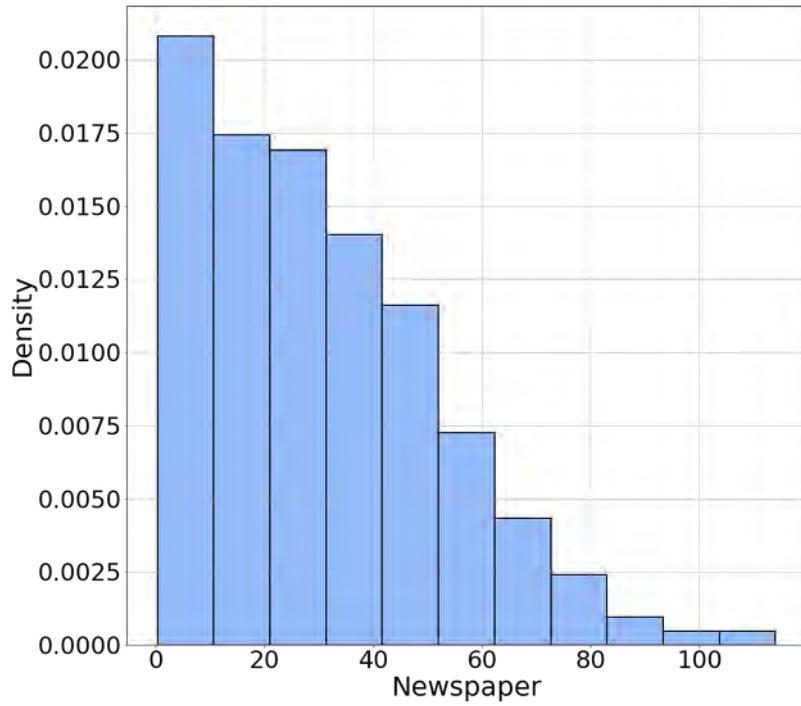
Histograms

It represents a density function

- It is positive
- Area under the curve is 1

But...

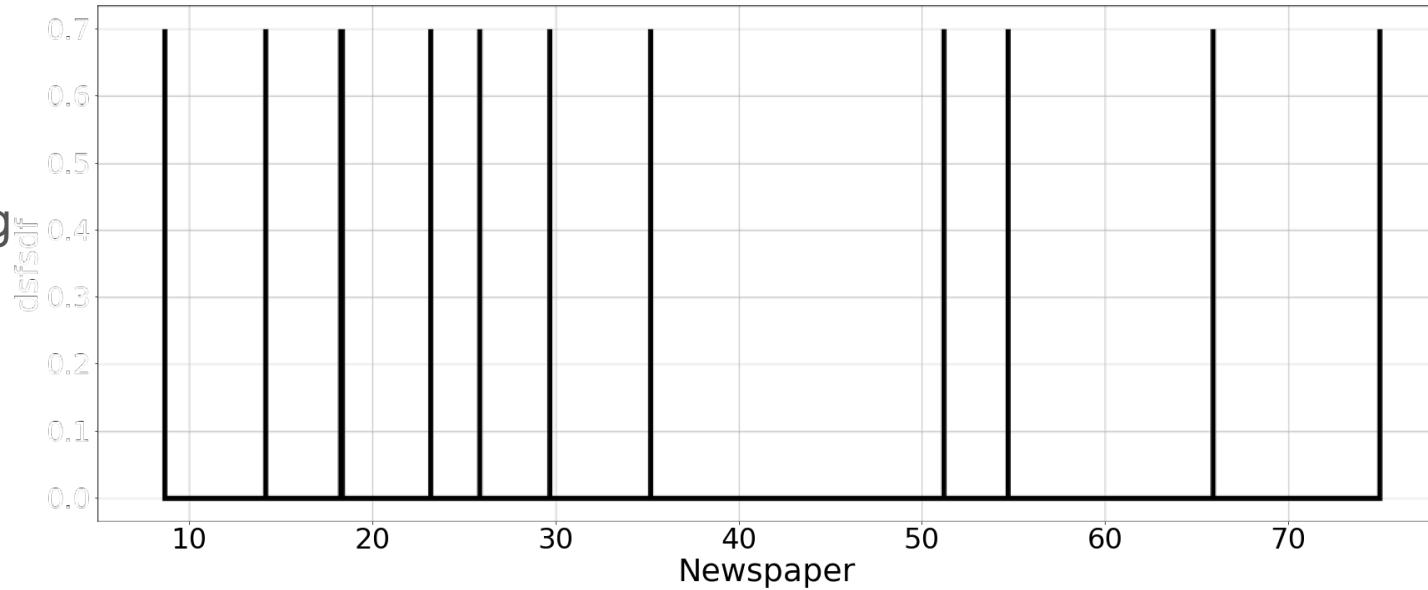
- PDFs are usually smooth function
- The discontinuities come from the method and not the data



Kernel Density Estimation

Kernel Density Estimation

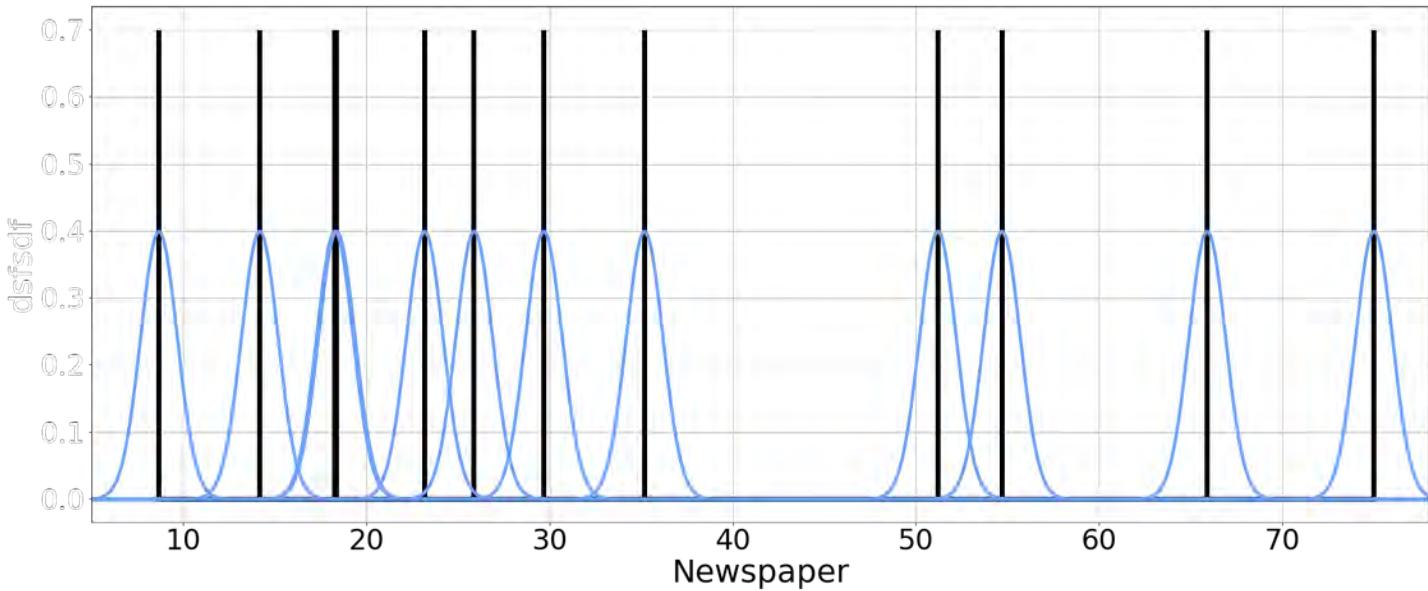
First: draw your observations along the x axis



Kernel Density Estimation

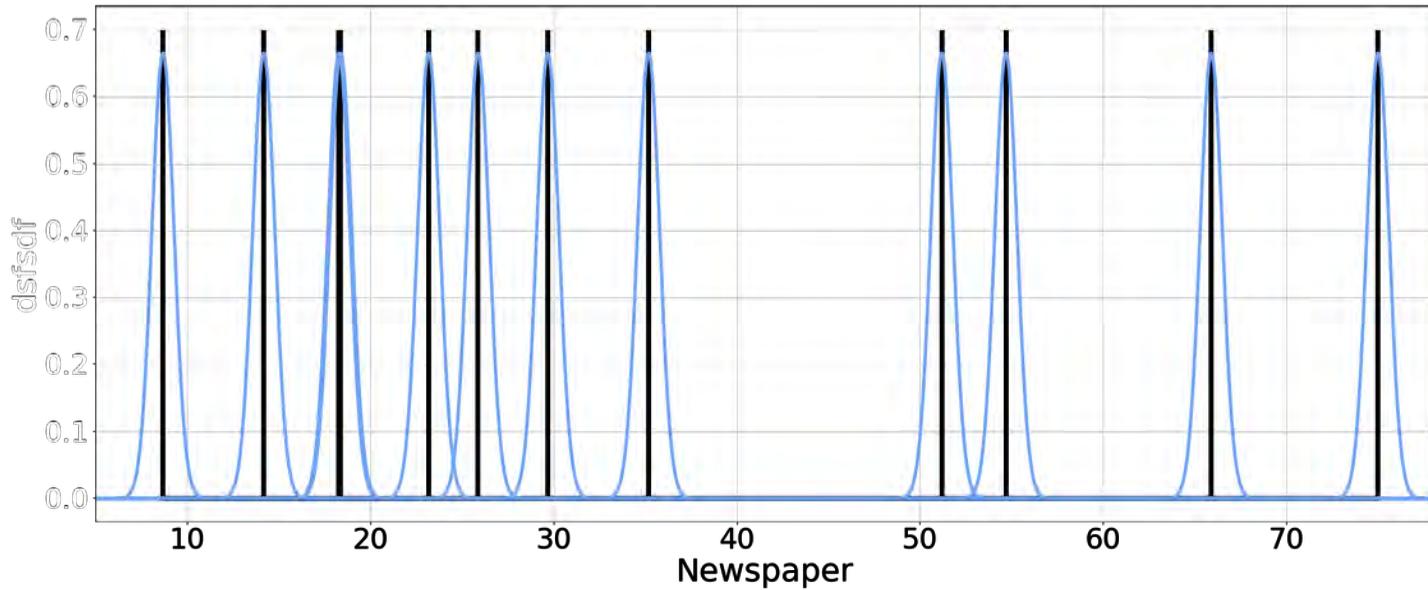
Kernel Density Estimation

Second: draw a gaussian centered at each observation



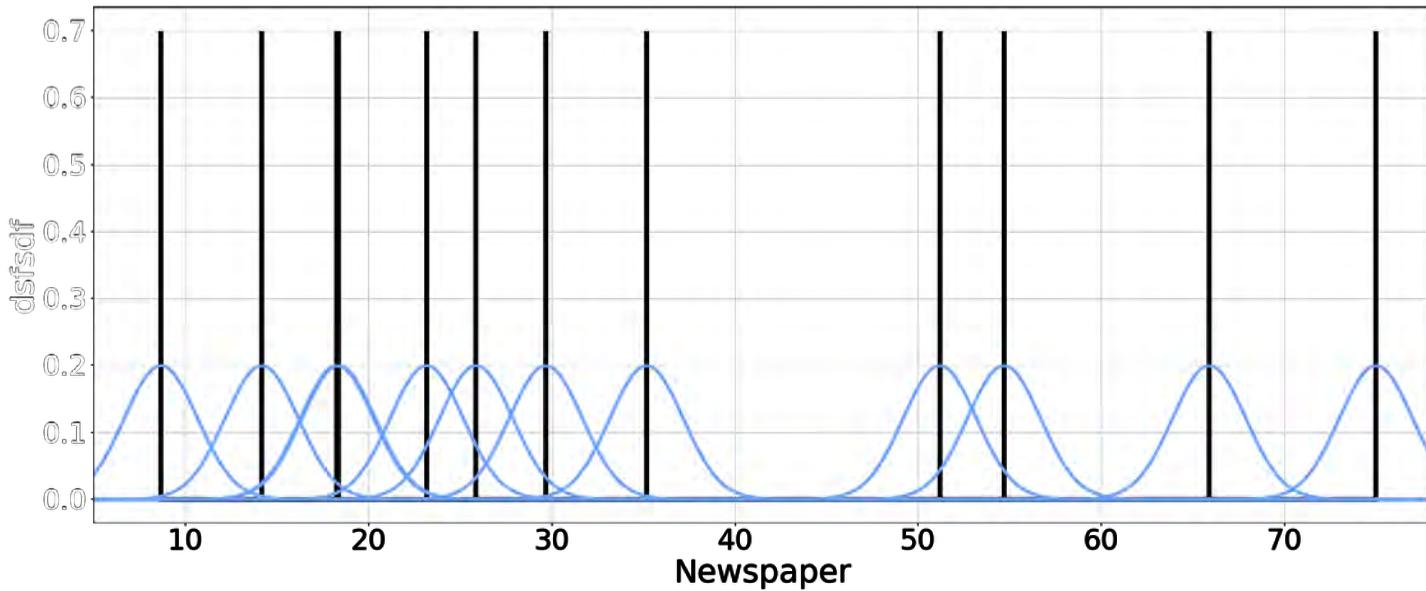
Kernel Density Estimation

Second: draw a gaussian centered at each observation



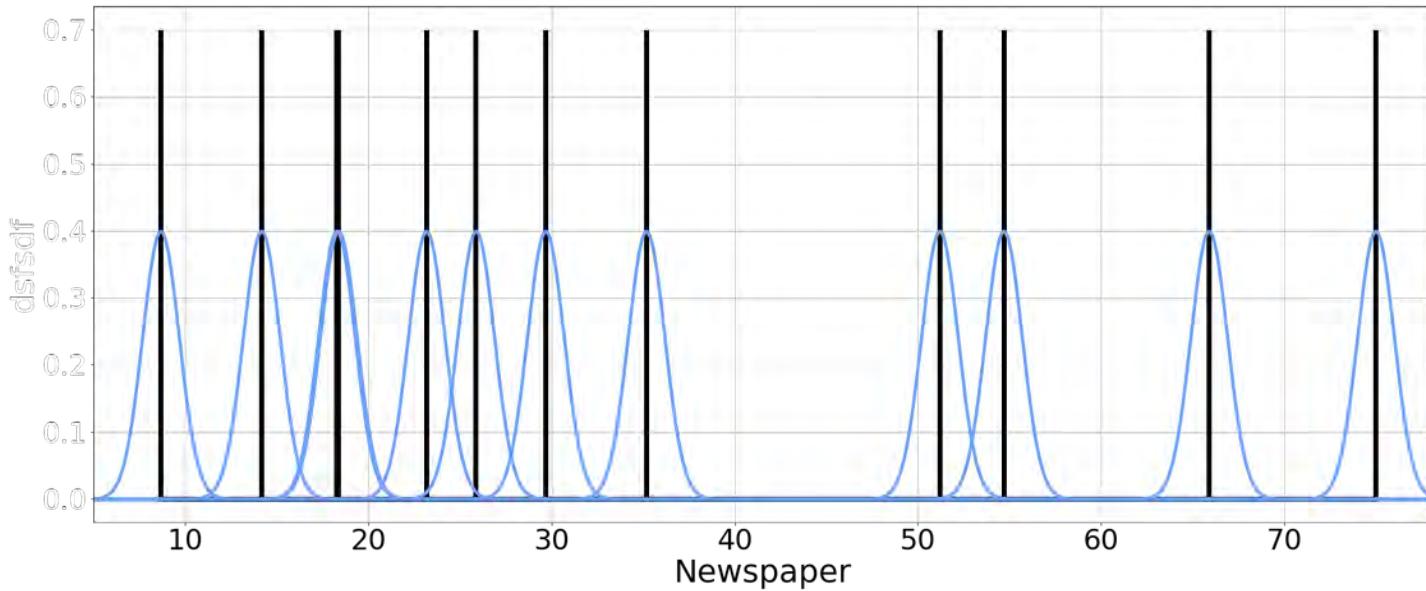
Kernel Density Estimation

Second: draw a gaussian centered at each observation

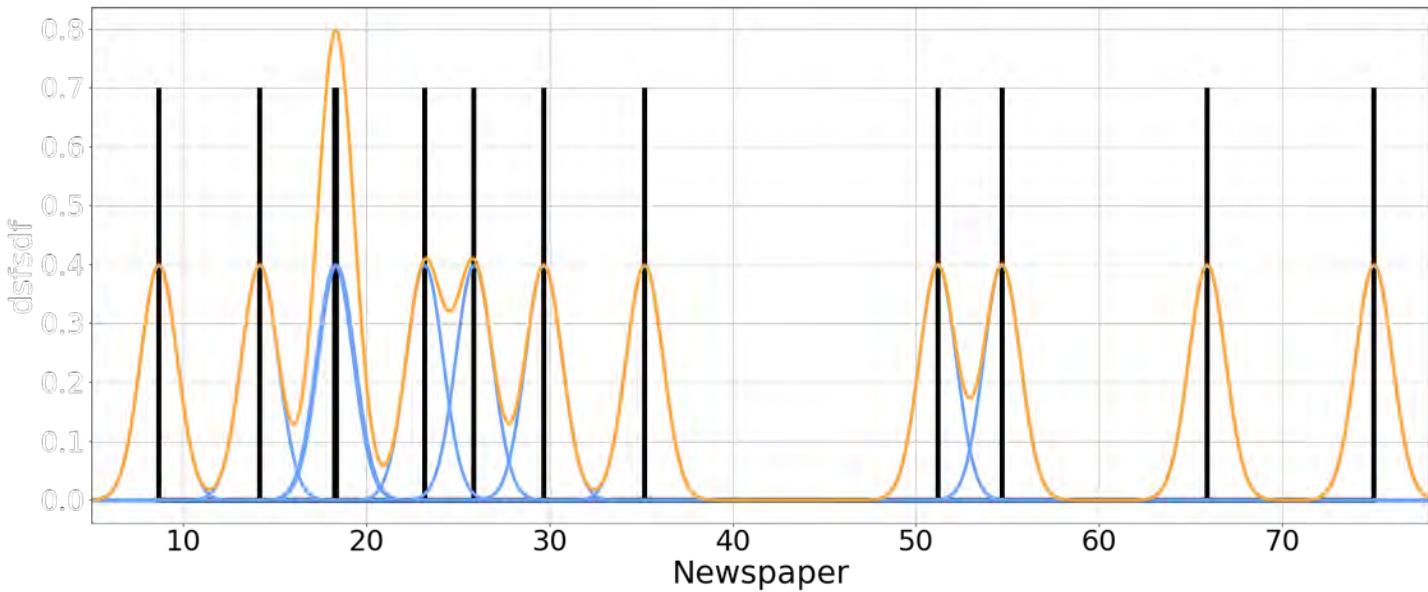


Kernel Density Estimation

Second: draw a gaussian centered at each observation

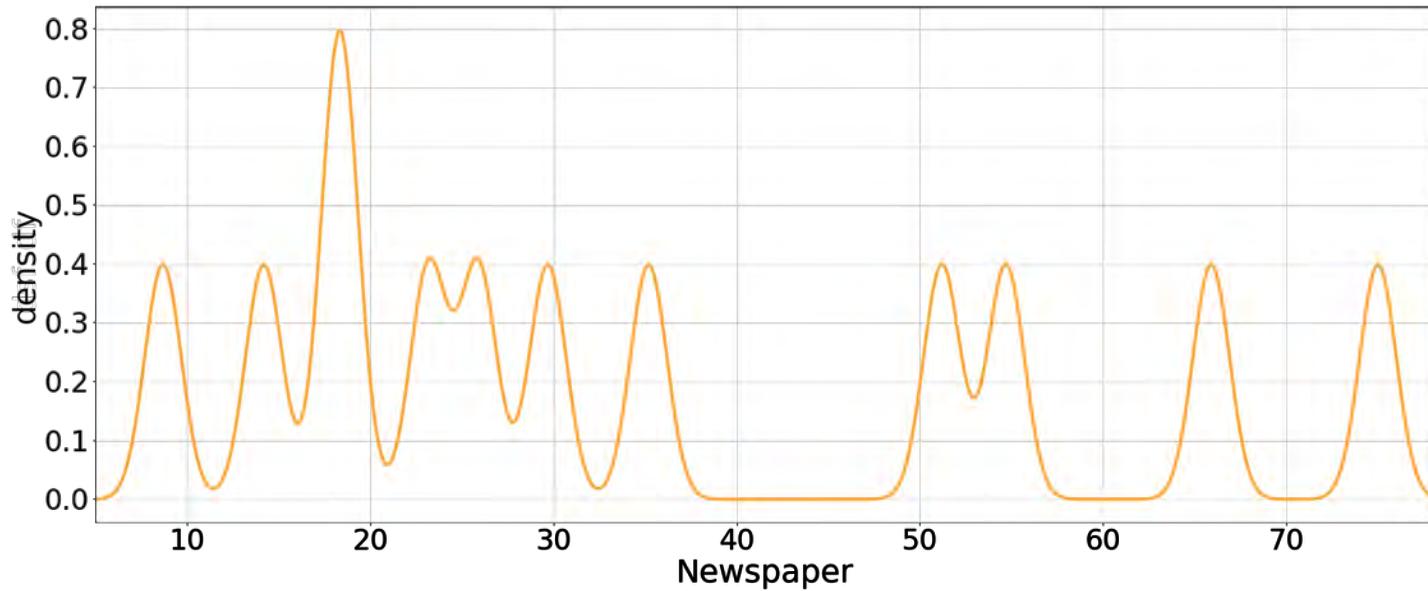


Kernel Density Estimation



Kernel Density Estimation

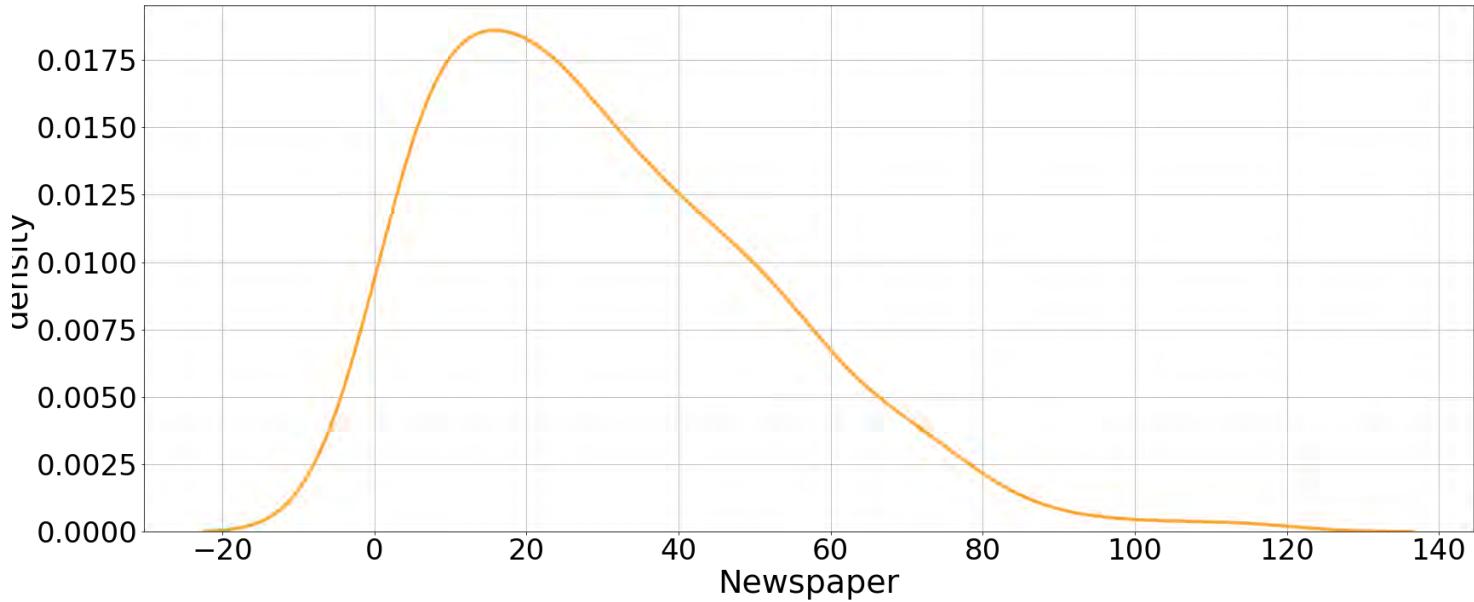
Third: multiply everything by $1/n$ and sum the curves



Kernel Density Estimation

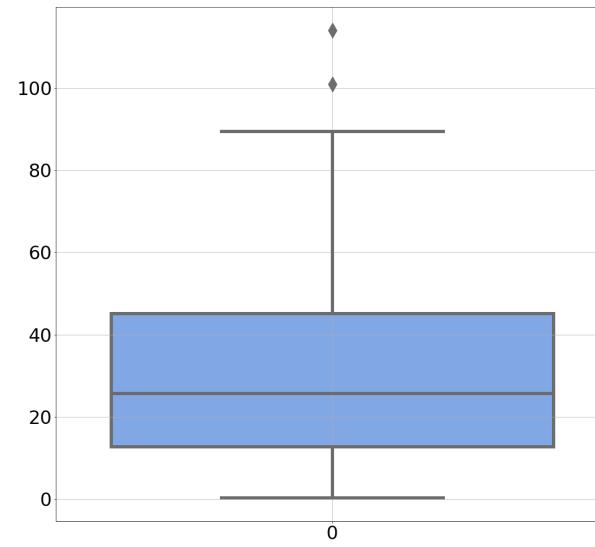
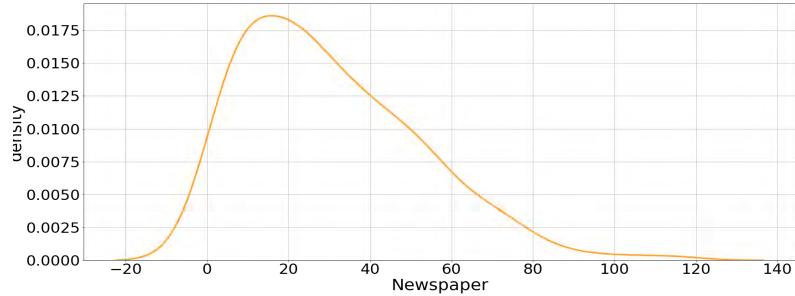
Kernel Density Estimation

What if
you used
all the
dataset?

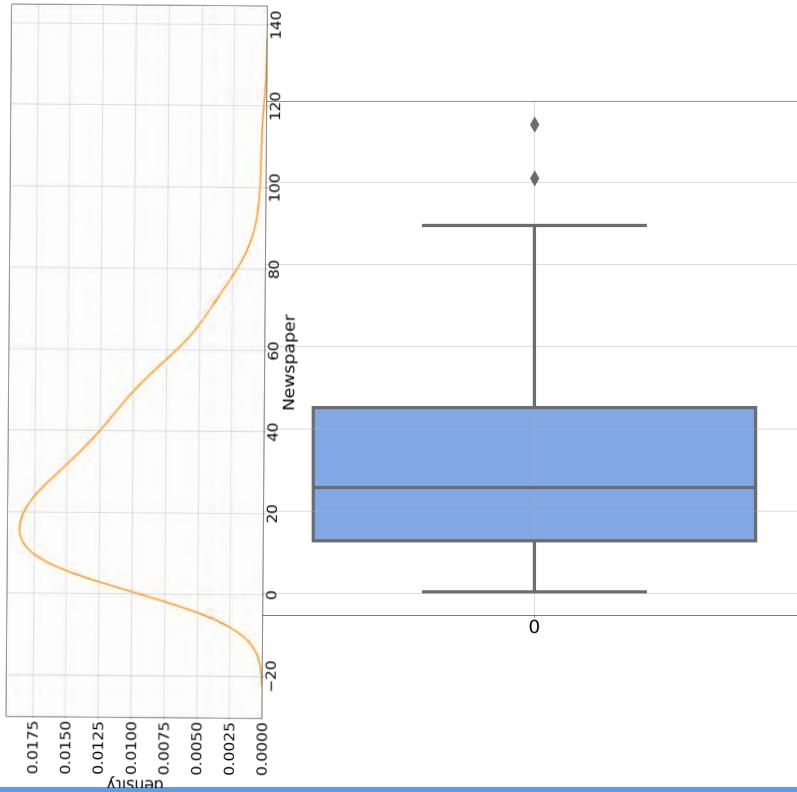


Violin Plots

Violin Plots

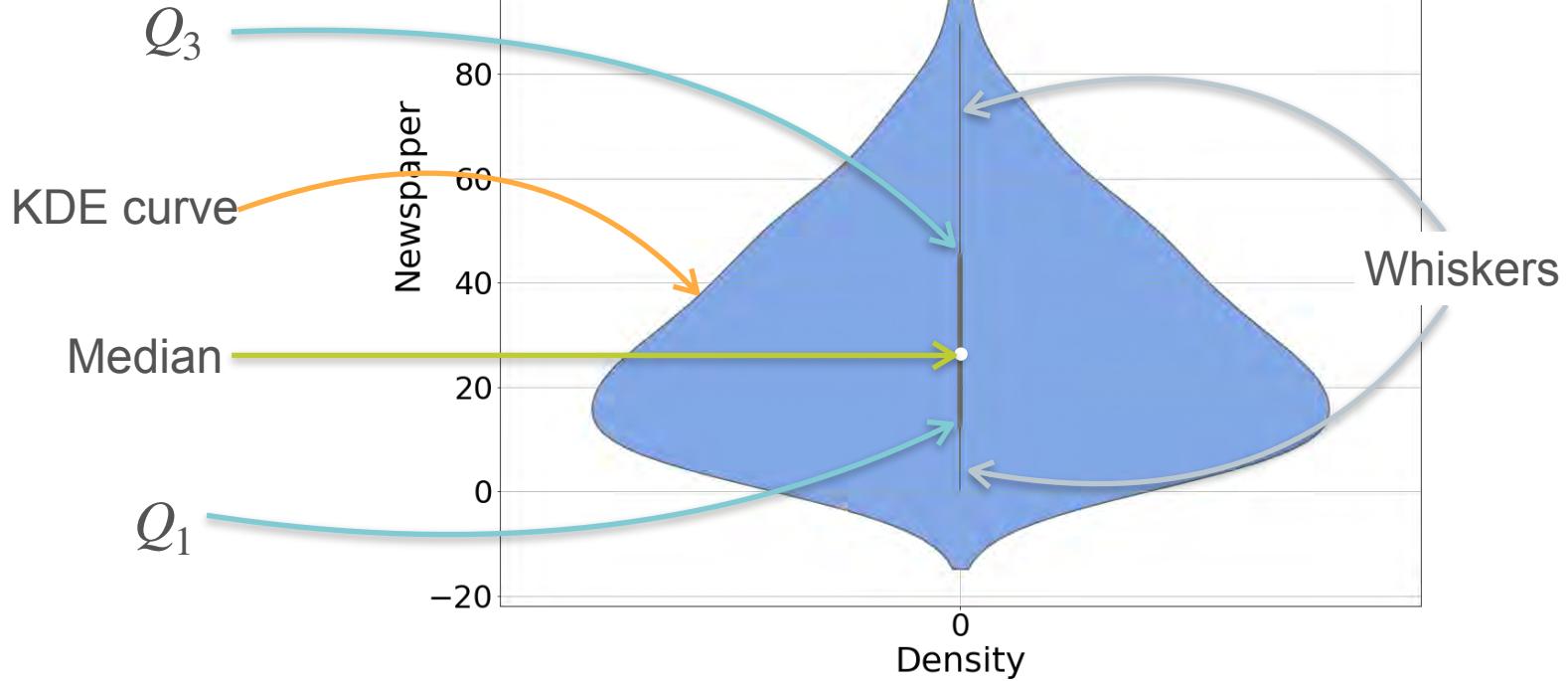


Violin Plots



Violin Plots

Violin Plots



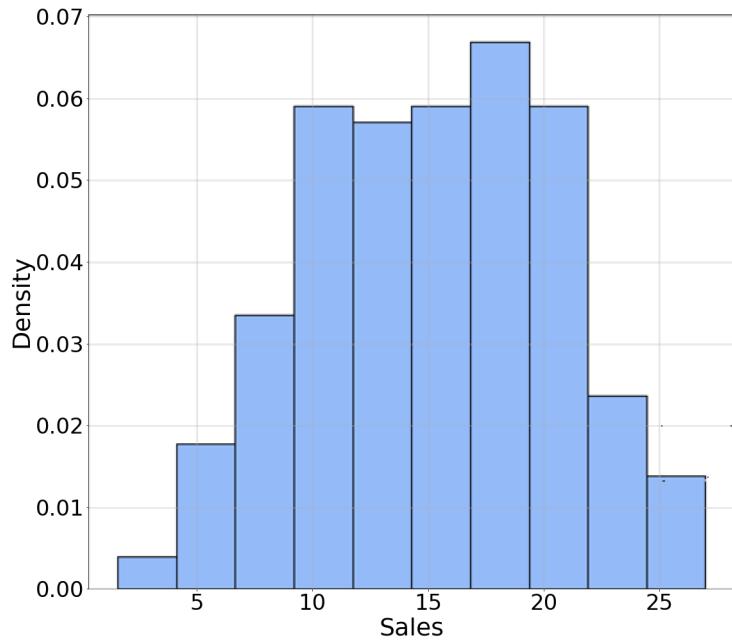
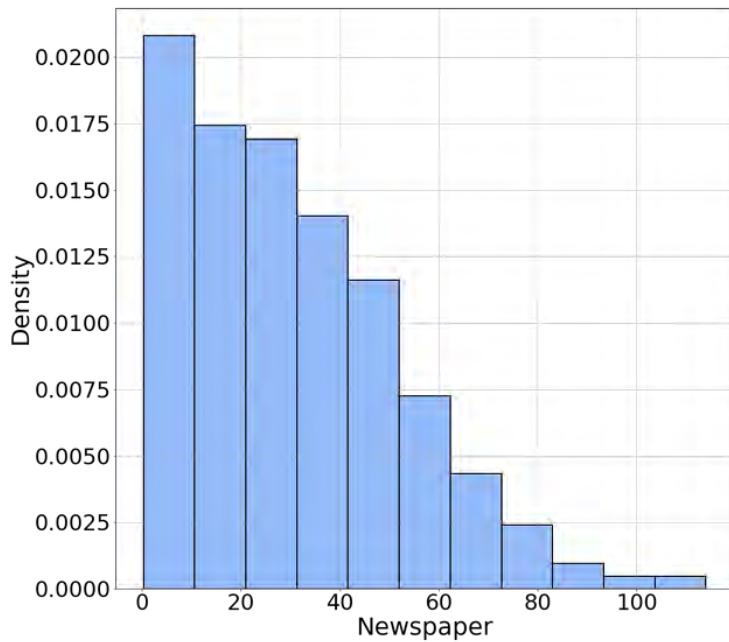
Assessing Normality of Data

Some models assume normally distributed data

- Linear regression
- Logistic regression
- Gaussian Naive Bayes
- Others

Some tests used in Data Science also assume normality.

Assessing Normality of Data



Newspaper

QQ Plots

QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

- Compute quantiles

QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

- Compute quantiles
- Compare to gaussian quantiles

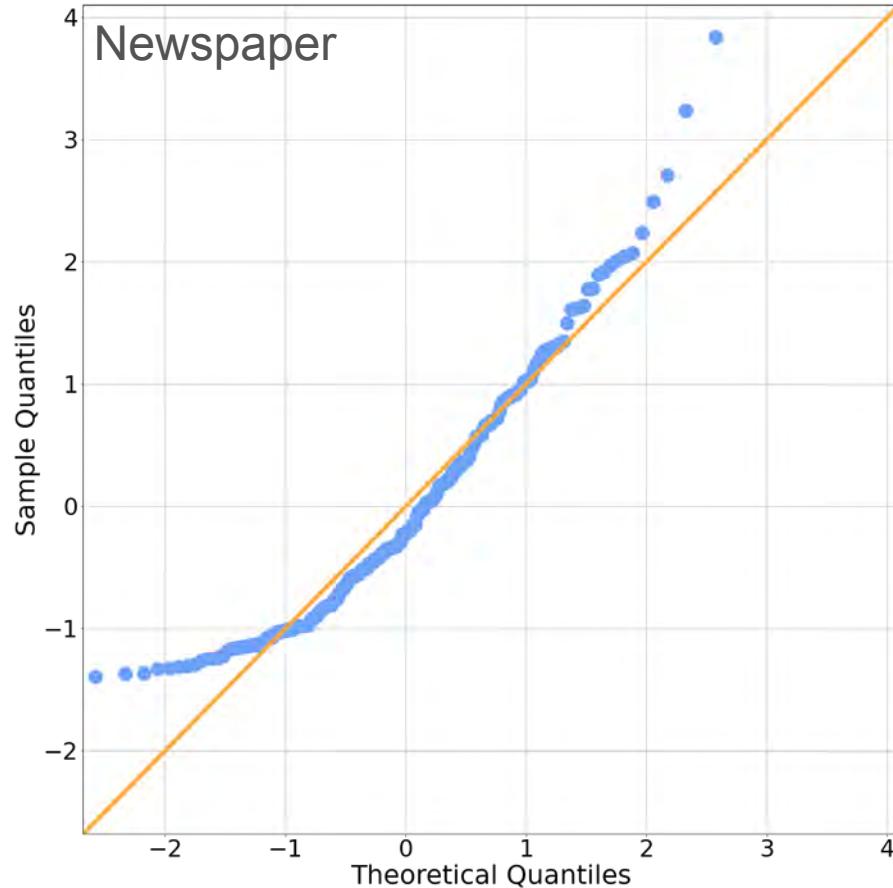
QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

- Compute quantiles
- Compare to gaussian quantiles



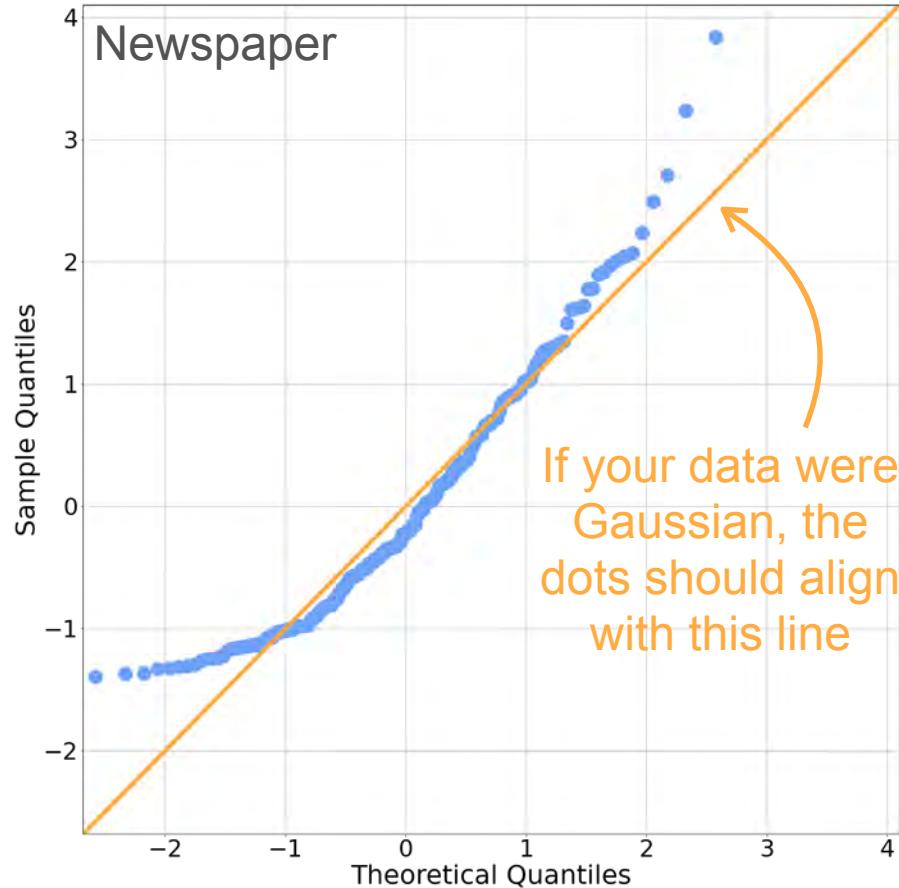
QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

- Compute quantiles
- Compare to gaussian quantiles



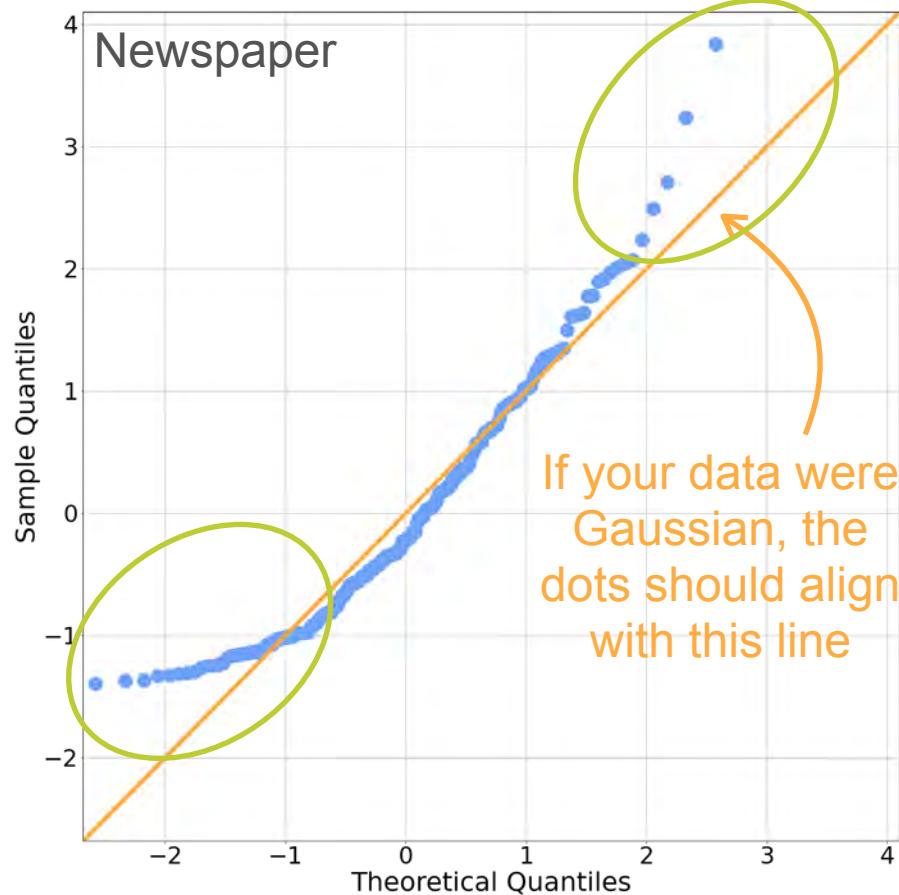
QQ Plots

Quantile-Quantile plots (QQ Plots) compare quantiles

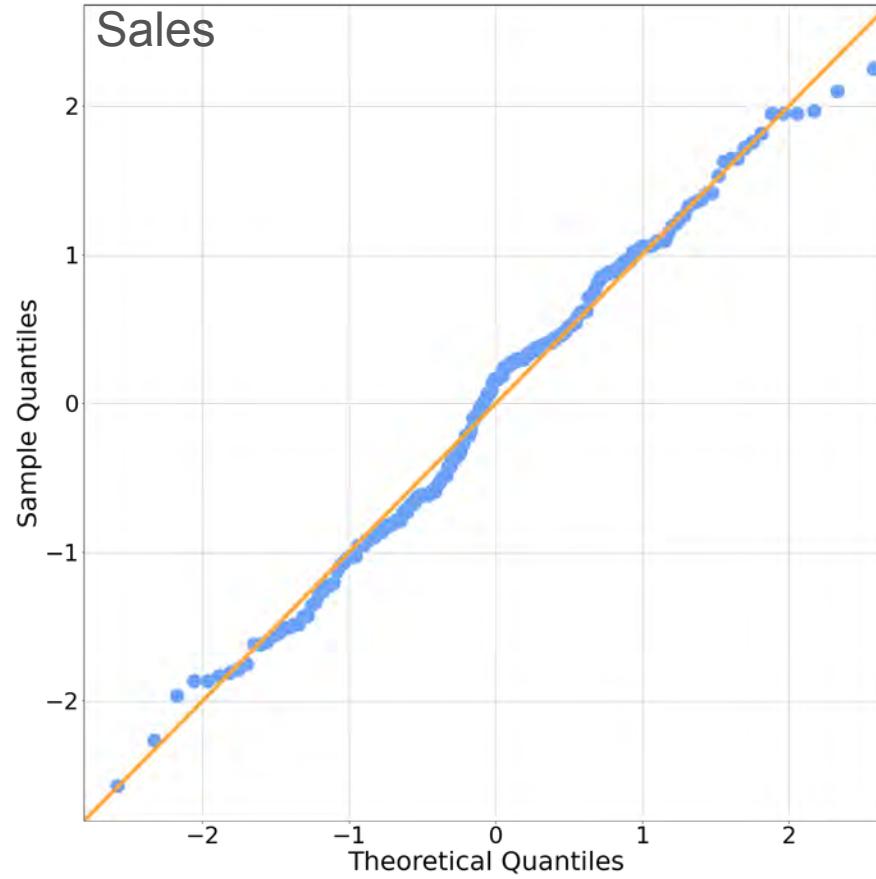
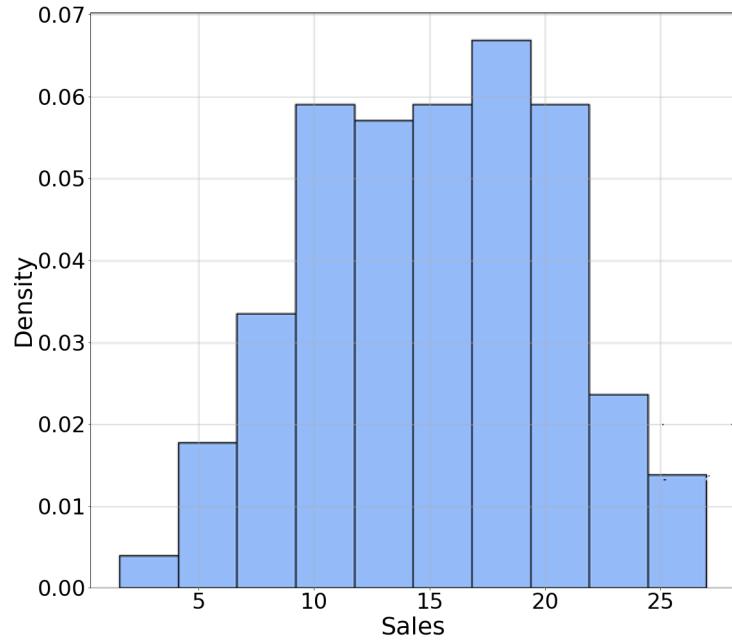
- Standardize your data:

$$\left(\frac{x - \mu}{\sigma} \right)$$

- Compute quantiles
- Compare to gaussian quantiles



QQ Plots



W2 Lesson 2



DeepLearning.AI

Probability Distributions with Multiple Variables

Joint Distribution (Discrete)

Joint Distributions (Discrete): Example 1

Joint Distributions (Discrete): Example 1

Age (Year)	Count
7	3
8	2
9	4
10	1

Joint Distributions (Discrete): Example 1

Age (Year)	Count
7	3
8	2
9	4
10	1



Joint Distributions (Discrete): Example 1

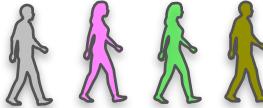
Age (Year)	Count
7	3
8	2
9	4
10	1



Joint Distributions (Discrete): Example 1

Age (Year)	Count	
7	3	
8	2	
9	4	
10	1	

Joint Distributions (Discrete): Example 1

Age (Year)	Count	
7	3	
8	2	
9	4	
10	1	

Joint Distributions (Discrete): Example 1

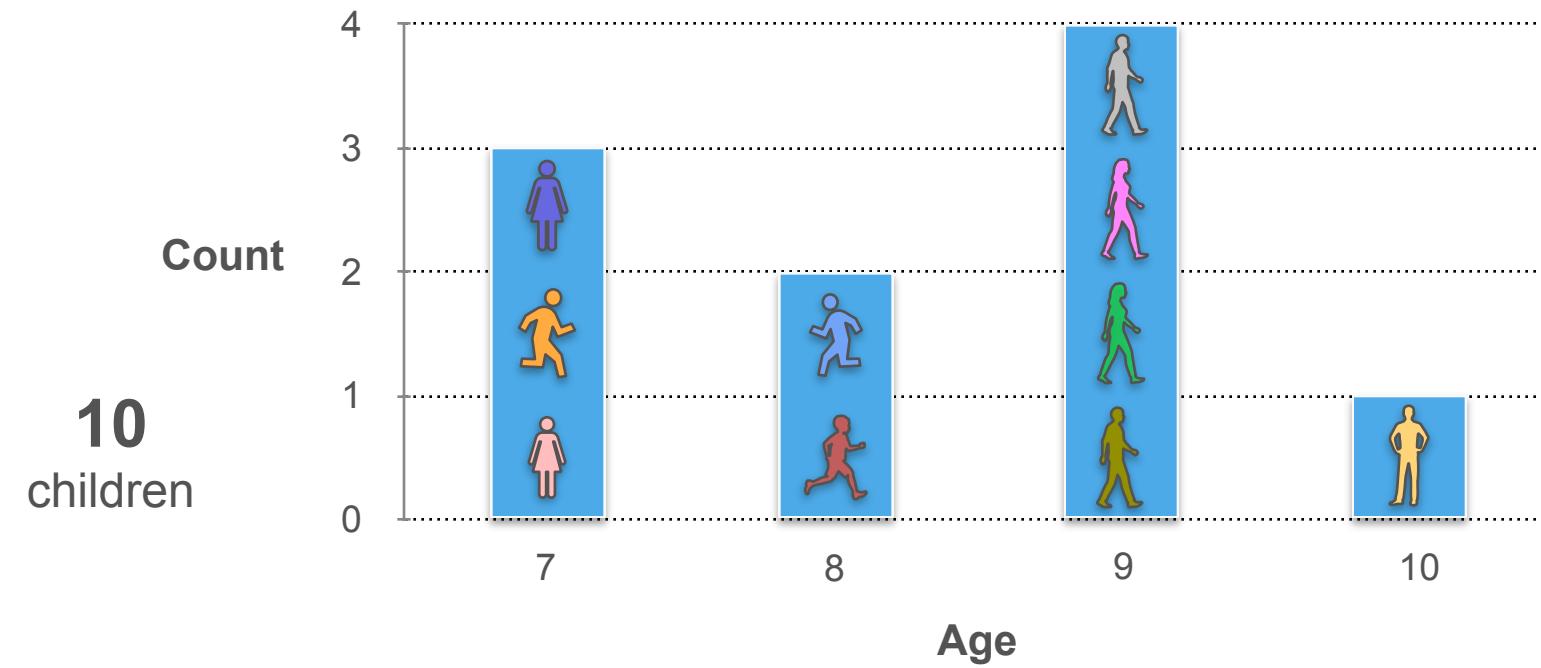
Age (Year)	Count
7	3
8	2
9	4
10	1



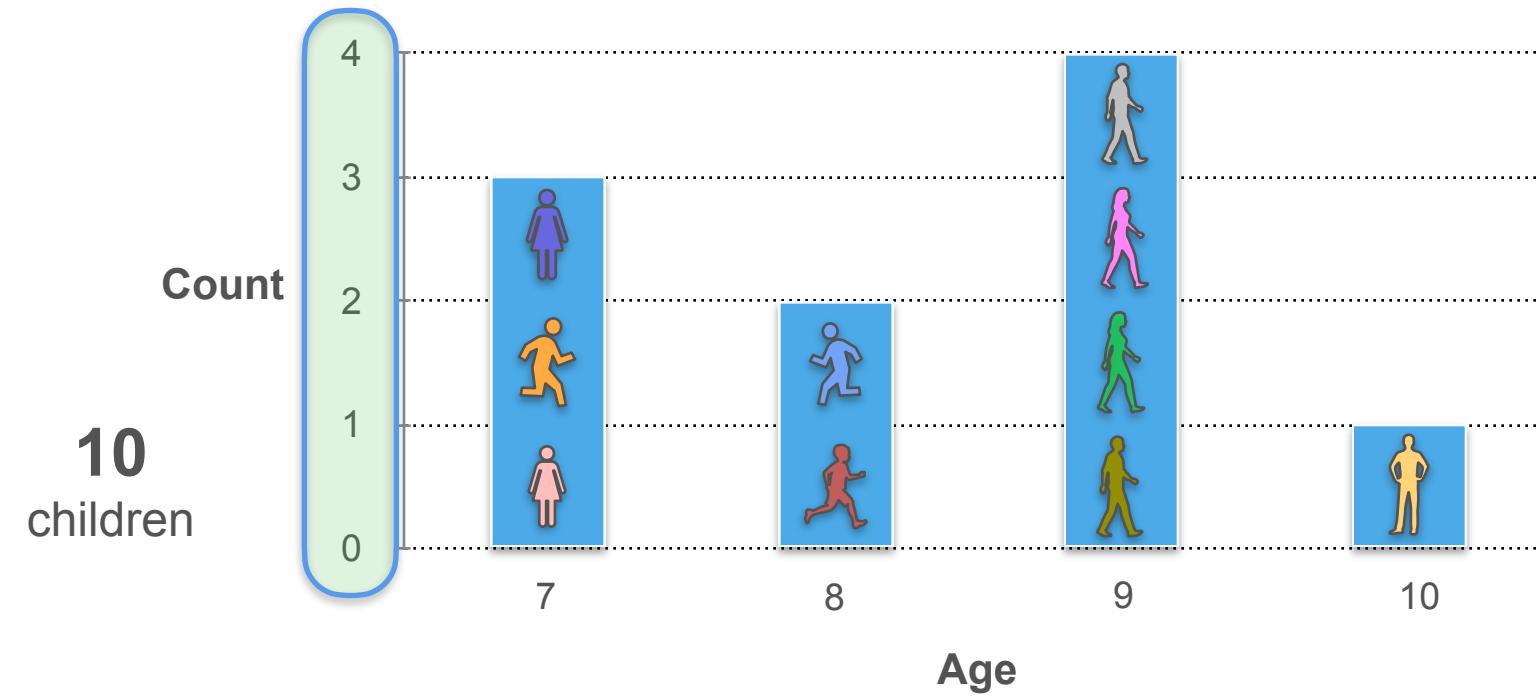
10

children

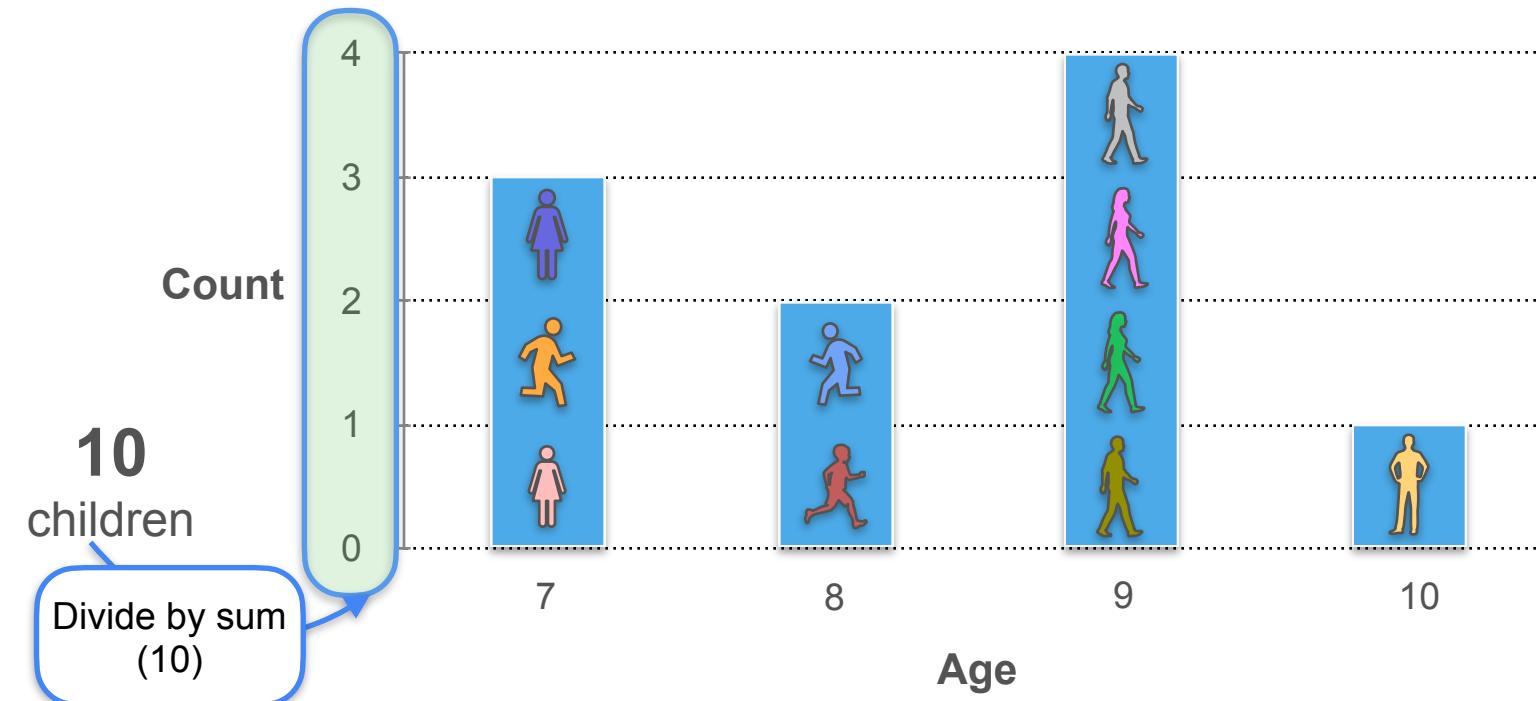
Joint Distributions (Discrete): Example 1



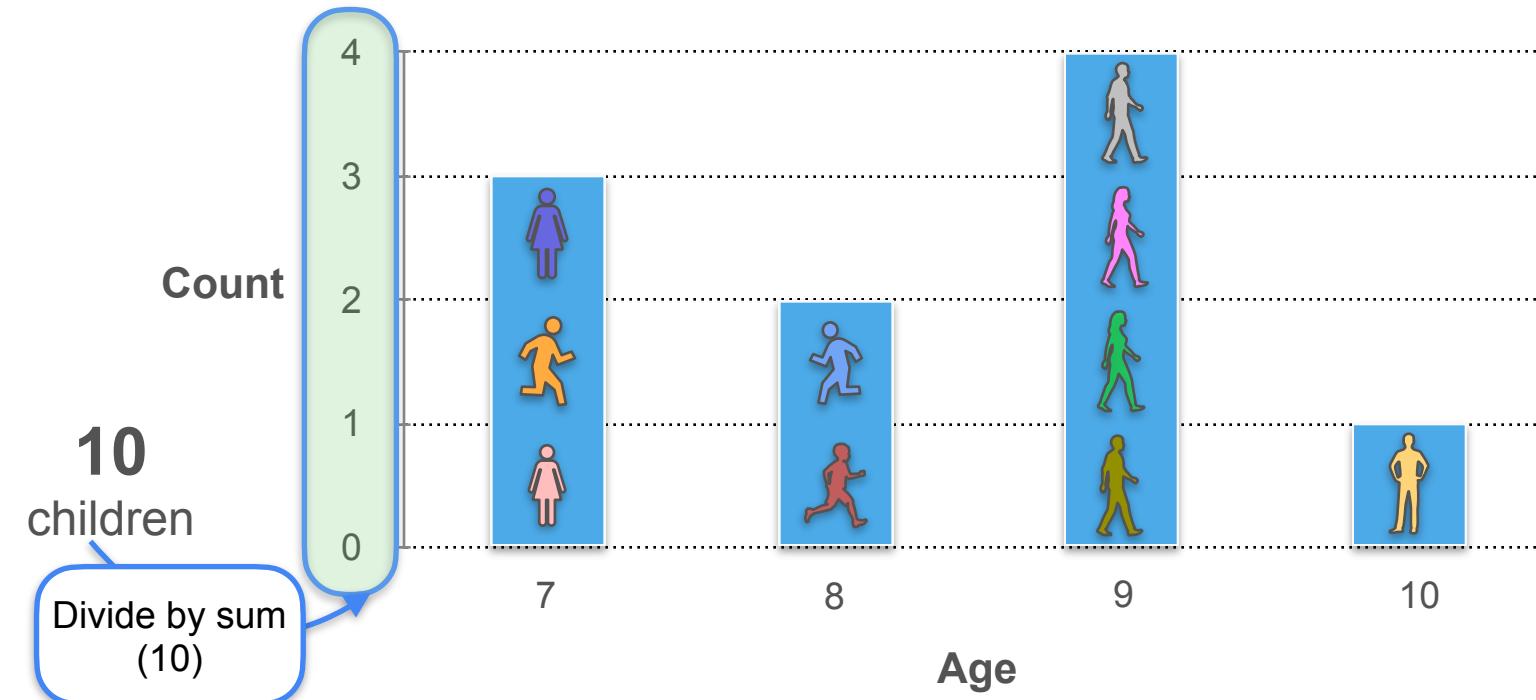
Joint Distributions (Discrete): Example 1



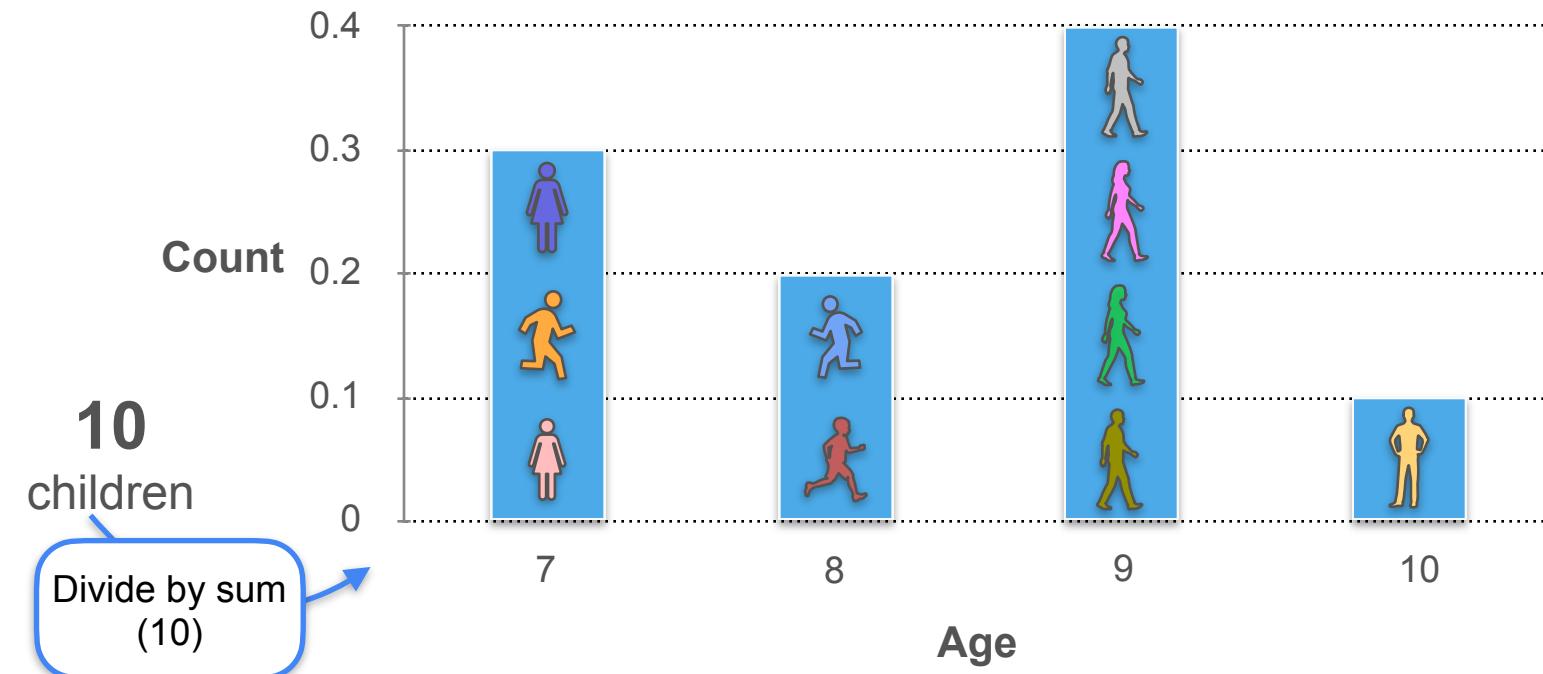
Joint Distributions (Discrete): Example 1



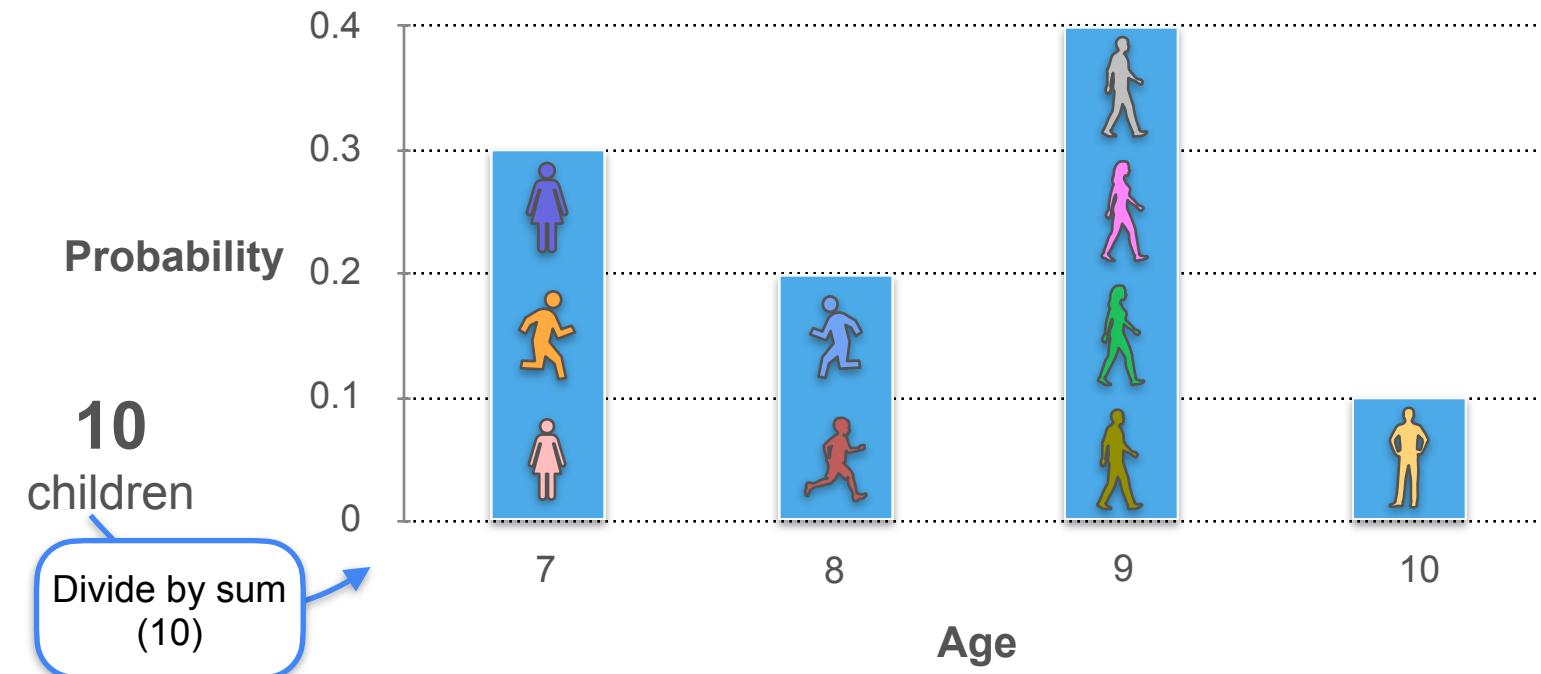
Joint Distributions (Discrete): Example 1



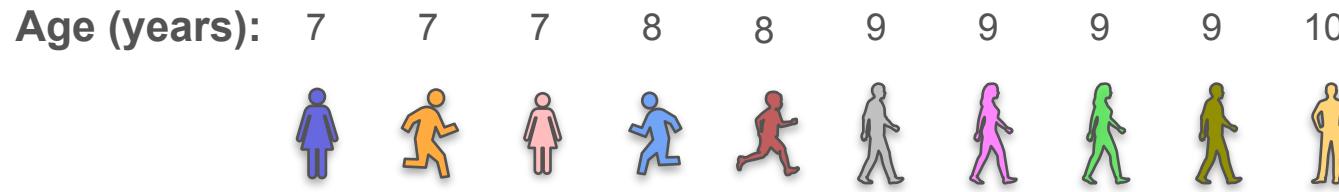
Joint Distributions (Discrete): Example 1



Joint Distributions (Discrete): Example 1



Joint Distributions (Discrete): Example 1



Joint Distributions (Discrete): Example 1

Age (years): 7 7 7 8 8 9 9 9 9 10



Height (in):

Joint Distributions (Discrete): Example 1

Age (years):	7	7	7	8	8	9	9	9	9	10
										
Height (in):	45	46	46	47	47	49	49	49	49	50

Joint Distributions (Discrete): Example 1

Age (years):

7 7 7 8 8 9 9 9 9 10

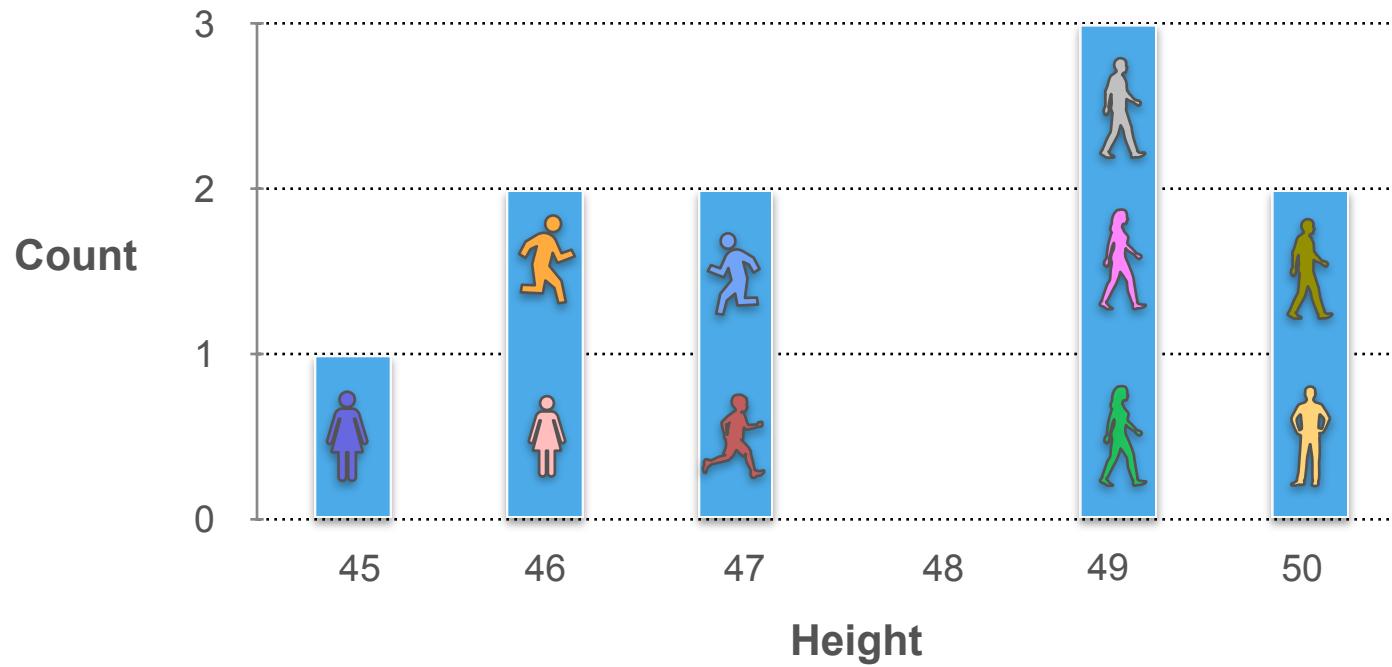


Height (in):

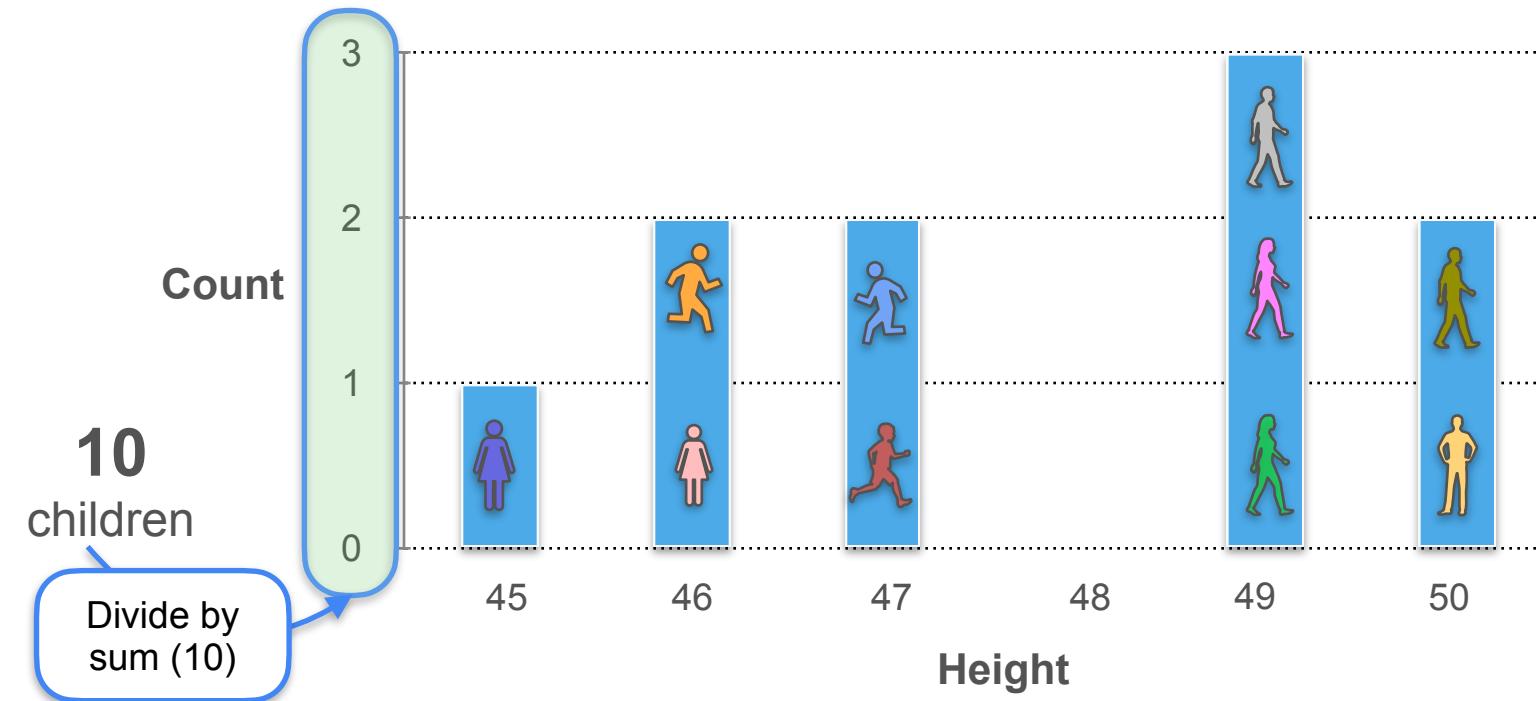
45 46 46 47 47 49 49 49 49 50

Height (in)	Count
45	1
46	2
47	2
48	0
49	4
50	1

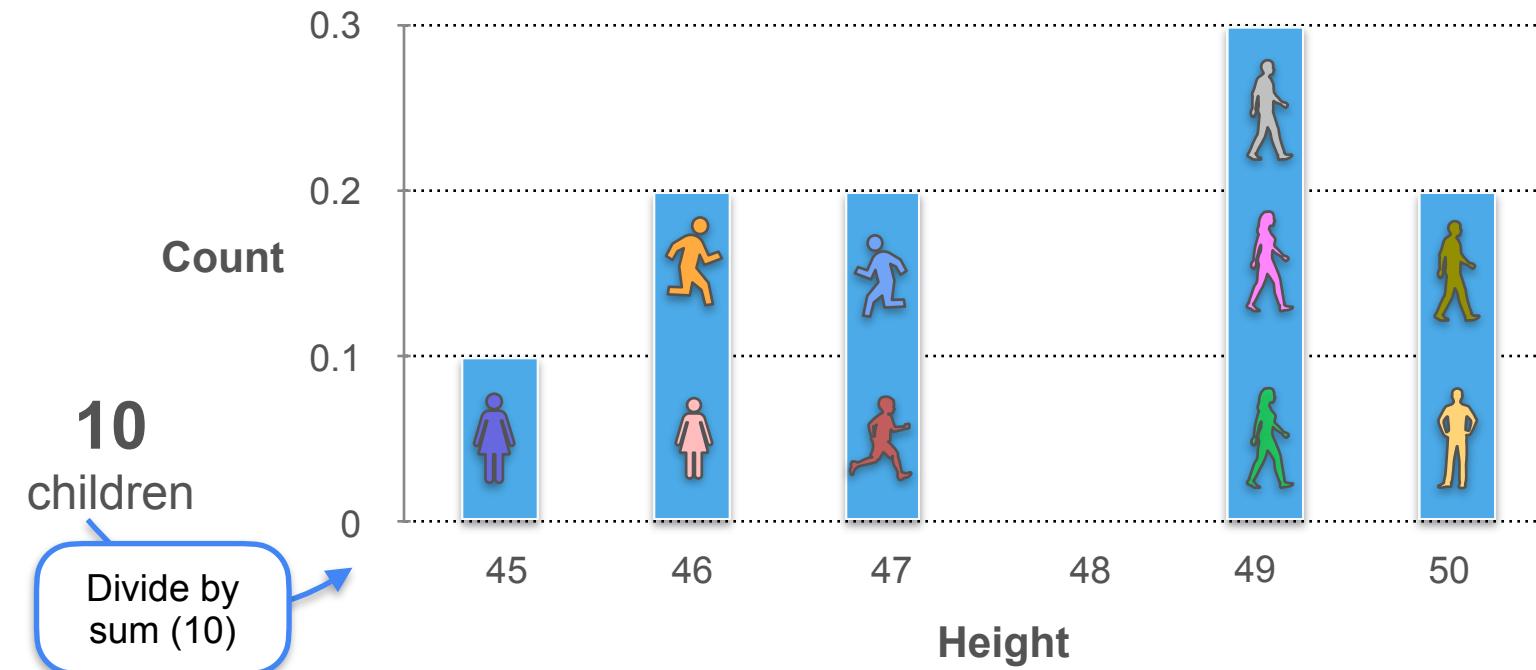
Joint Distributions (Discrete): Example 1



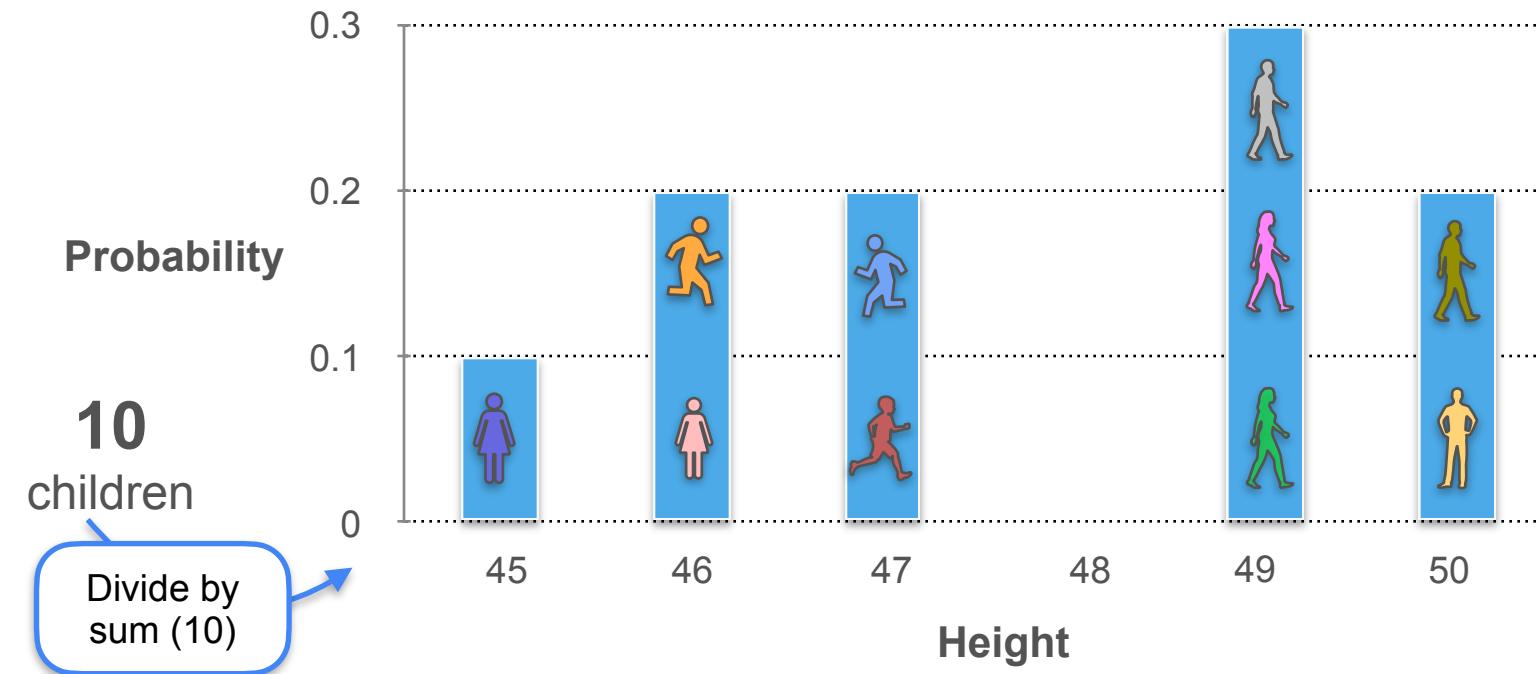
Joint Distributions (Discrete): Example 1



Joint Distributions (Discrete): Example 1

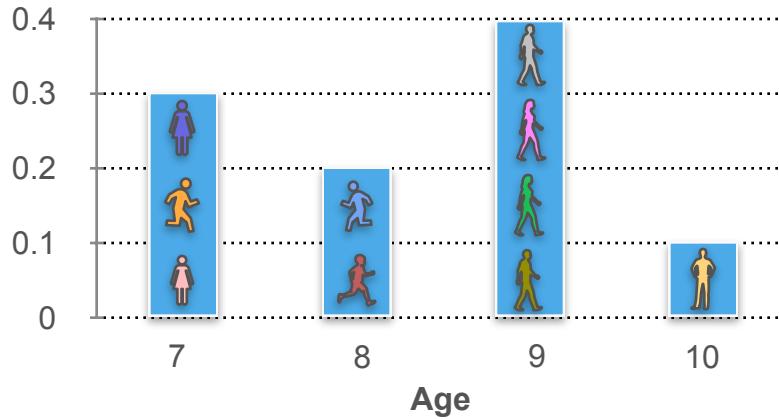


Joint Distributions (Discrete): Example 1

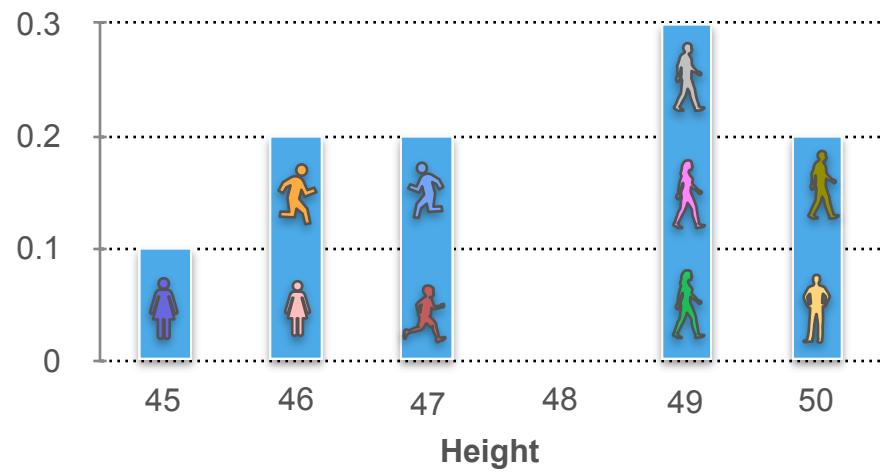
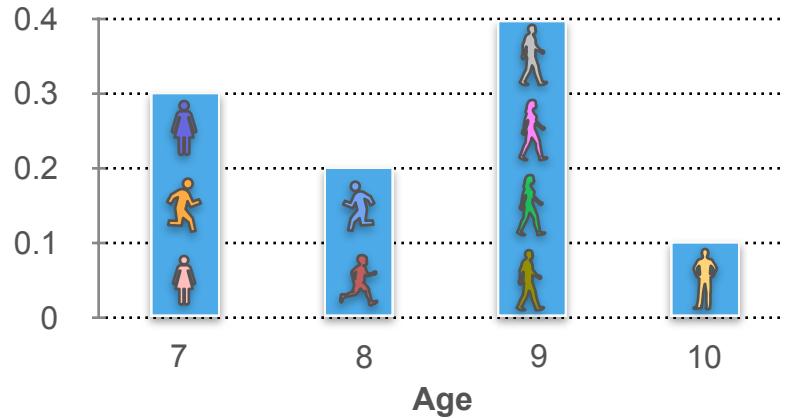


Joint Distributions (Discrete): Example 1

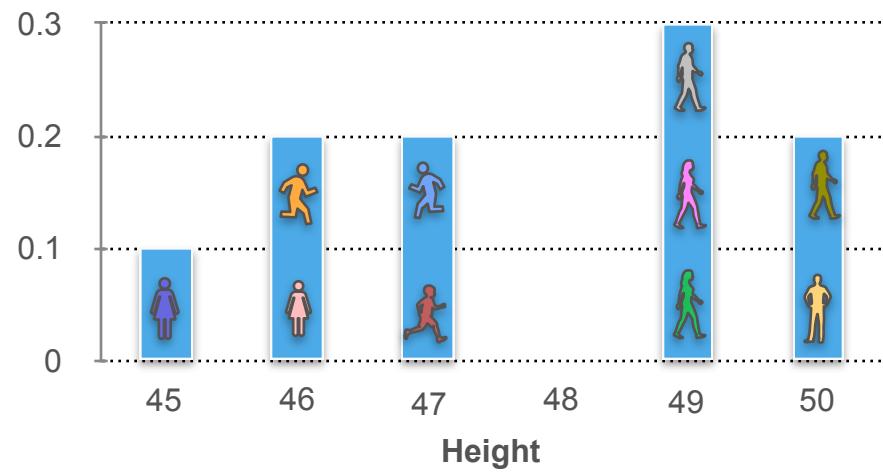
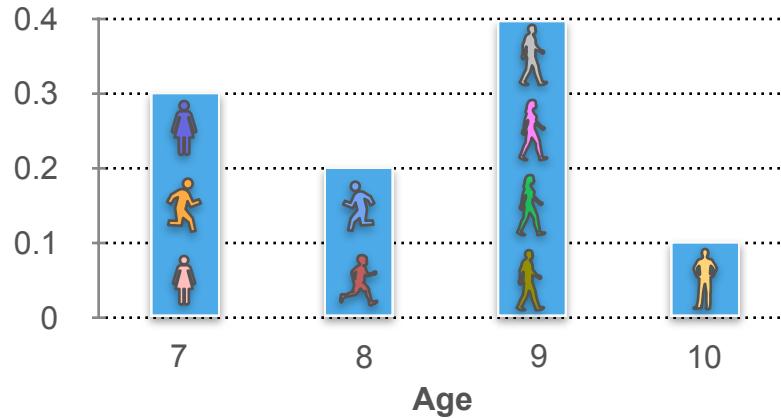
Joint Distributions (Discrete): Example 1



Joint Distributions (Discrete): Example 1

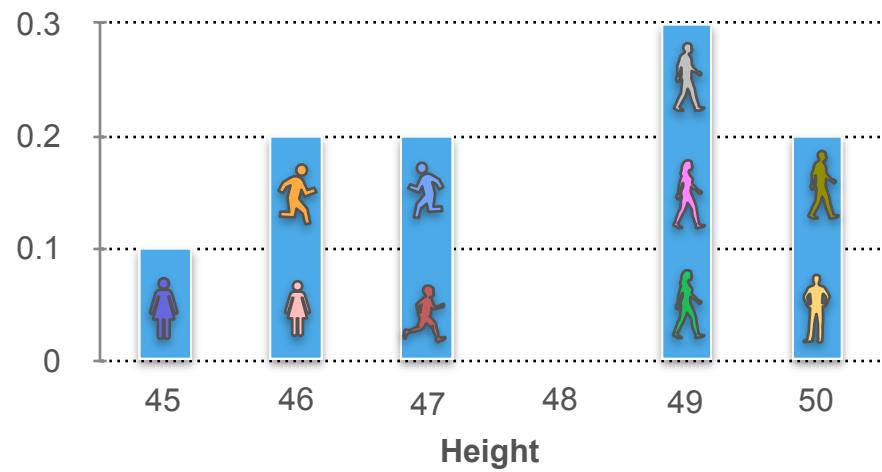
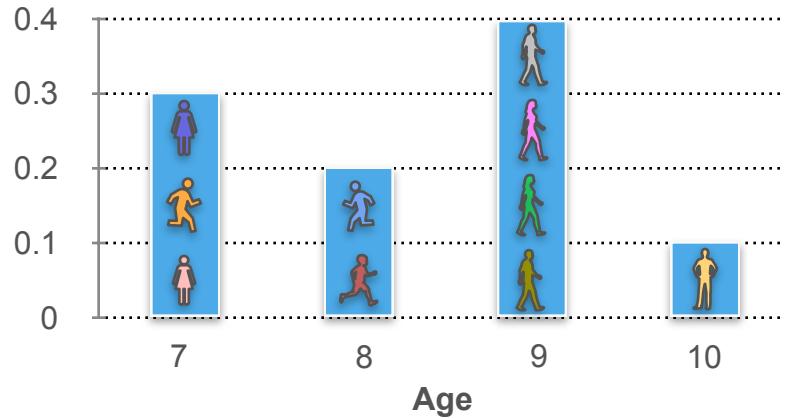


Joint Distributions (Discrete): Example 1



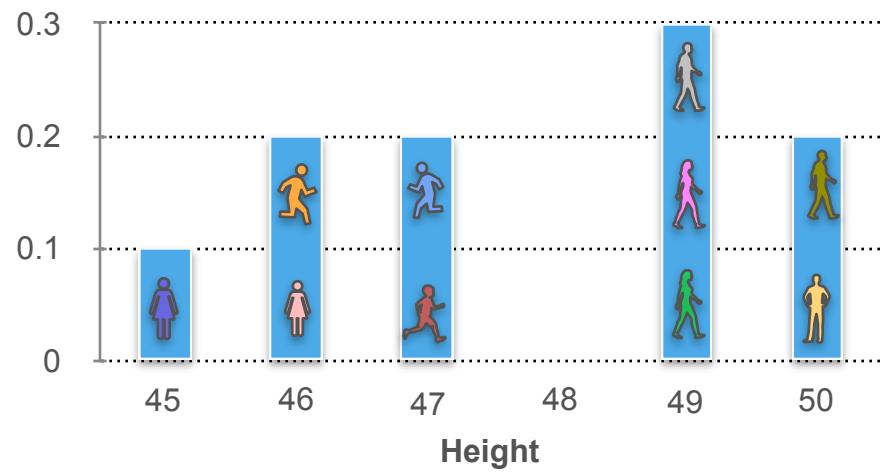
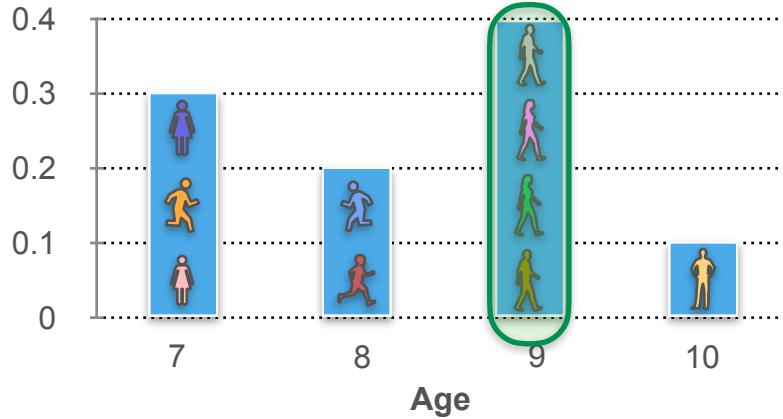
What is the probability that a child is **9 years** old and **49 inches** tall?

Joint Distributions (Discrete): Example 1



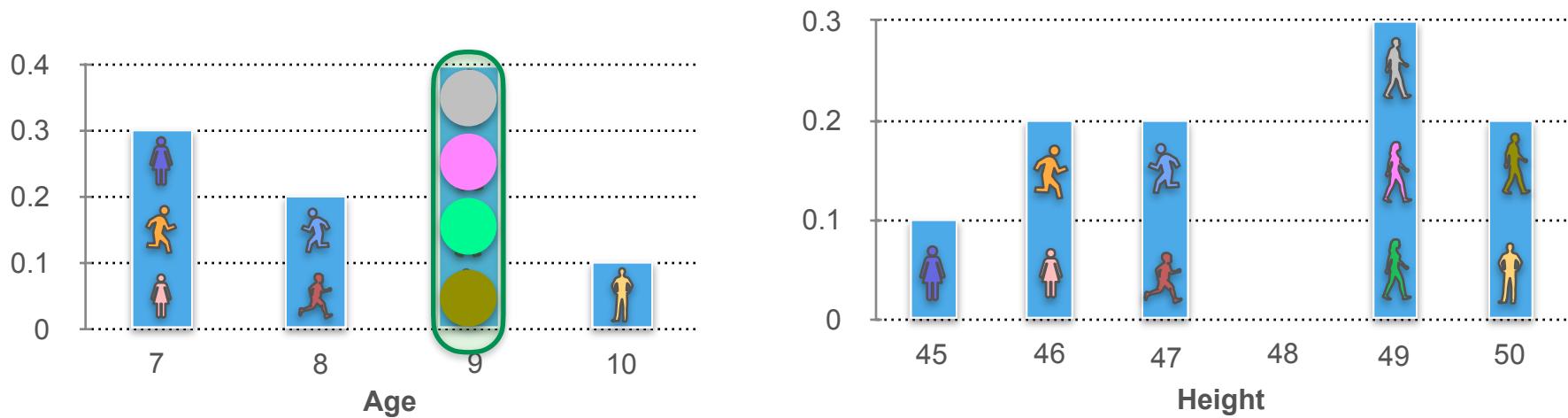
What is the probability that a child is **9 years** old and **49 inches** tall?

Joint Distributions (Discrete): Example 1



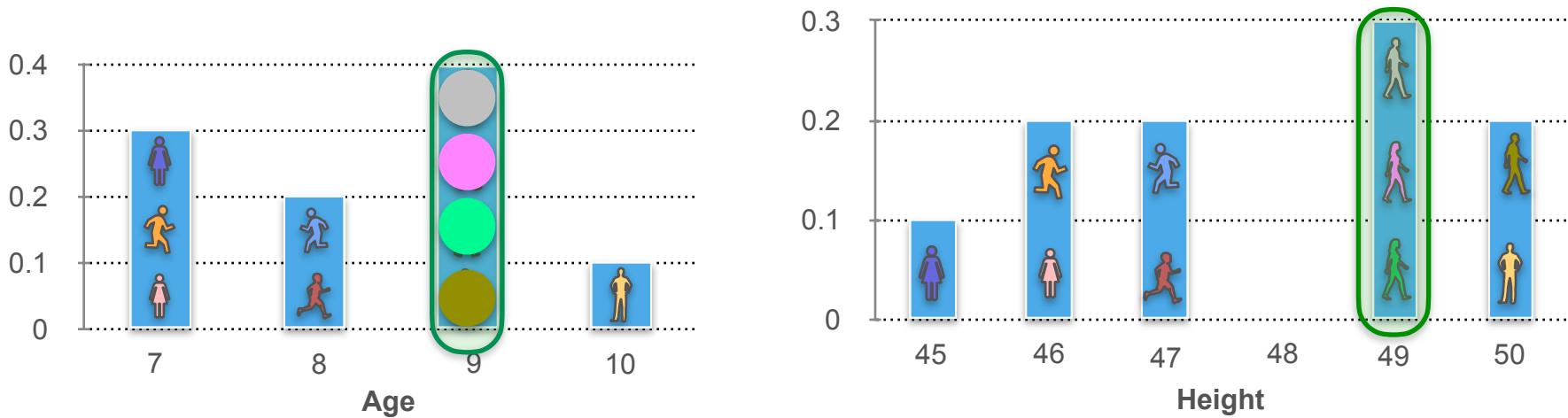
What is the probability that a child is **9 years** old and **49 inches** tall?

Joint Distributions (Discrete): Example 1



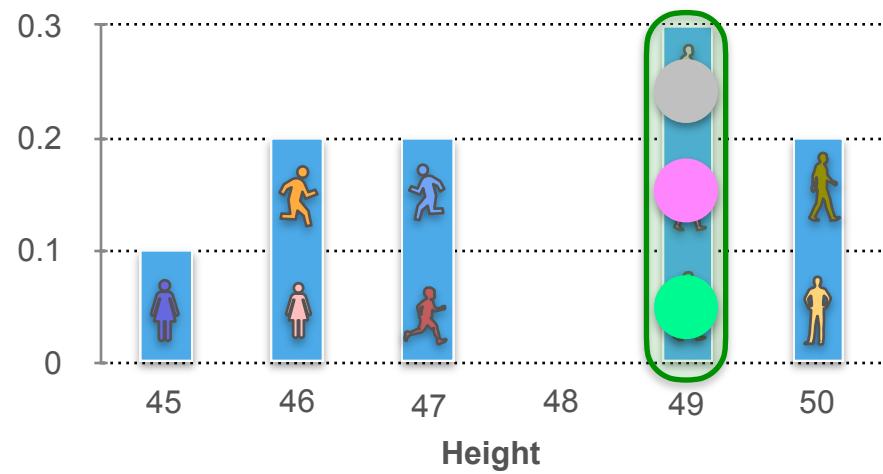
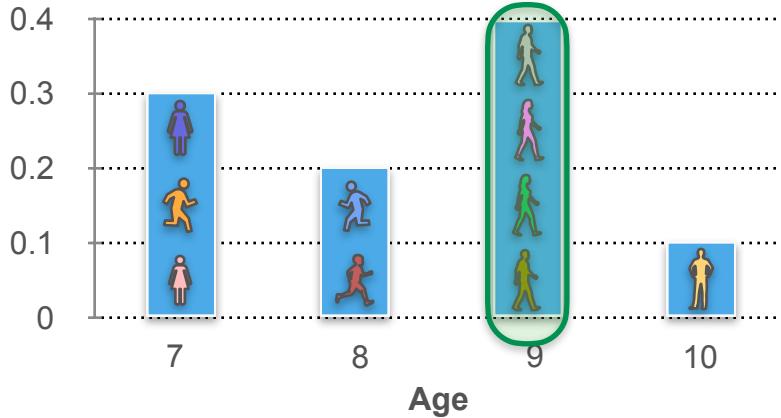
What is the probability that a child is **9 years** old and **49 inches** tall?

Joint Distributions (Discrete): Example 1



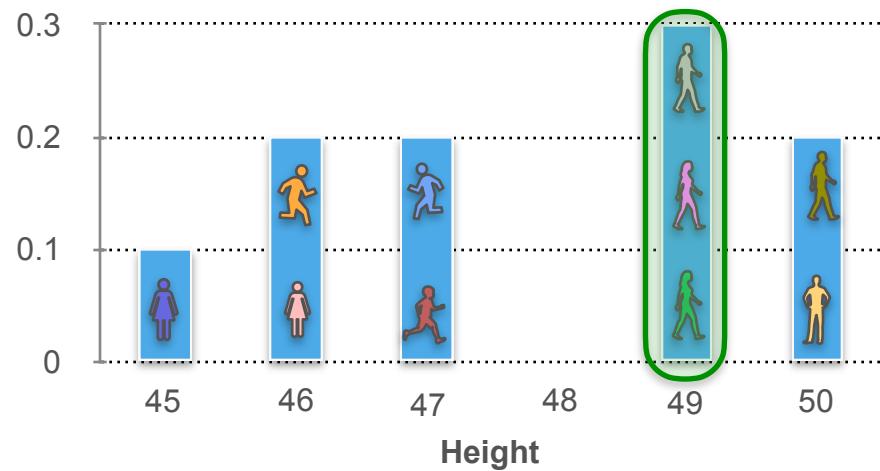
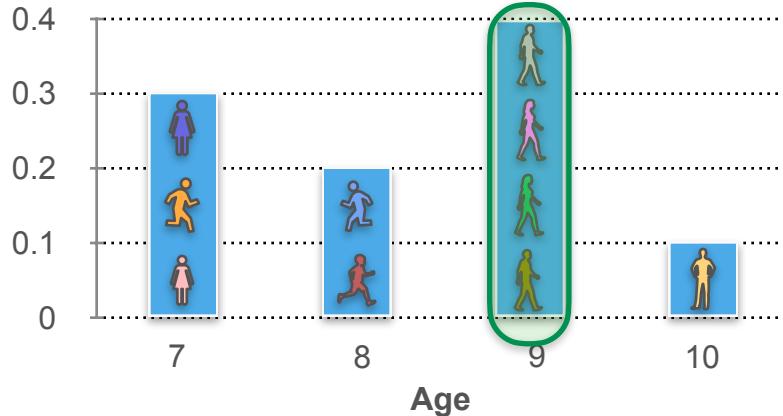
What is the probability that a child is **9 years** old and **49 inches** tall?

Joint Distributions (Discrete): Example 1



What is the probability that a child is **9 years** old and **49 inches** tall?

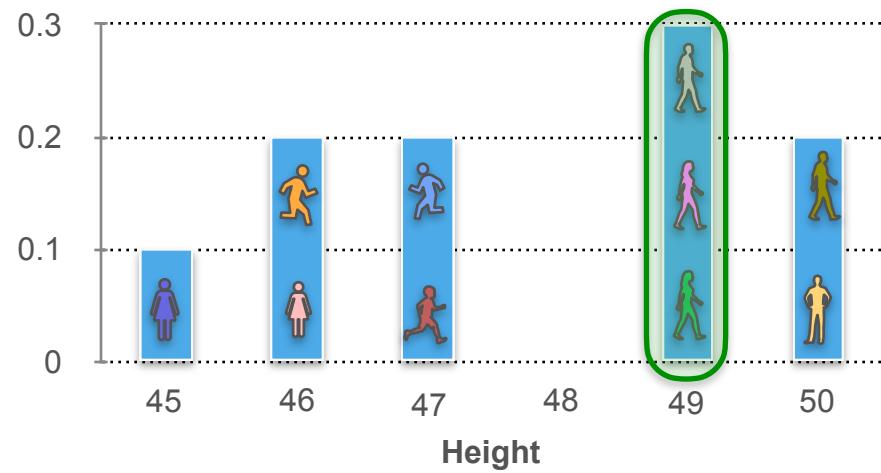
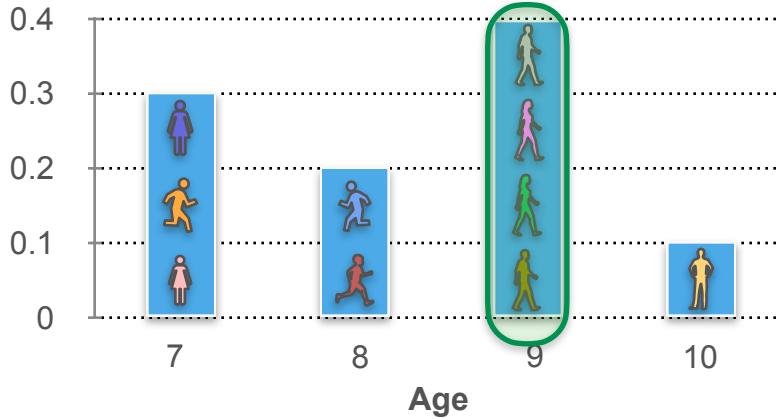
Joint Distributions (Discrete): Example 1



What is the probability that a child is **9 years** old and **49 inches** tall?



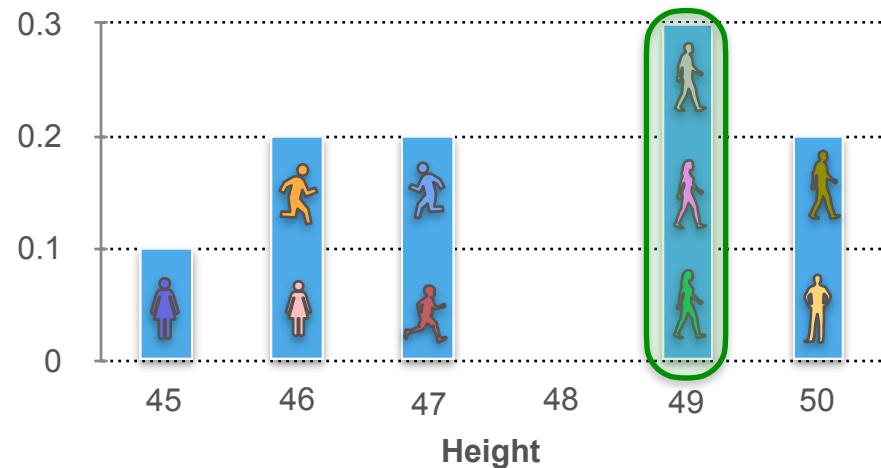
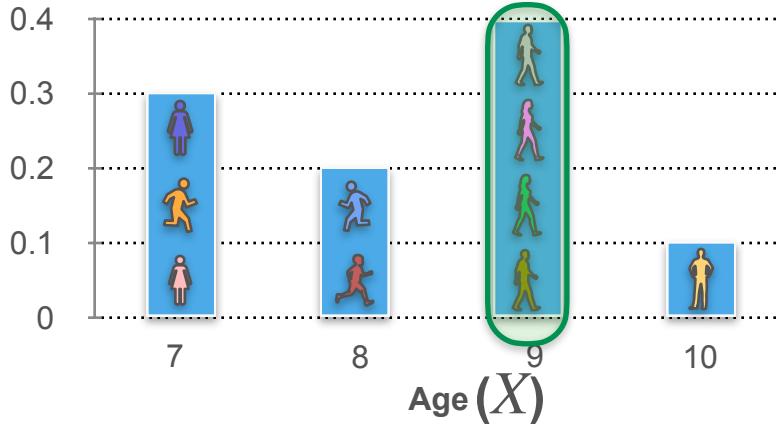
Joint Distributions (Discrete): Example 1



What is the probability that a child is **9 years** old and **49 inches** tall?

$$\frac{\text{Grey circle} \times \text{Pink circle} \times \text{Green circle}}{10} = \frac{3}{10}$$

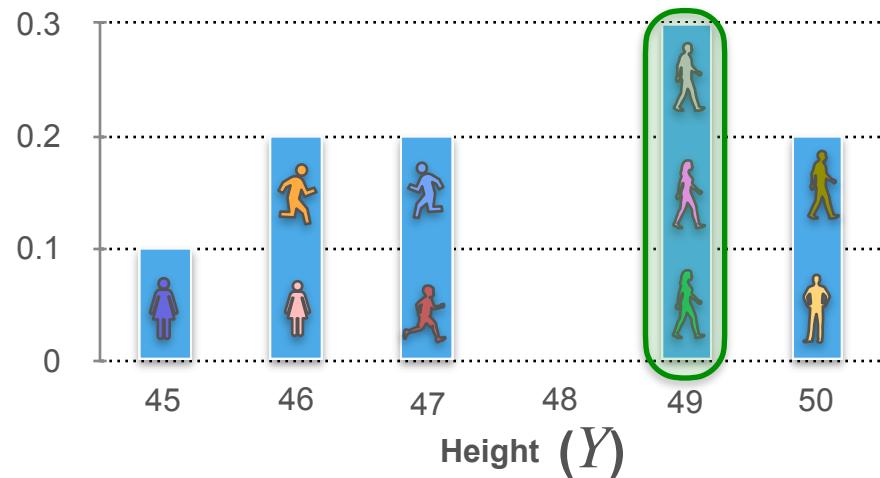
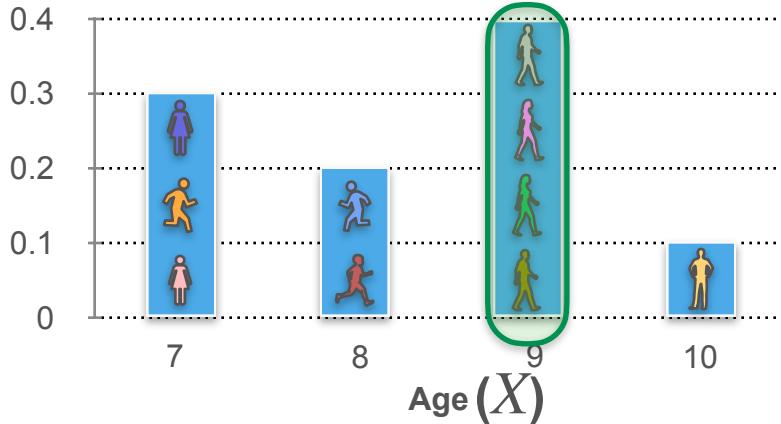
Joint Distributions (Discrete): Example 1



What is the probability that a child is **9 years** old and **49 inches** tall?

$$\frac{\text{Grey circle} \times \text{Pink circle} \times \text{Green circle}}{10} = \frac{3}{10}$$

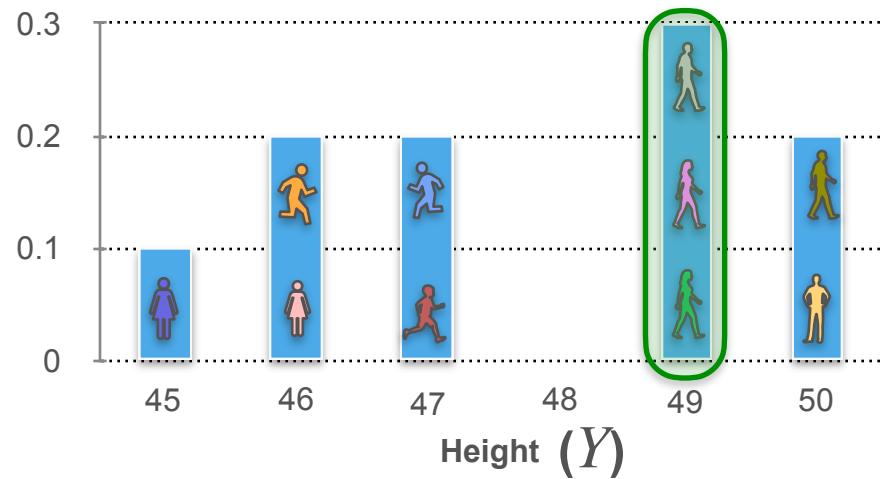
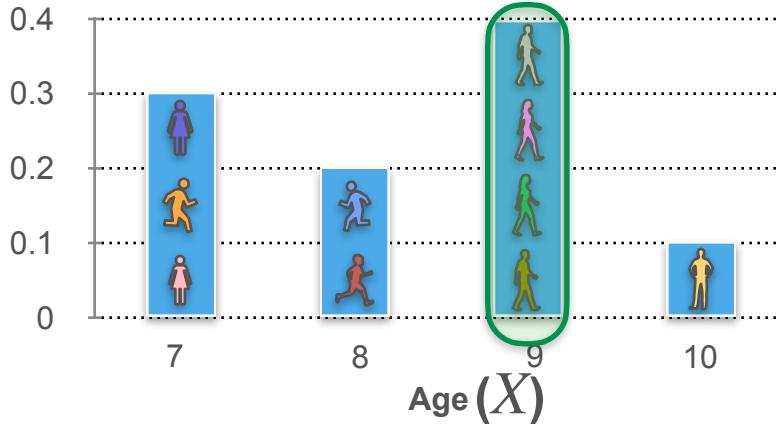
Joint Distributions (Discrete): Example 1



What is the probability that a child is **9 years** old and **49 inches** tall?

$$\frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \frac{3}{10}$$

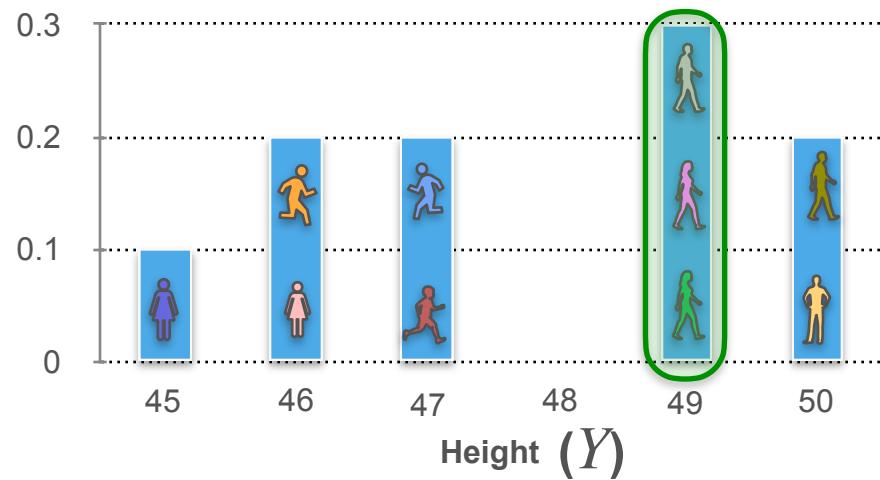
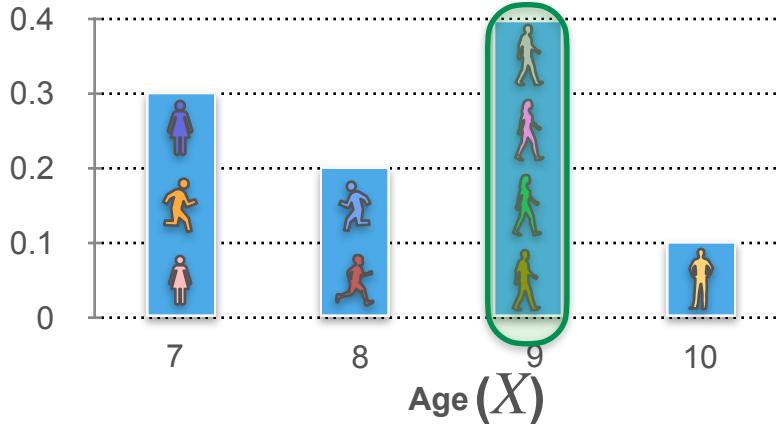
Joint Distributions (Discrete): Example 1



What is the probability that a child is 9 years old and 49 inches tall?

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

Joint Distributions (Discrete): Example 1

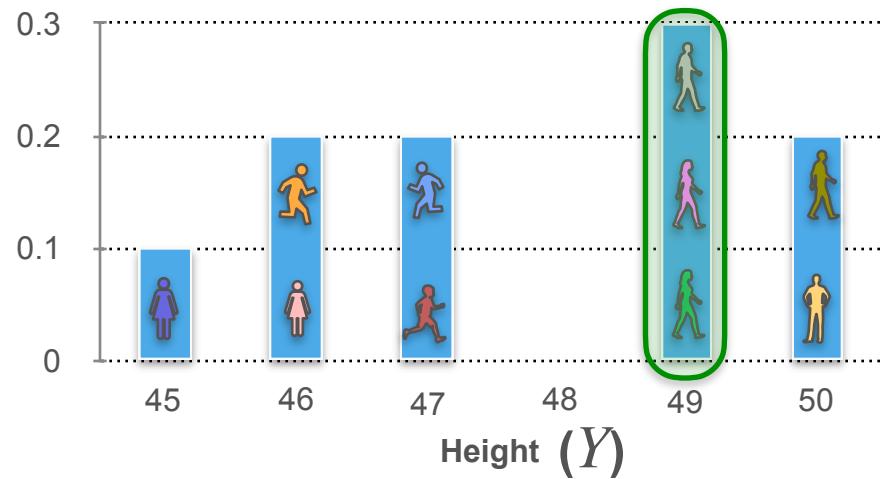
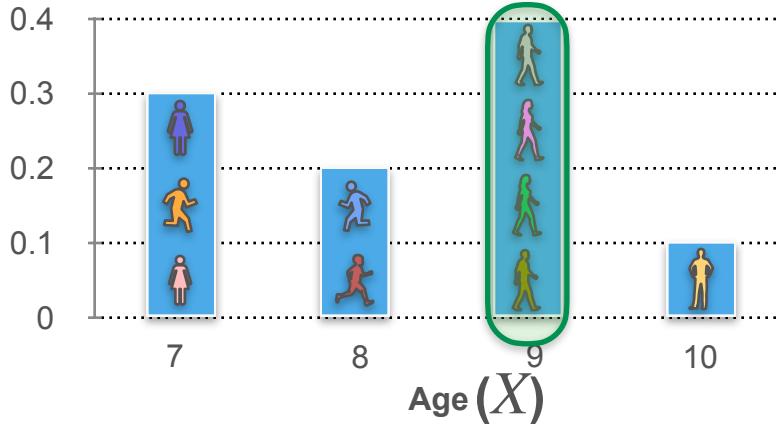


What is the probability that a child is 9 years old and 49 inches tall?

$$p_{XY}(9, 49)$$

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

Joint Distributions (Discrete): Example 1



What is the probability that a child is 9 years old and 49 inches tall?

$$p_{XY}(9, 49) = \mathbf{P}(X = 9, Y = 49) = \frac{\text{Grey circle}}{10} = \frac{3}{10}$$

Joint Distributions (Discrete): Example 1

What is the probability that a child is **9 years** old and **49 inches** tall?

$$p_{XY}(9, 49) = \mathbf{P}(X = 9, Y = 49) = \frac{\text{_____}}{10} = \frac{3}{10}$$

Joint Distributions (Discrete): Example 1

What is the probability that a child is **9 years** old and **49 inches** tall?

$$p_{XY}(9, 49) = \mathbf{P}(X = 9, Y = 49) = \frac{\text{_____}}{10} = \frac{3}{10}$$

$$p_{XY}(x, y)$$

Joint Distributions (Discrete): Example 1

What is the probability that a child is **9 years** old and **49 inches** tall?

$$p_{XY}(9, 49) = \mathbf{P}(X = 9, Y = 49) = \frac{\text{_____}}{10} = \frac{3}{10}$$

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

Joint Distributions: Example 1

										
Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

		Height						
		45	46	47	48	49	50	
Age	7							
	8							
	9							
	10							

Joint Distributions: Example 1

												
Age (years):	7	7	7	8	8	9	9	9	9	10		
Height (in):	45	46	46	47	47	49	49	49	50	50		

		Height						
		45	46	47	48	49	50	
Age	7	1						
	8							
	9							
	10							

Joint Distributions: Example 1

										
Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

		Height					
		45	46	47	48	49	50
Age	7	1	2				
	8						
	9						
	10						

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8						
9						
10						

Joint Distributions: Example 1



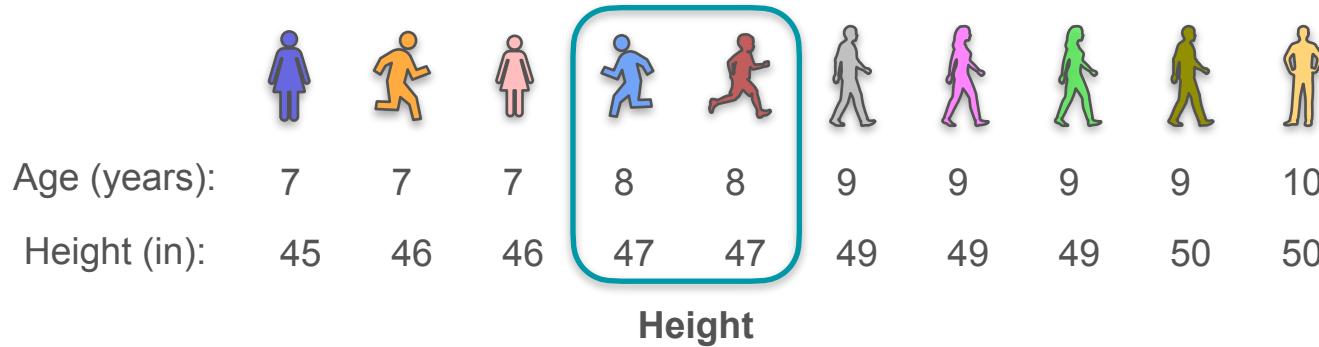
Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0				
9						
10						

Joint Distributions: Example 1



	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2			
9						
10						

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2	0	0	0
9						
10						

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2	0	0	0
9	0	0	0	0		
10						

Joint Distributions: Example 1

Age (years):	7	7	7	8	8	9	9	9	9
Height (in):	45	46	46	47	47	49	49	49	50

		Height					
		45	46	47	48	49	50
Age	7	1	2	0	0	0	0
	8	0	0	2	0	0	0
	9	0	0	0	0	3	
	10						

Joint Distributions: Example 1

Age (years):	7	7	7	8	8	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50

		Height						
		45	46	47	48	49	50	
Age	7	1	2	0	0	0	0	
	8	0	0	2	0	0	0	
	9	0	0	0	0	3	1	
	10							

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2	0	0	0
9	0	0	0	0	3	1
10	0	0	0	0	0	

Joint Distributions: Example 1

										
Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

		Height						
		45	46	47	48	49	50	
Age	7	1	2	0	0	0	0	
	8	0	0	2	0	0	0	
	9	0	0	0	0	3	1	
	10	0	0	0	0	0	1	

Joint Distributions: Example 1



Age (years): 7 7 7 8 8 9 9 9 9 10

Height (in): 45 46 46 47 47 49 49 49 50 50

Height

	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2	0	0	0
9	0	0	0	0	3	1
10	0	0	0	0	0	1

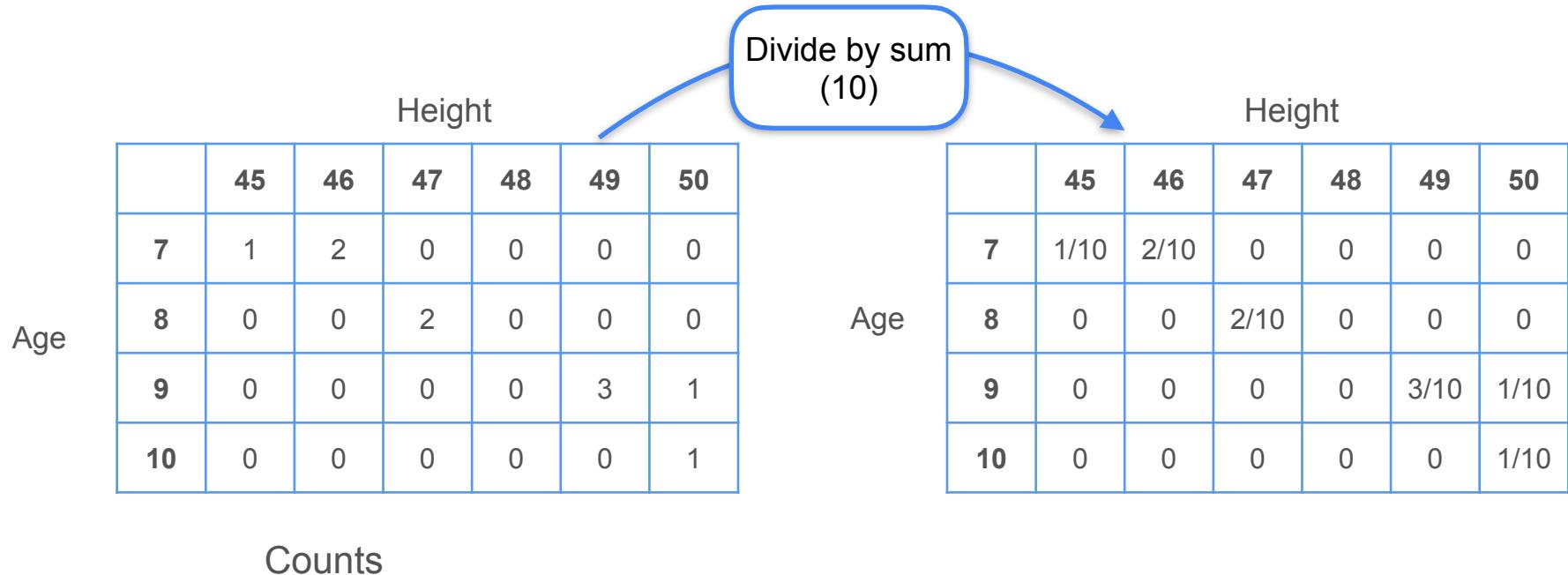
Joint Distributions: Example 1

Height

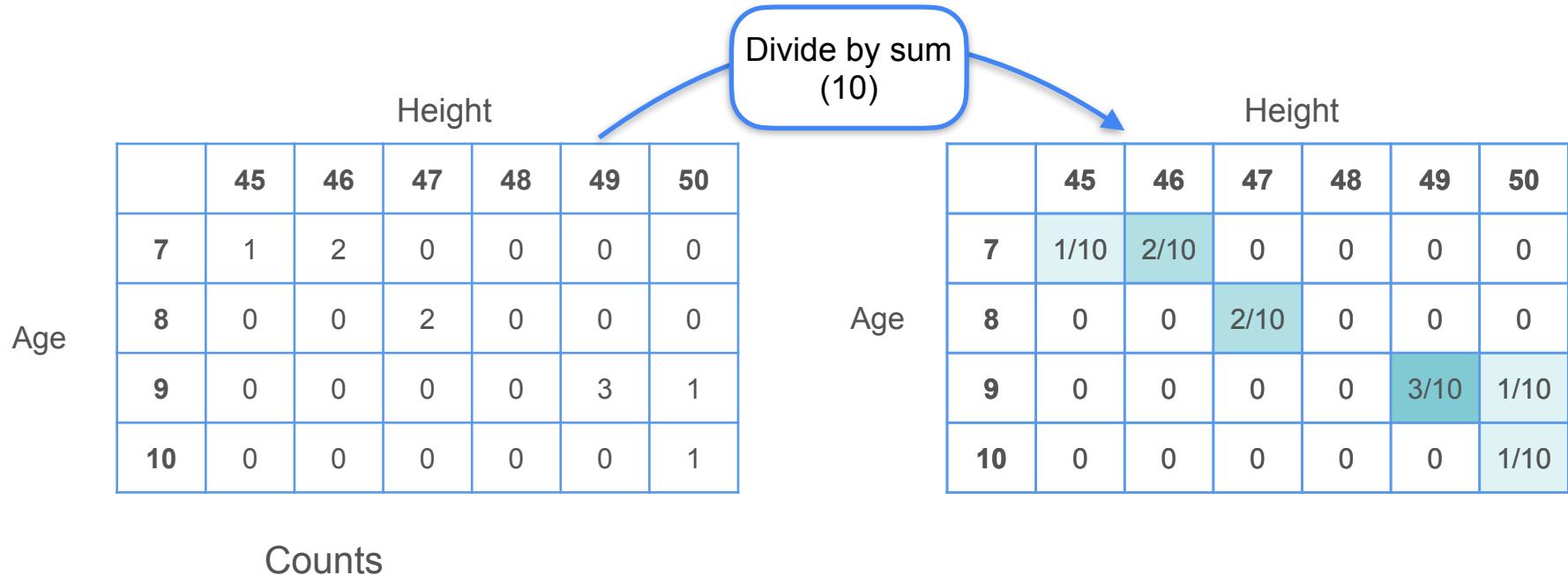
	45	46	47	48	49	50
7	1	2	0	0	0	0
8	0	0	2	0	0	0
9	0	0	0	0	3	1
10	0	0	0	0	0	1

Counts

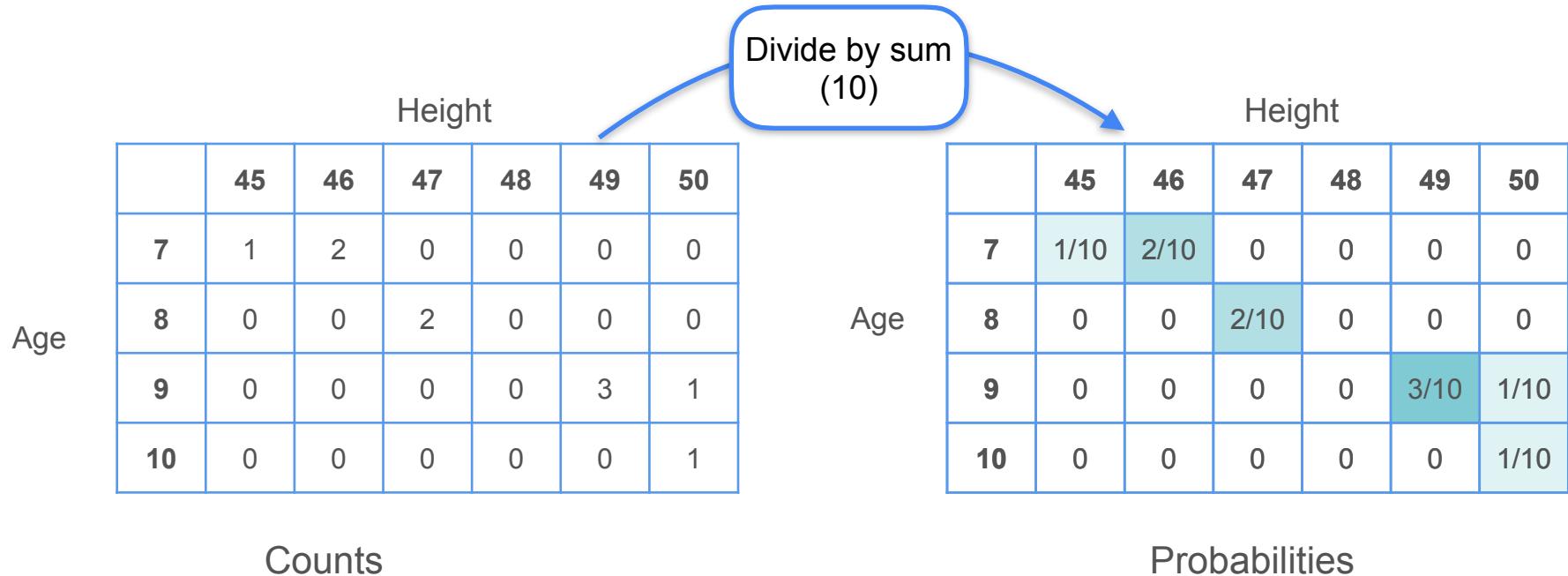
Joint Distributions: Example 1



Joint Distributions: Example 1



Joint Distributions: Example 1



Joint Distributions: Example 1

		Height						
		45	46	47	48	49	50	
Age	7	1/10	2/10	0	0	0	0	
	8	0	0	2/10	0	0	0	
	9	0	0	0	0	3/10	1/10	
	10	0	0	0	0	0	1/10	

Probabilities

Joint Distributions: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10

Probabilities

Joint Distributions: Example 1

All probabilities for all possible combinations of X and Y

Age
(X)

Height (Y)

	45	46	47	48	49	50
7	1/10	2/10	0	0	0	0
8	0	0	2/10	0	0	0
9	0	0	0	0	3/10	1/10
10	0	0	0	0	0	1/10

Probabilities

Joint Distributions: Example 1

Joint Distribution

All probabilities for all possible combinations of X and Y

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		Probabilities					

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10

Probabilities

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10

Probabilities

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
	Probabilities						

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

	Height (Y)					
	45	46	47	48	49	50
7	1/10	2/10	0		0	0
8	0	0	2/10	0	0	0
9	0	0	0	0	3/10	1/10
10	0	0	0	0	0	1/10

Age (X)

Probabilities

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

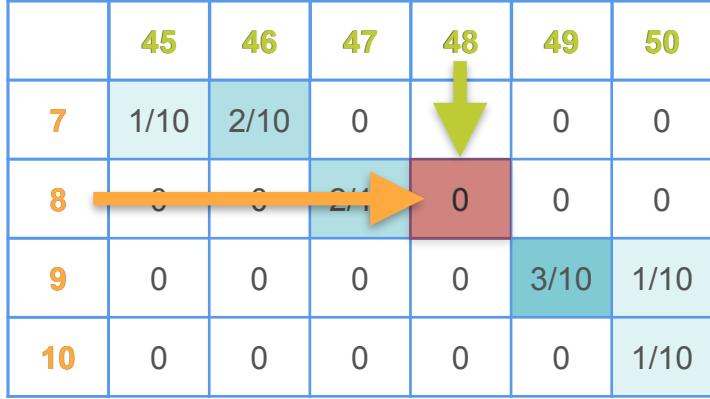
$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

$$p_{XY}(8, 48) = 0$$

	Height (Y)					
	45	46	47	48	49	50
7	1/10	2/10	0		0	0
8	0	0	2/10	0	0	0
9	0	0	0	0	3/10	1/10
10	0	0	0	0	0	1/10

Age (X)

Probabilities



Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

$$p_{XY}(8, 48) = 0$$

$$p_{XY}(7, 46) =$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
	Probabilities						

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

$$p_{XY}(8, 48) = 0$$

$$p_{XY}(7, 46) =$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
	Probabilities						

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

$$p_{XY}(8, 48) = 0$$

$$p_{XY}(7, 46) =$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/1	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
	Probabilities						

Joint Distributions: Example 1

What is the probability that a child is 8 and 48 inches tall?

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$$

$$p_{XY}(8, 48) = \mathbf{P}(X = 8, Y = 48)$$

$$p_{XY}(8, 48) = 0$$

$$p_{XY}(7, 46) = \frac{2}{10}$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/1	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
	Probabilities						

Joint Distributions (Discrete): Example 2

Joint Distributions (Discrete): Example 2

X

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice

Y

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice

Y

the number rolled on the 2nd dice

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice



Y

the number rolled on the 2nd dice

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice



$$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$$

Y

the number rolled on the 2nd dice

Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice



$$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$$

Y

the number rolled on the 2nd dice



Joint Distributions (Discrete): Example 2

X

the number rolled on the 1st dice



$$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$$

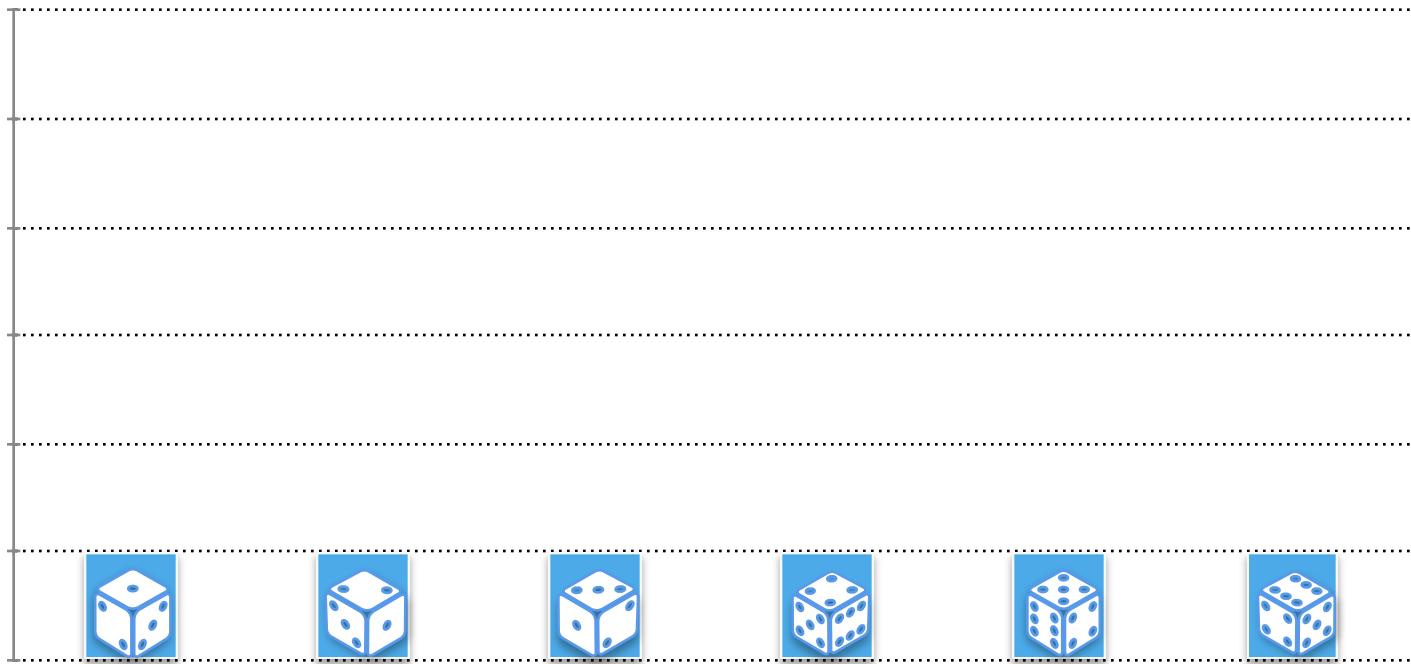
Y

the number rolled on the 2nd dice

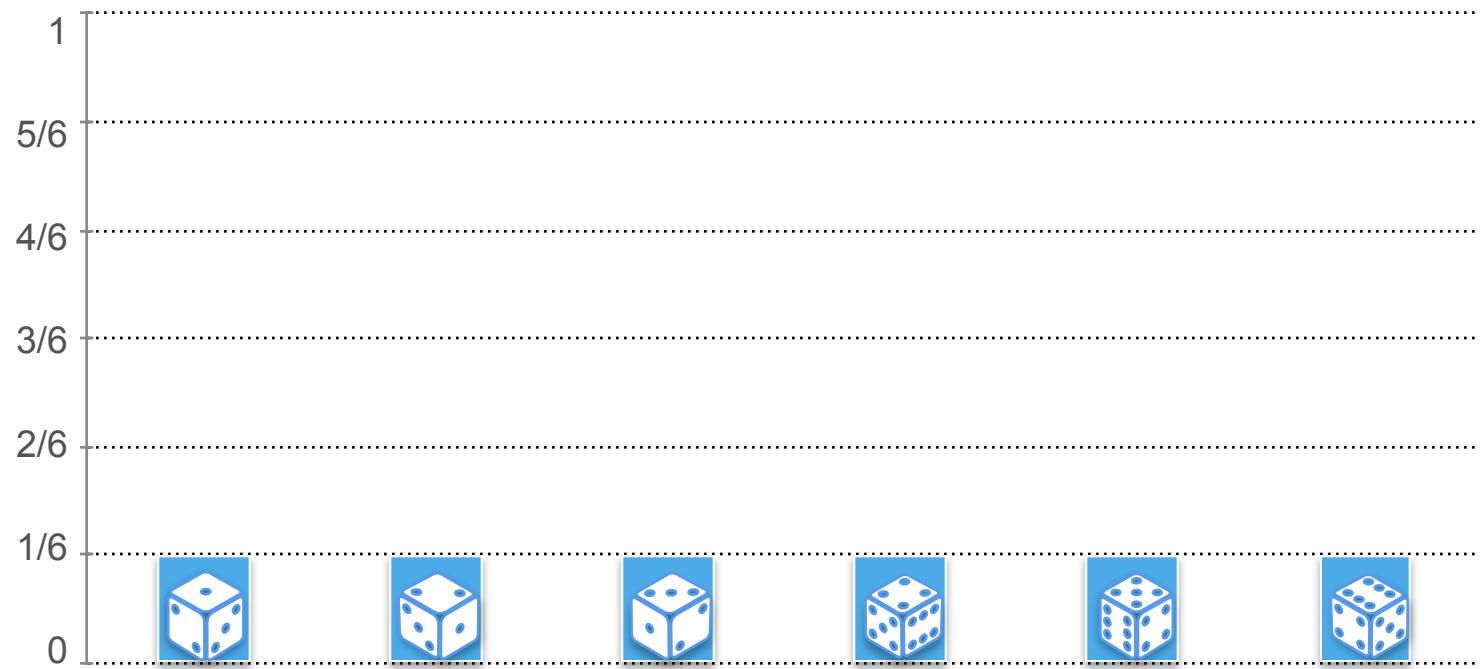


$$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$$

Joint Distributions: Example 2



Joint Distributions: Example 2



Joint Distributions: Example 2

Joint Distributions: Example 2

X

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

						
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6
$P(y)$	1/6	1/6	1/6	1/6	1/6	1/6

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice



$$\mathbf{P}(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$\mathbf{P}(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

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

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent



$P(x)$ 1/6 1/6 1/6 1/6 1/6 1/6

$P(y)$ 1/6 1/6 1/6 1/6 1/6 1/6

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

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent



$$\mathbf{P}(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$\mathbf{P}(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice 1}, \text{dice 2})$$

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

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

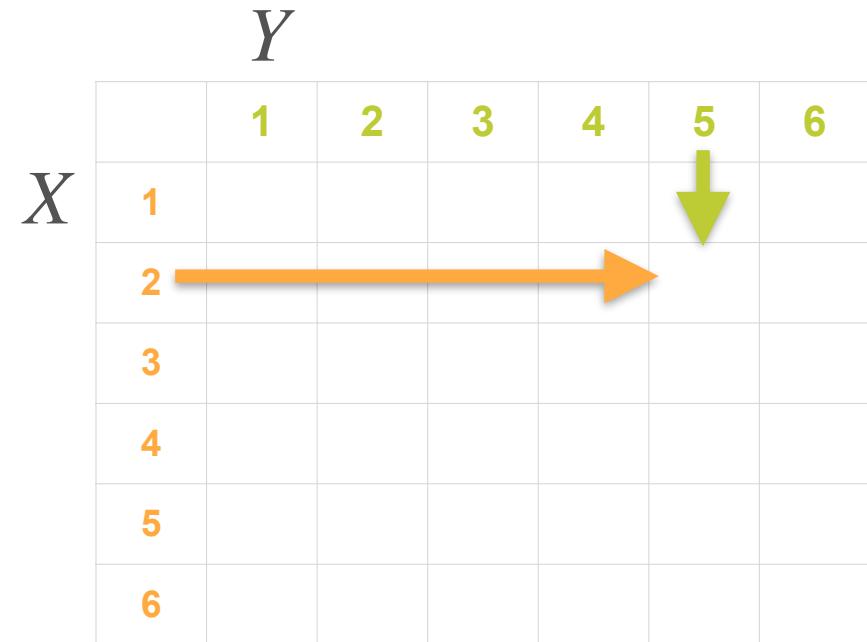
X and Y are independent



$$\mathbf{P}(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$\mathbf{P}(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{orange die}, \text{yellow die})$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

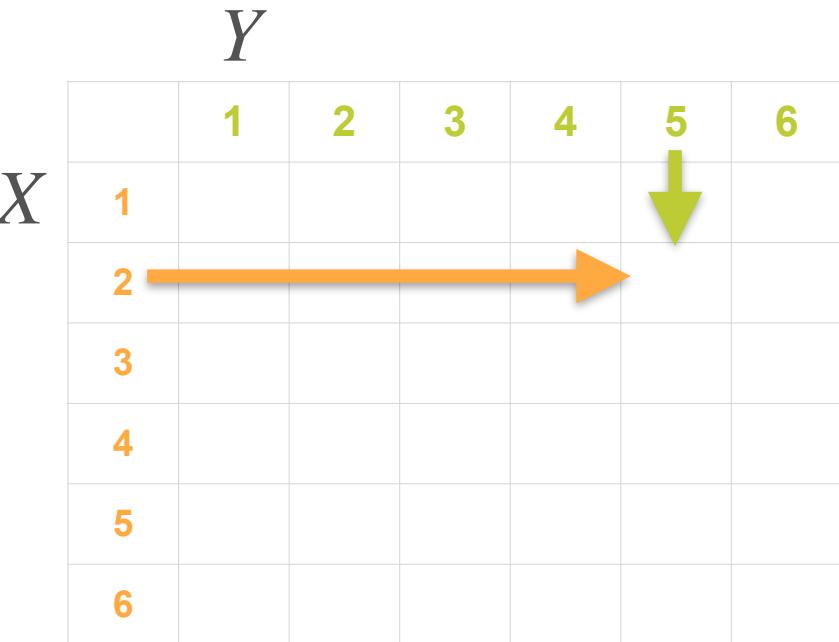
X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2)$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

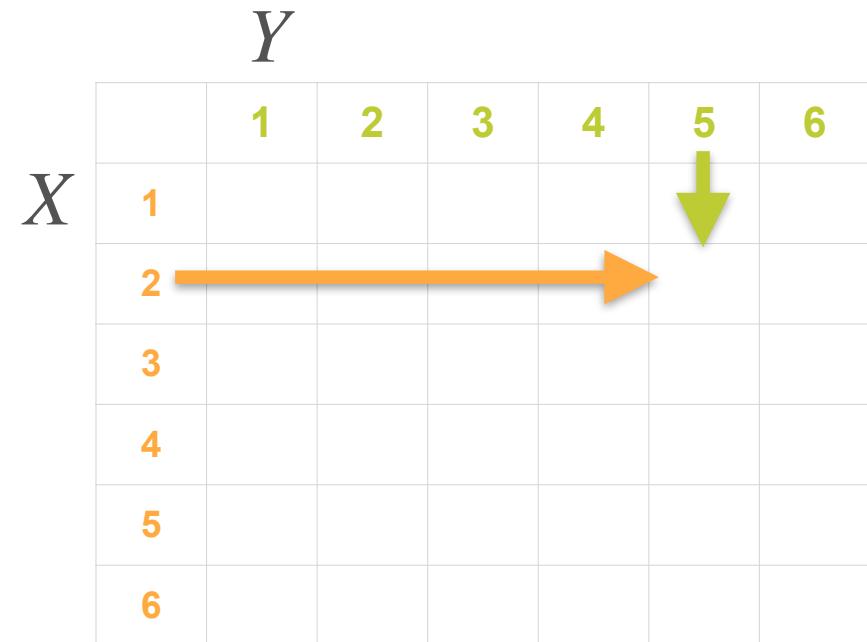
X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6}$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

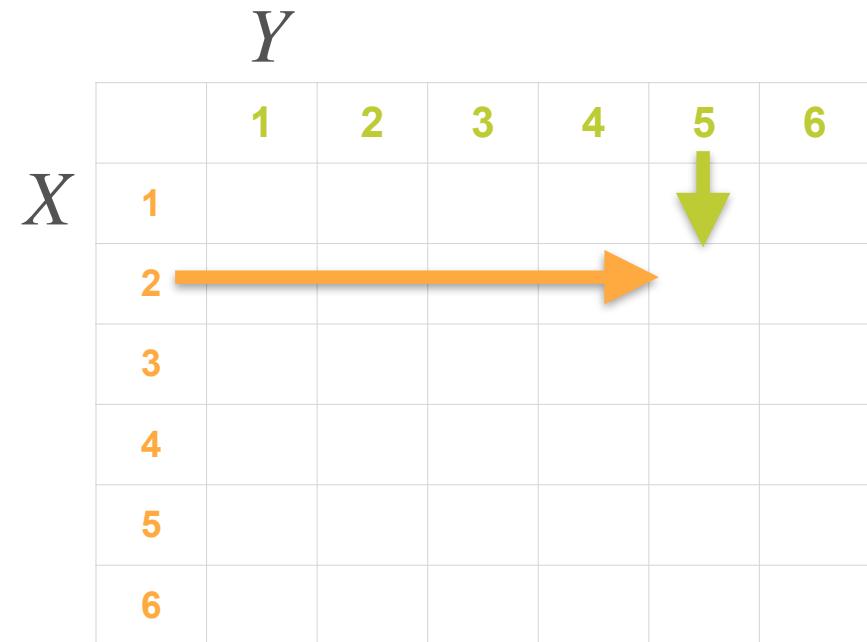
X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

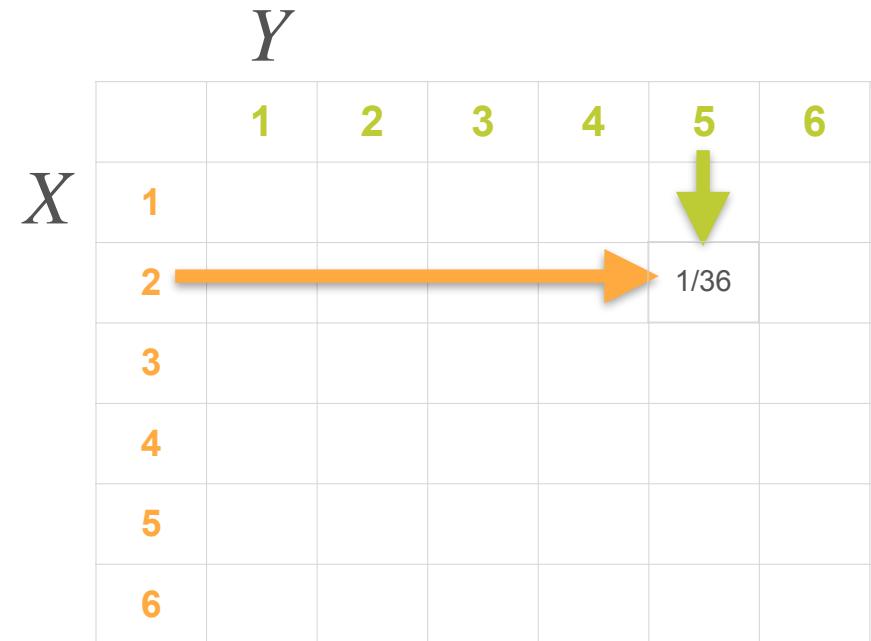
X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent

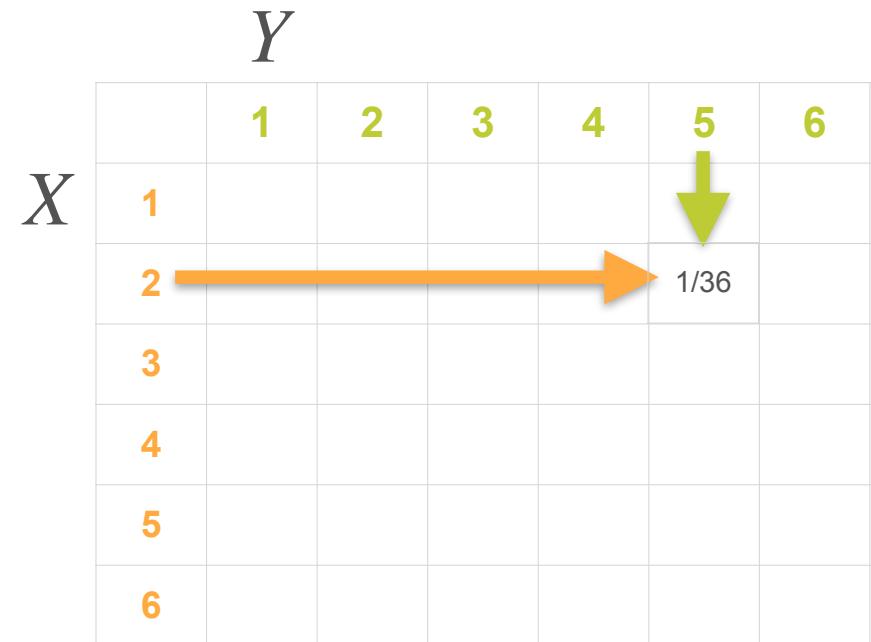


$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

$$p_{XY}(x, y) = P(X = x, Y = y)$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent

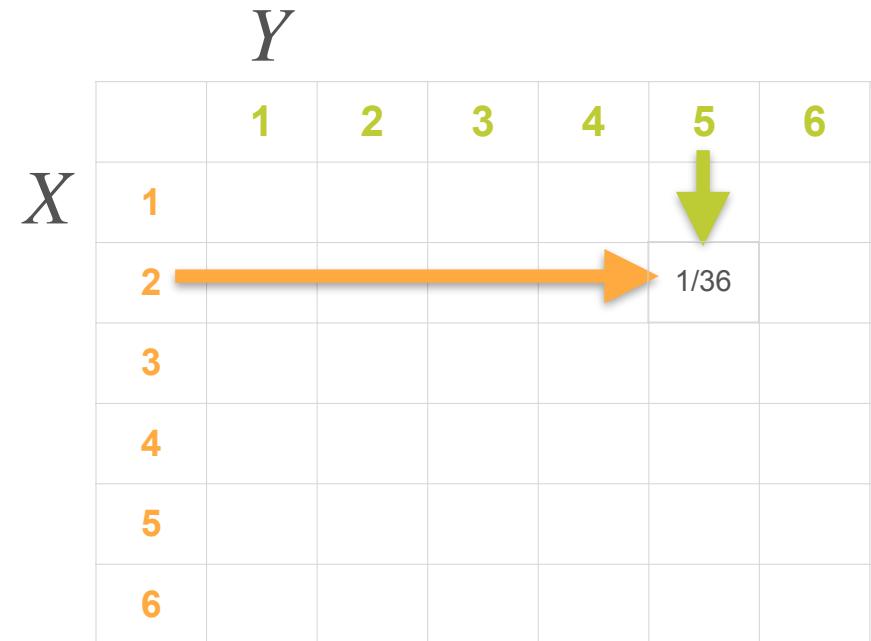


$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

$$p_{XY}(x, y) = P(X = x, Y = y) = \frac{1}{36}$$



Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } 1, \text{dice } 2) = P(\text{dice } 1) \cdot P(\text{dice } 2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

$$p_{XY}(x, y) = P(X = x, Y = y) = \frac{1}{36}$$

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

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X and Y are independent



$$P(x) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$P(y) \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6$$

$$p_{XY}(\text{dice } x, \text{dice } y) = P(\text{dice } x) \cdot P(\text{dice } y) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

$$p_{XY}(x, y) = P(X = x, Y = y) = \frac{1}{36}$$

	Y	1	2	3	4	5	6
X	1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36	1/36

Joint Distributions: Example 2

Joint Distributions: Example 2

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y) = \mathbf{P}(x) \cdot \mathbf{P}(y)$$

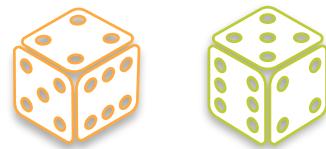
Joint Distributions: Example 2

Thus for independent discrete random variables:

$$p_{XY}(x, y) = \mathbf{P}(X = x, Y = y) = \mathbf{P}(x) \cdot \mathbf{P}(y)$$

Joint Distributions (Discrete): Example 2

Joint Distributions (Discrete): Example 2



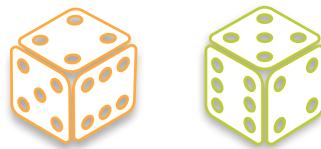
Joint Distributions (Discrete): Example 2

X



Joint Distributions (Discrete): Example 2

X



the number rolled on the 1st dice

Joint Distributions (Discrete): Example 2

X

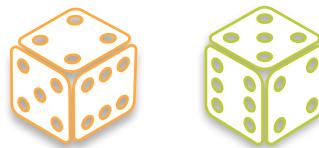


the number rolled on the 1st dice



Joint Distributions (Discrete): Example 2

X



the number rolled on the 1st dice



$$X = 4$$

Joint Distributions (Discrete): Example 2

X



Y

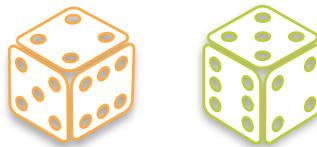
the number rolled on the 1st dice



$$X = 4$$

Joint Distributions (Discrete): Example 2

X



Y

the number rolled on the 1st dice

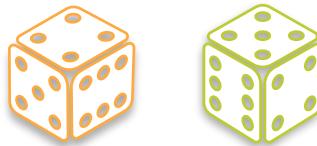
sum of the two dice



$$X = 4$$

Joint Distributions (Discrete): Example 2

X



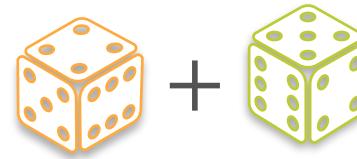
Y

the number rolled on the 1st dice

sum of the two dice



$$X = 4$$



$$Y = 4 + 5$$

Joint Distributions: Example 2

Joint Distributions: Example 2

X

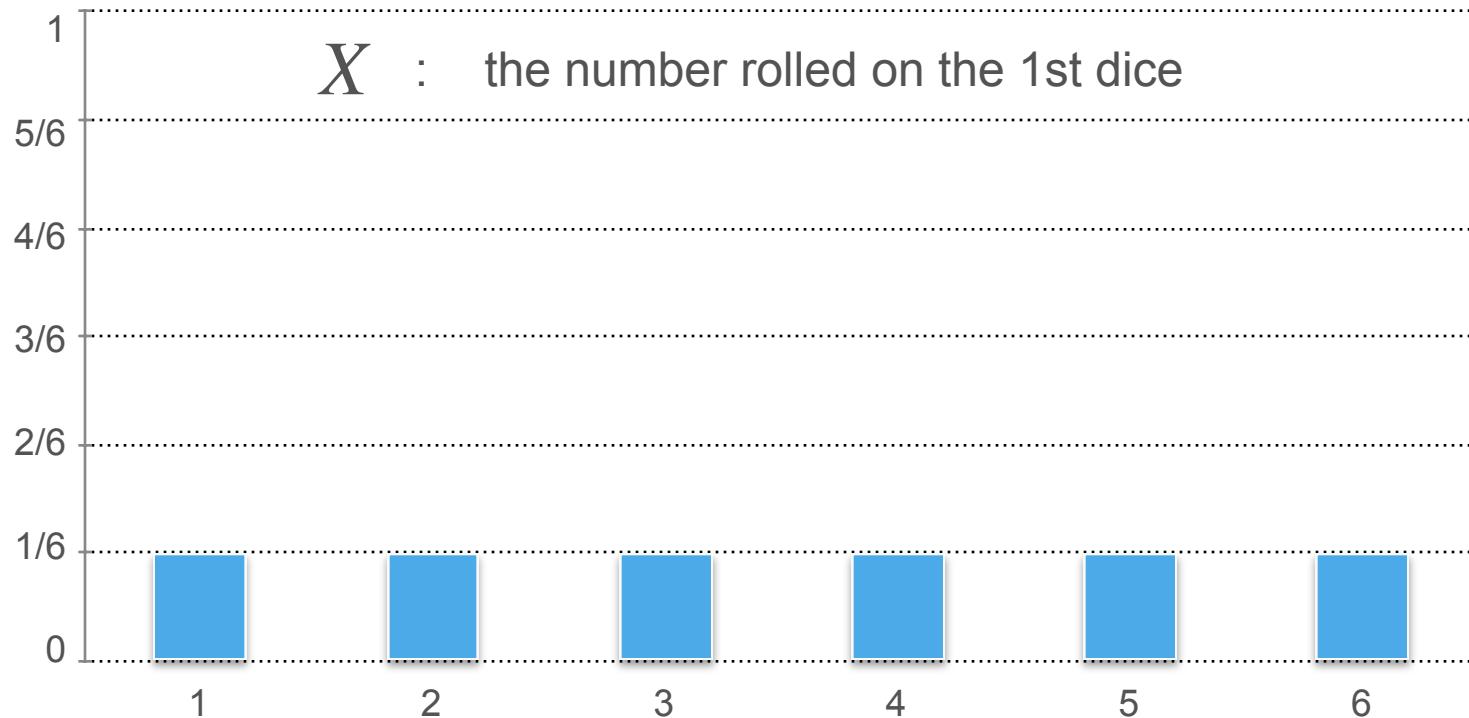
Joint Distributions: Example 2

X the number rolled on the 1st dice

Joint Distributions: Example 2

X : the number rolled on the 1st dice

Joint Distributions: Example 2

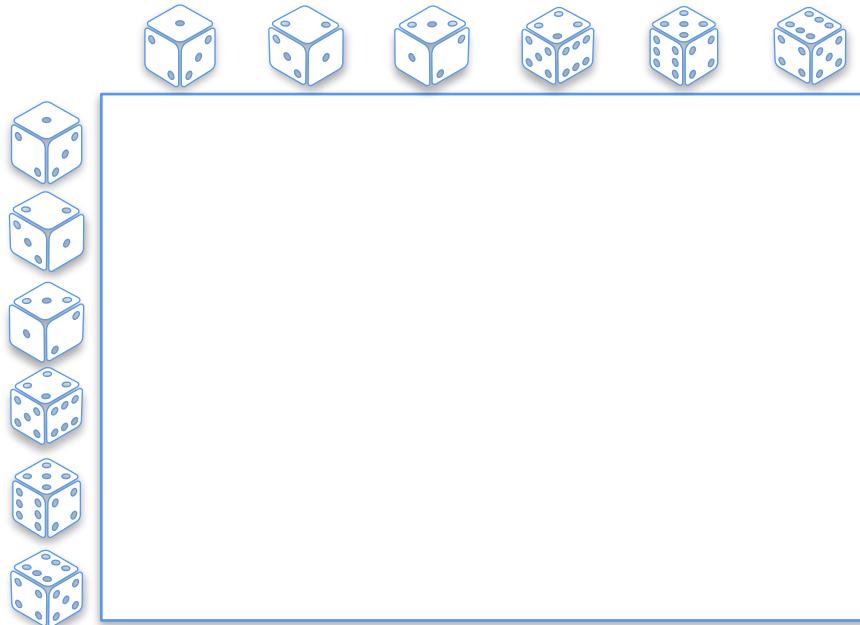


Joint Distributions - Example 3

Y : Sum of both dice

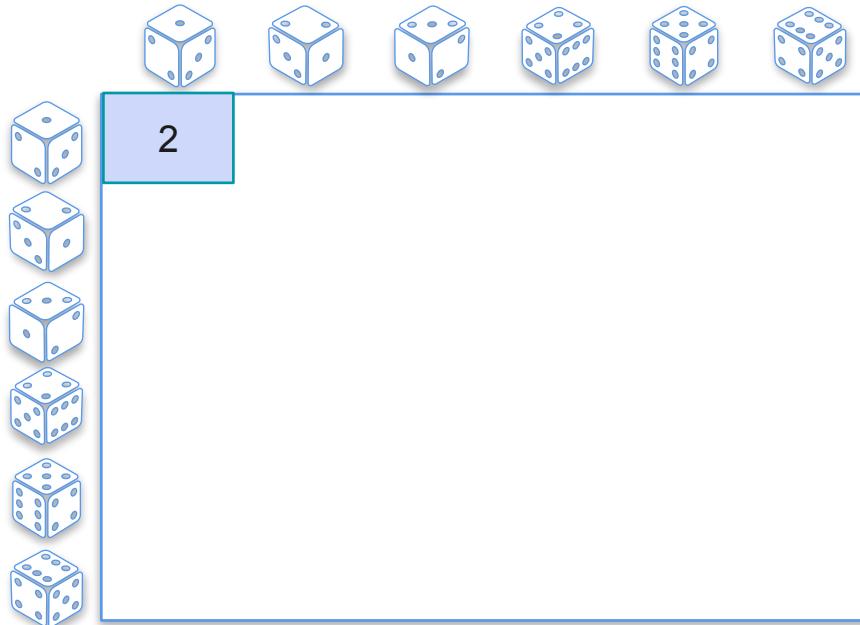
Joint Distributions - Example 3

Y : Sum of both dice



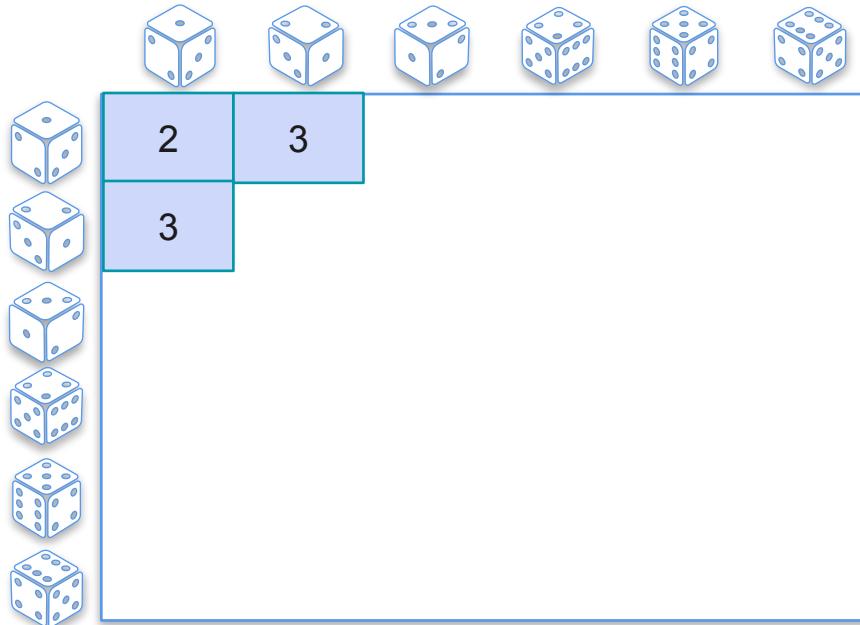
Joint Distributions - Example 3

Y : Sum of both dice



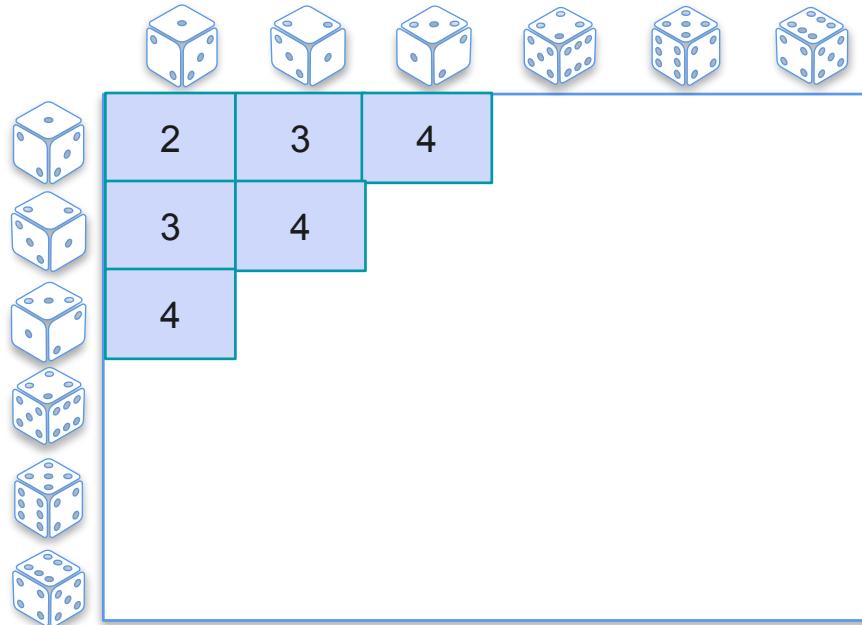
Joint Distributions - Example 3

Y : Sum of both dice



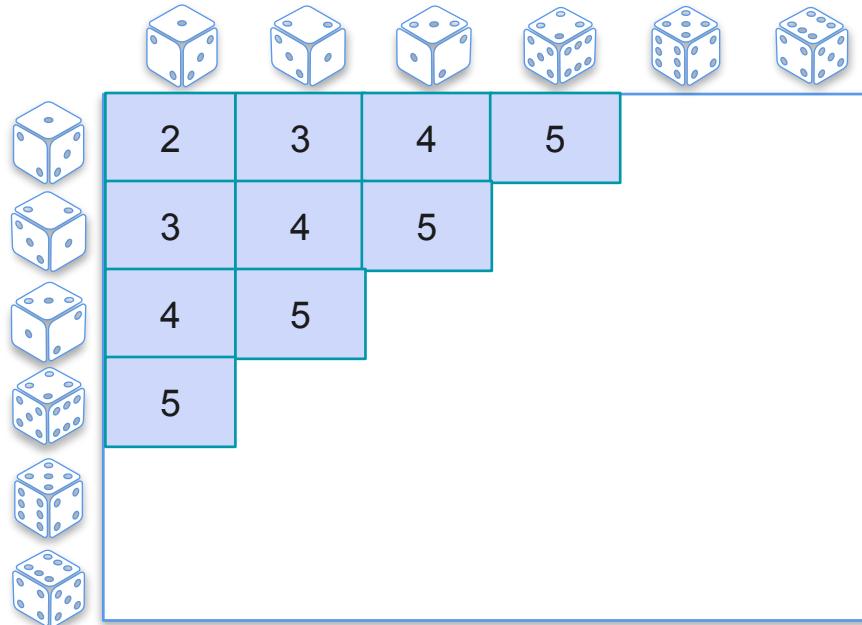
Joint Distributions - Example 3

Y : Sum of both dice



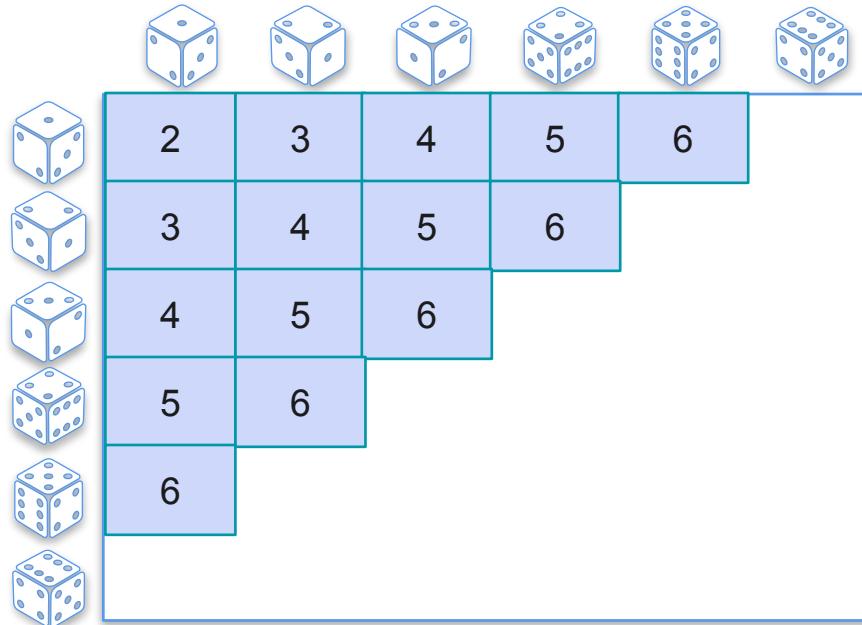
Joint Distributions - Example 3

Y : Sum of both dice



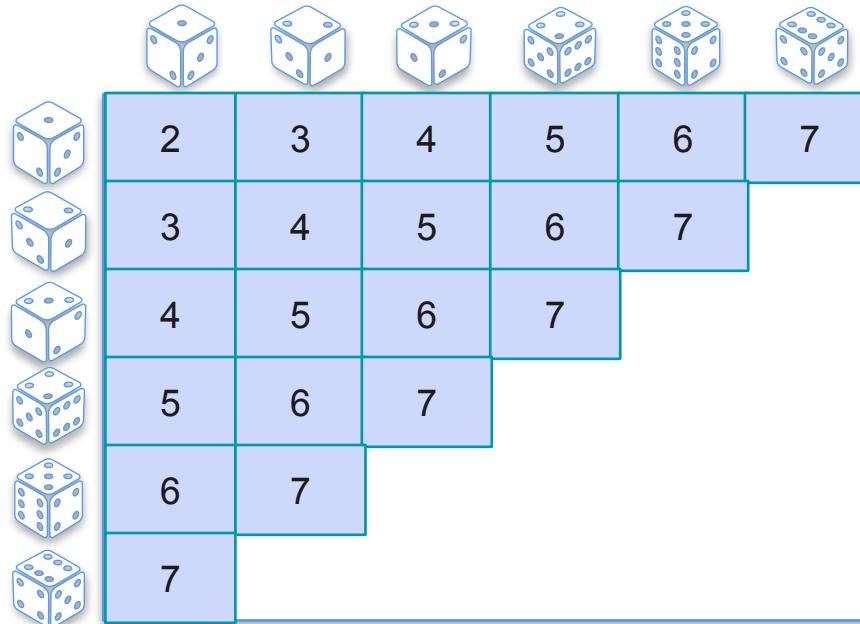
Joint Distributions - Example 3

Y : Sum of both dice



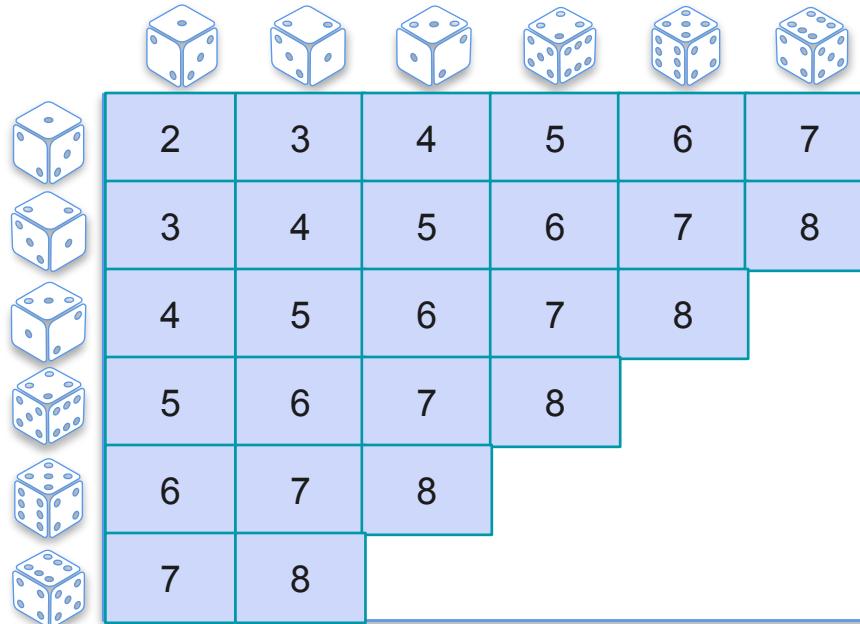
Joint Distributions - Example 3

Y : Sum of both dice



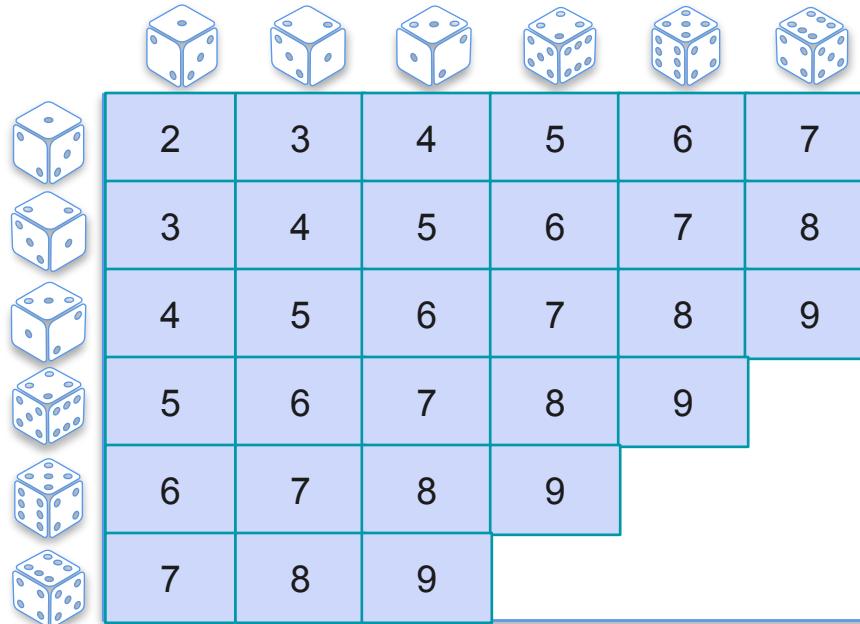
Joint Distributions - Example 3

Y : Sum of both dice



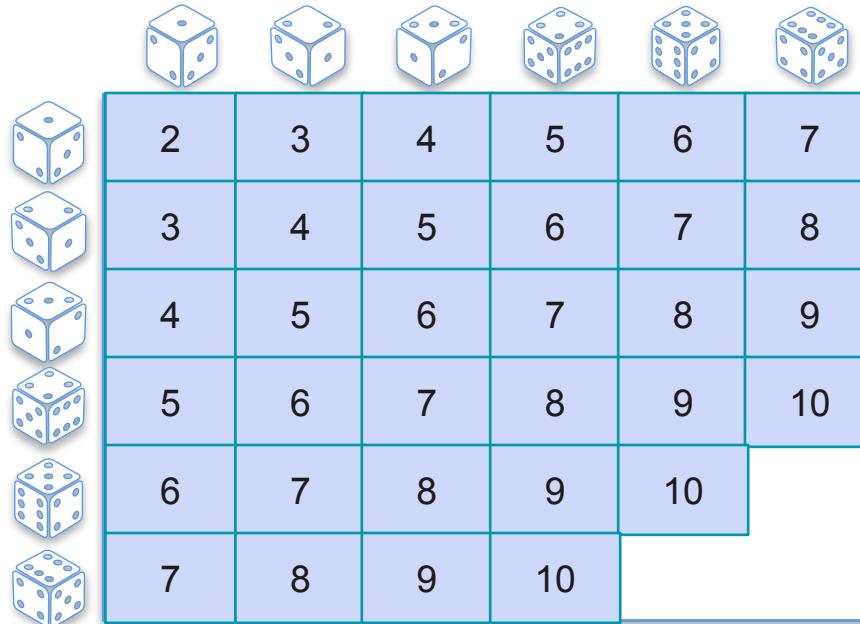
Joint Distributions - Example 3

Y : Sum of both dice



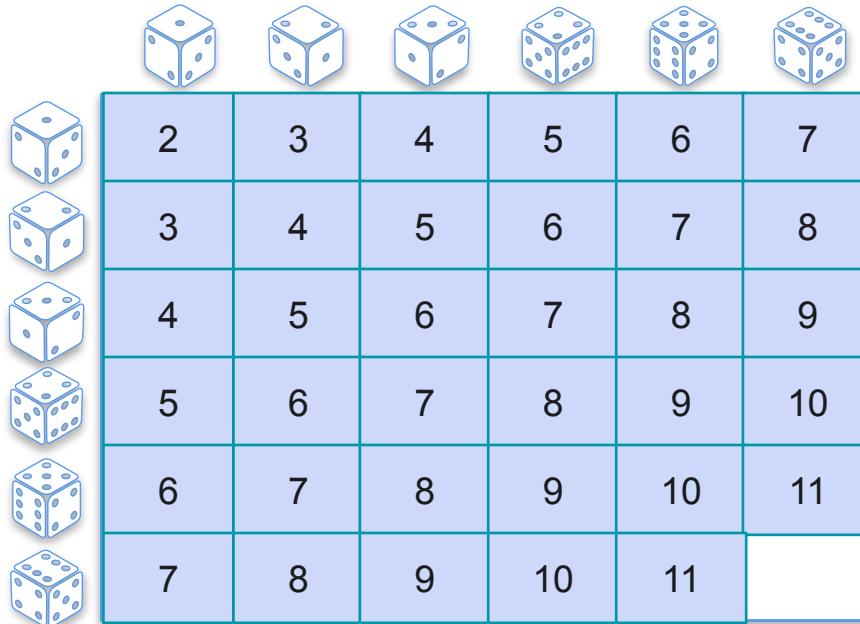
Joint Distributions - Example 3

Y : Sum of both dice



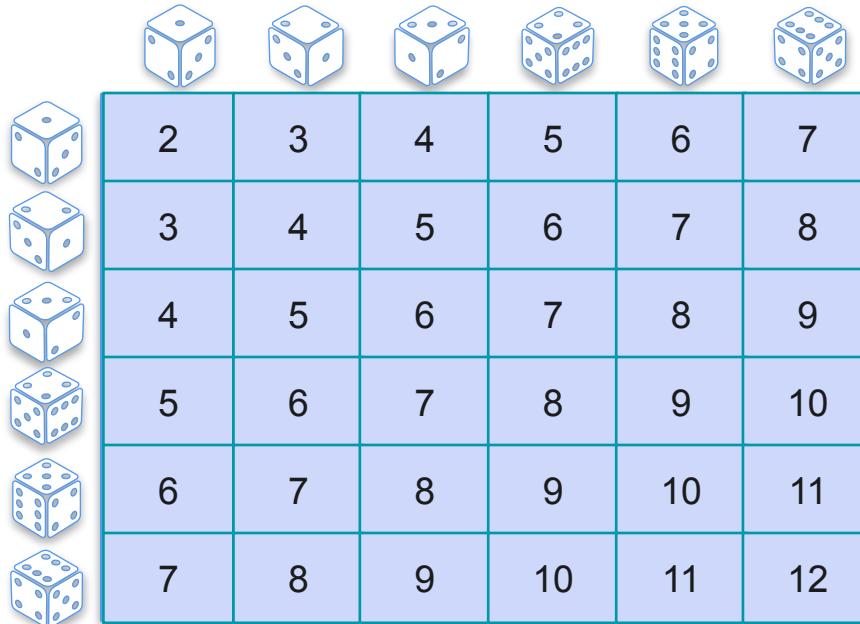
Joint Distributions - Example 3

Y : Sum of both dice



Joint Distributions - Example 3

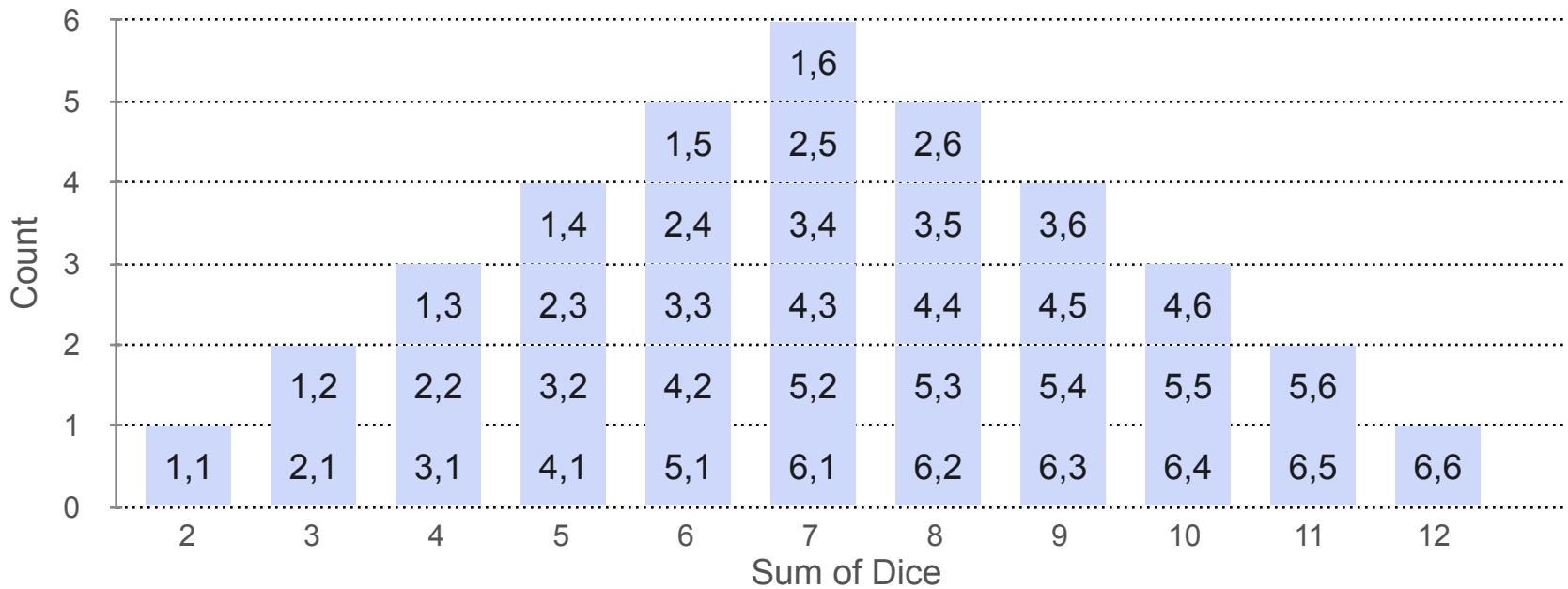
Y : Sum of both dice



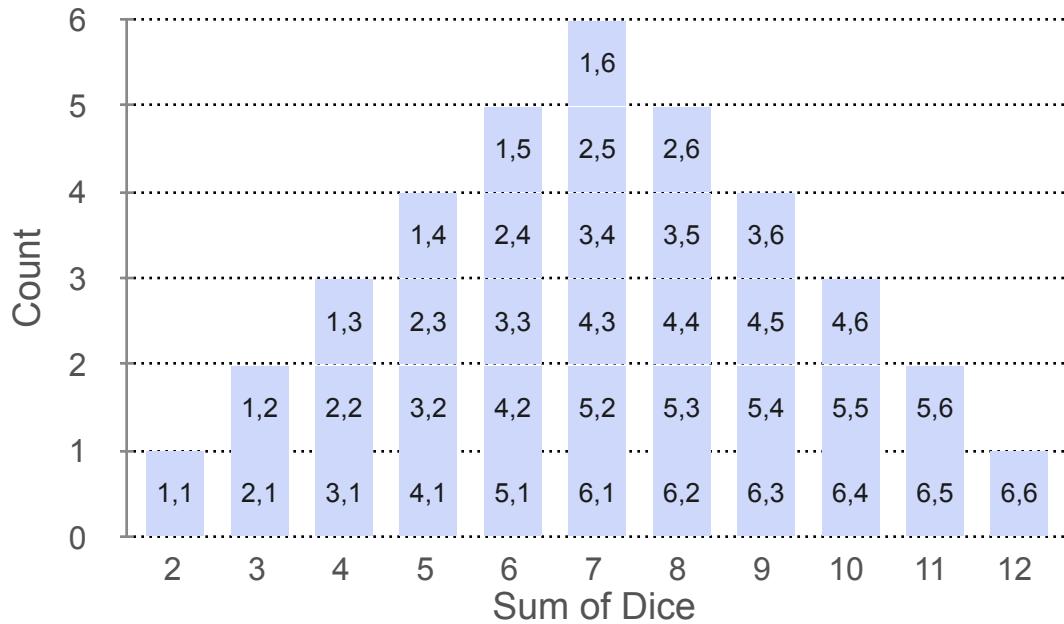
Joint Distributions: Example 3



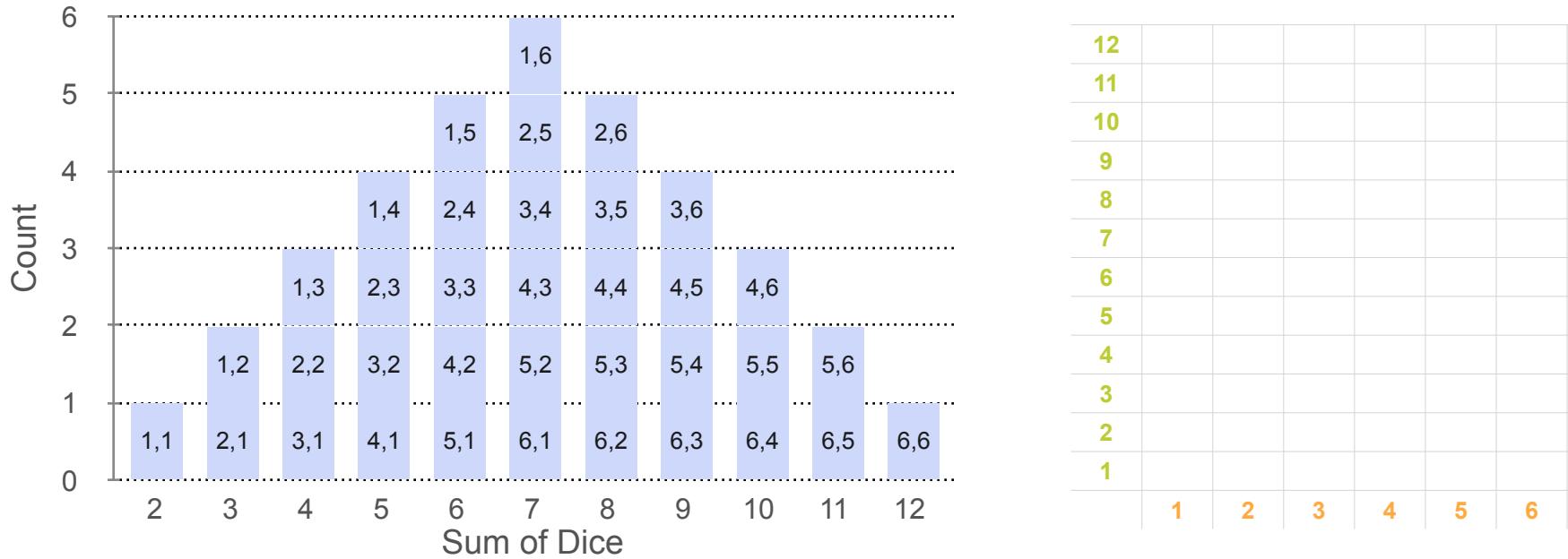
Joint Distributions: Example 3



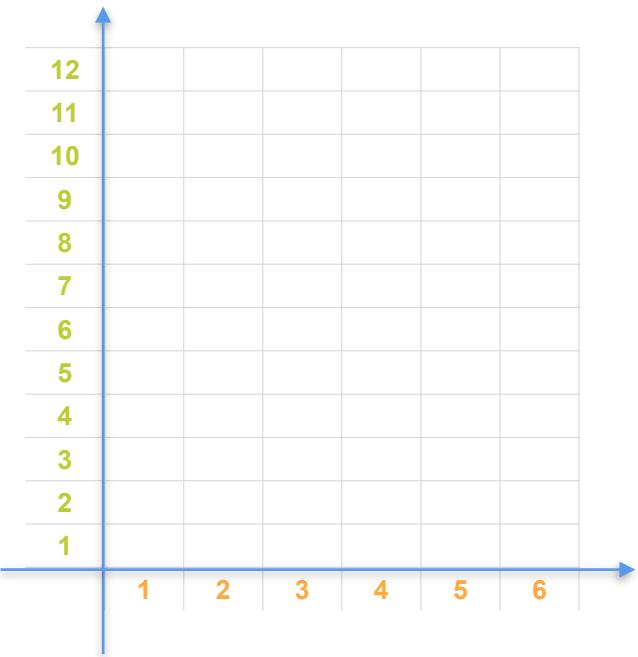
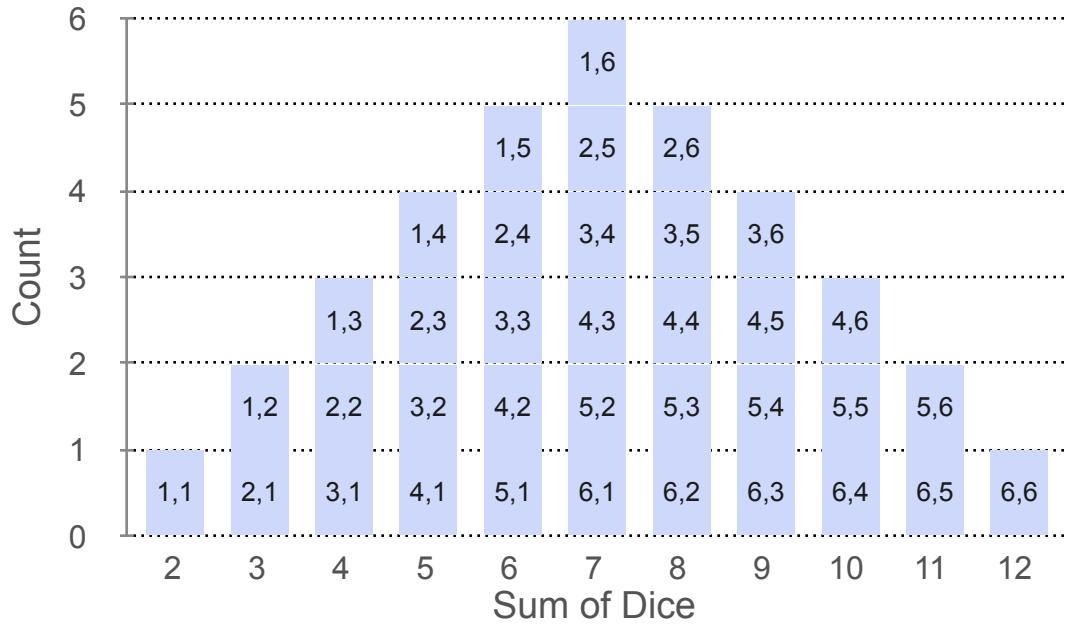
Joint Distributions: Example 3



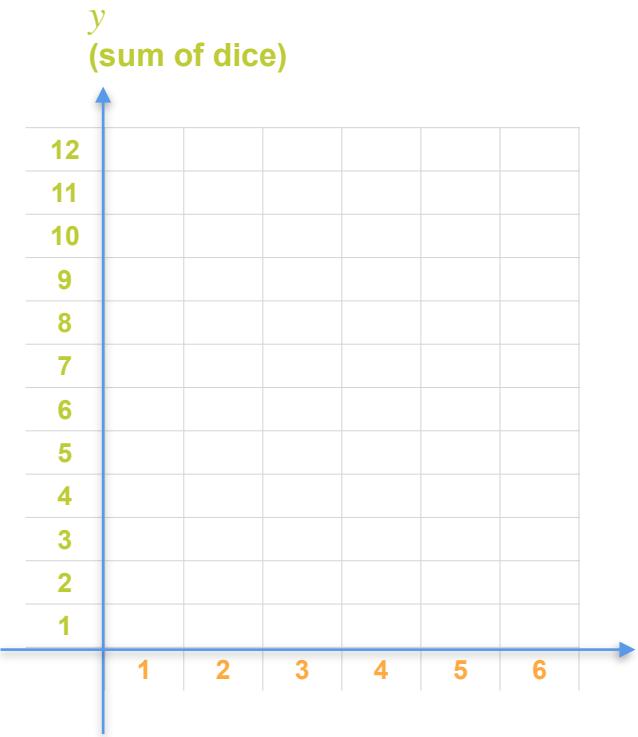
Joint Distributions: Example 3



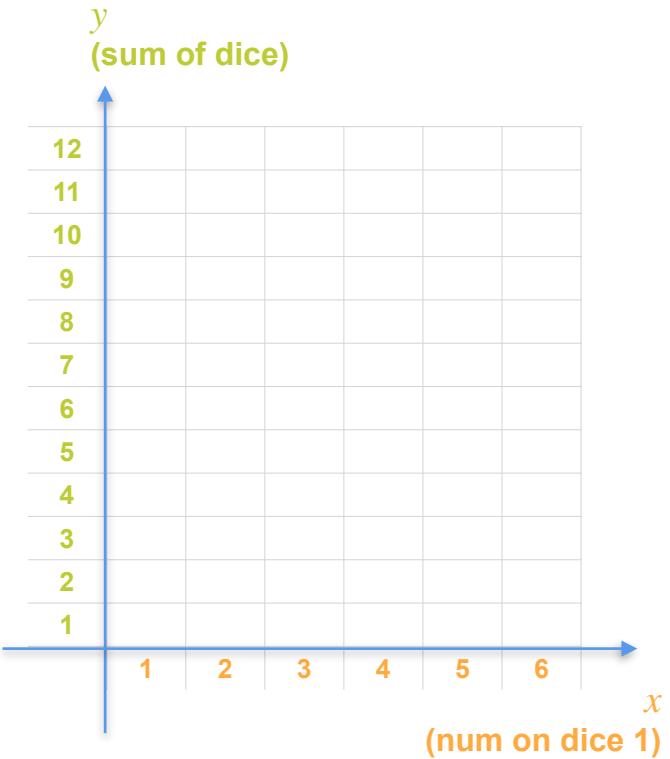
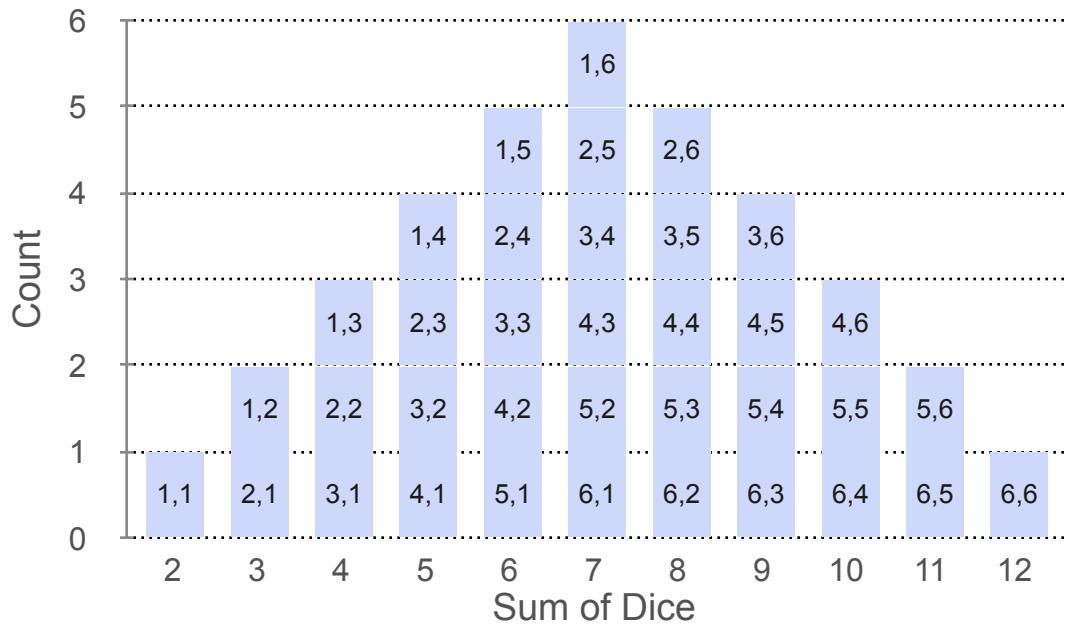
Joint Distributions: Example 3



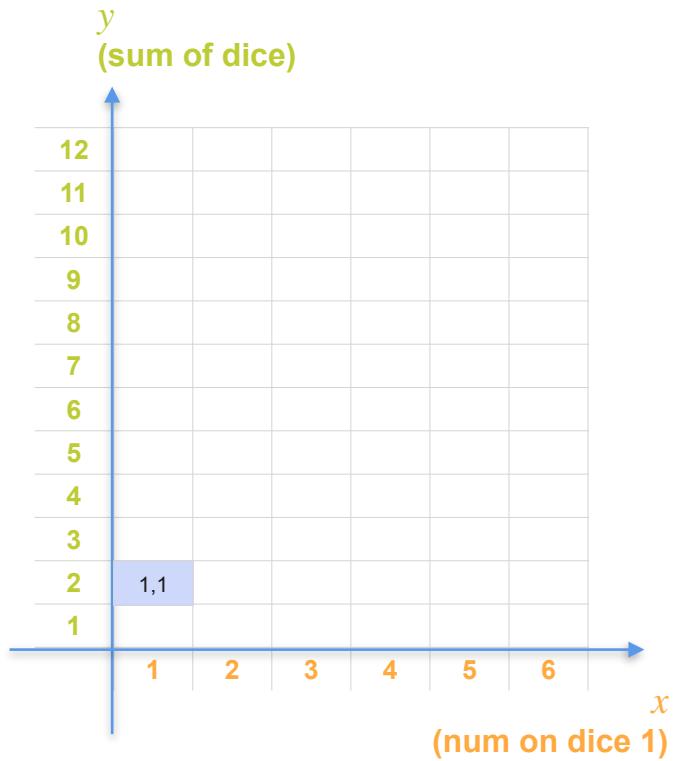
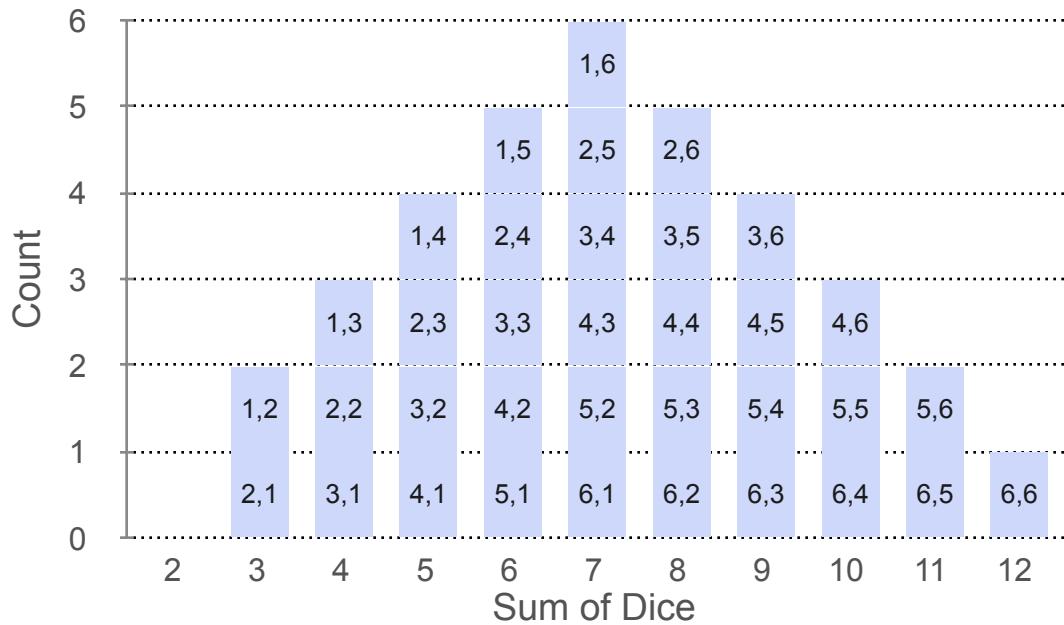
Joint Distributions: Example 3



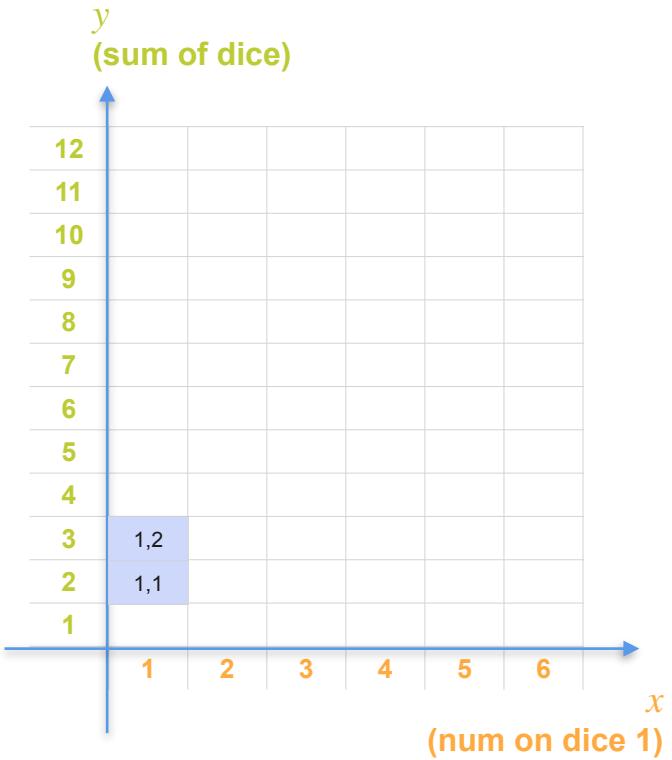
Joint Distributions: Example 3



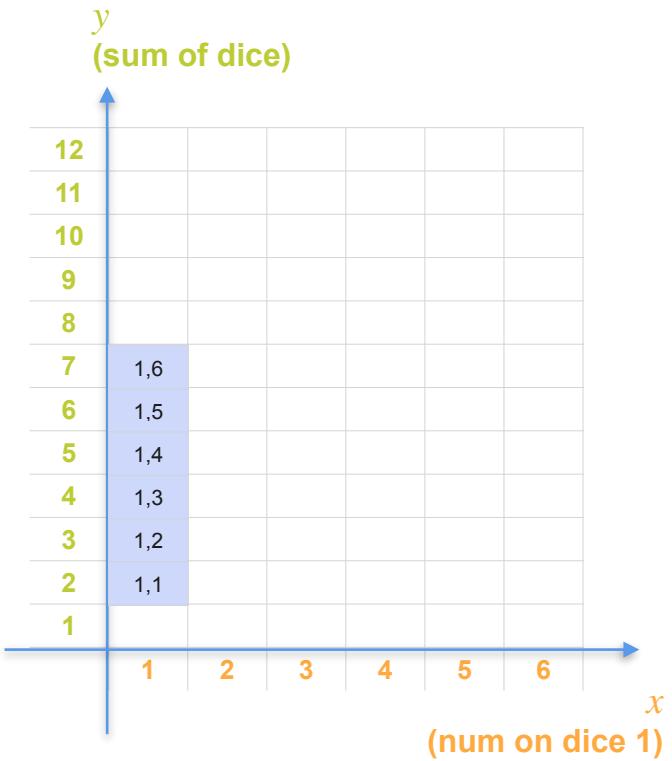
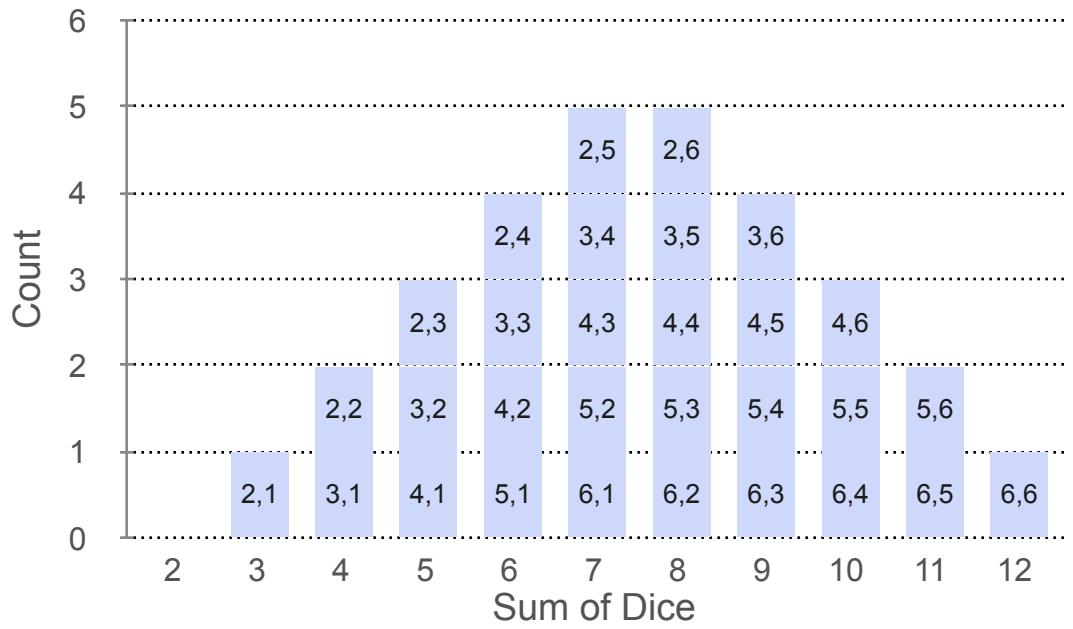
Joint Distributions: Example 3



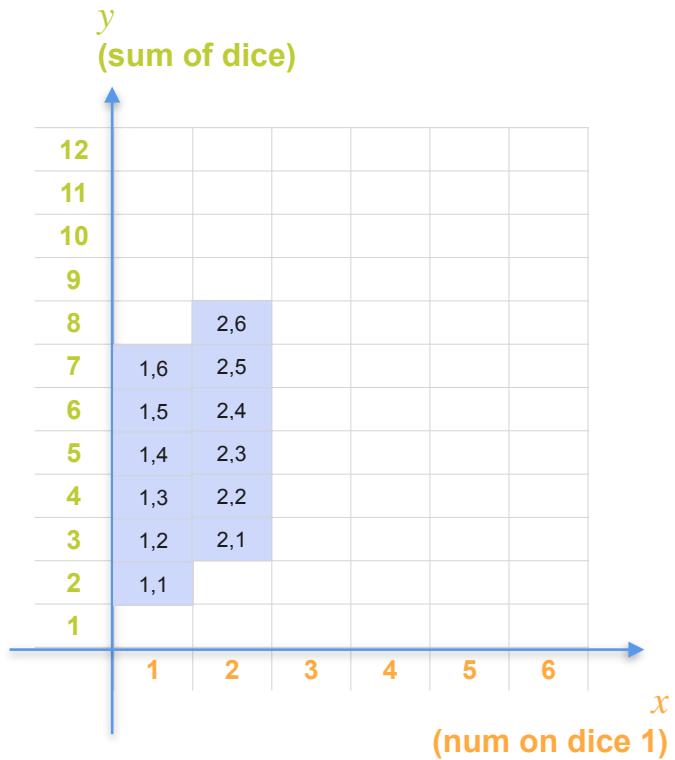
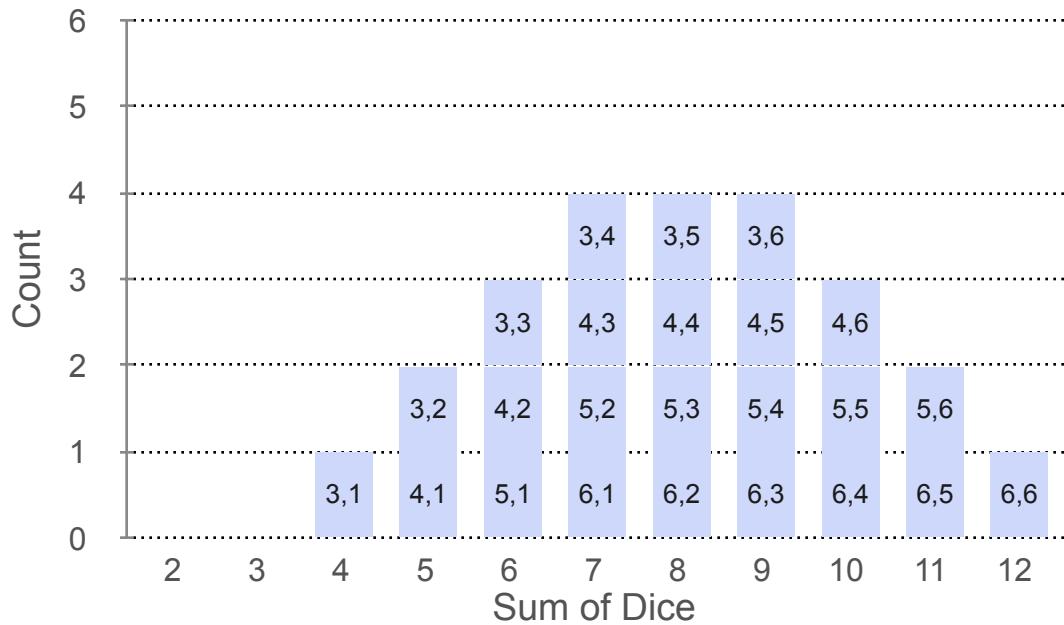
Joint Distributions: Example 3



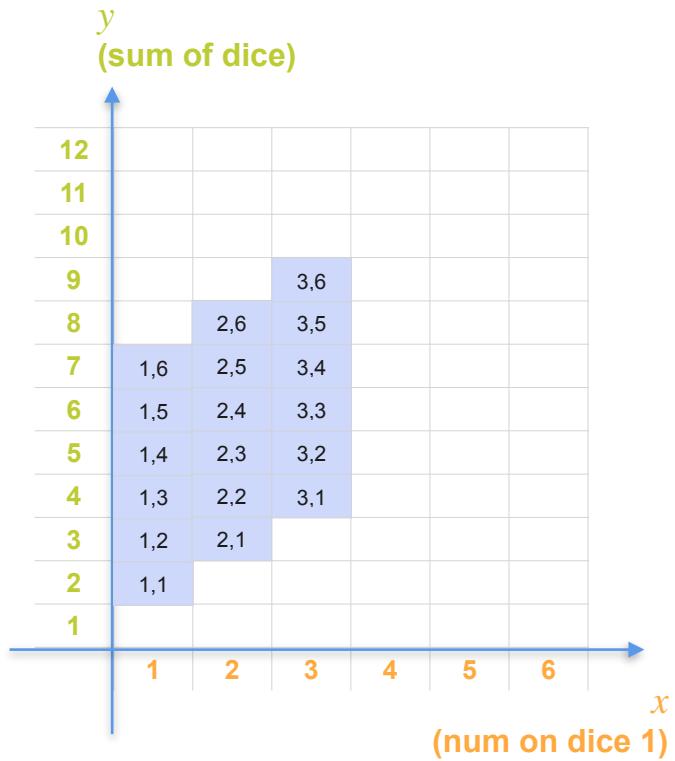
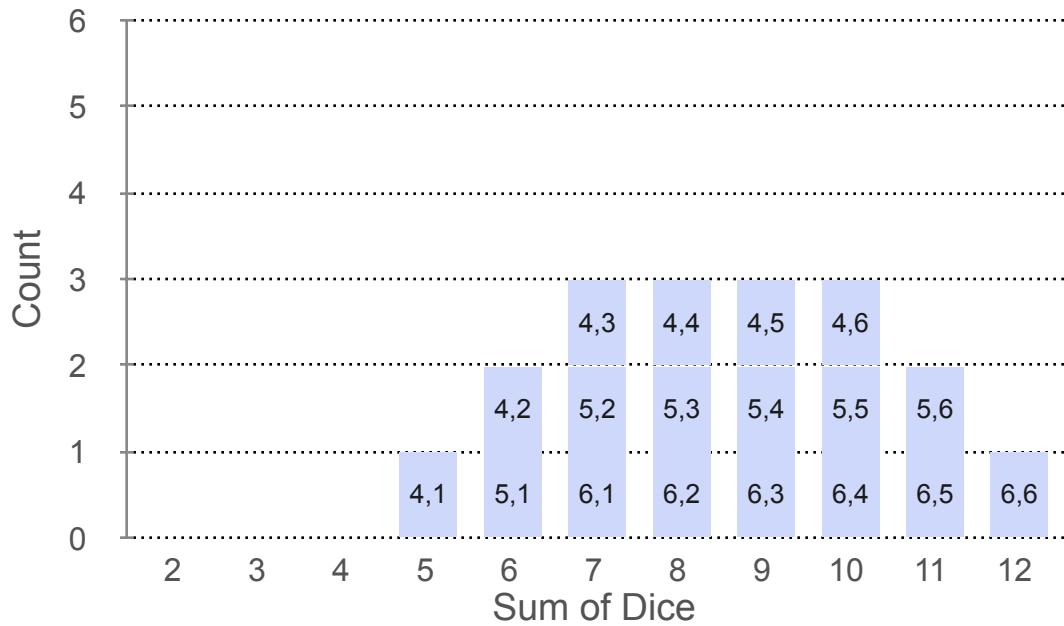
Joint Distributions: Example 3



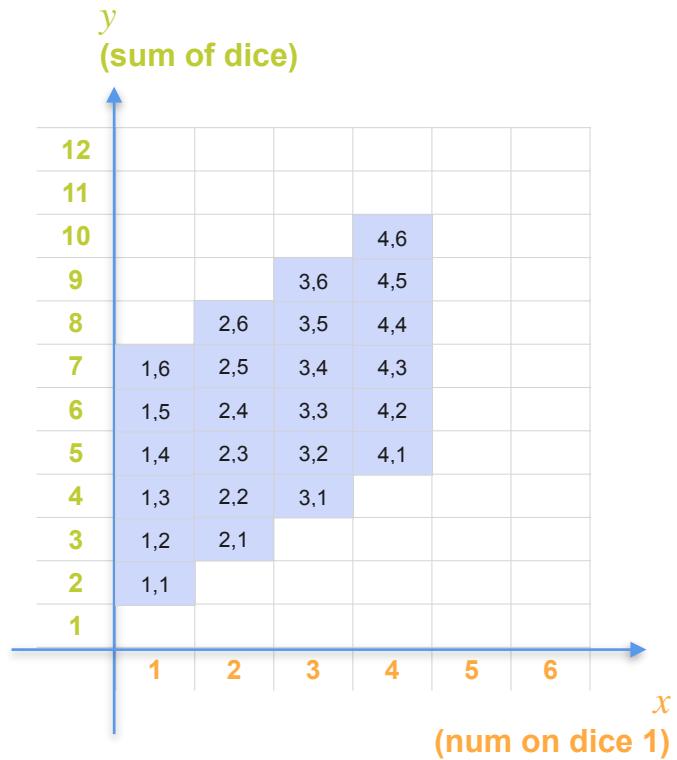
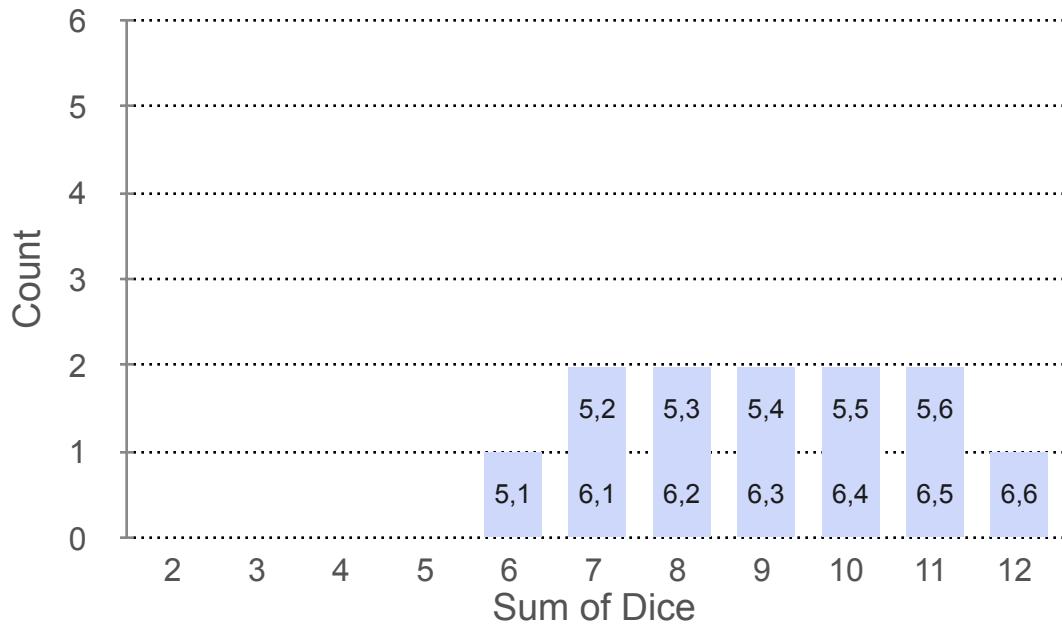
Joint Distributions: Example 3



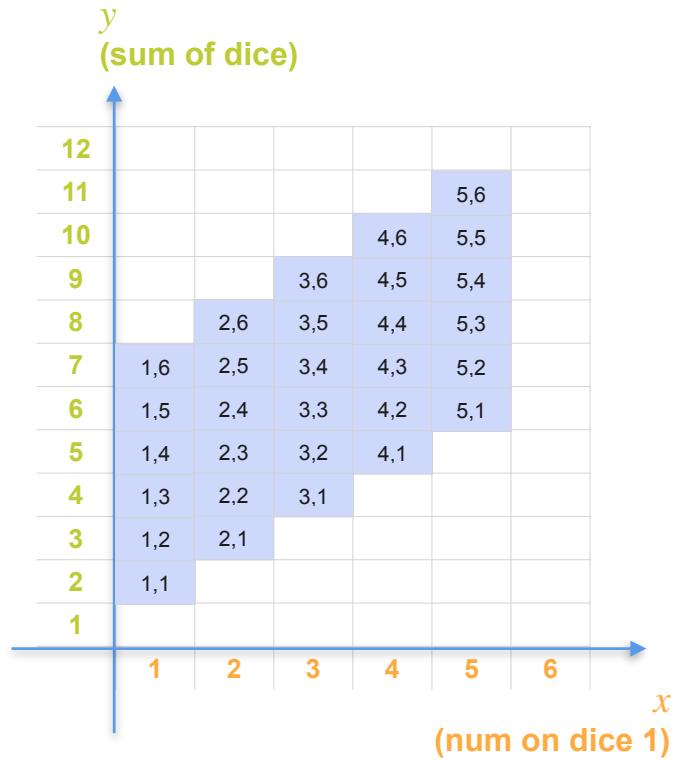
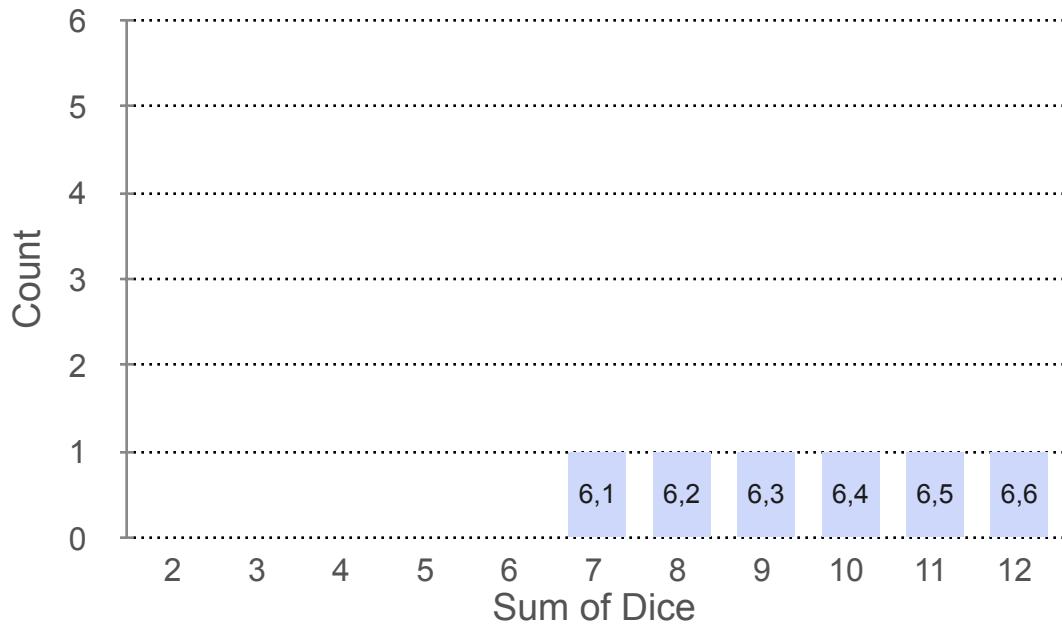
Joint Distributions: Example 3



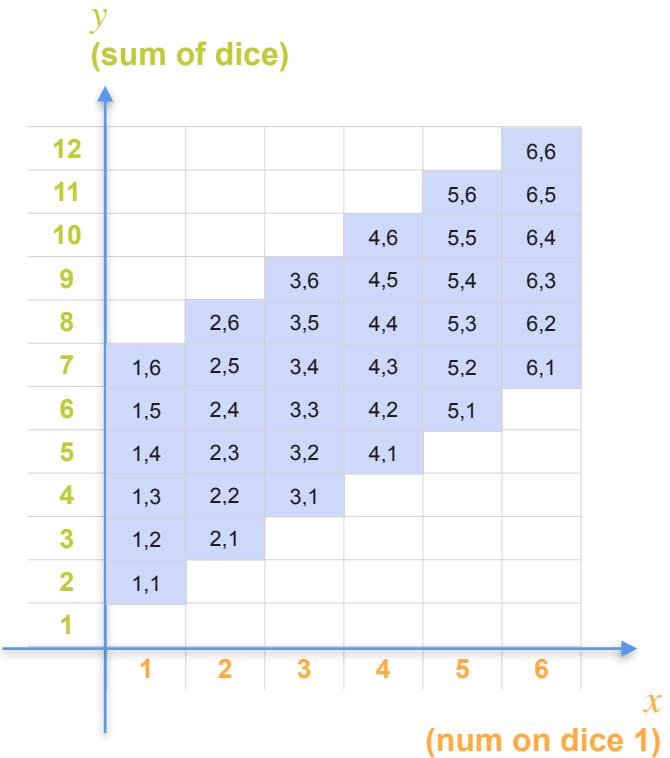
Joint Distributions: Example 3



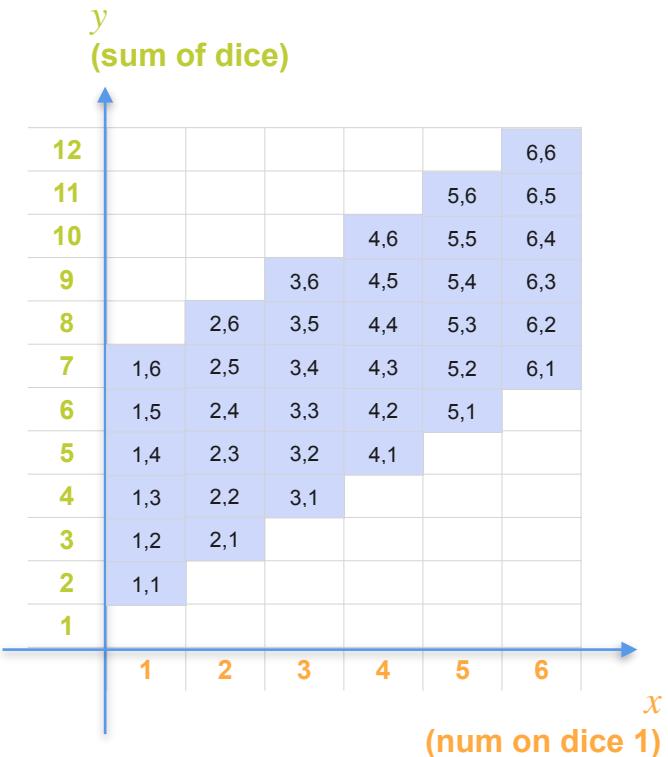
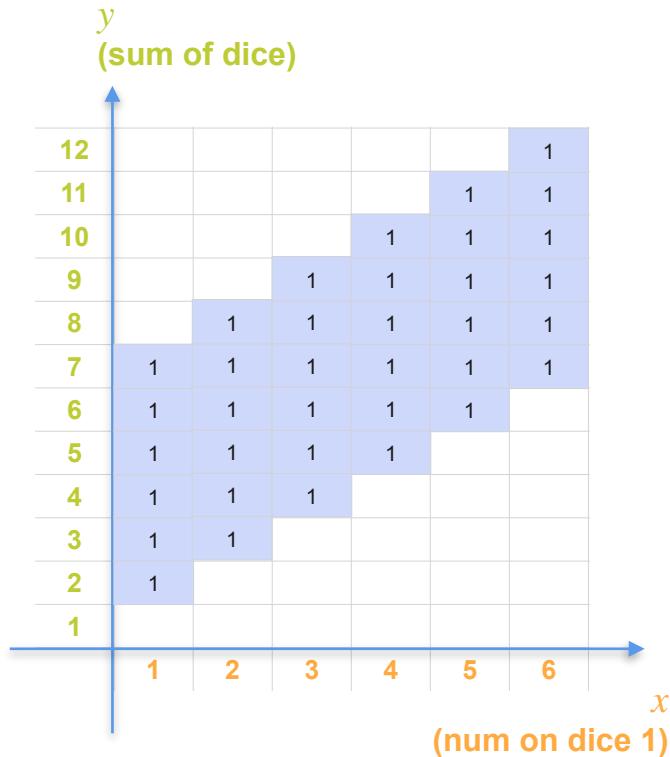
Joint Distributions: Example 3



Joint Distributions: Example 3



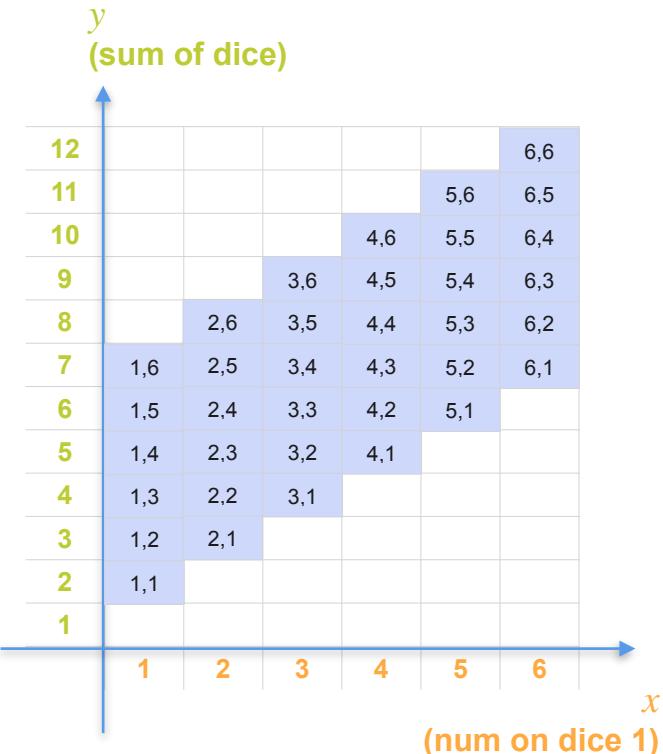
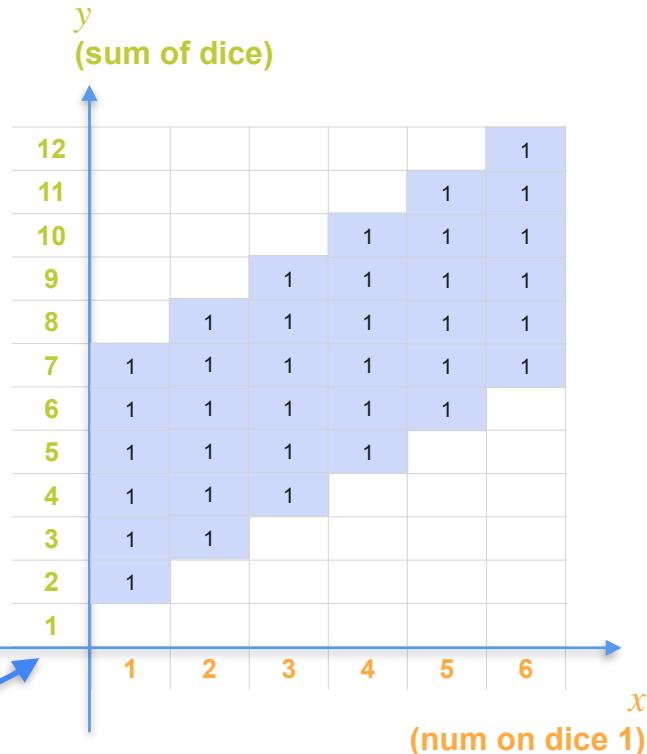
Joint Distributions: Example 3



Joint Distributions: Example 3

36
possible
outcomes

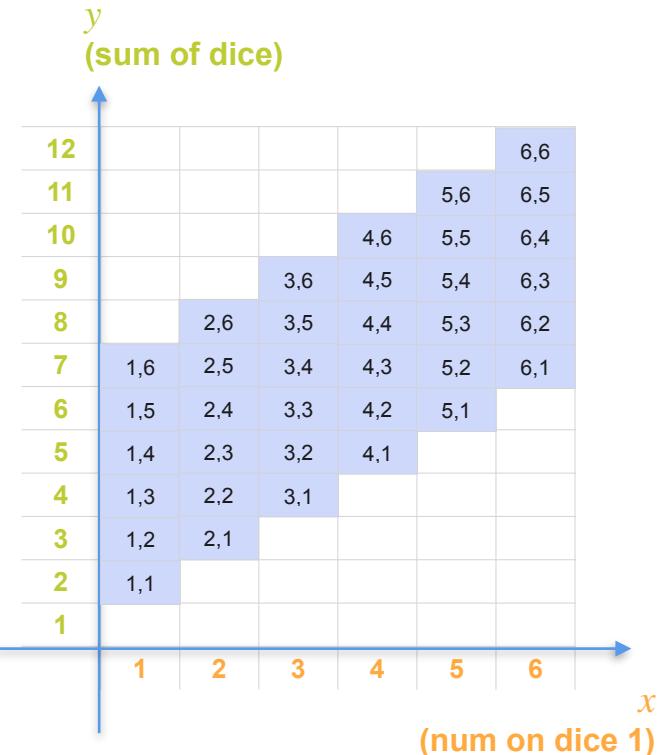
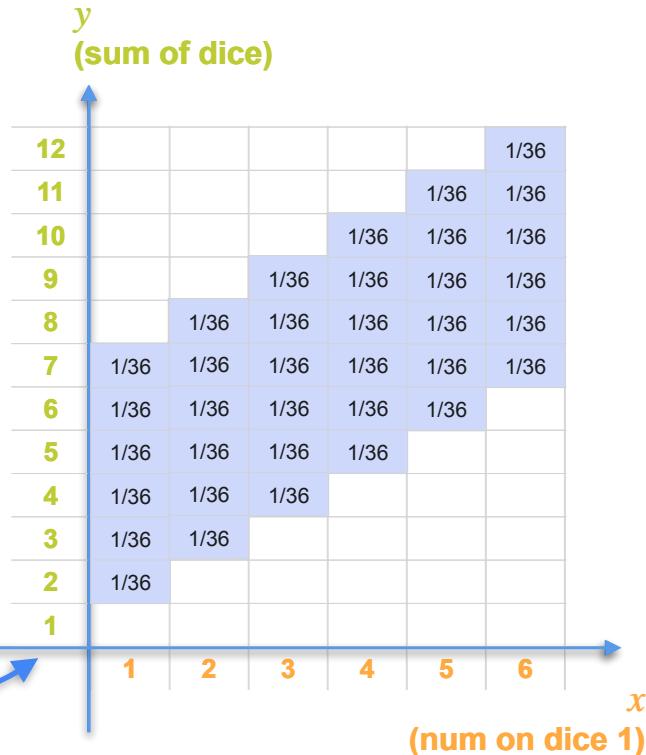
Divide by sum
(36)



Joint Distributions: Example 3

36
possible
outcomes

Divide by sum
(36)

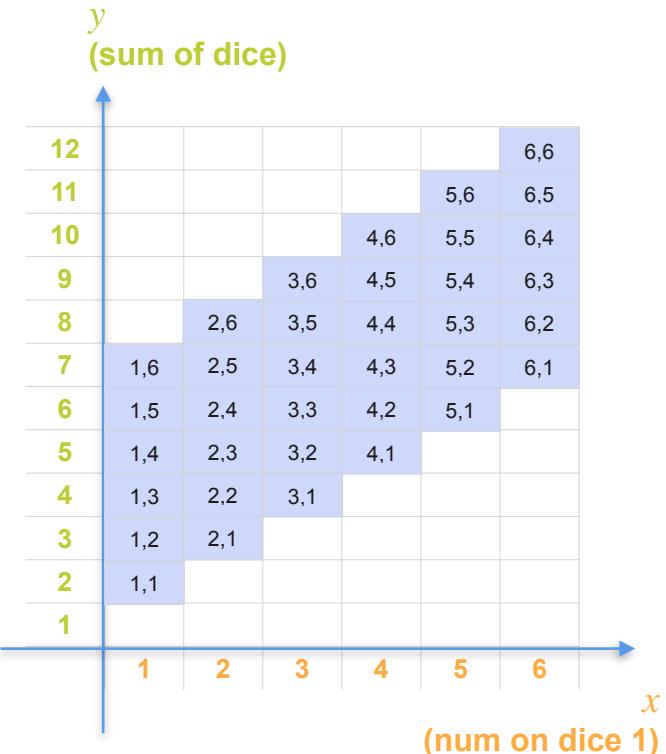
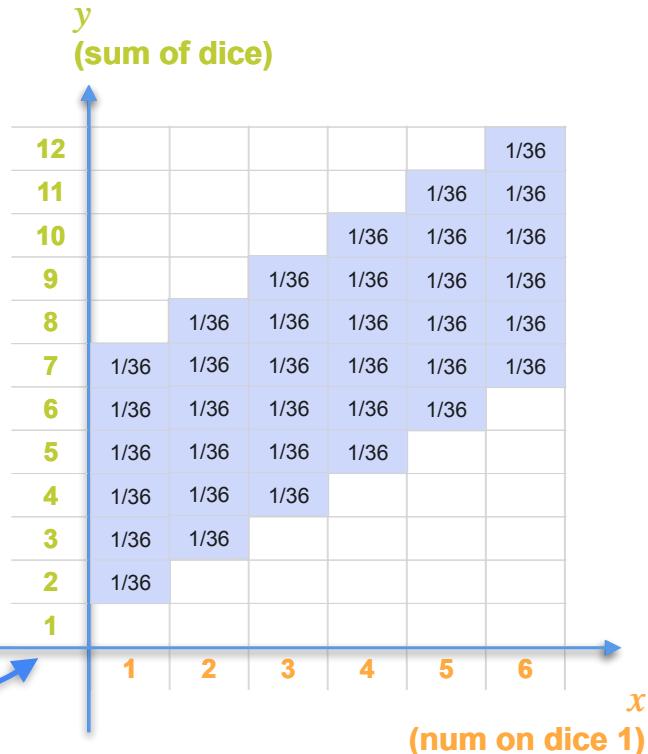


Joint Distributions: Example 3

Joint Distribution for
 X and Y

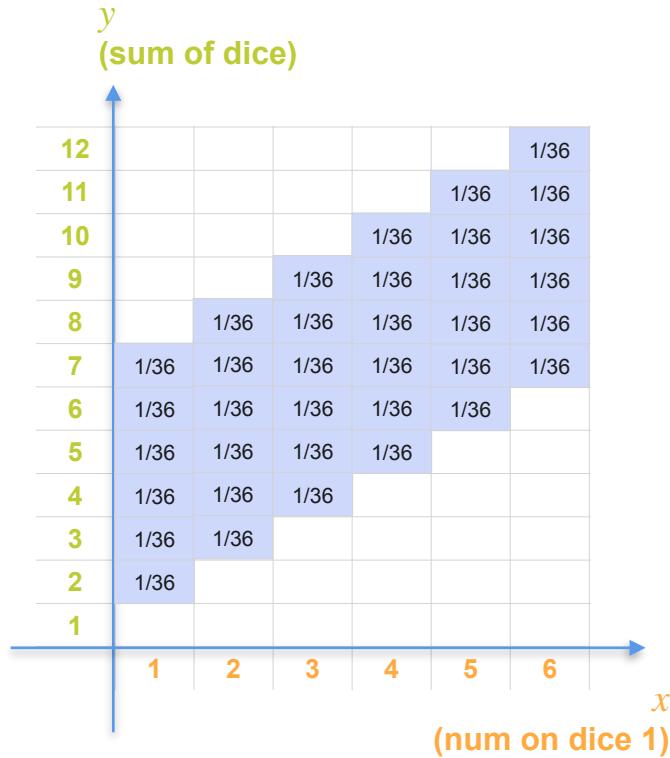
36
possible
outcomes

Divide by sum
(36)



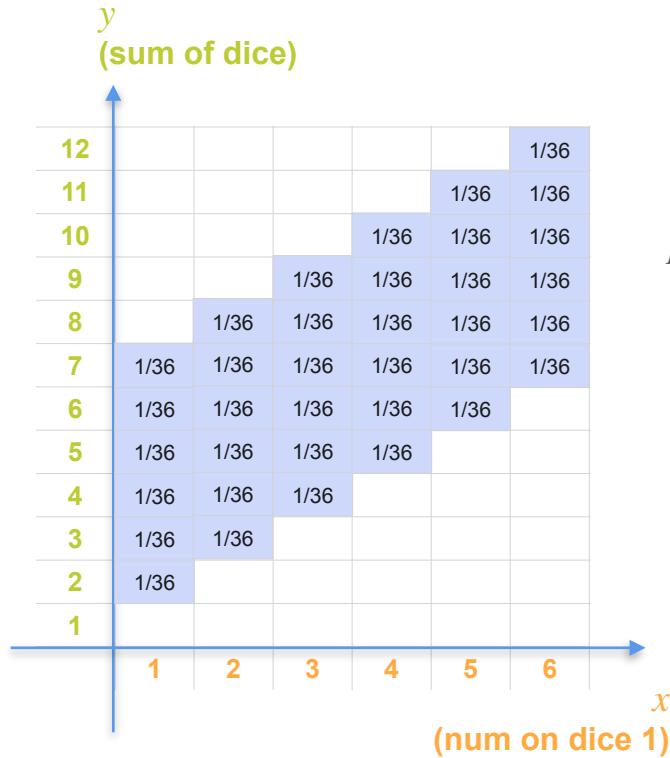
Joint Distributions: Example 3

Joint Distribution for
 X and Y



Joint Distributions: Example 3

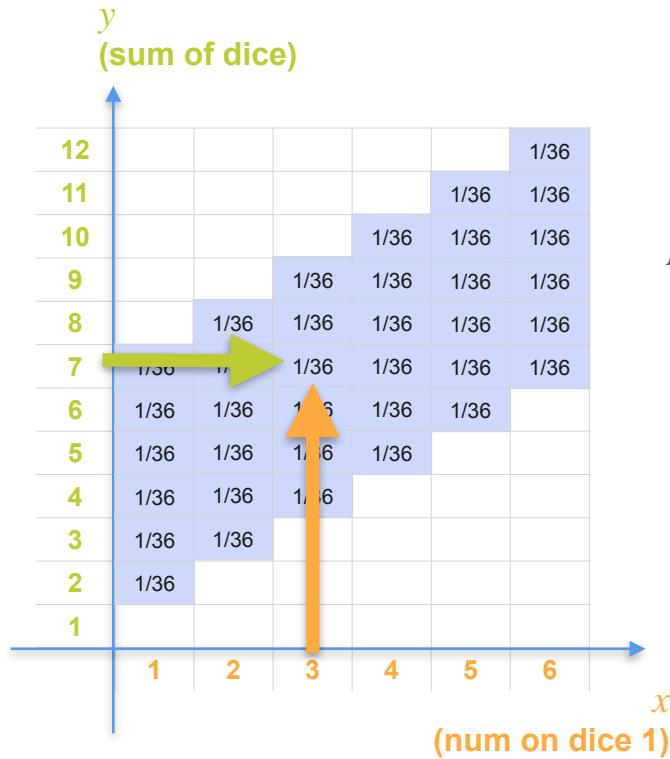
Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7)$$

Joint Distributions: Example 3

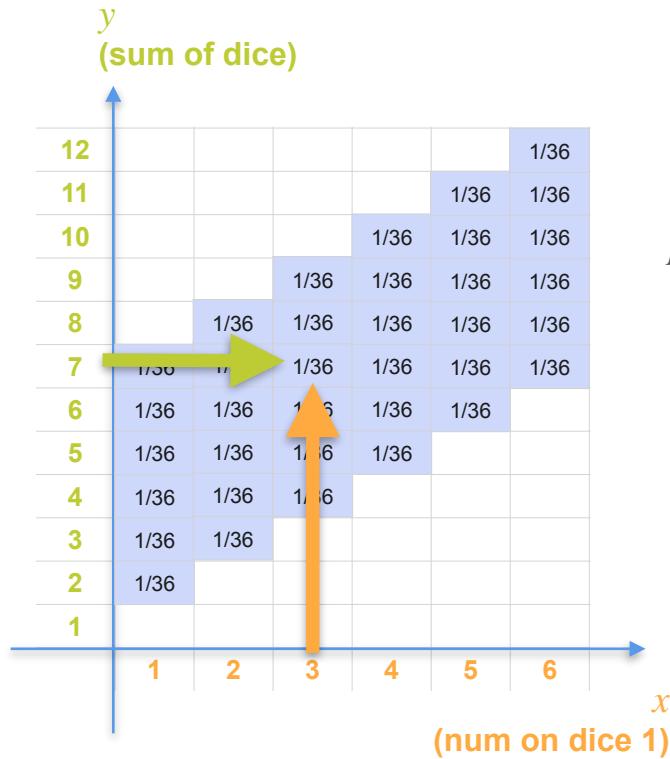
Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7)$$

Joint Distributions: Example 3

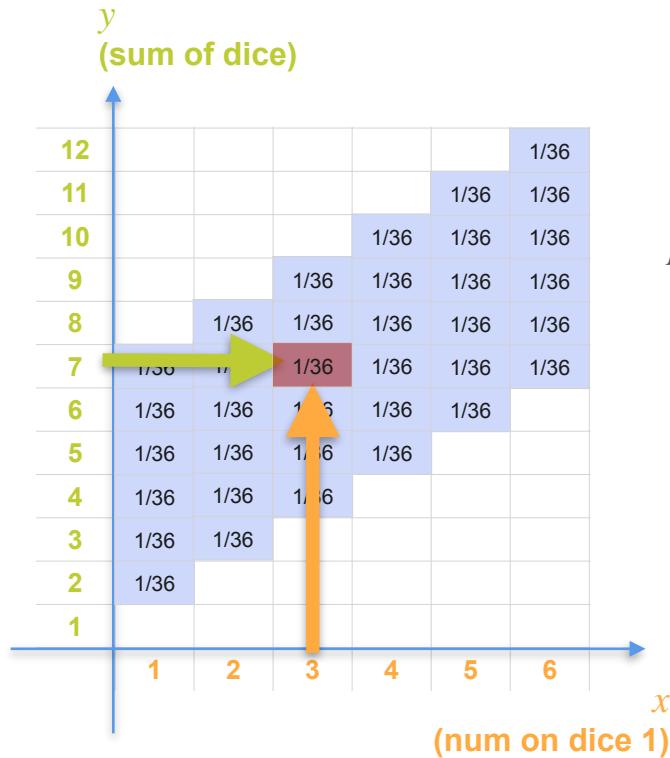
Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

Joint Distributions: Example 3

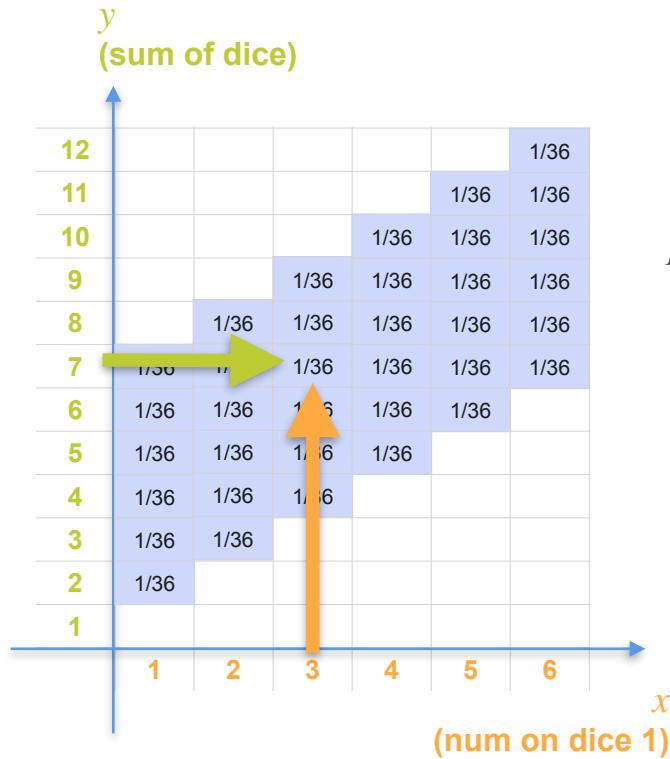
Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

Joint Distributions: Example 3

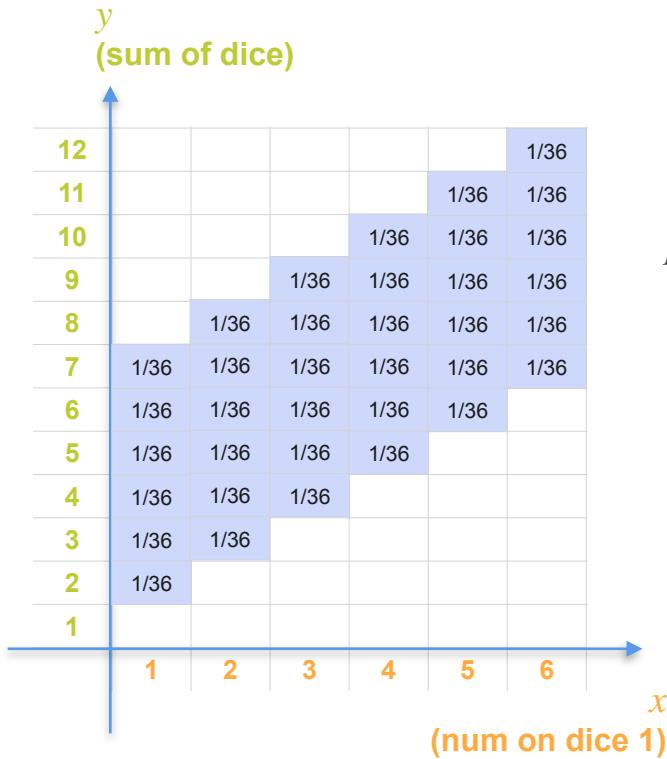
Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

Joint Distributions: Example 3

Joint Distribution for
X and Y

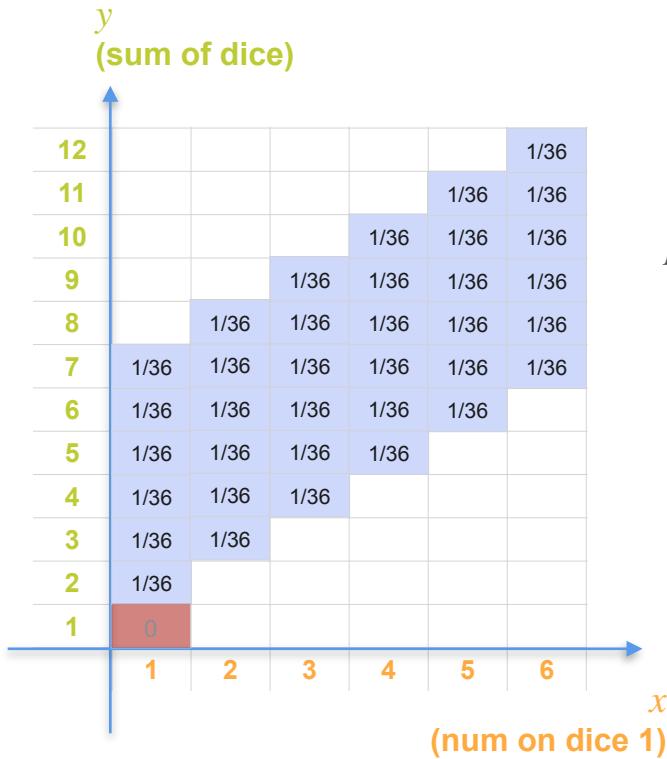


$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

$$p_{XY}(1, 1) = \mathbf{P}(X = 1, Y = 1)$$

Joint Distributions: Example 3

Joint Distribution for
X and Y

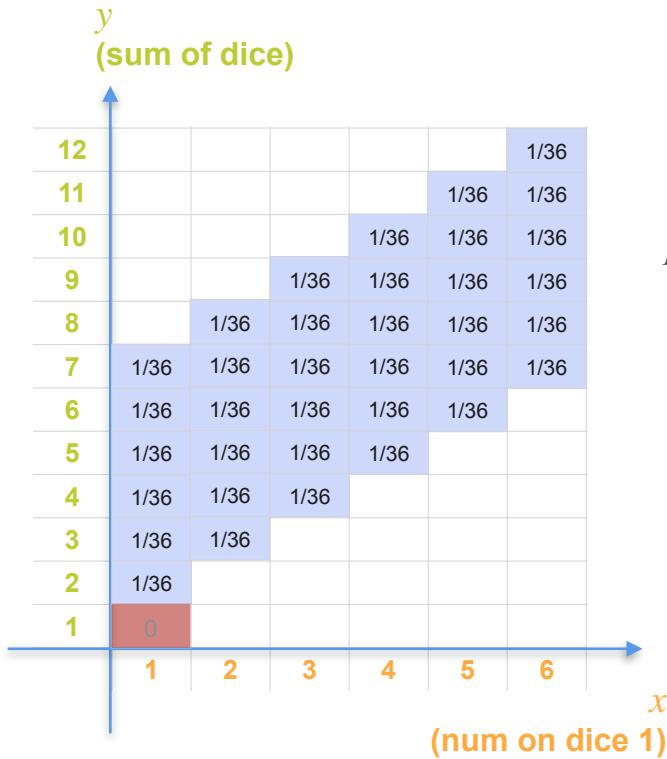


$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

$$p_{XY}(1, 1) = \mathbf{P}(X = 1, Y = 1)$$

Joint Distributions: Example 3

Joint Distribution for
X and Y



$$p_{XY}(3, 7) = \mathbf{P}(X = 3, Y = 7) = \frac{1}{36}$$

$$p_{XY}(1, 1) = \mathbf{P}(X = 1, Y = 1) = 0$$



DeepLearning.AI

Probability Distributions with Multiple Variables

**Joint Distribution
(Continuous)**

Joint Continuous Distributions

Joint Continuous Distributions

X : age of a child in year

Y : discrete values of height of child in inches

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X : the number rolled on the 1st dice

Y : sum of both dice

Joint Continuous Distributions

X : age of a child in year

Y : discrete values of height of child in inches

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X : the number rolled on the 1st dice

Y : sum of both dice

X and Y are
Discrete Random Variables

Joint Continuous Distributions

X : age of a child in year

Y : discrete values of height of child in inches

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

X : the number rolled on the 1st dice

Y : sum of both dice

X and Y are
Discrete Random Variables

What about when X and Y are
Continuous Random Variables?

Joint Continuous Distributions

Joint Continuous Distributions



Joint Continuous Distributions

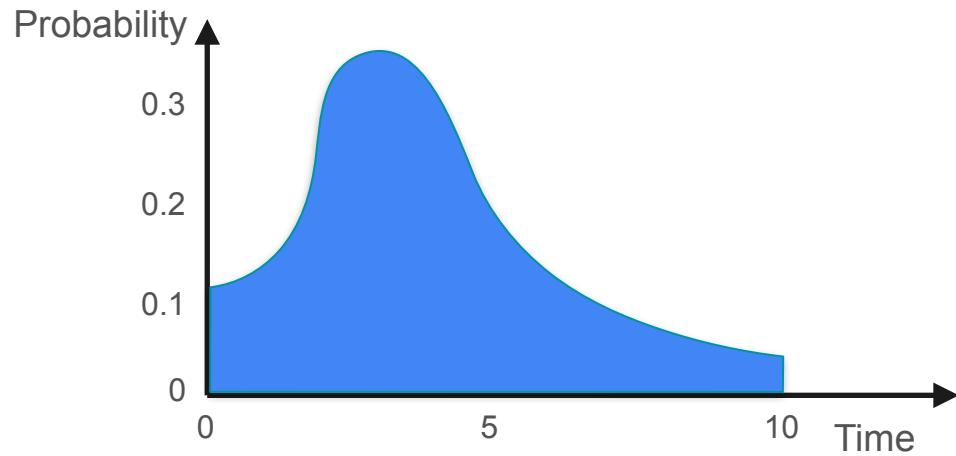


X variable: Waiting time

Joint Continuous Distributions



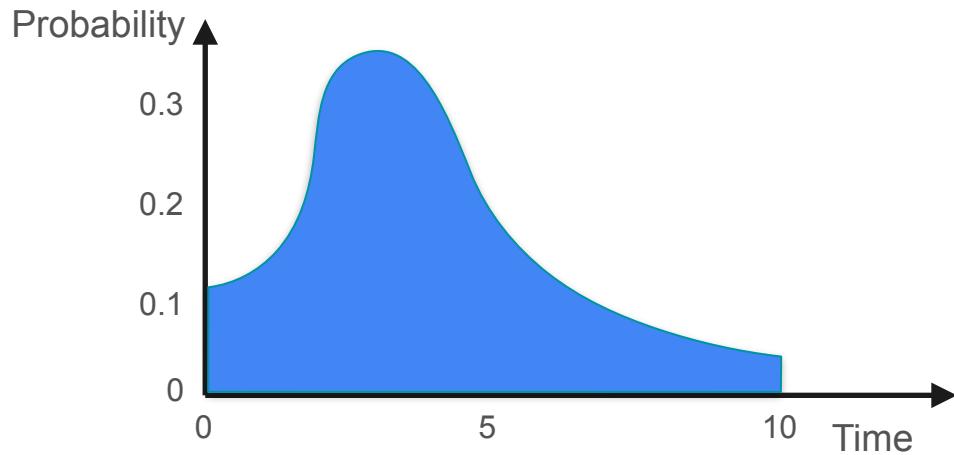
X variable: Waiting time



Joint Continuous Distributions



X variable: Waiting time

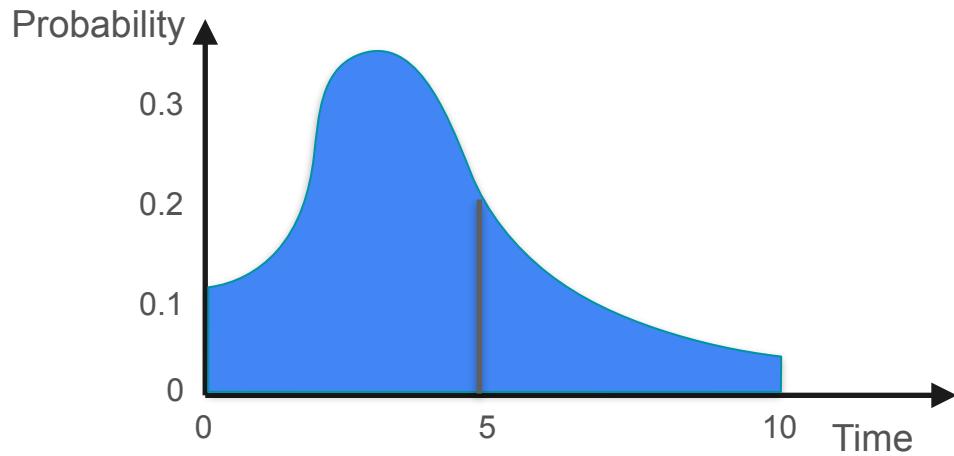


$P(X \text{ between } 0 \text{ and } 5 \text{ mins})$

Joint Continuous Distributions



X variable: Waiting time

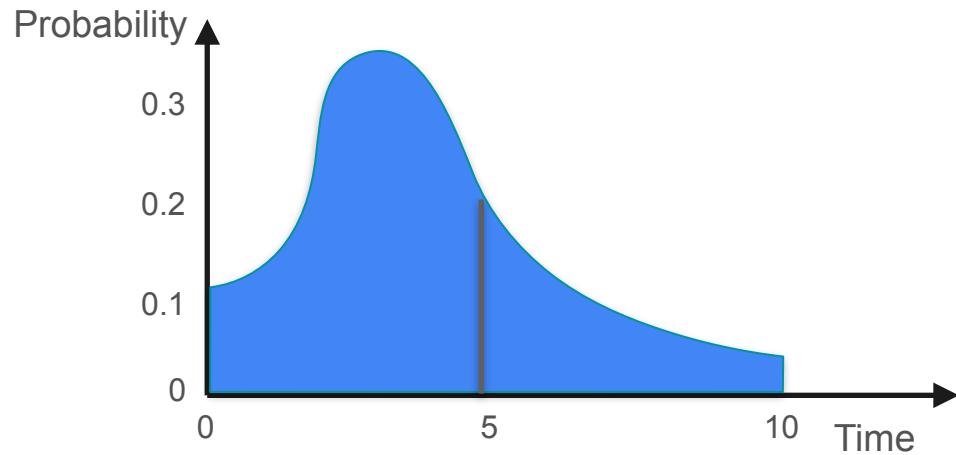


$P(X \text{ between } 0 \text{ and } 5 \text{ mins})$

Joint Continuous Distributions



X variable: Waiting time

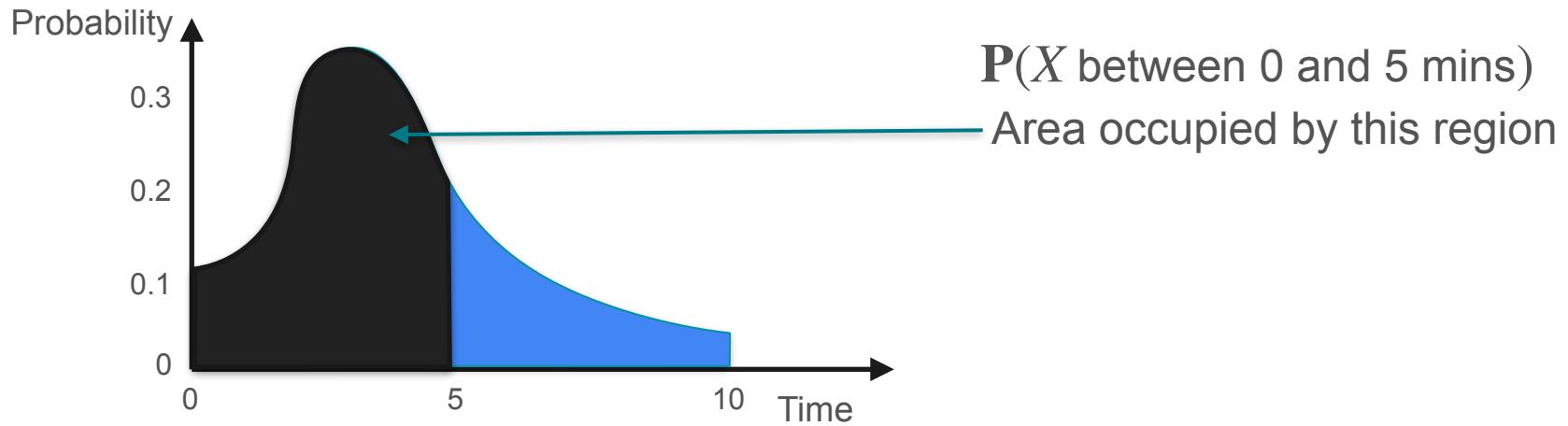


$P(X \text{ between } 0 \text{ and } 5 \text{ mins})$
Area occupied by this region

Joint Continuous Distributions



X variable: Waiting time



Joint Continuous Distributions

Joint Continuous Distributions

X

Y

Joint Continuous Distributions

X

Waiting time
before a call is picked up
[0 - 10 minutes]



Y

Joint Continuous Distributions

X

Waiting time
before a call is picked up
[0 - 10 minutes]



Y

Customer
satisfaction rating
[0 - 10]



Joint Continuous Distributions

X

Waiting time
before a call is picked up
[0 - 10 minutes]



Y

Customer
satisfaction rating
[0 - 10]



**Both variables are
continuous**

Joint Continuous Distributions

X

Waiting time
before a call is picked up
[0 - 10 minutes]



2.4 minutes

1.5 minutes

Y

Customer
satisfaction rating
[0 - 10]



Both variables are
continuous

Joint Continuous Distributions

X

Waiting time
before a call is picked up
[0 - 10 minutes]



2.4 minutes

1.5 minutes

Y

Customer
satisfaction rating
[0 - 10]



0.0

5.7

Both variables are
continuous

Joint Continuous Distributions: Dataset

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)

0 - 10 mins

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)

0 - 10 mins

1000 customers

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)

0 - 10 mins

1000 customers



Joint Continuous Distributions: Dataset

Joint Continuous Distributions: Dataset

Y variable: Satisfaction rating

0 - 10

Joint Continuous Distributions: Dataset

Y variable: Satisfaction rating

0 - 10

1000 customers

Joint Continuous Distributions: Dataset

Y variable: Satisfaction rating

0 - 10

1000 customers



Joint Continuous Distributions: Dataset

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)
0 - 10 mins

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

1000 customers

Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

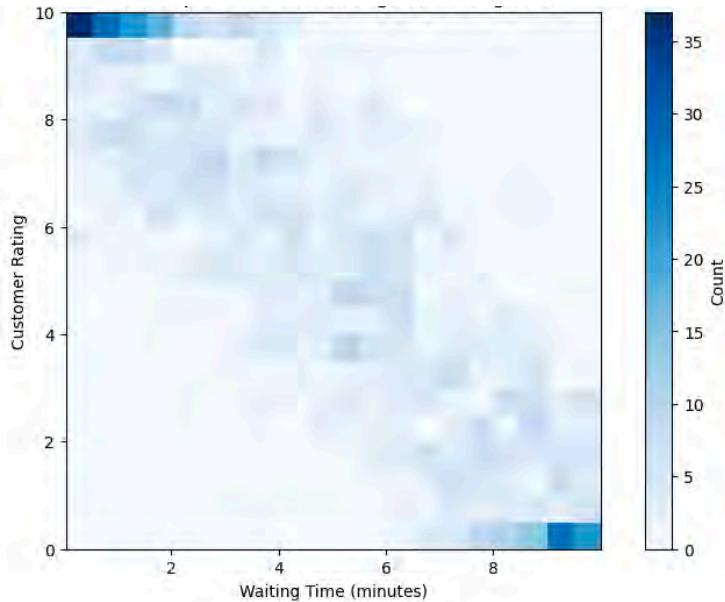
1000 customers



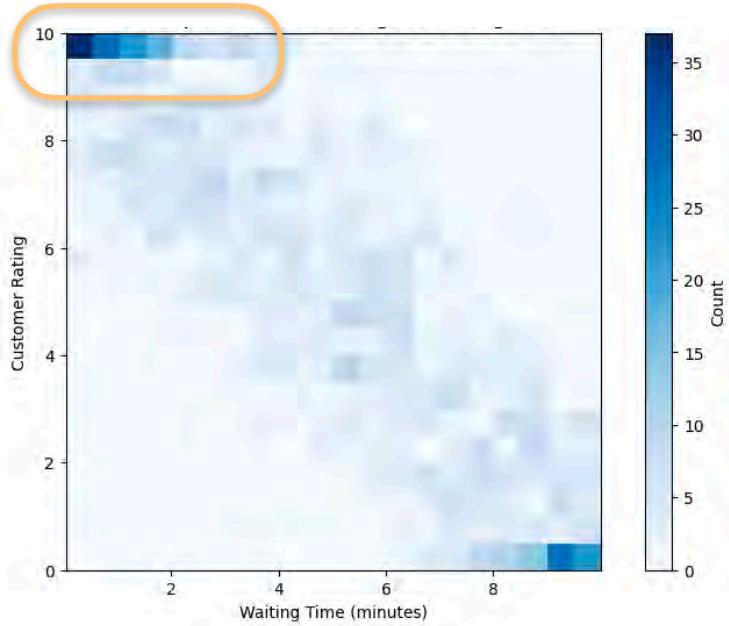
Joint Continuous Distributions: Dataset



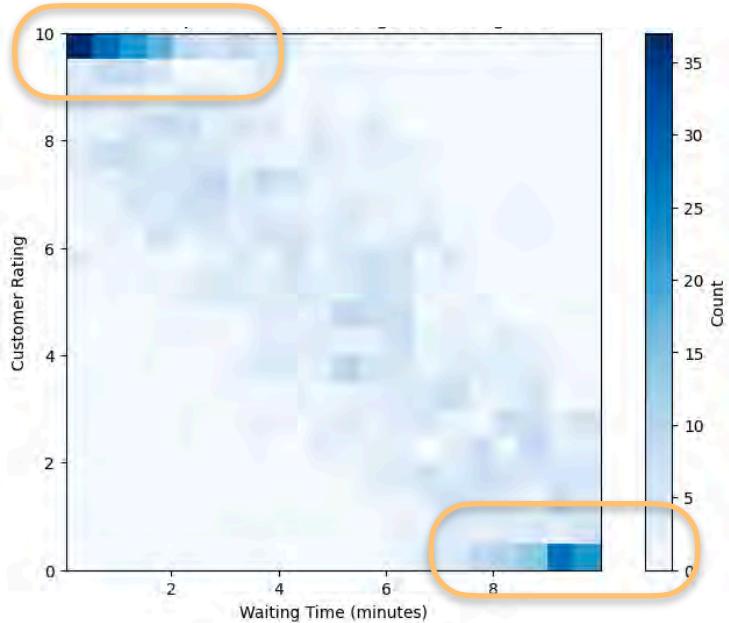
Joint Continuous Distributions: Dataset



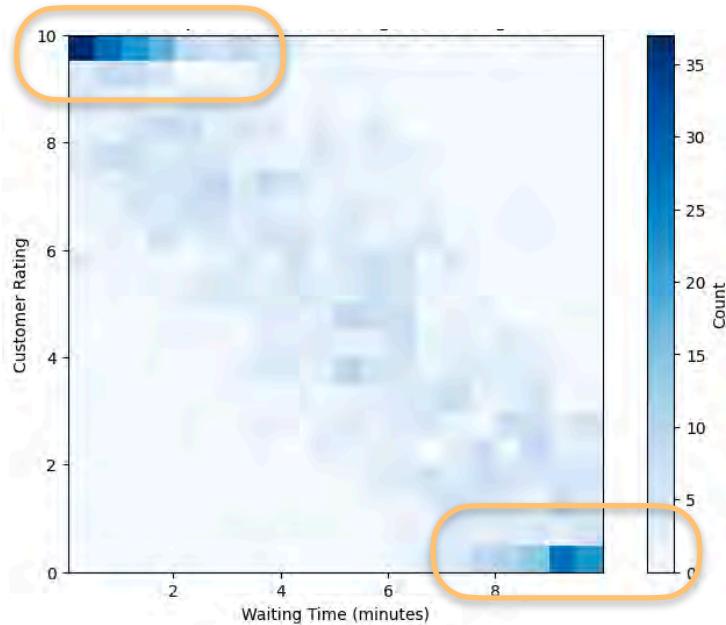
Joint Continuous Distributions: Dataset



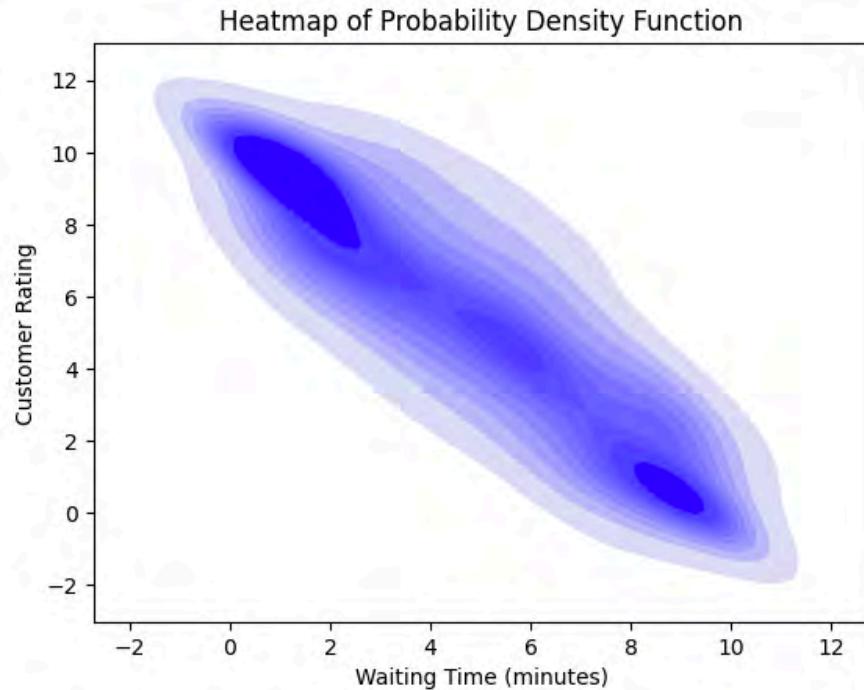
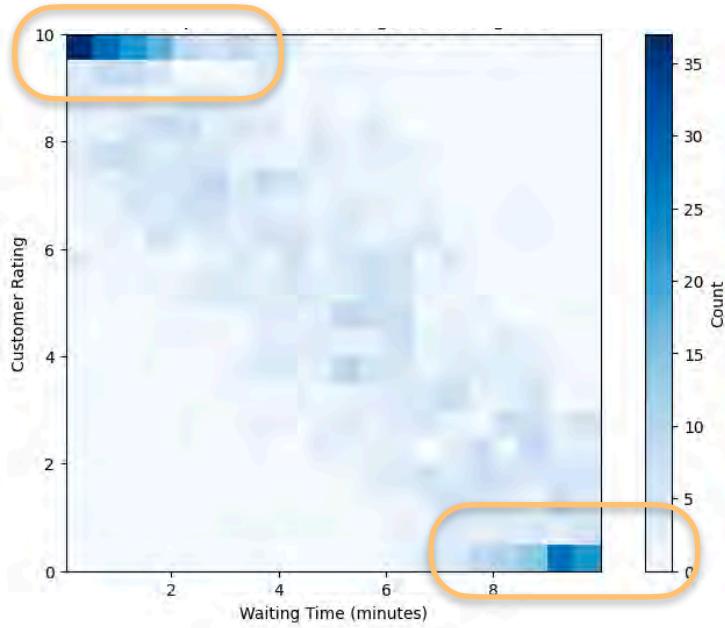
Joint Continuous Distributions: Dataset



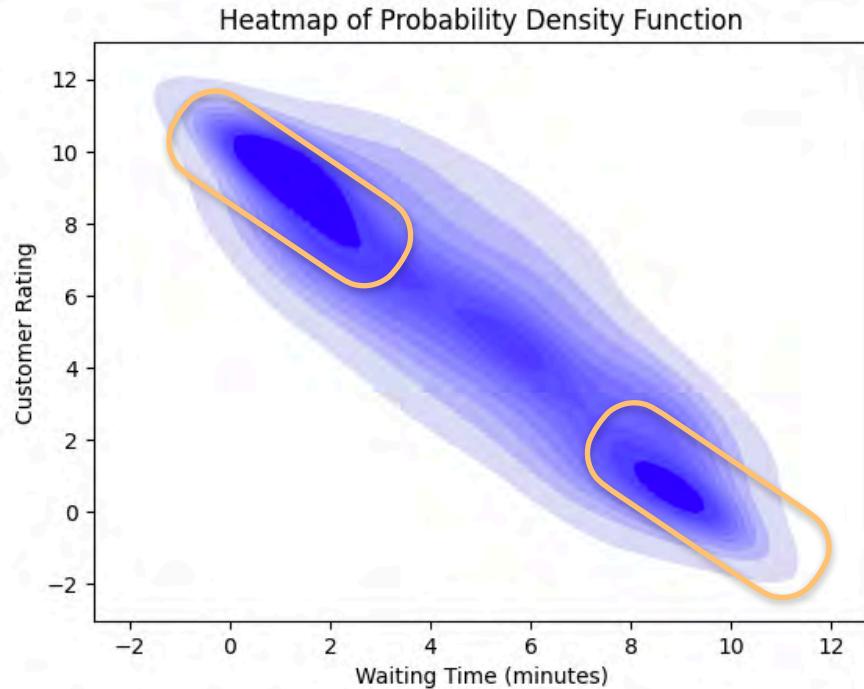
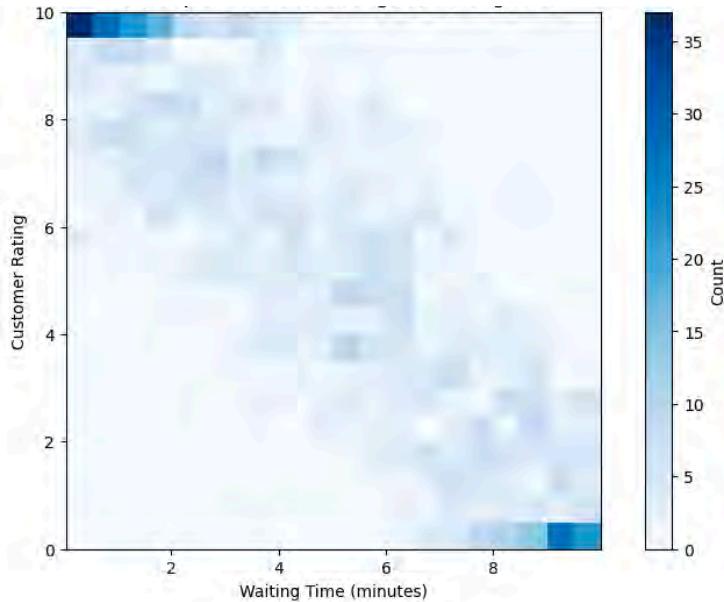
Joint Continuous Distributions: Dataset



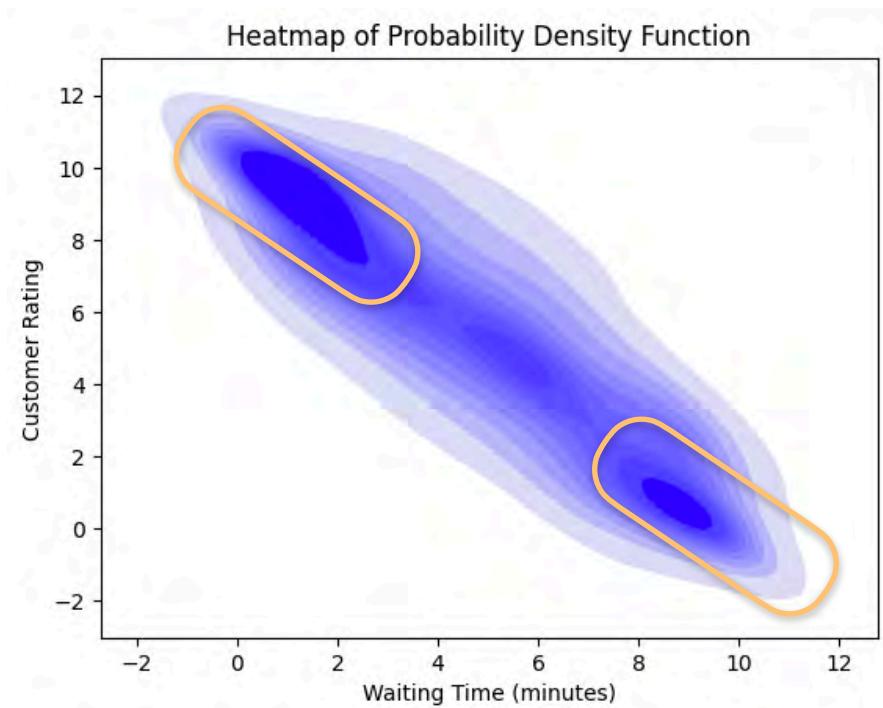
Joint Continuous Distributions: Dataset



Joint Continuous Distributions: Dataset

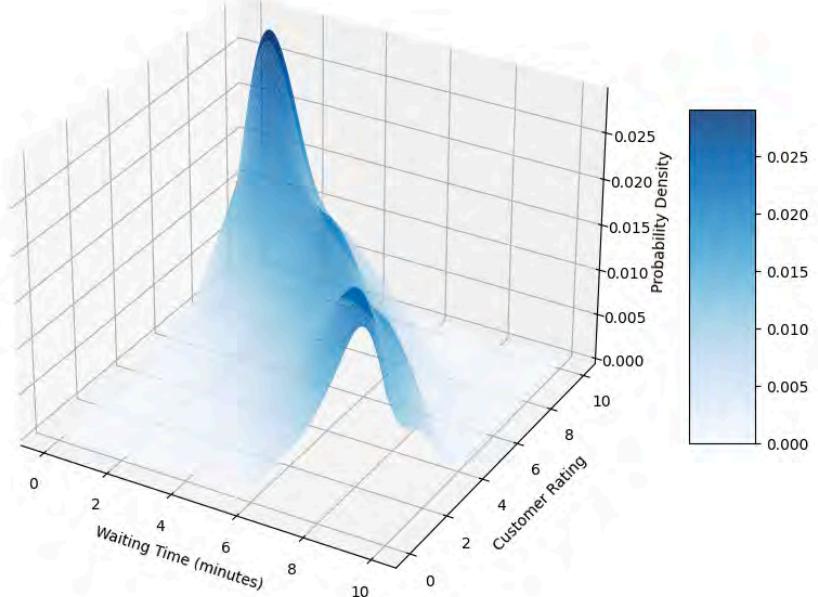


Joint Continuous Distributions: Dataset

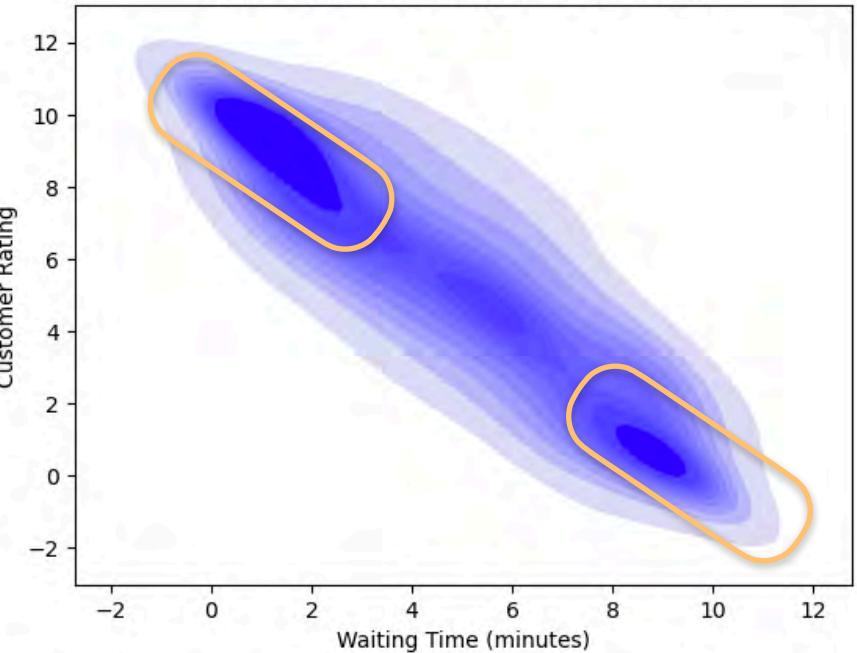


Joint Continuous Distributions: Dataset

3D Probability Density Distribution for Customer Ratings vs Waiting Time

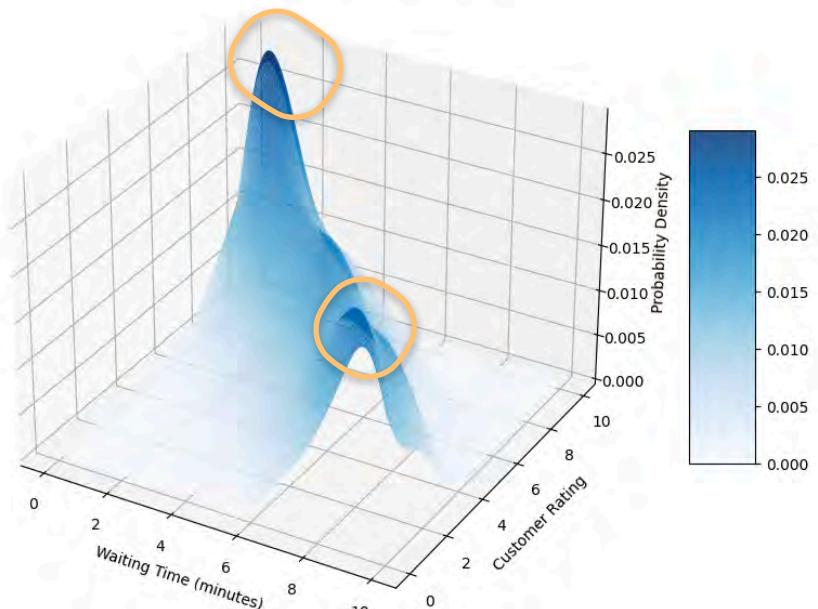


Heatmap of Probability Density Function

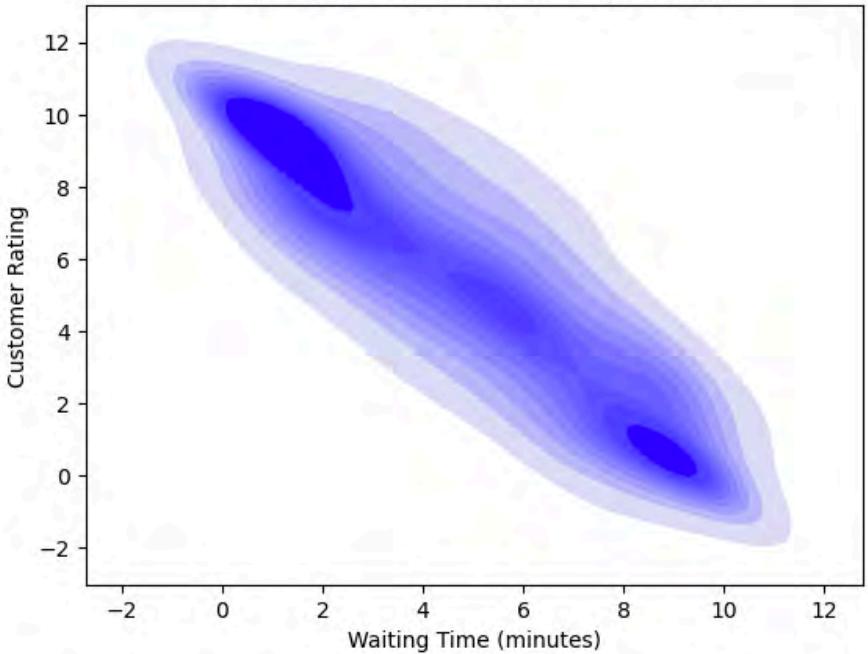


Joint Continuous Distributions: Dataset

3D Probability Density Distribution for Customer Ratings vs Waiting Time



Heatmap of Probability Density Function



Joint Continuous Distributions: Dataset

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

1000 customers



Expected Value

X variable: Waiting time (mins)

0 - 10 mins

Y variable: Satisfaction rating

0 - 10

1000 customers



Expected Value

X variable: Waiting time (mins)

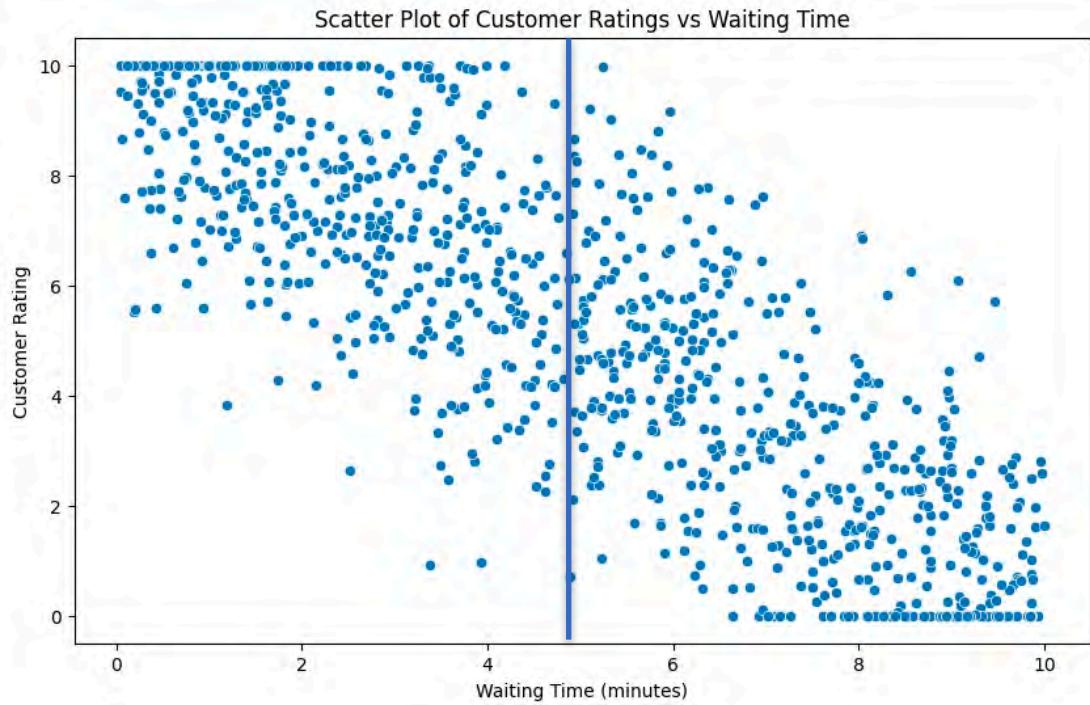
0 - 10 mins

Y variable: Satisfaction rating

0 - 10

1000 customers

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$



Expected Value

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

1000 customers

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

$$\mathbb{E}[Y] = 5.280$$



Expected Value

X variable: Waiting time (mins)
0 - 10 mins

Y variable: Satisfaction rating
0 - 10

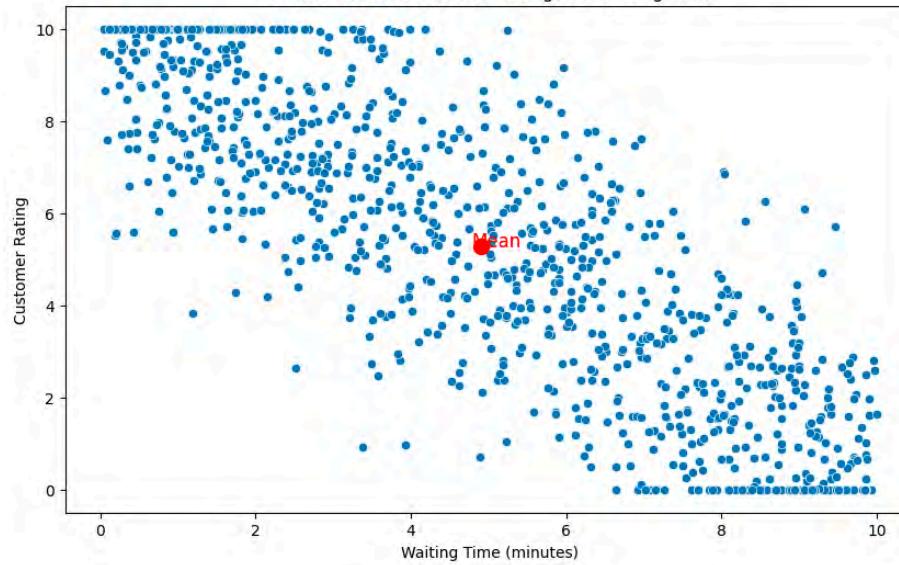
1000 customers

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

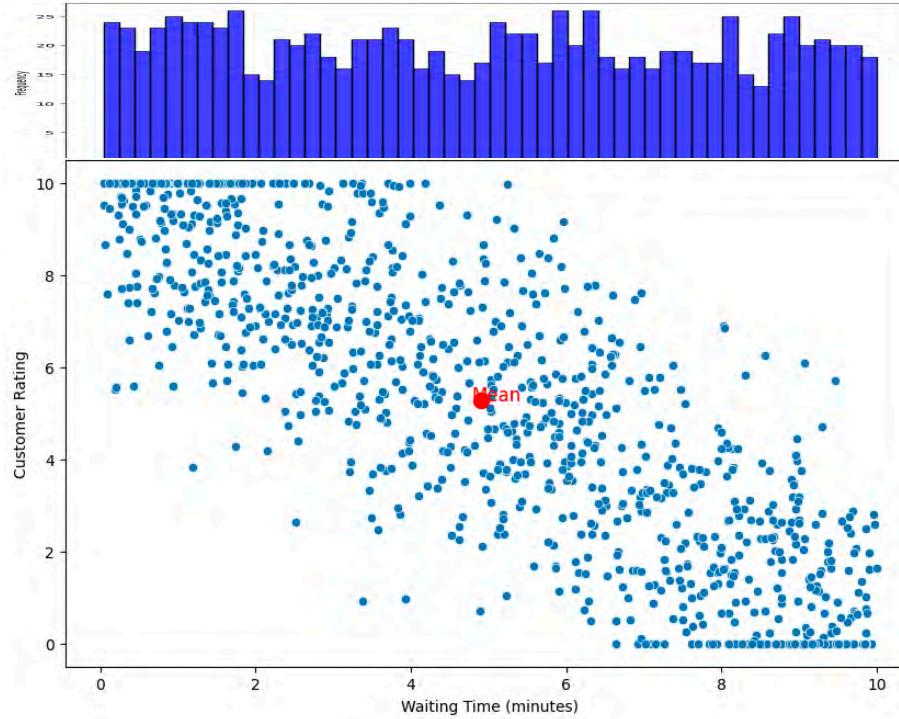
$$\mathbb{E}[Y] = 5.280$$



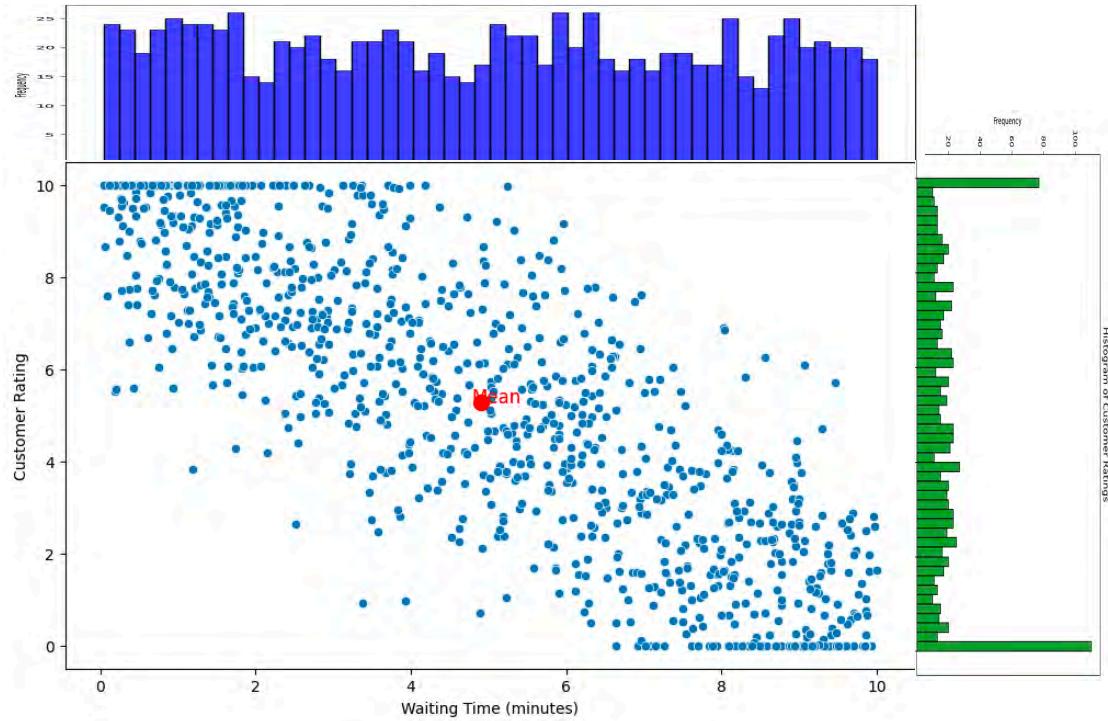
Variances



Variances

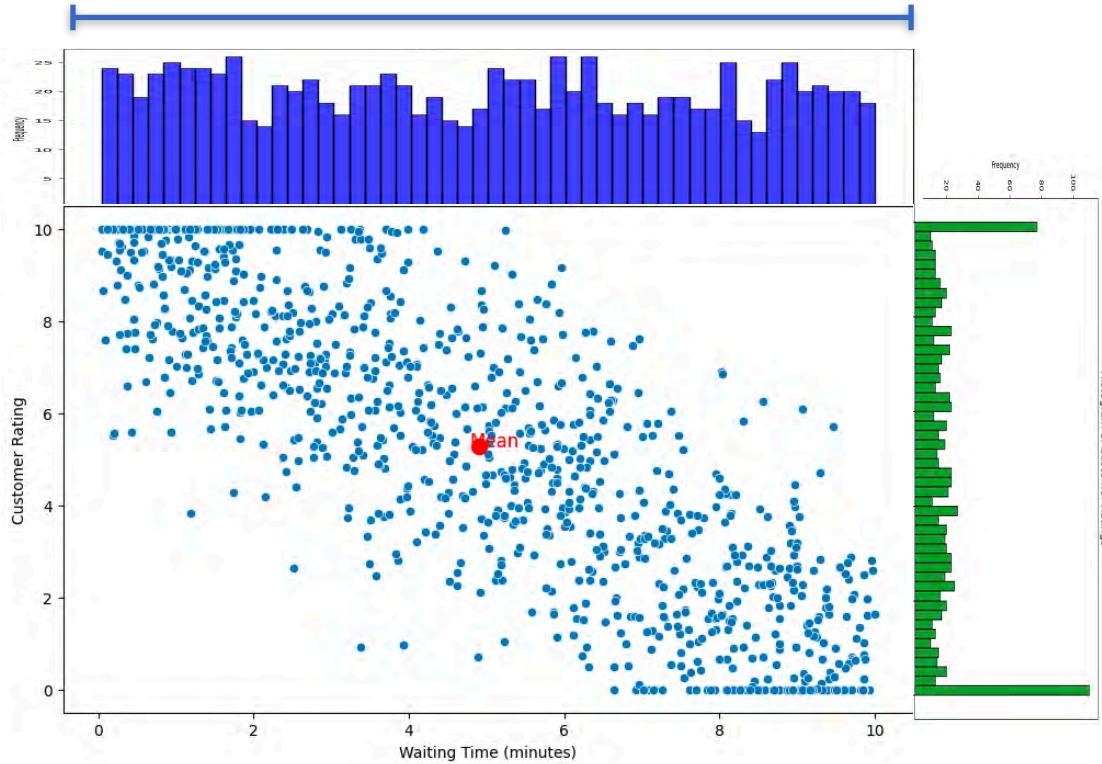


Variances



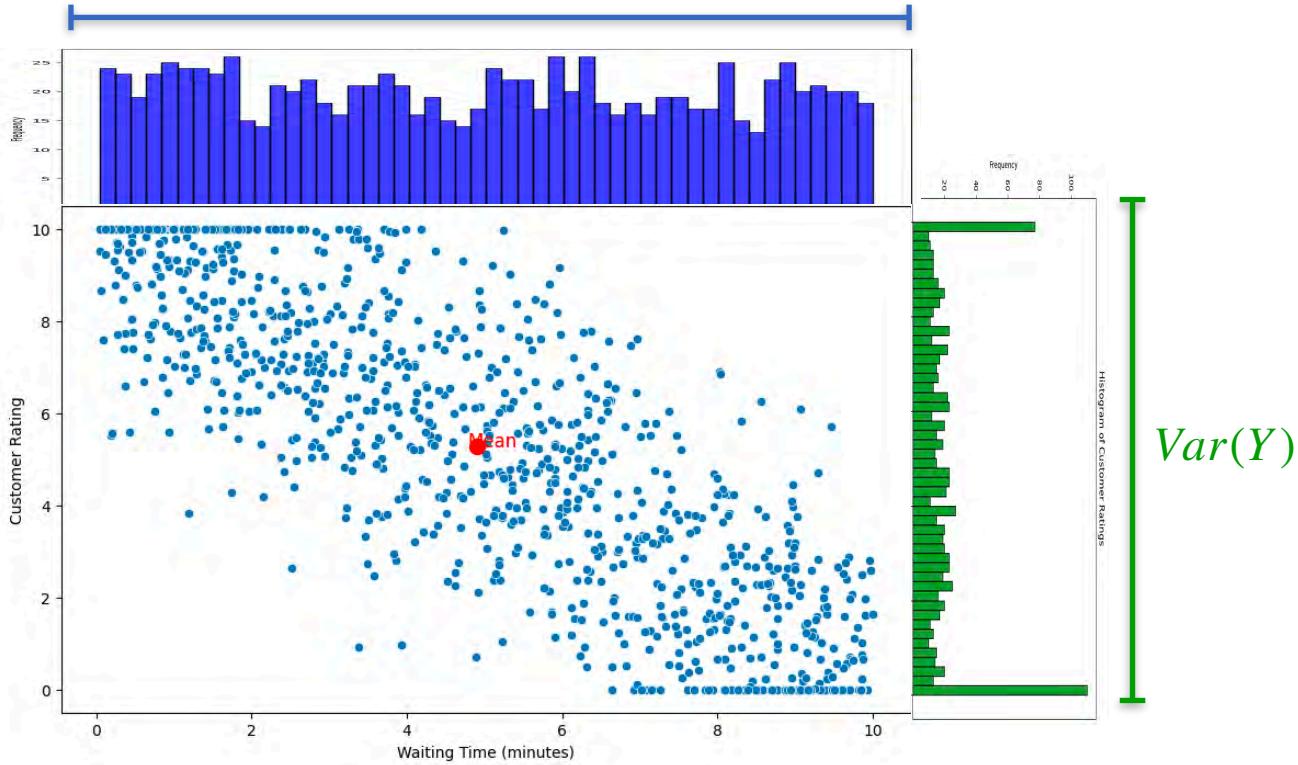
Variances

$$Var(X)$$



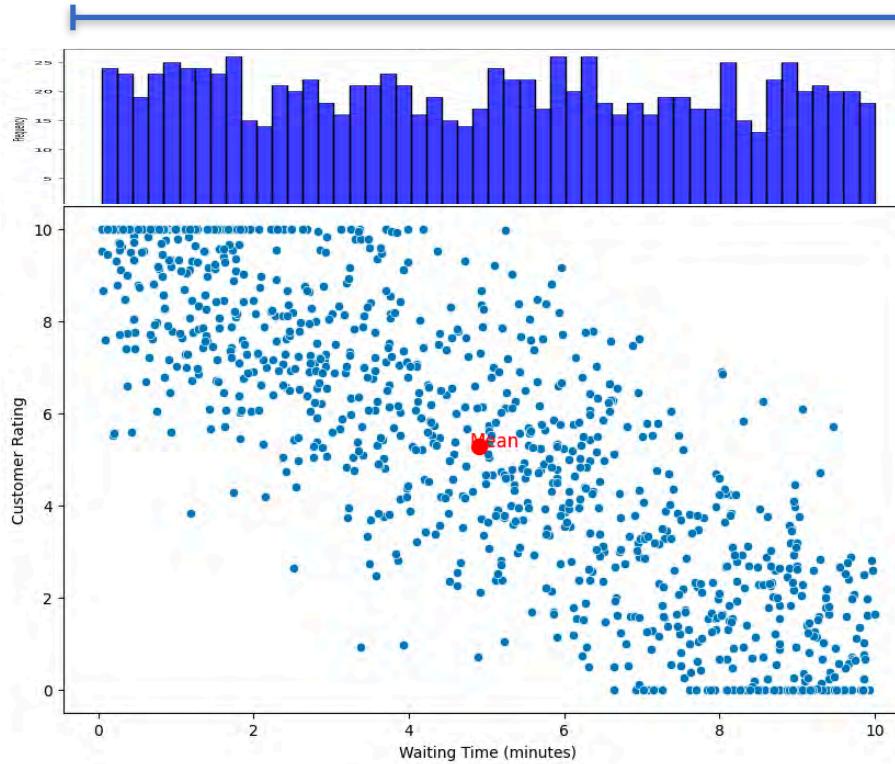
Variances

$$Var(X)$$



Variances

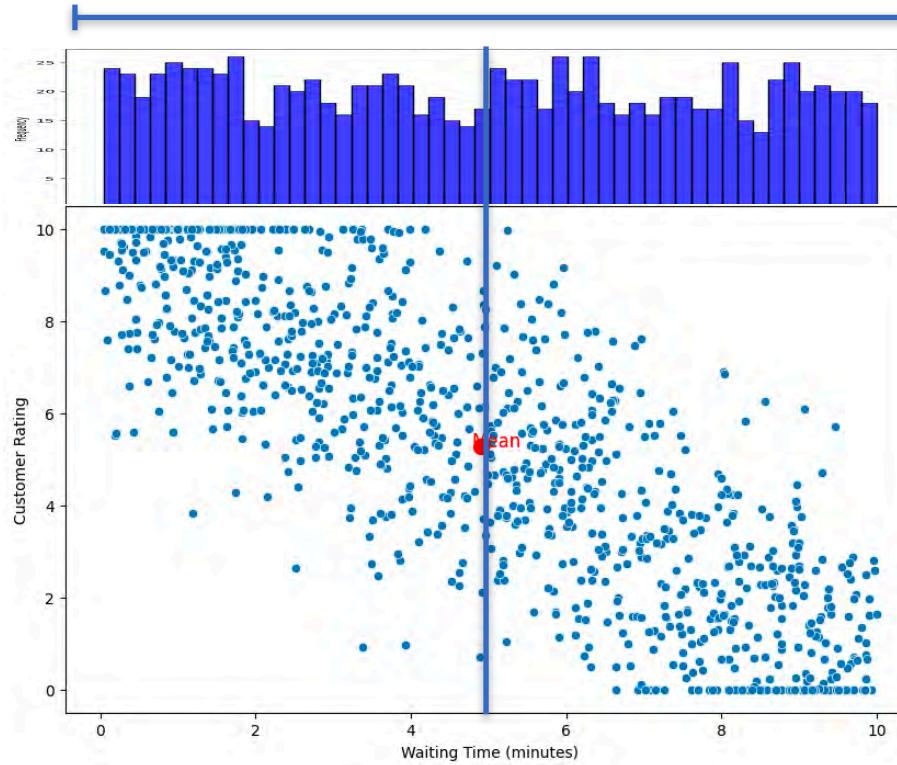
$$Var(X)$$



Variances

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

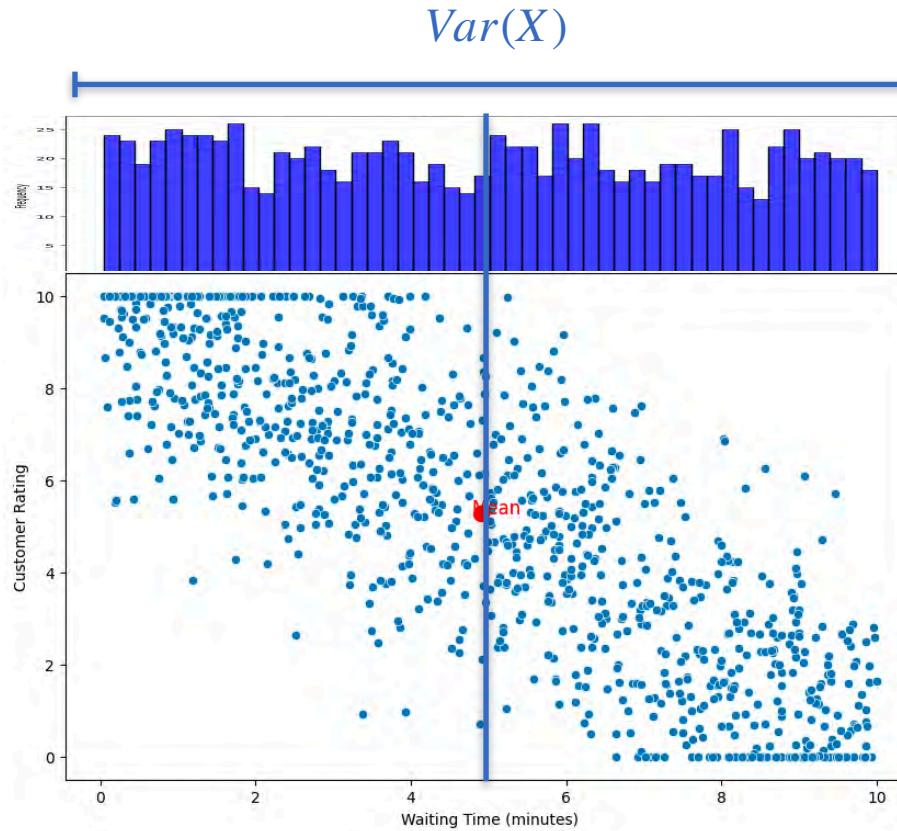
$$Var(X)$$



Variances

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

$$\mathbb{E}[X^2] = 32.561$$

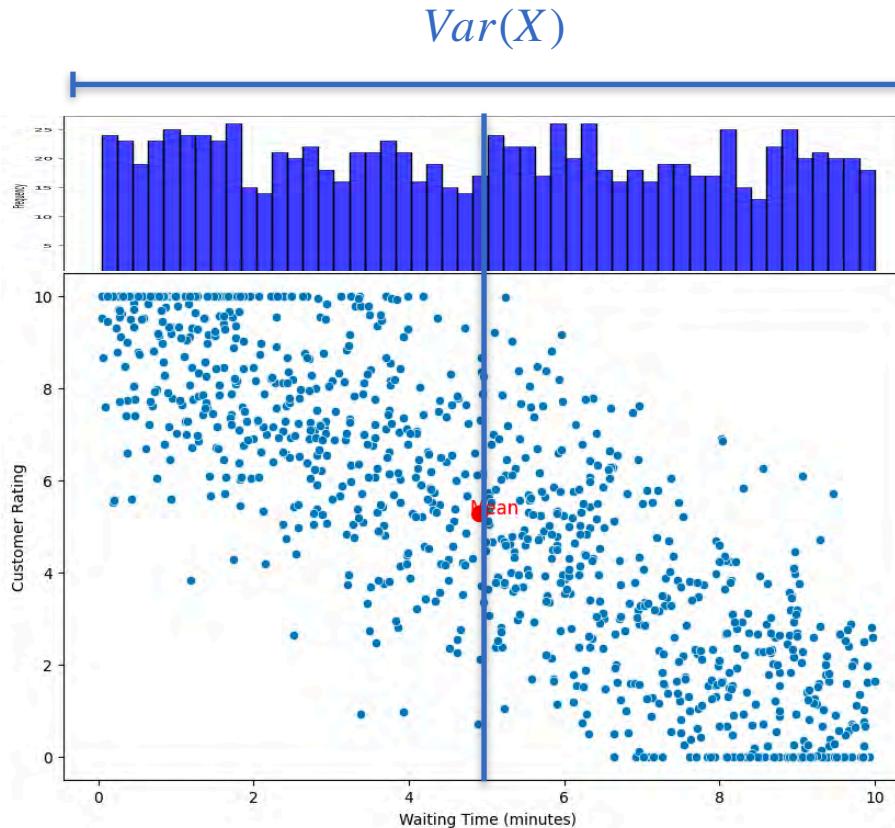


Variances

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

$$\mathbb{E}[X^2] = 32.561$$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

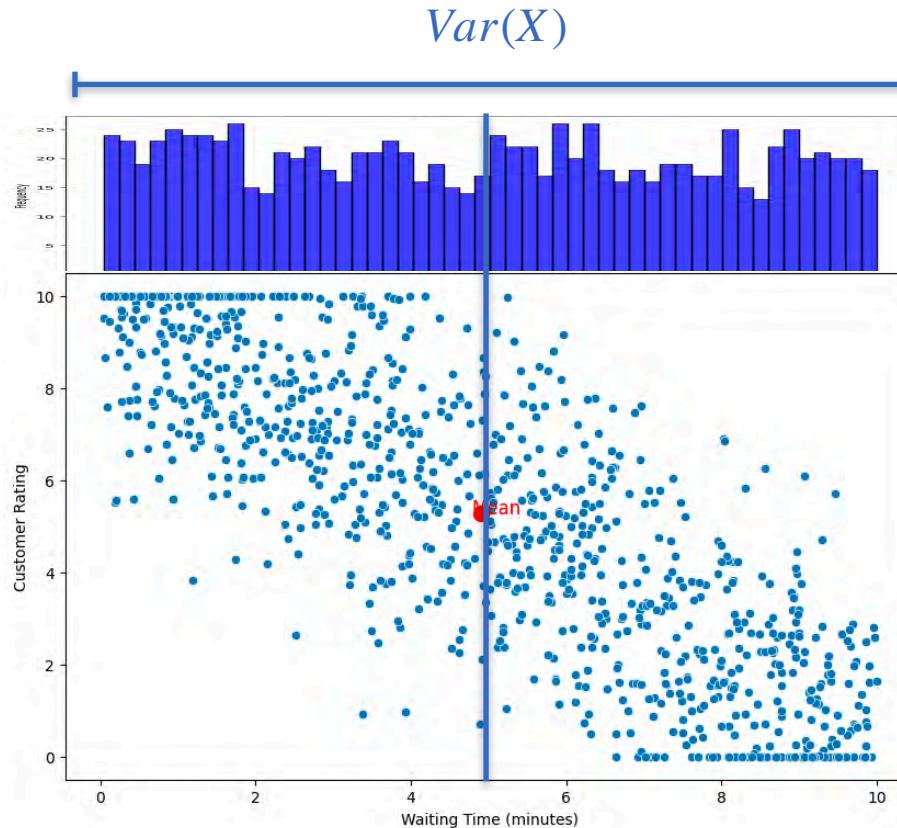


Variances

$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

$$\mathbb{E}[X^2] = 32.561$$

$$\begin{aligned}Var(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\&= 32.561 - 4.903^2\end{aligned}$$

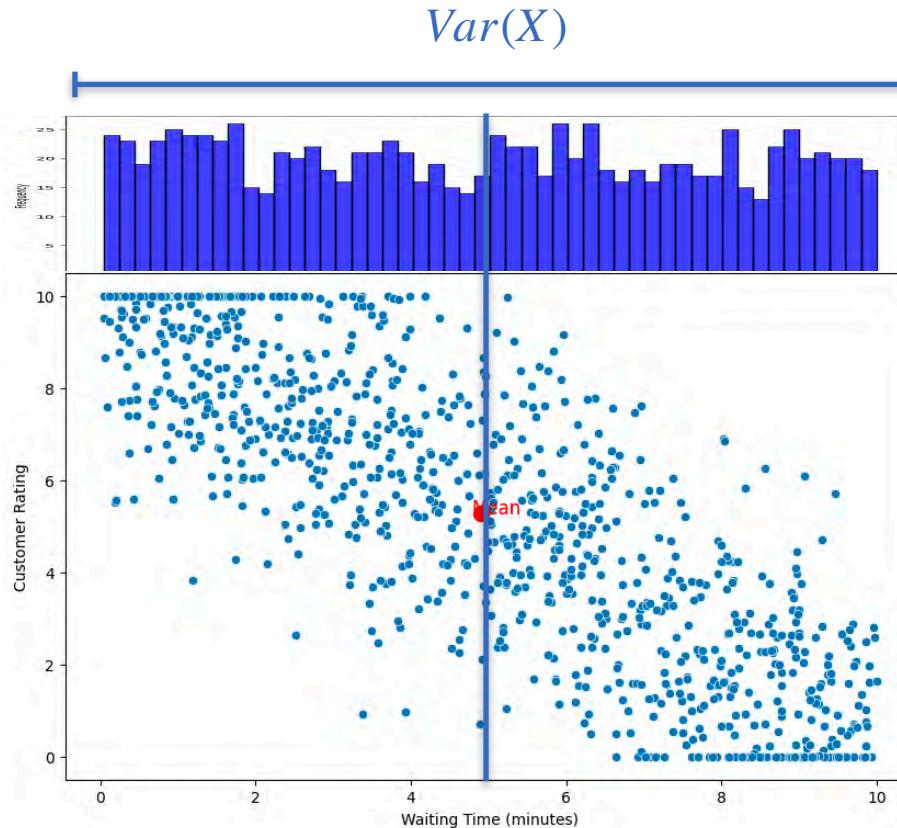


Variances

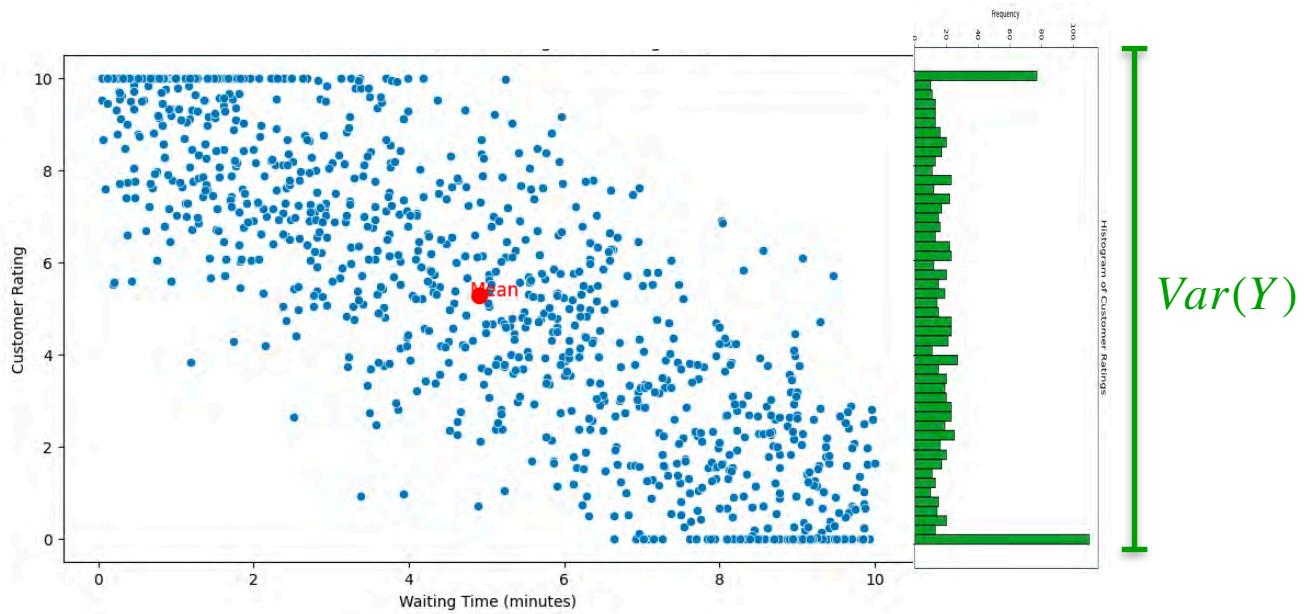
$$\mathbb{E}[X] = 4.903 \text{ minutes}$$

$$\mathbb{E}[X^2] = 32.561$$

$$\begin{aligned}Var(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\&= 32.561 - 4.903^2 \\&= 8.526\end{aligned}$$

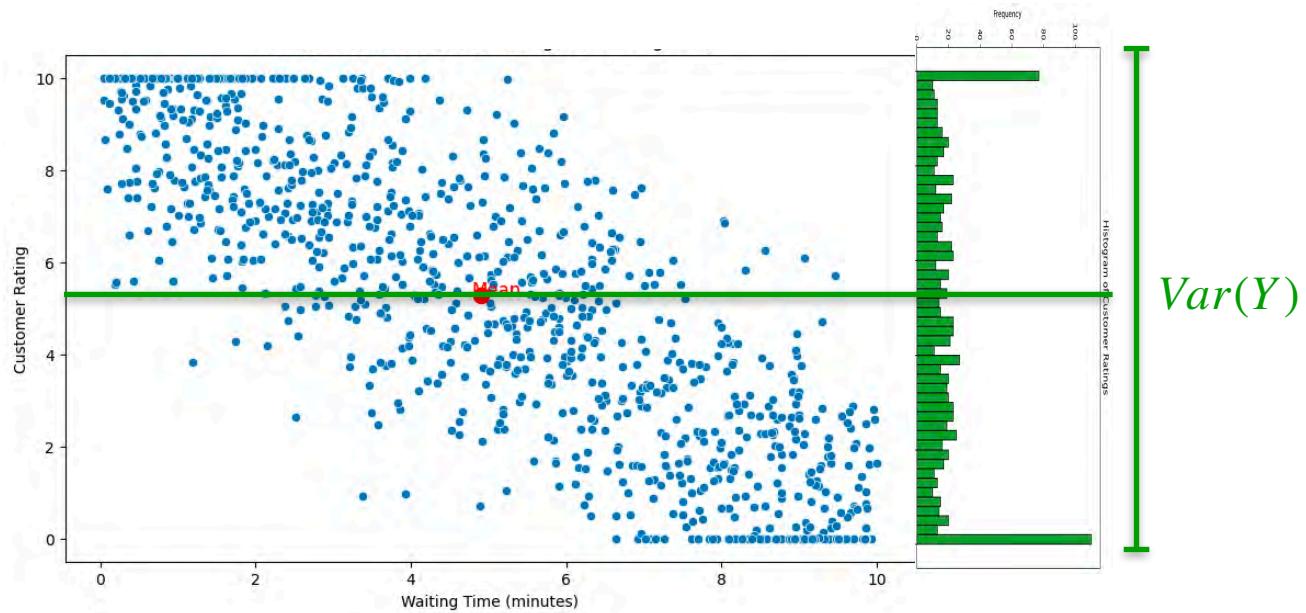


Variances



Variances

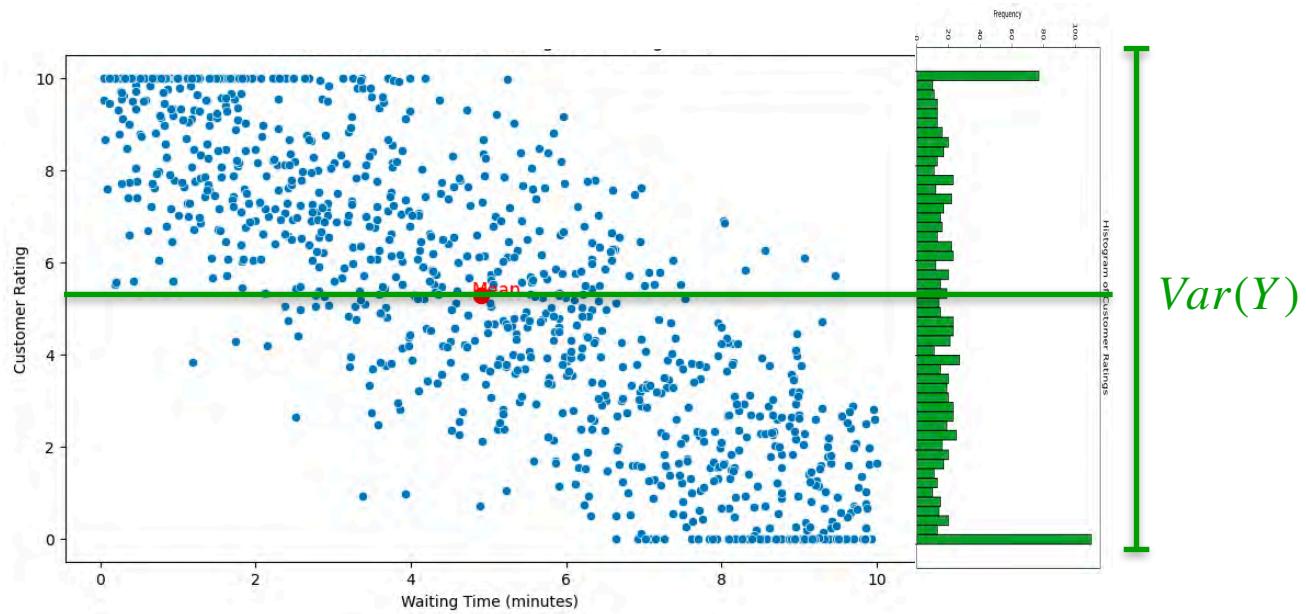
$$\mathbb{E}[Y] = 5.280$$



Variances

$$\mathbb{E}[Y] = 5.280$$

$$\mathbb{E}[Y^2] = 38.037$$

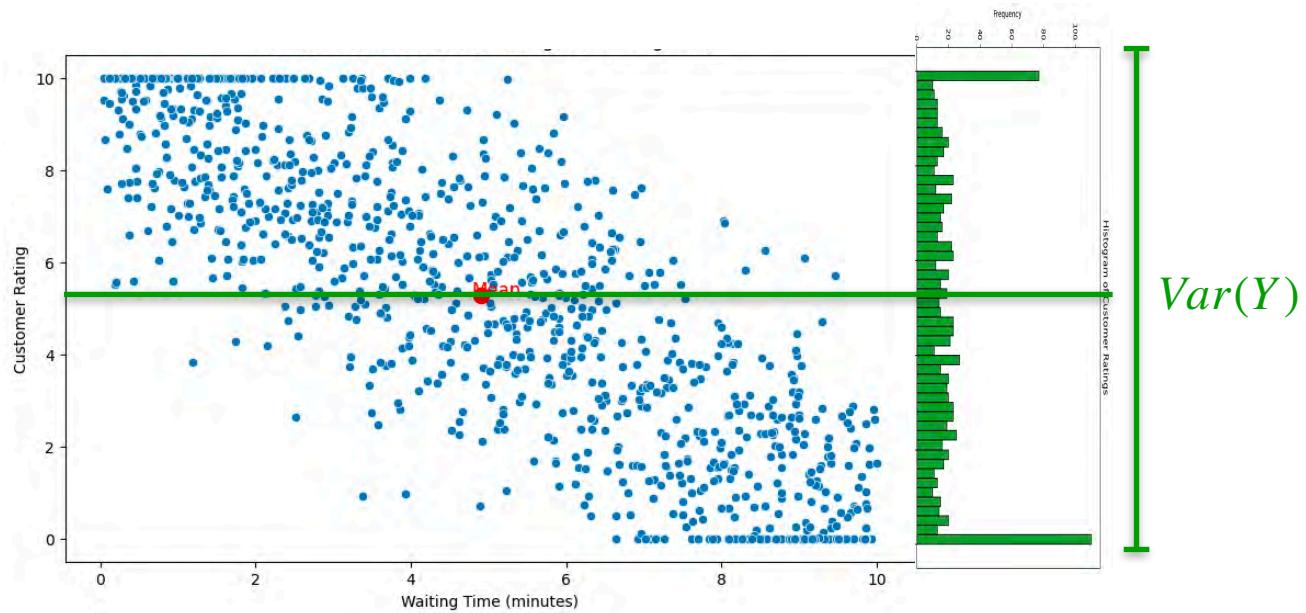


Variances

$$\mathbb{E}[Y] = 5.280$$

$$\mathbb{E}[Y^2] = 38.037$$

$$Var(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$$

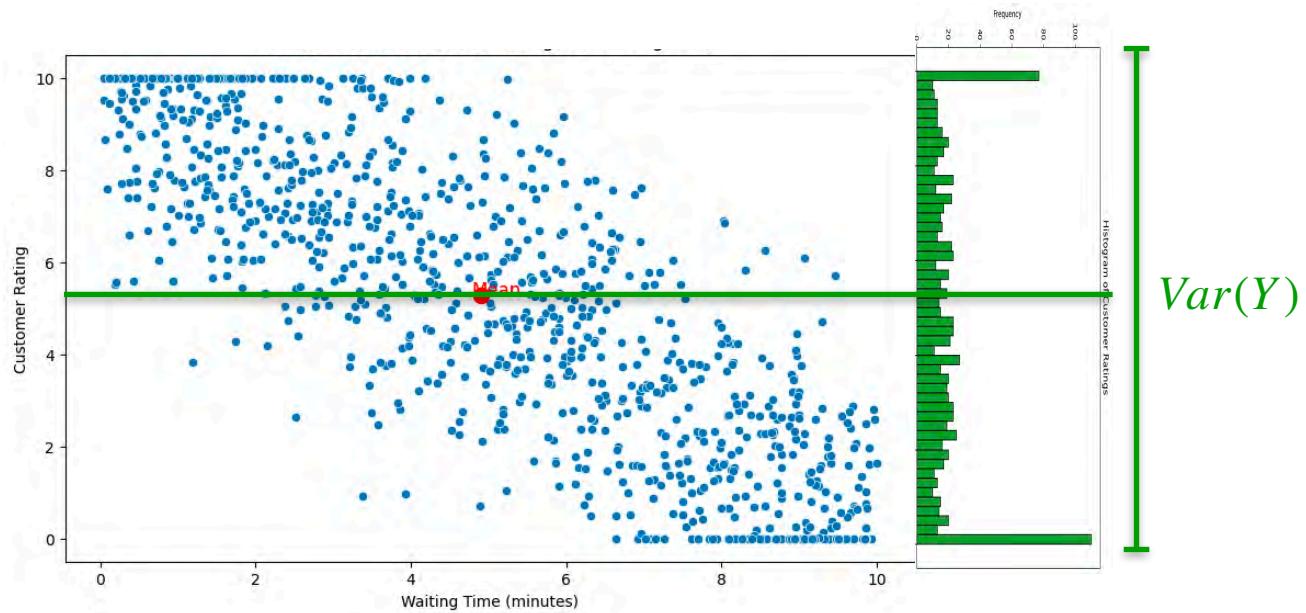


Variances

$$\mathbb{E}[Y] = 5.280$$

$$\mathbb{E}[Y^2] = 38.037$$

$$\begin{aligned} \text{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &= 38.037 - 5.280^2 \end{aligned}$$

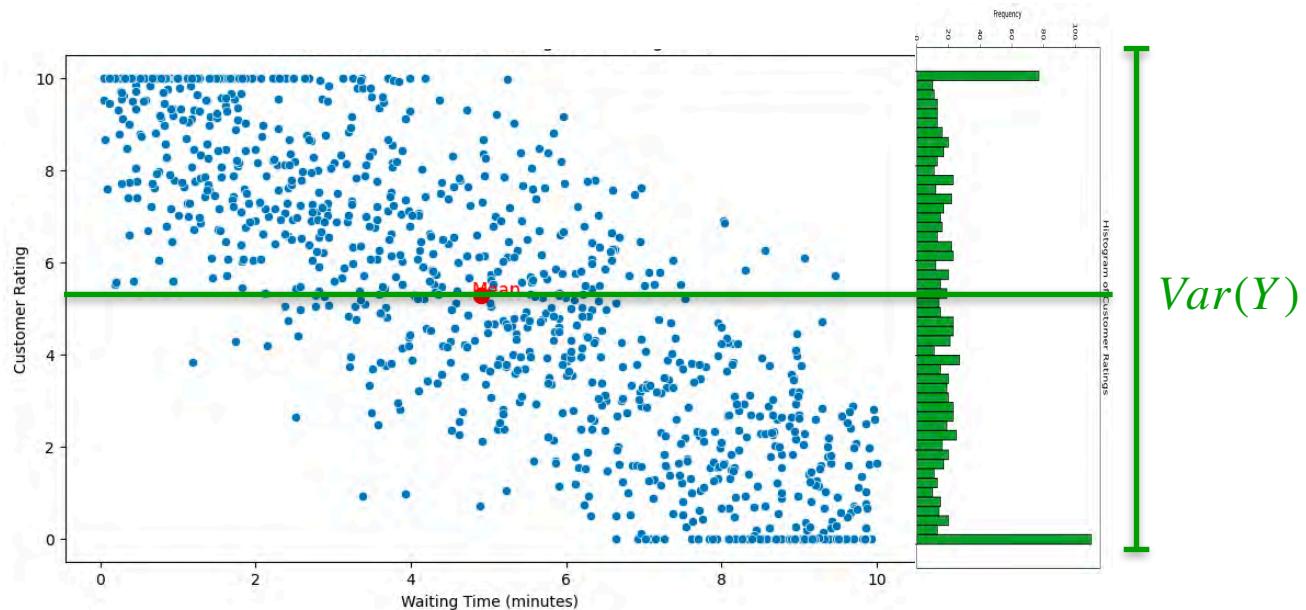


Variances

$$\mathbb{E}[Y] = 5.280$$

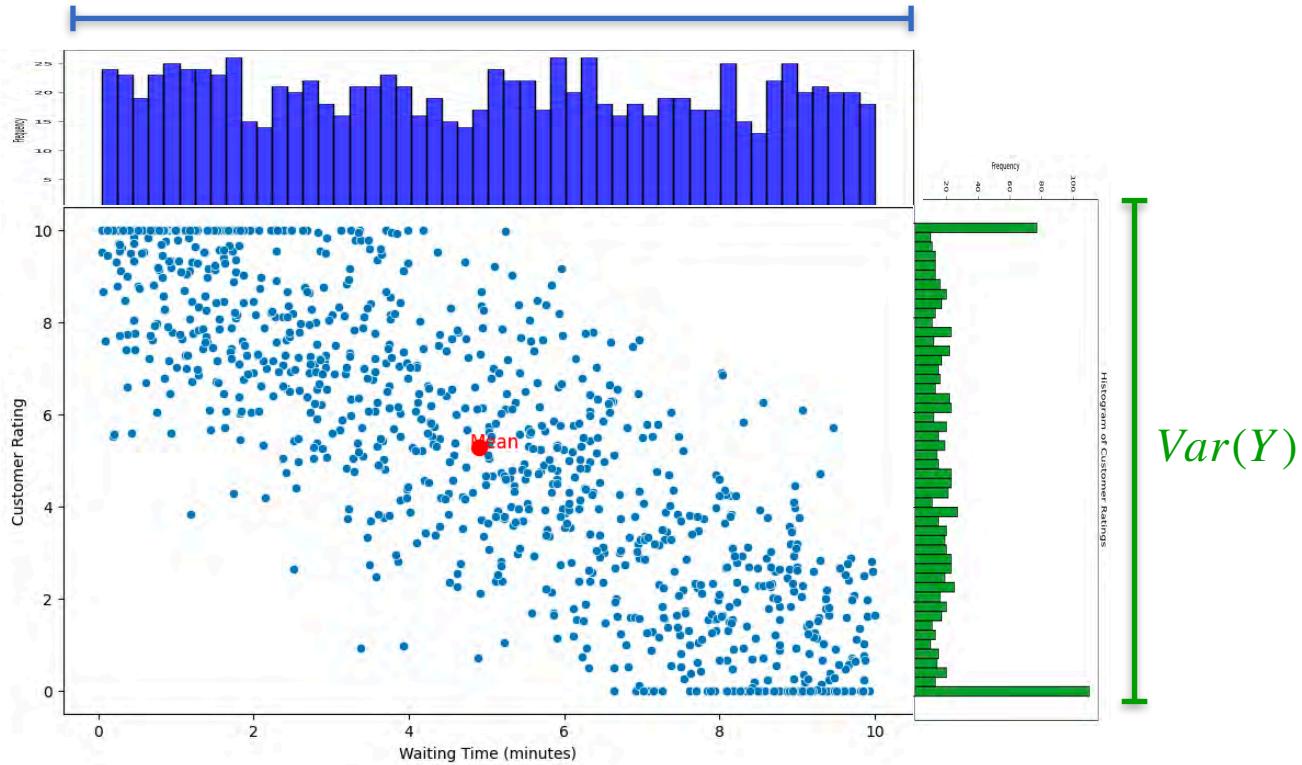
$$\mathbb{E}[Y^2] = 38.037$$

$$\begin{aligned} \text{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &= 38.037 - 5.280^2 \\ &= 10.163 \end{aligned}$$



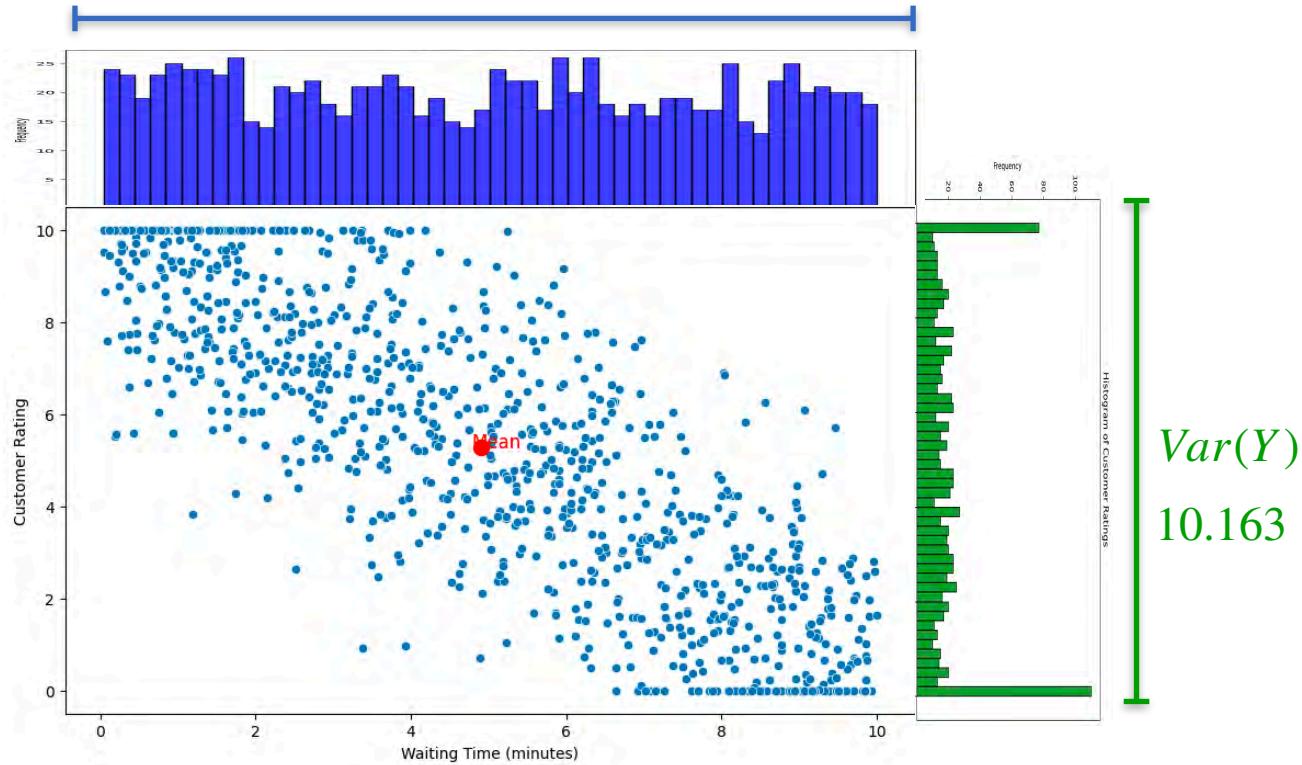
Variances

$$Var(X)$$



Variances

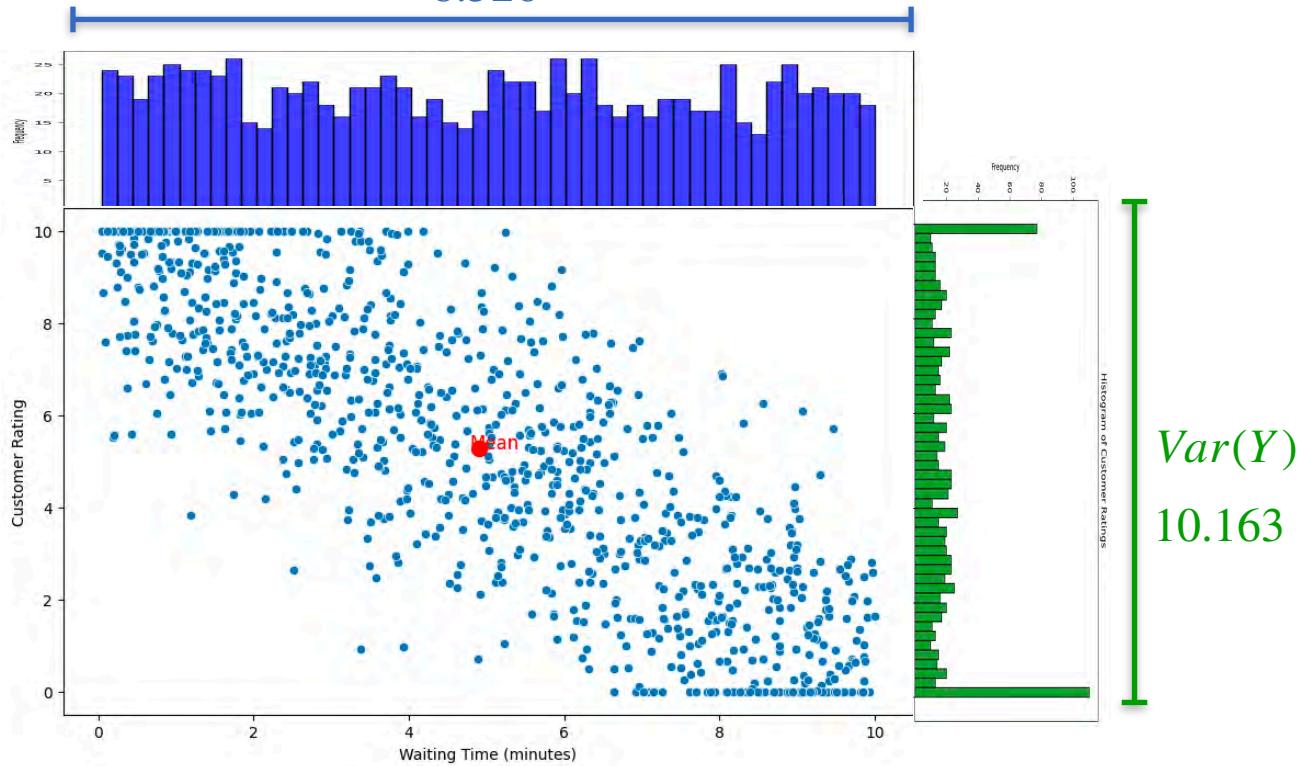
$$Var(X)$$



Variances

$$Var(X)$$

8.526





DeepLearning.AI

Probability Distributions with Multiple Variables

Marginal and Conditional Distribution

Marginal Distribution: Example 1

Marginal Distribution: Example 1



Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution

Distribution of one variable while ignoring the other

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



To find the marginal distribution of height:

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



To find the marginal distribution of height:

sum the joint probability distribution over all values of age

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

$$p_Y(50) =$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

$$p_Y(50) = \sum_i p_{XY}(x_i, 50)$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j)$$

$$p_Y(50) = \sum_i p_{XY}(x_i, 50)$$

$$p_Y(50) = \frac{2}{10}$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10



Marginal Distribution: Example 1

		Height (Y)						Age (years): 7 7 7 8 8 9 9 9 9 10									
		45	46	47	48	49	50	Height (in): 45 46 46 47 47 49 49 49 49 50 50									
Age (X)	7	1/10	2/10	0	0	0	0	3/10									
	8	0	0	2/10	0	0	0	2/10									
	9	0	0	0	0	3/10	1/10	4/10									
	10	0	0	0	0	0	1/10	1/10									
	1/10		2/10	2/10	0	3/10	2/10										

$$p_X(x_i) = \sum_j p_{XY}(x_i, y_j)$$

Marginal Distribution: Example 1

		Height (Y)						Age (years): 7 7 7 8 8 9 9 9 9 10									
		45	46	47	48	49	50	Height (in): 45 46 46 47 47 49 49 49 49 50 50									
Age (X)	7	1/10	2/10	0	0	0	0	3/10									
	8	0	0	2/10	0	0	0	2/10									
	9	0	0	0	0	3/10	1/10	4/10									
	10	0	0	0	0	0	1/10	1/10									
	1/10		2/10	2/10	0	3/10	2/10										

$$p_X(x_i) = \sum_j p_{XY}(x_i, y_j)$$

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10

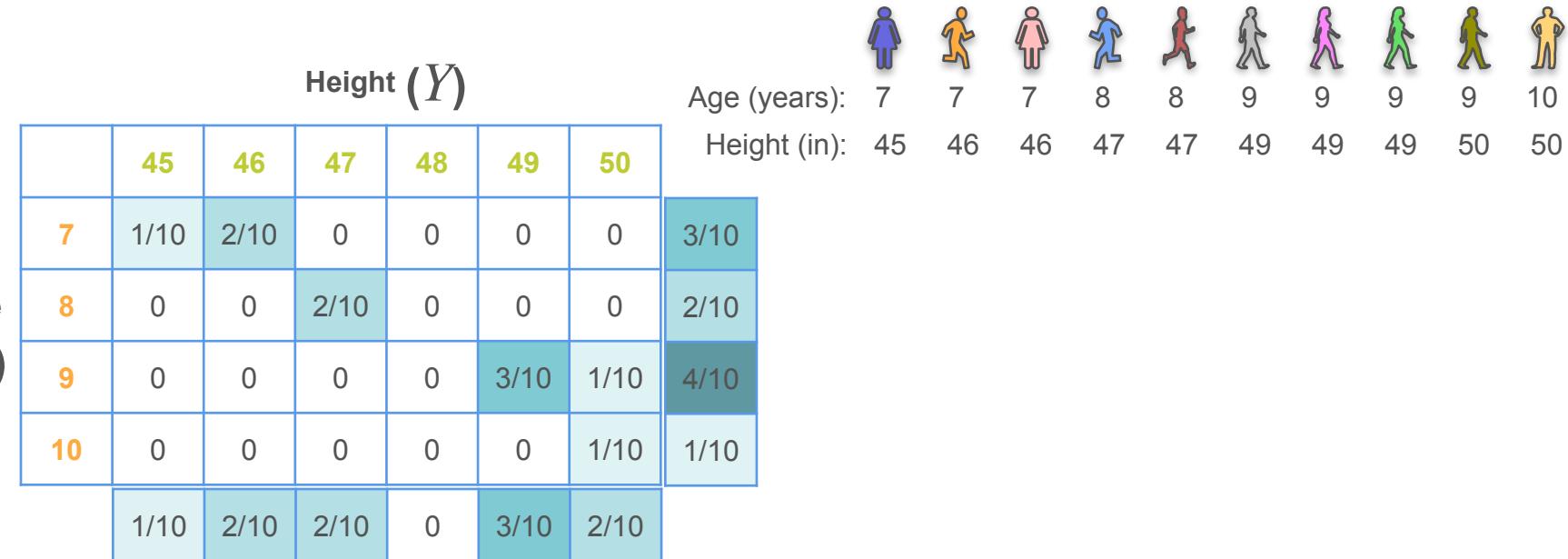


$$p_X(x_i) = \sum_j p_{XY}(x_i, y_j)$$

$$p_X(7) = \sum_j p_{XY}(7, y_j)$$

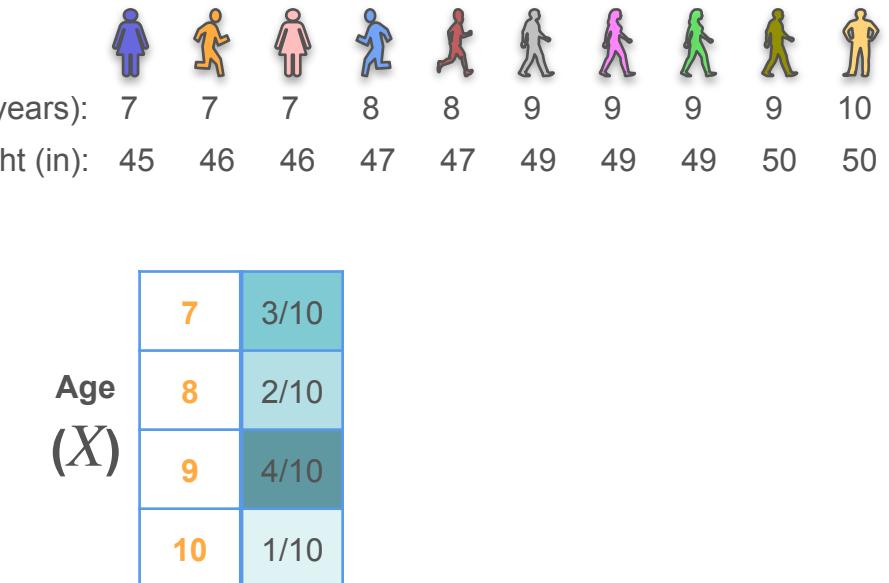
$$p_X(7) = \frac{3}{10}$$

Marginal Distribution: Example 1



Marginal Distribution: Example 1

		Height (Y)						Age (years): 7 7 7 8 8 9 9 9 9 10									
		45	46	47	48	49	50	Height (in): 45 46 46 47 47 47 49 49 49 49 50 50									
Age (X)	7	1/10	2/10	0	0	0	0	3/10									
	8	0	0	2/10	0	0	0	2/10									
	9	0	0	0	0	3/10	1/10	4/10									
	10	0	0	0	0	0	1/10	1/10									
	1/10		2/10	2/10	0	3/10	2/10										

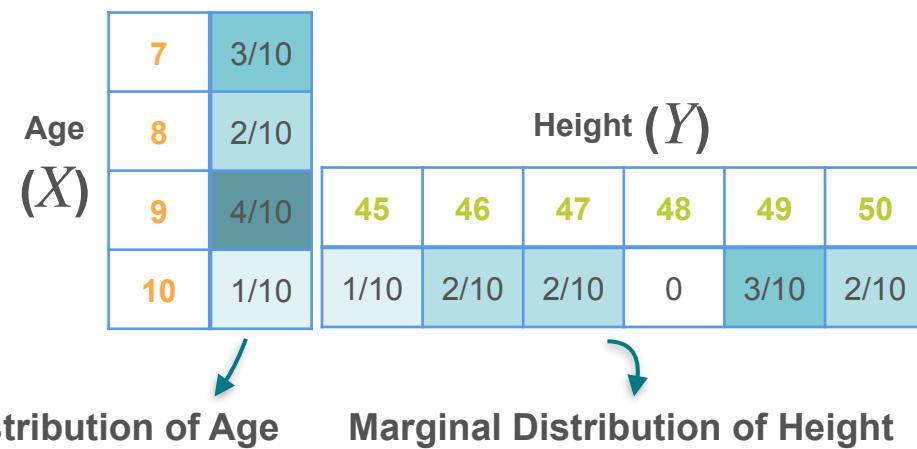


Marginal Distribution of Age

Marginal Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10

Age (years):	7	7	7	8	8	9	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	49	50	50



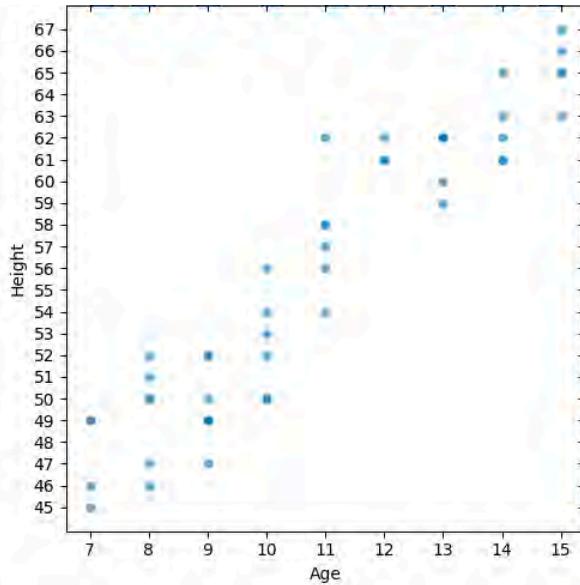
Marginal Distribution: Example 1

Marginal Distribution: Example 1

Age and Height Dataset
for 50 children

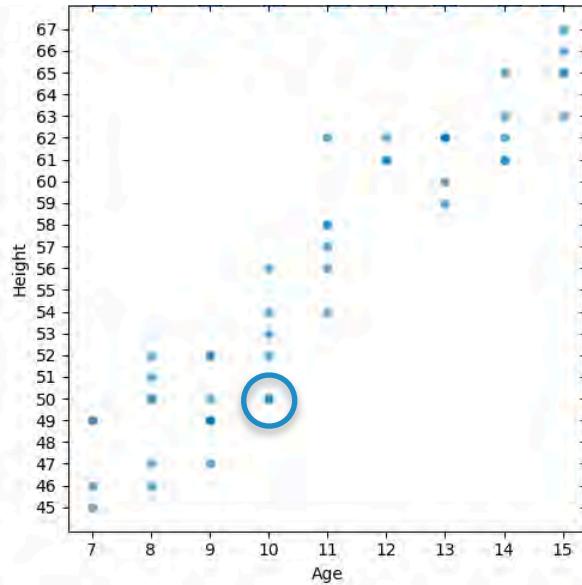
Marginal Distribution: Example 1

Age and Height Dataset
for 50 children



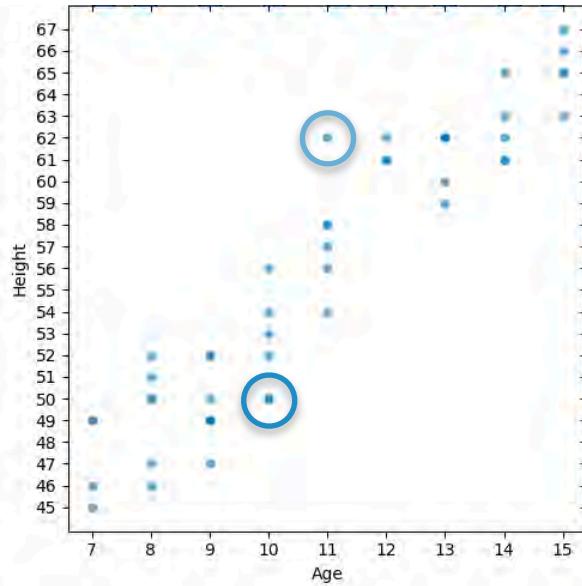
Marginal Distribution: Example 1

Age and Height Dataset
for 50 children



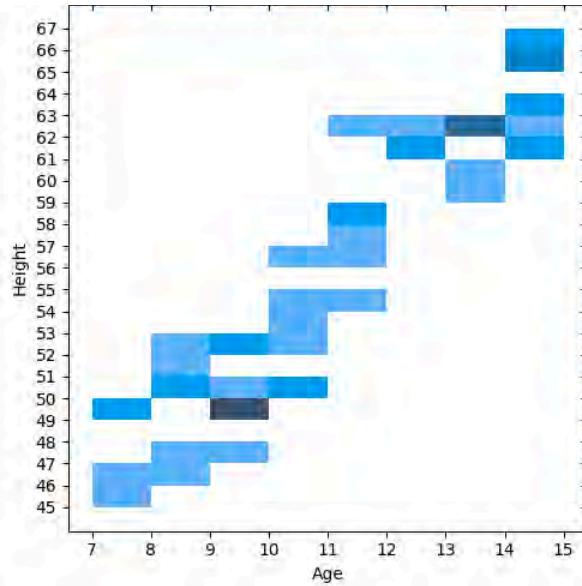
Marginal Distribution: Example 1

Age and Height Dataset
for 50 children



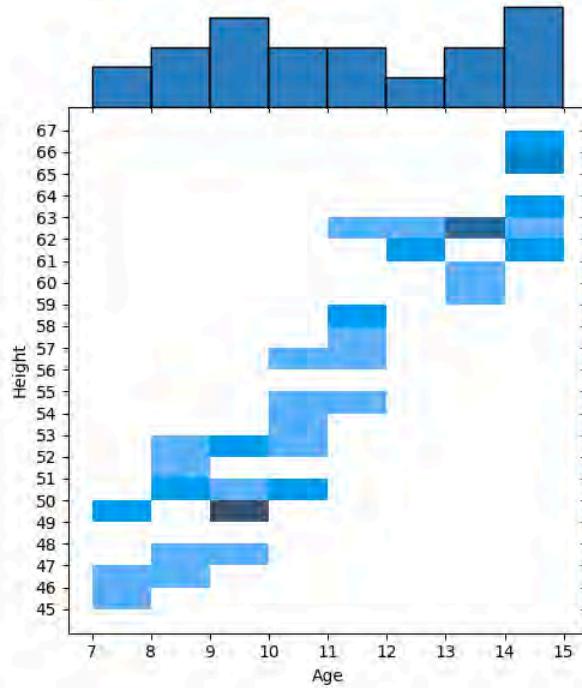
Marginal Distribution: Example 1

Age and Height Dataset
for 50 children



Marginal Distribution: Example 1

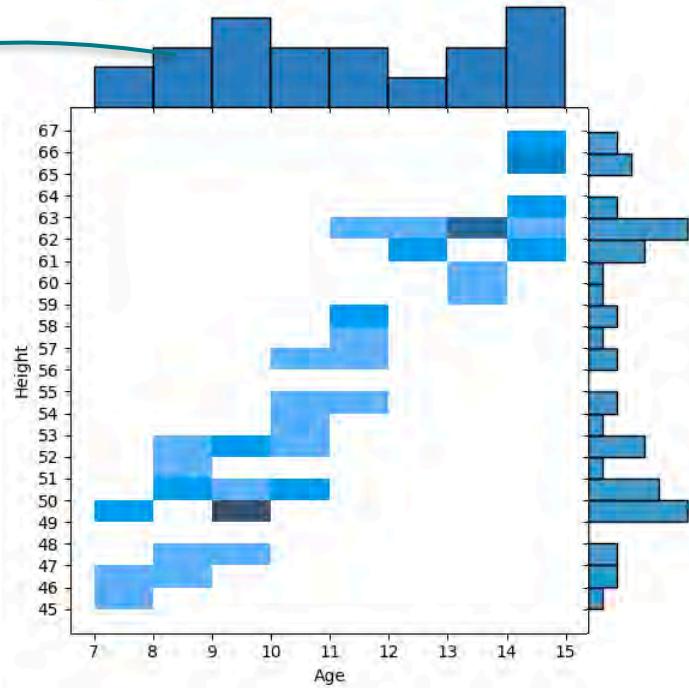
Age and Height Dataset
for 50 children



Marginal Distribution: Example 1

Marginal Distribution of Age

Age and Height Dataset
for 50 children



Marginal Distribution of Height

Marginal Distributions: Example 2

X : the number rolled on the 1st dice



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

Y : the number rolled on the 2nd dice



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

Marginal Distributions: Example 2

		Y						
		1	2	3	4	5	6	
X		1	1/36	1/36	1/36	1/36	1/36	1/36
2		1/36	1/36	1/36	1/36	1/36	1/36	
3		1/36	1/36	1/36	1/36	1/36	1/36	
4		1/36	1/36	1/36	1/36	1/36	1/36	
5		1/36	1/36	1/36	1/36	1/36	1/36	
6		1/36	1/36	1/36	1/36	1/36	1/36	

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

Marginal Distributions: Example 2

Y

X

	1	2	3	4	5	6	
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

Marginal Distributions: Example 2

		Y												
		1	2	3	4	5	6							
X		1	1/36	1/36	1/36	1/36	1/36	1/36						
1		1	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
2		2	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
3		3	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
4		4	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
5		5	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
6		6	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
		1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6

X : the number rolled on the 1st dice
 Y : the number rolled on the 2nd dice

Marginal Distributions: Example 2

		Y												
		1	2	3	4	5	6							
X		1	1/36	1/36	1/36	1/36	1/36	1/36						
1		1	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
2		2	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
3		3	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
4		4	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
5		5	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
6		6	1/36	1/36	1/36	1/36	1/36	1/36	1/6	1/6	1/6	1/6	1/6	1/6
		1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6	1/6

X : the number rolled on the 1st dice

Y : the number rolled on the 2nd dice

$$p_X(x_i) = \frac{1}{6}$$

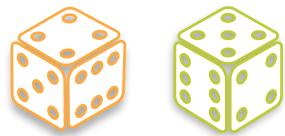
Marginal Distributions: Example 2

		Y							
		1	2	3	4	5	6		
X		1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
1		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6		1/36	1/36	1/36	1/36	1/36	1/36	1/36	1/6
		1/6	1/6	1/6	1/6	1/6	1/6		

X : the number rolled on the 1st dice
 Y : the number rolled on the 2nd dice

$$p_X(x_i) = \frac{1}{6}$$
$$p_Y(y_j) = \frac{1}{6}$$

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

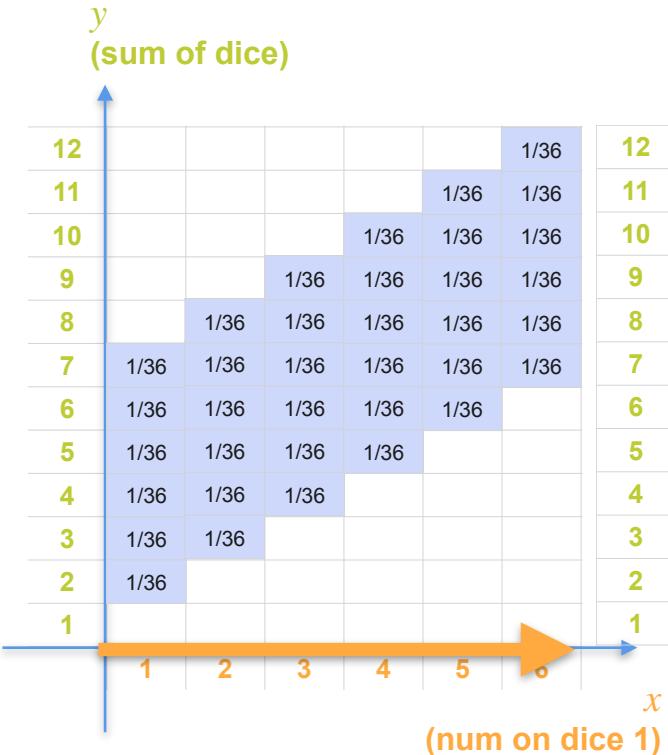
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

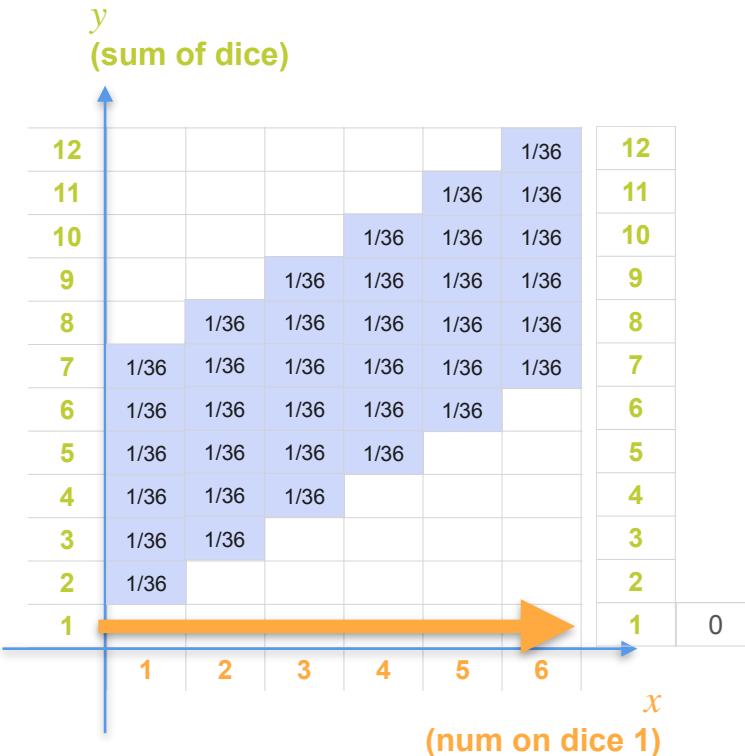
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

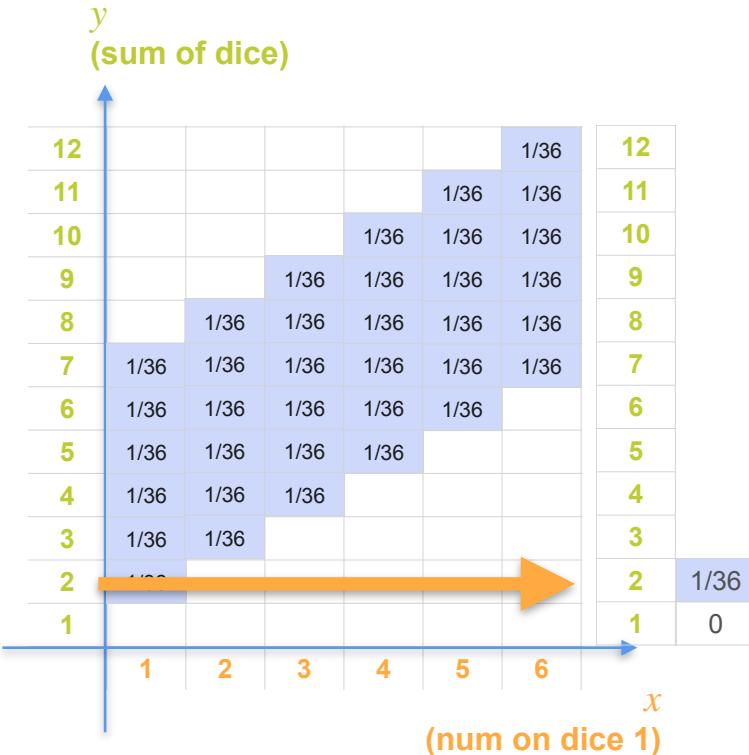
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

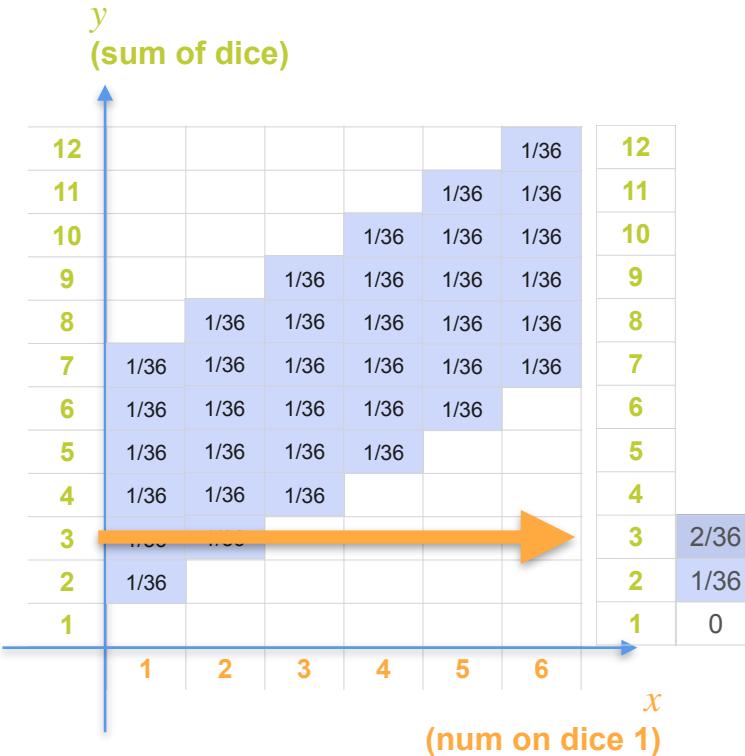
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

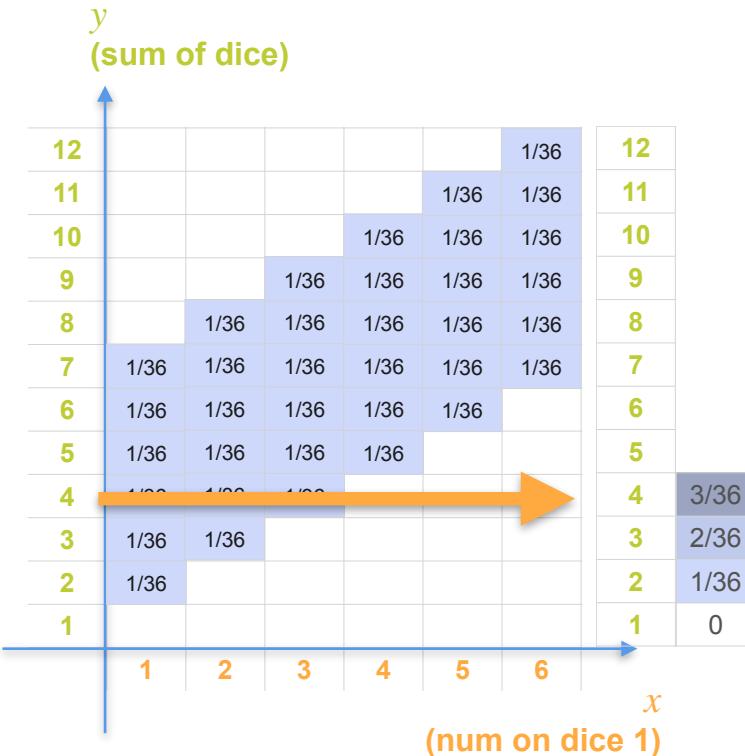
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

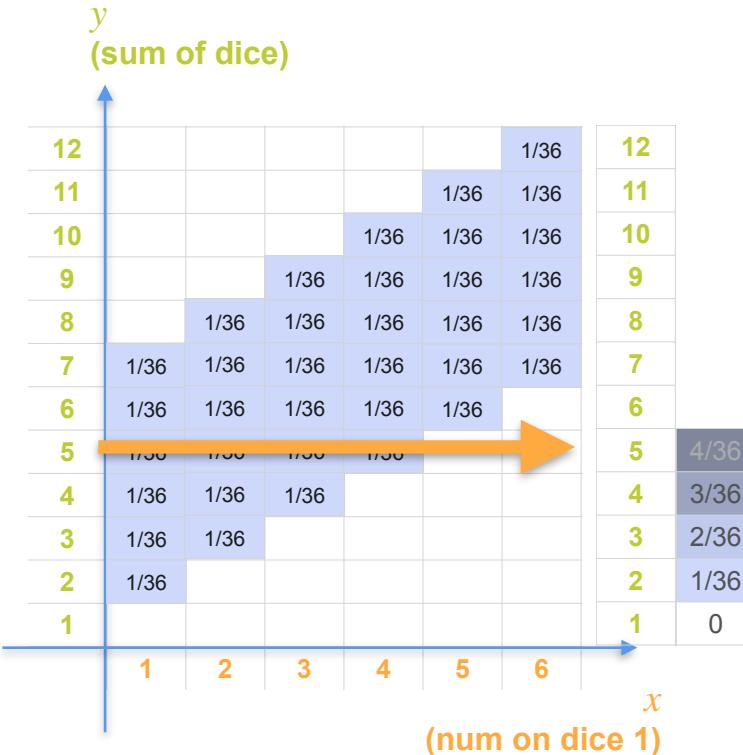
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

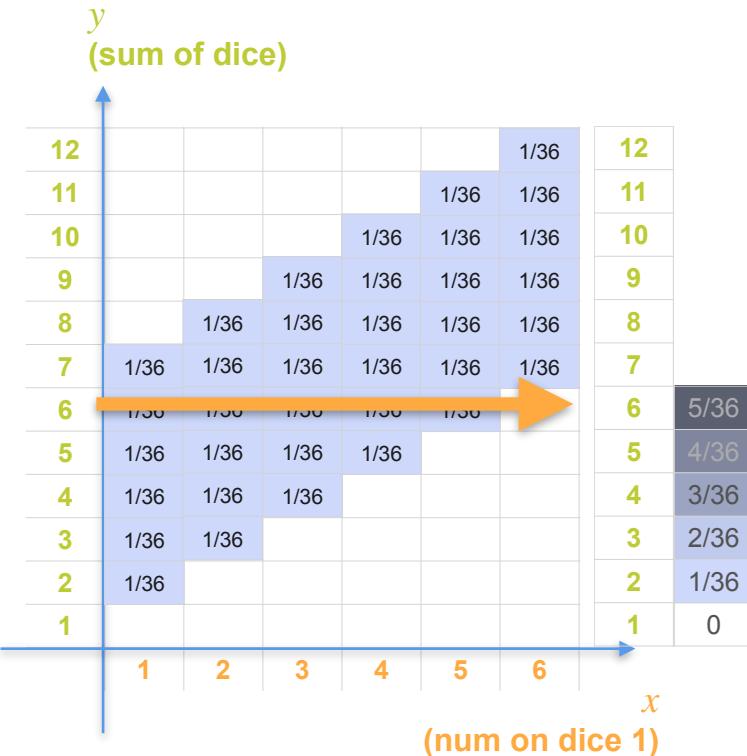
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

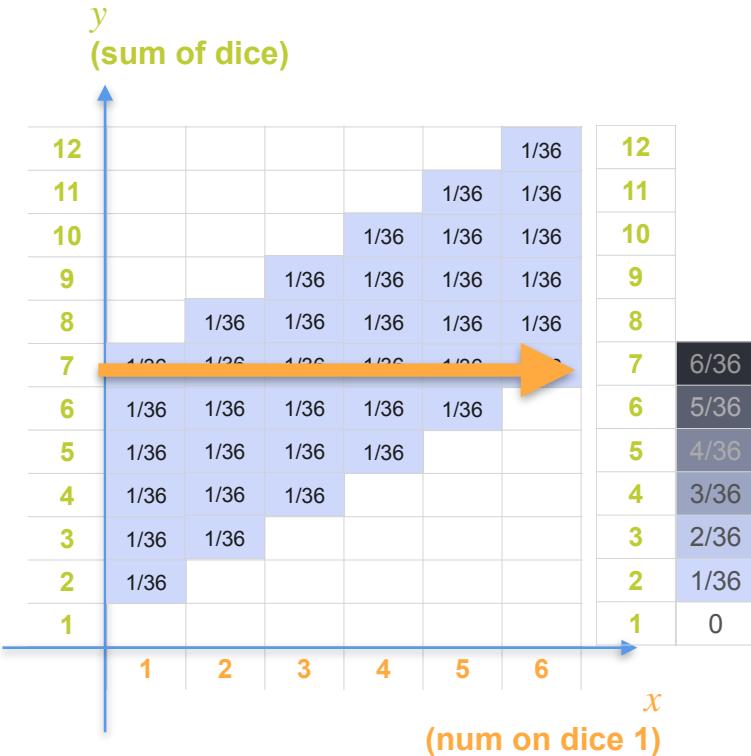
Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j) = ?$$

Marginal Distributions: Example 3



X : the number rolled on the 1st dice

Y : sum of the two dice

$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j) = ?$$

$$p_Y(10) = ?$$

Marginal Distributions: Example 3



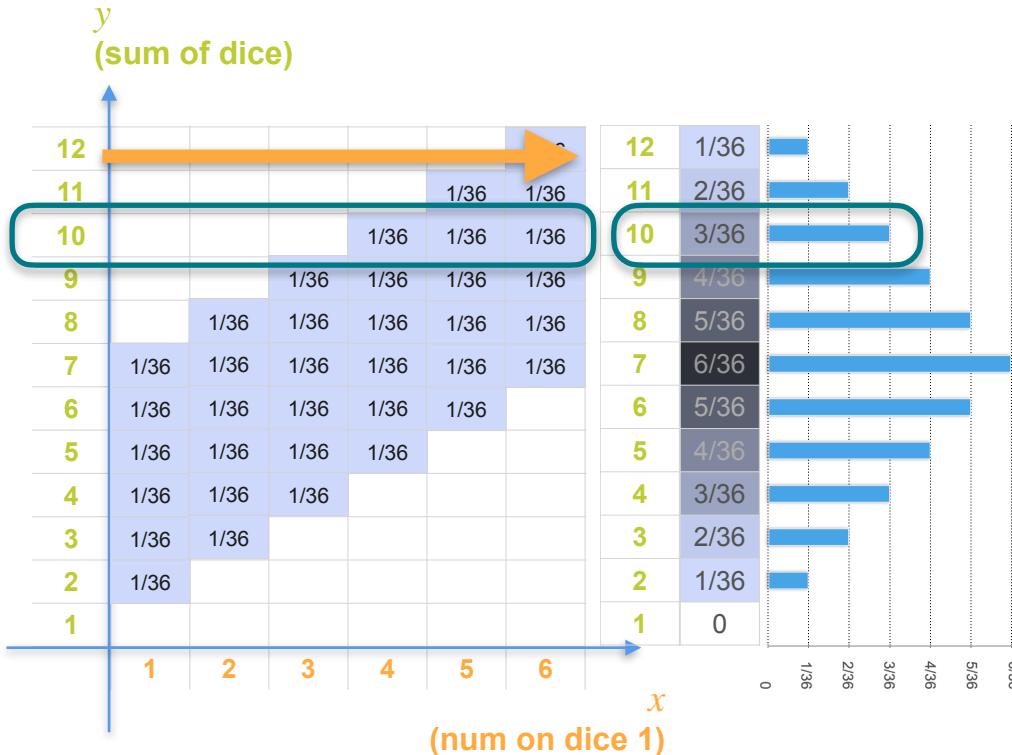
X : the number rolled on the 1st dice

Y : sum of the two dice

$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j) = ?$$

$$p_Y(10) = ?$$

Marginal Distributions: Example 3



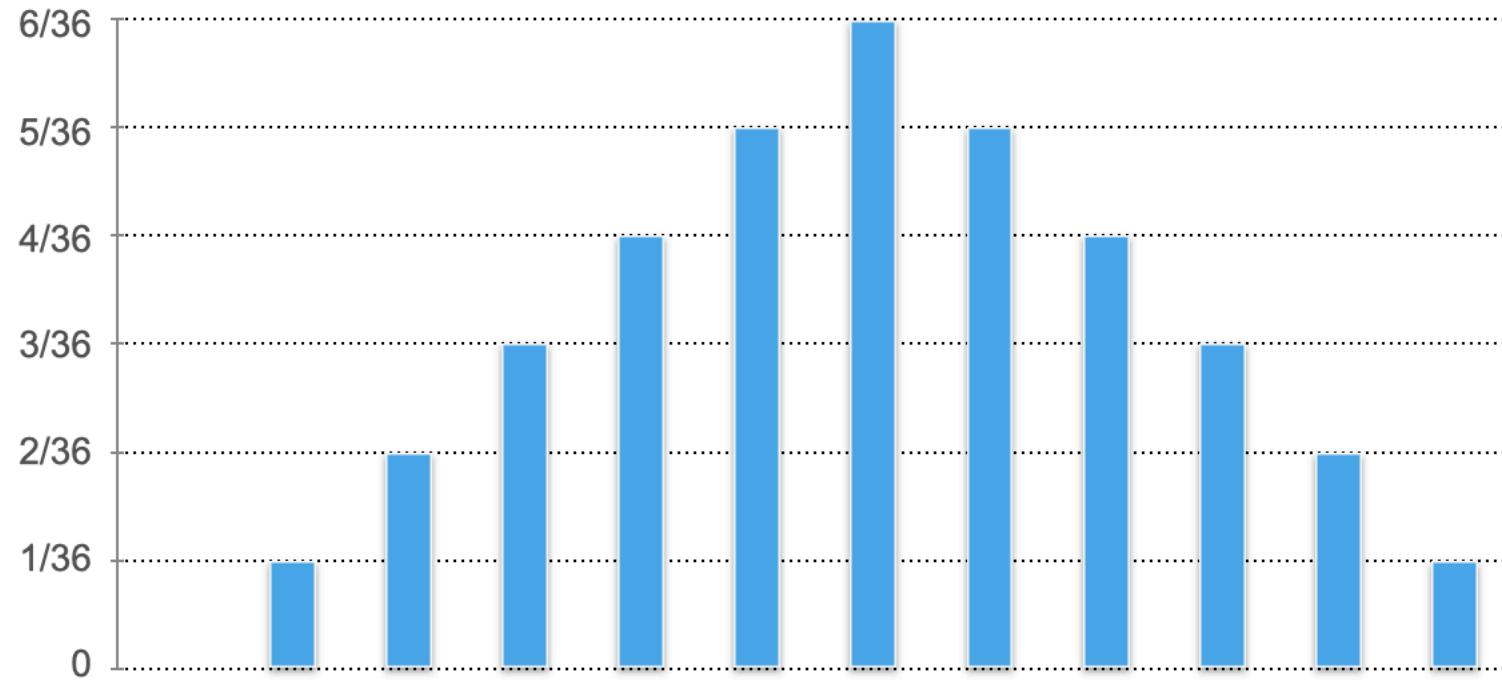
X : the number rolled on the 1st dice

Y : sum of the two dice

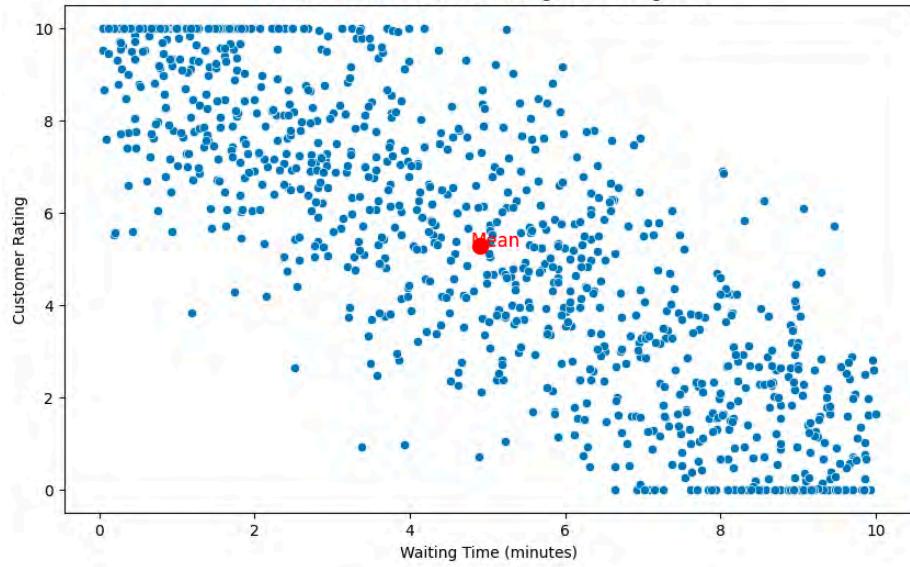
$$p_Y(y_j) = \sum_i p_{XY}(x_i, y_j) = ?$$

$$p_Y(10) = ? = \frac{3}{36}$$

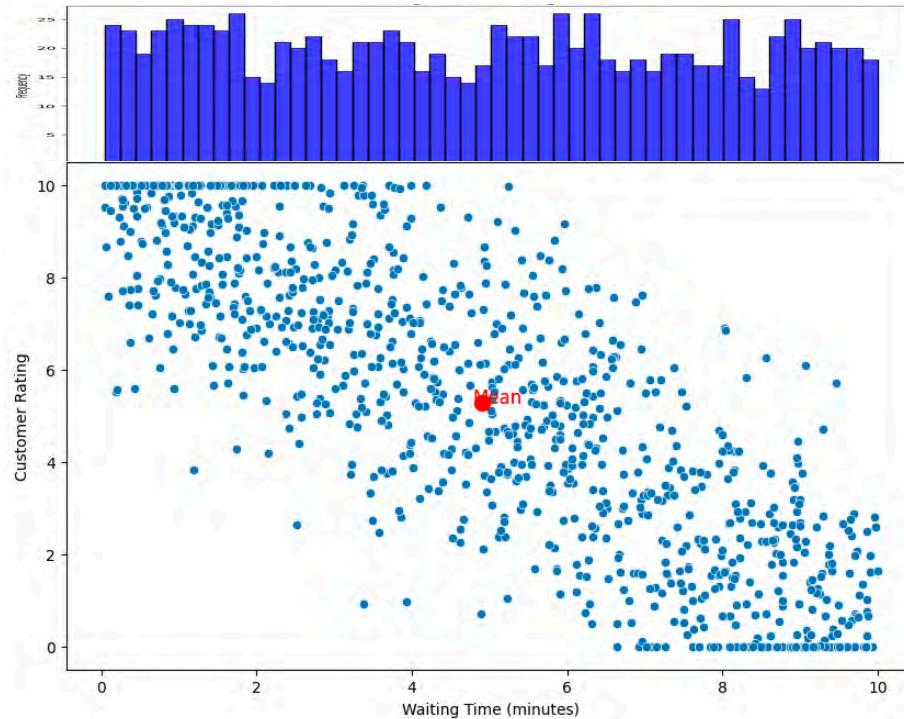
Marginal Distributions: Example 2



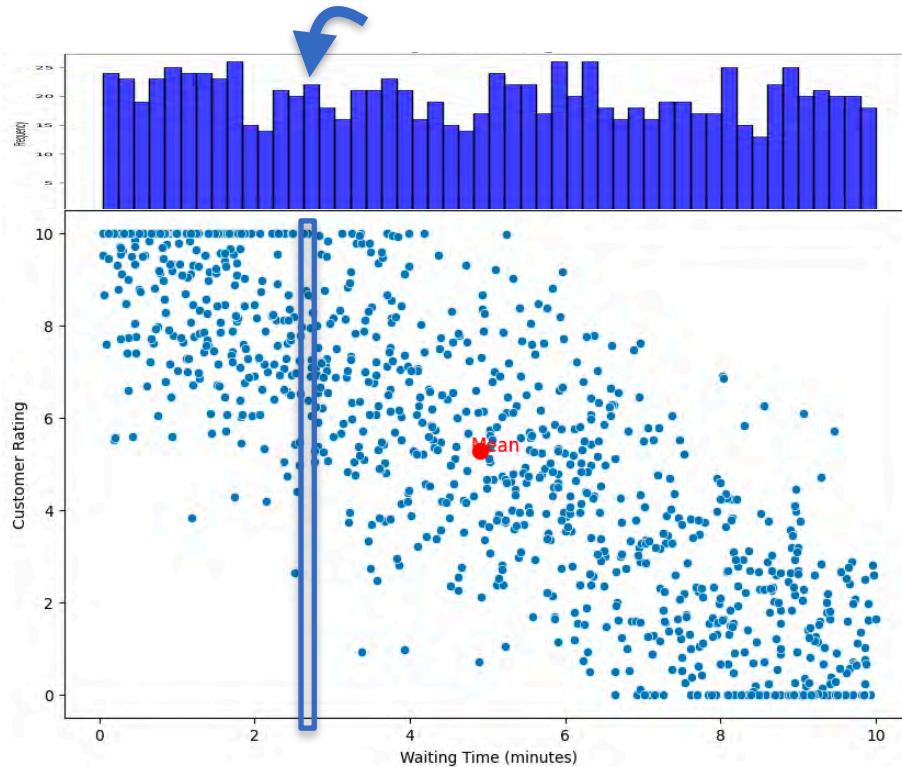
Marginal Distributions



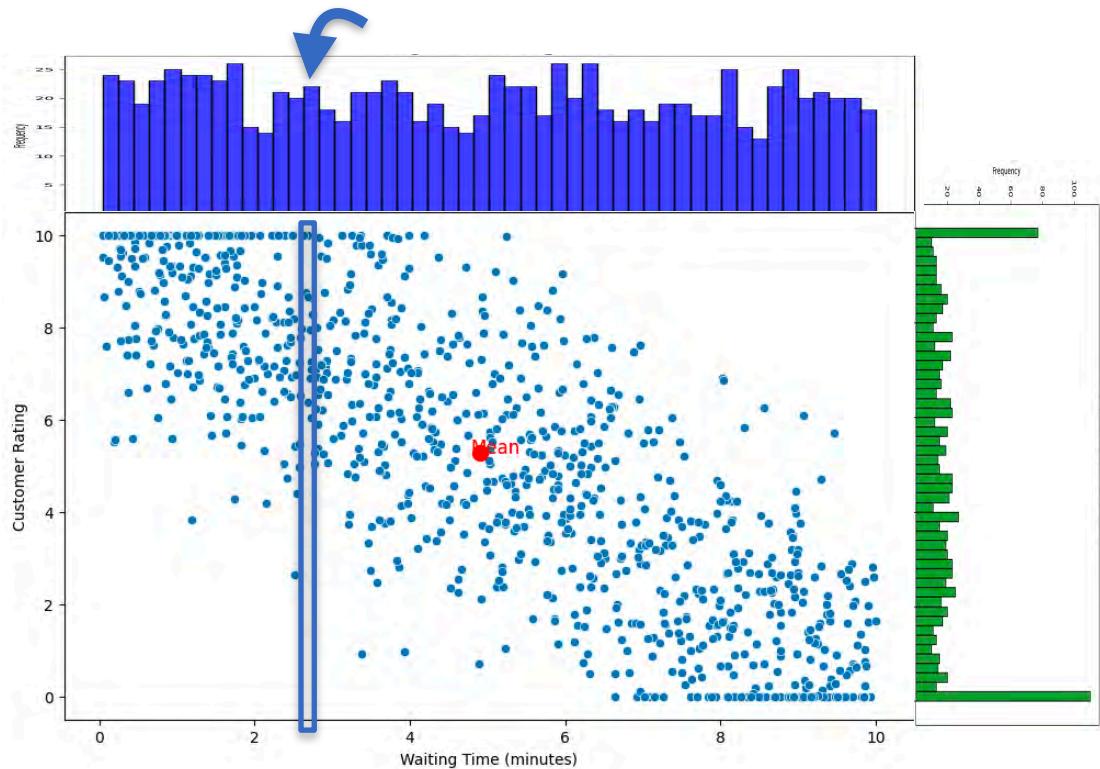
Marginal Distributions



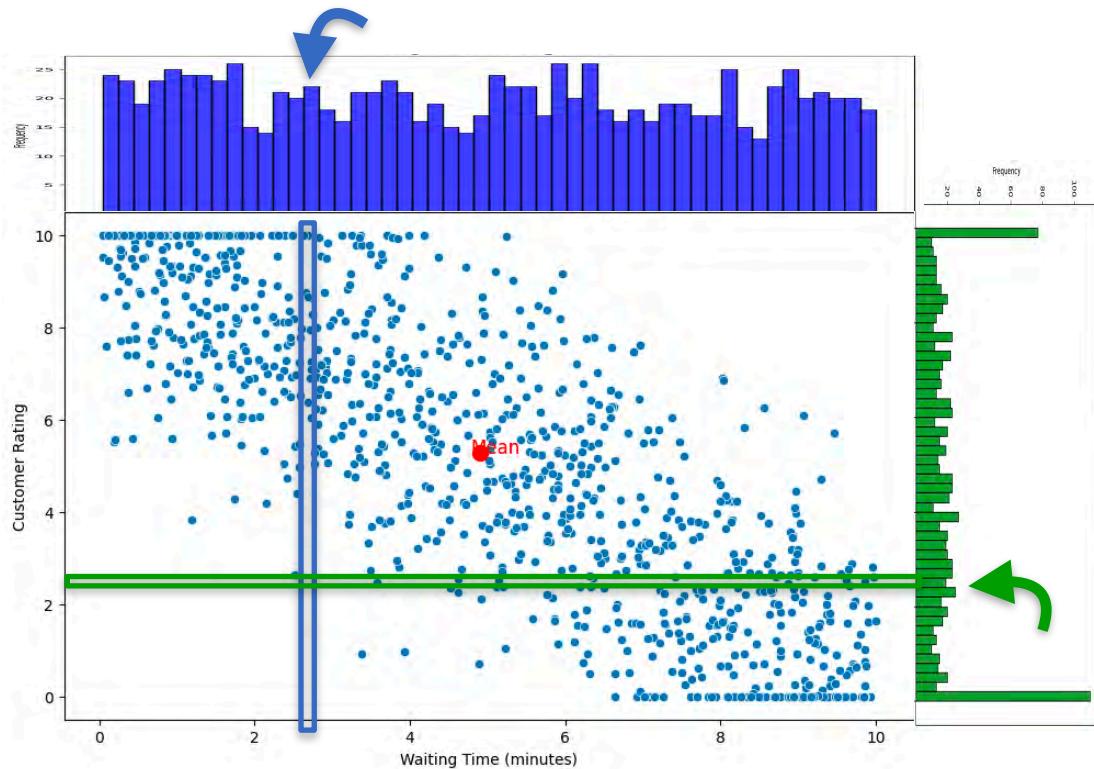
Marginal Distributions



Marginal Distributions



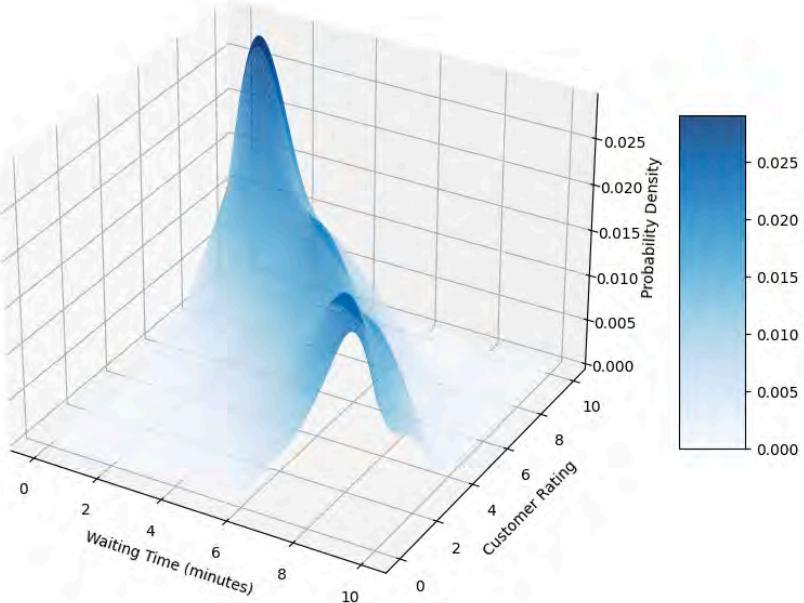
Marginal Distributions



Continuous Marginal Distribution

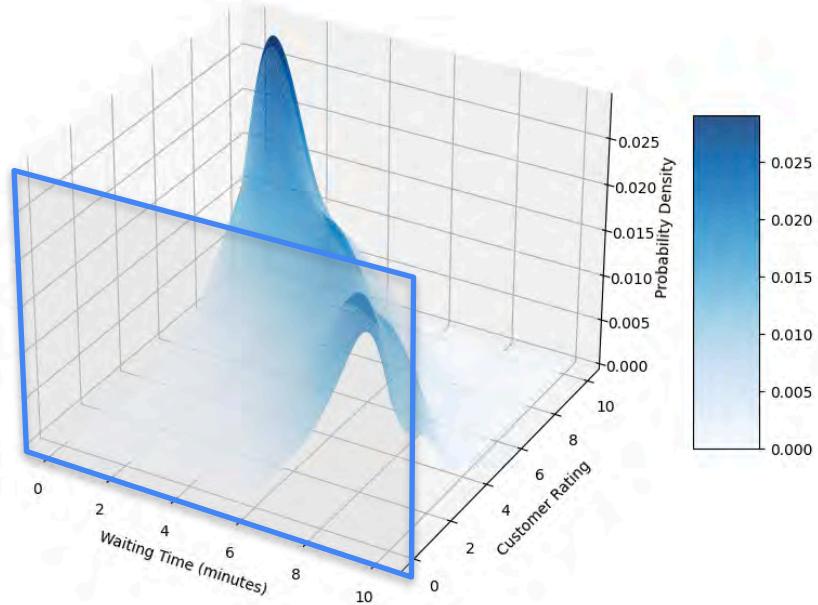
Continuous Marginal Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time



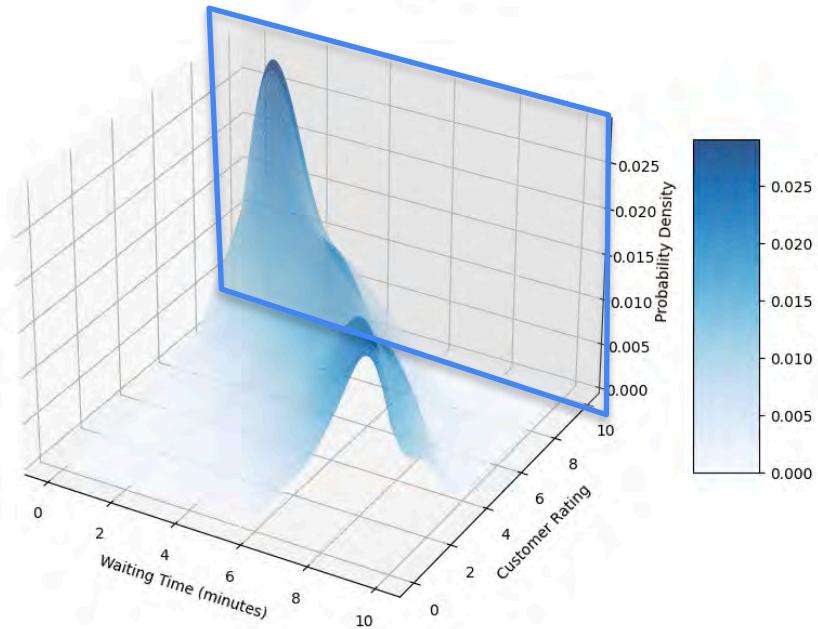
Continuous Marginal Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time



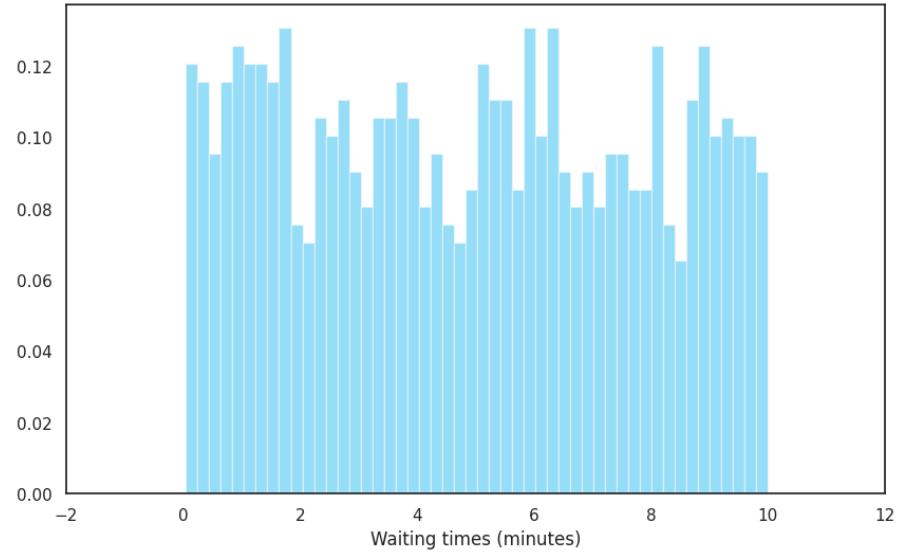
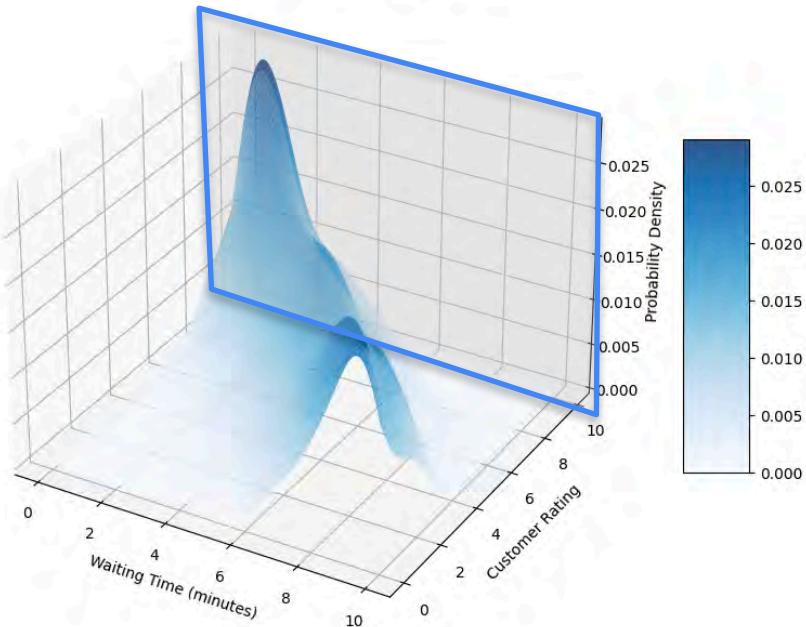
Continuous Marginal Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time



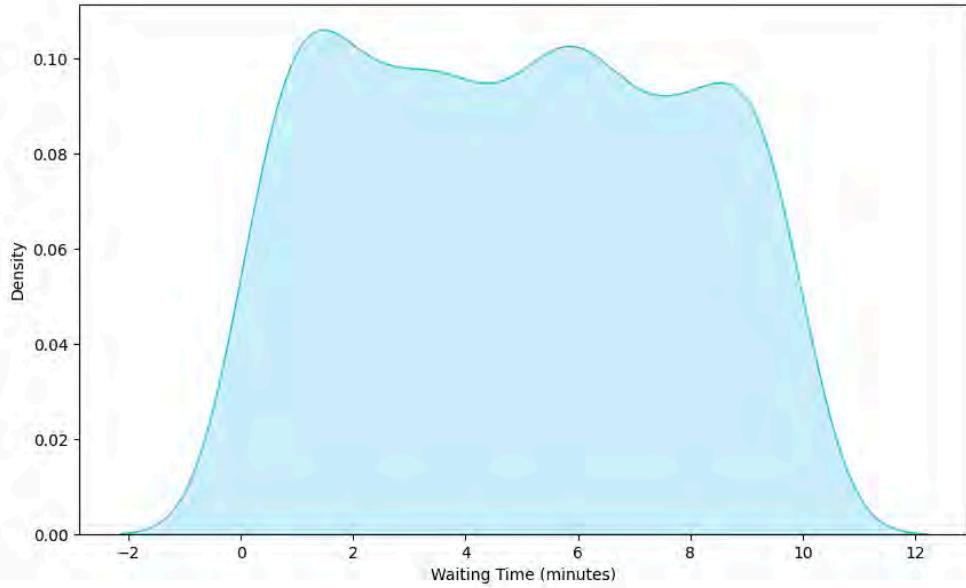
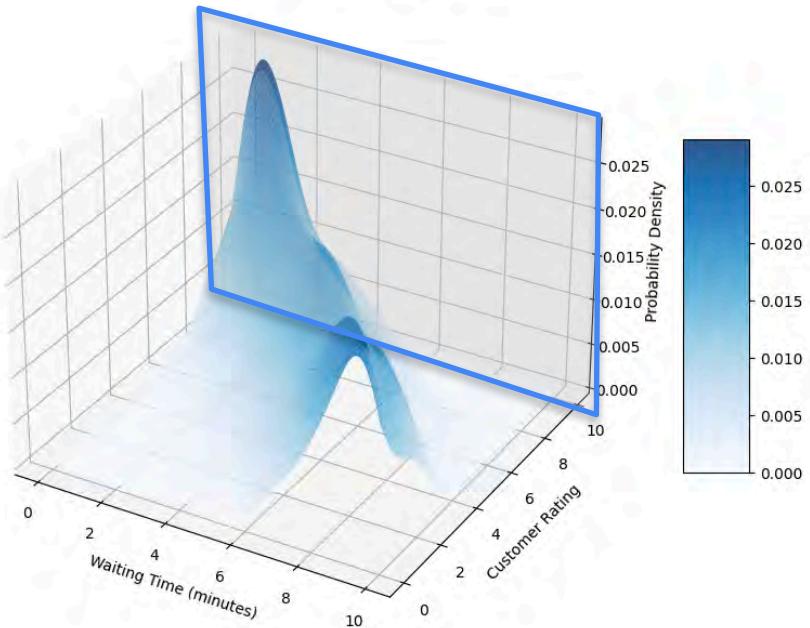
Continuous Marginal Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time

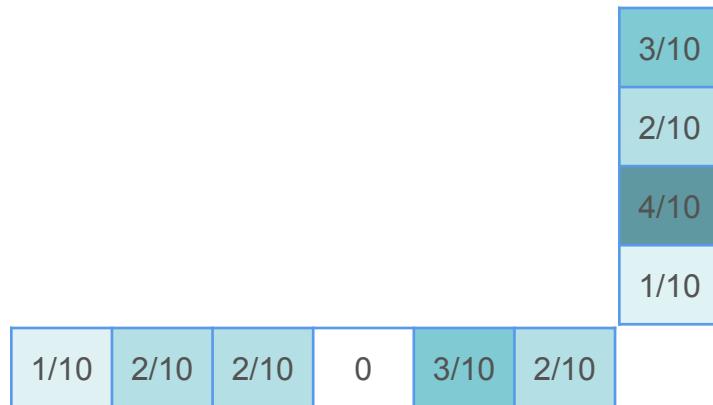


Continuous Marginal Distribution

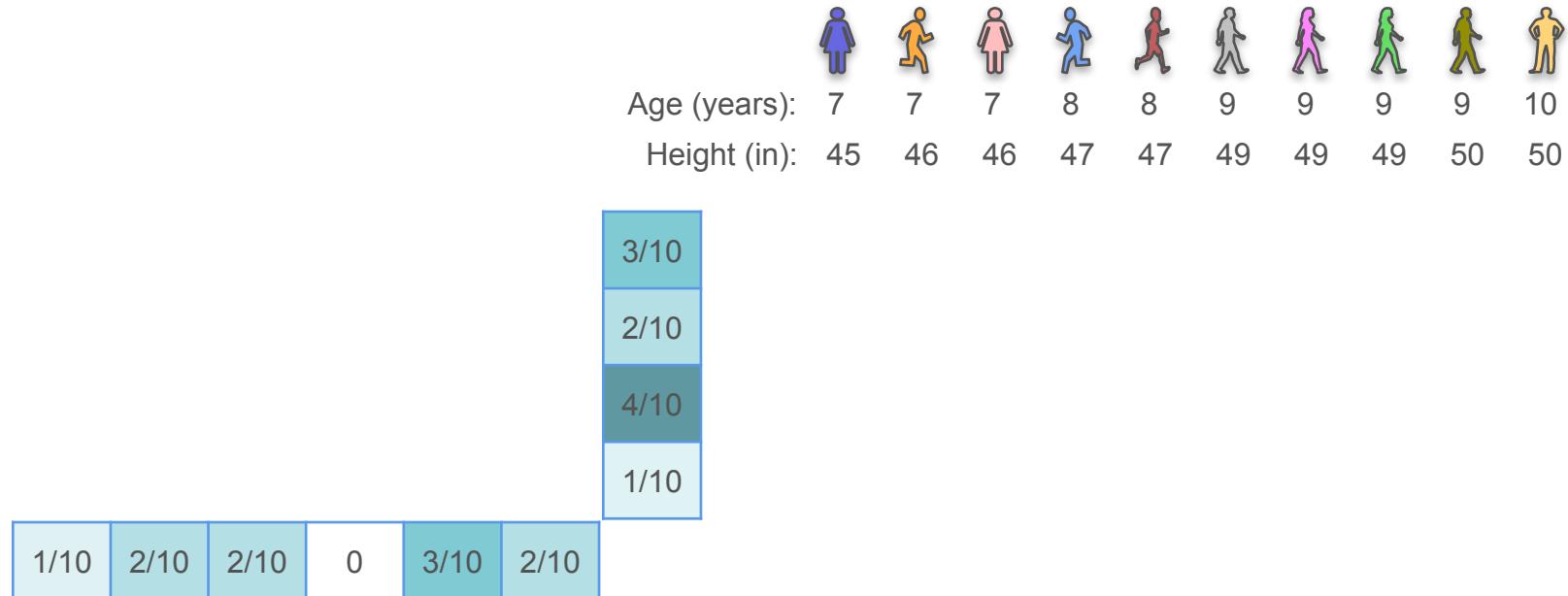
3D Probability Density Distribution for Customer Ratings vs Waiting Time



Conditional Distribution: Example 1



Conditional Distribution: Example 1



Conditional Distribution: Example 1

		Height (Y)						Age (years):		Height (in):									
		45	46	47	48	49	50	7	7	7	8	8	9	9	9	9	10	10	
Age (X)	7	1/10	2/10	0	0	0	0	3/10											
	8	0	0	2/10	0	0	0	2/10											
	9	0	0	0	0	3/10	1/10		4/10										
	10	0	0	0	0	0	1/10		1/10										
		1/10	2/10	2/10	0	3/10	2/10												

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

Age (X)	7	3/10
	8	2/10
	9	4/10
	10	1/10



Marginal Distribution of Age

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10
		1/10	2/10	2/10	0	3/10	2/10

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

Age (X)	Height (Y)					
	45	46	47	48	49	50
7	3/10					
8		2/10				
9			4/10			
10				1/10		
	1/10	2/10	2/10	0	3/10	2/10

Marginal Distribution of Age

Marginal Distribution of Height

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



Conditional Distribution: Example 1

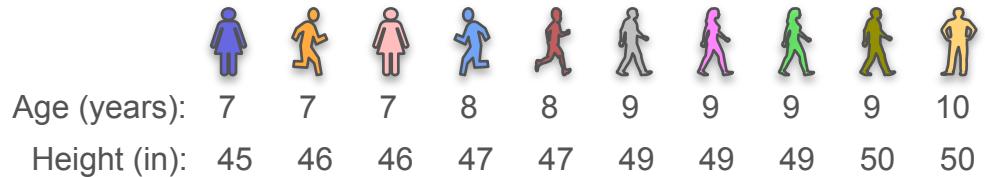
		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



If age = 9, what is the distribution across the height variable?

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



If age = 9, what is the distribution across the height variable?

Conditional Distribution

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



If age = 9, what is the distribution across the height variable?

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



If age = 9, what is the distribution across the height variable?

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10



If age = 9, what is the distribution across the height variable?

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

Conditional Distribution: Example 1

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

		Height (Y)					
Age (X)		45	46	47	48	49	50
	9	0	0	0	0	3/10	1/10

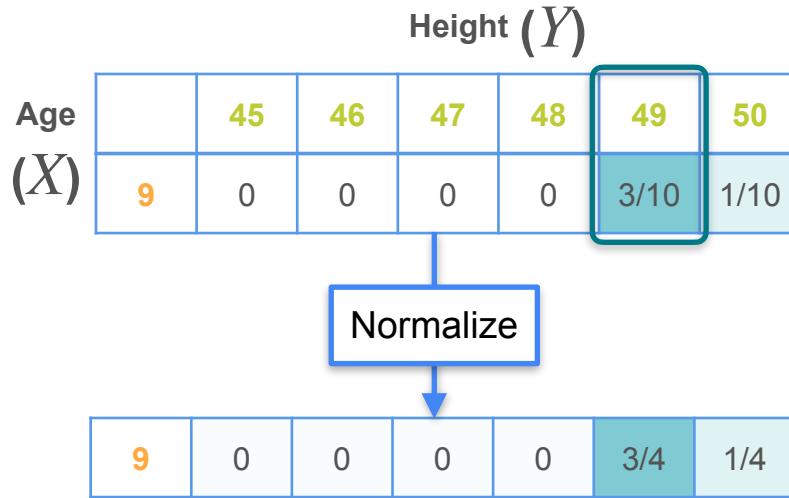
Conditional Distribution: Example 1

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

		Height (Y)					
Age (X)		45	46	47	48	49	50
	9	0	0	0	0	3/10	1/10

Conditional Distribution: Example 1

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$



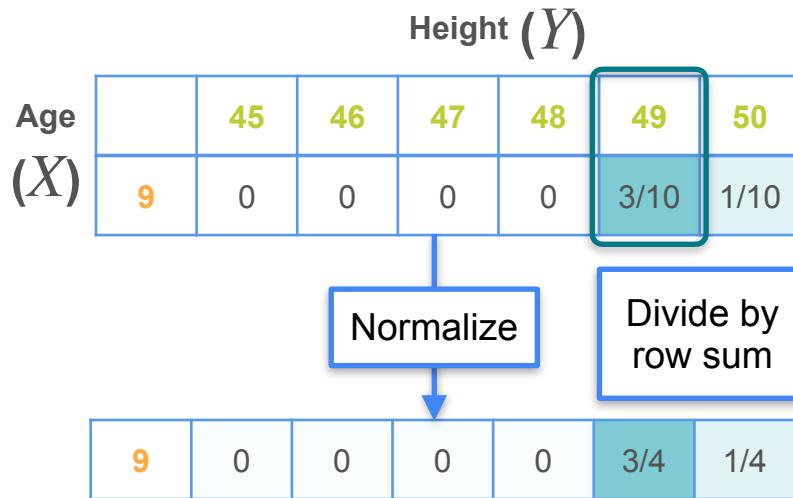
Conditional Distribution: Example 1

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

		Height (Y)					
		45	46	47	48	49	50
Age (X)	9	0	0	0	0	3/10	1/10
	9	0	0	0	0	3/10	1/10
		Normalize		Divide by row sum			
		9	0	0	0	0	3/4

Conditional Distribution: Example 1

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$



$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

Conditional Distribution: Example 1

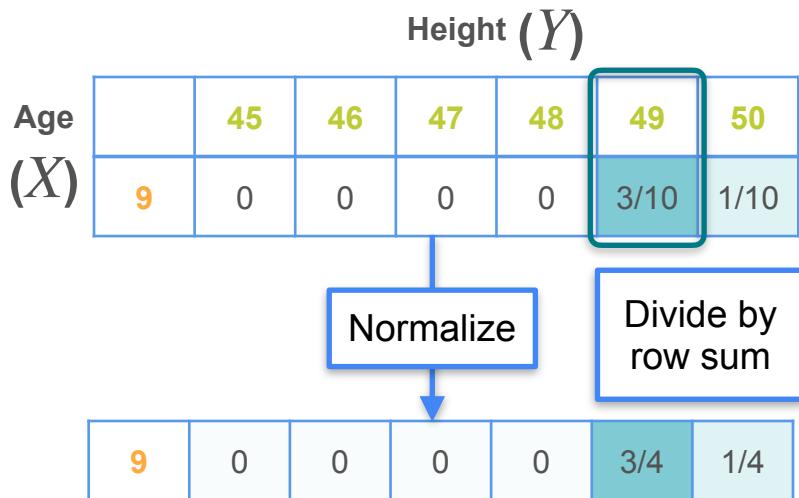
		Height (Y)					
		45	46	47	48	49	50
Age (X)	9	0	0	0	0	3/10	1/10
	9	0	0	0	0	3/10	1/10
		Normalize		Divide by row sum			
		9	0	0	0	0	3/4

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

Conditional Distribution: Example 1



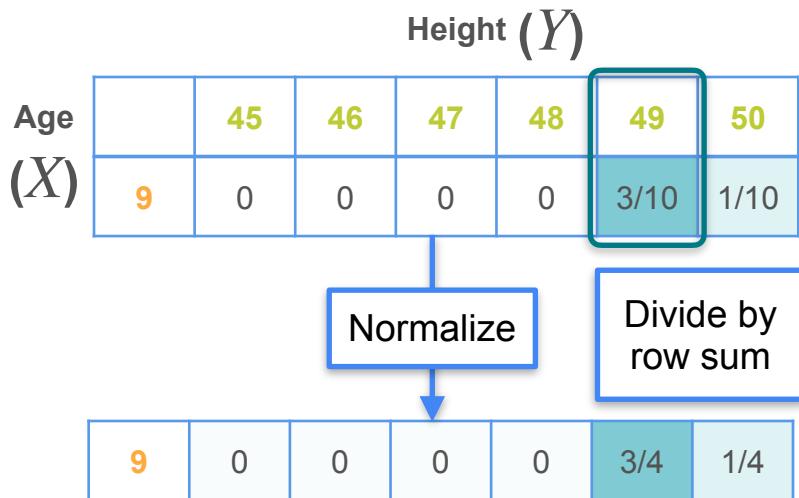
$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(X = 9, Y = 49) = \mathbf{P}(X = 9) \cdot \mathbf{P}(Y = 49 | X = 9)$$

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

Conditional Distribution: Example 1



$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

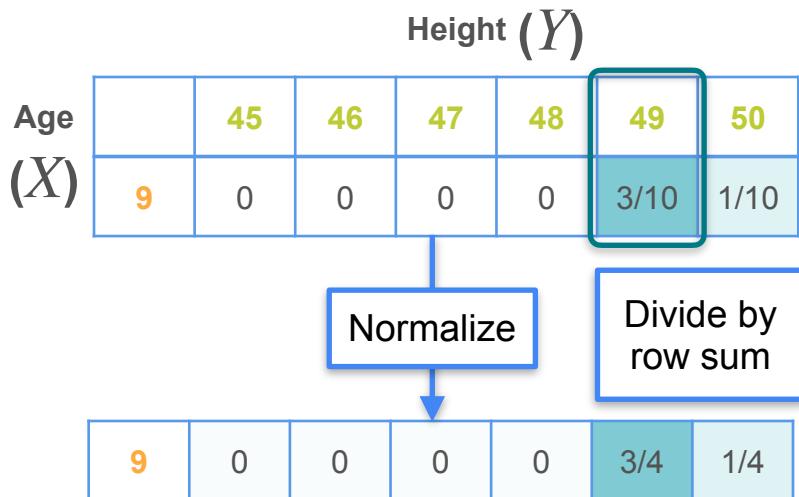
$$\mathbf{P}(A, B) = \mathbf{P}(A) \cdot \mathbf{P}(B | A)$$

$$\mathbf{P}(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

Conditional Distribution: Example 1



$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(X = 9, Y = 49) = \mathbf{P}(X = 9) \cdot \mathbf{P}(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{\mathbf{P}(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 | X = 9)$$

Conditional Distribution: Example 1

		Height (Y)					
		45	46	47	48	49	50
Age (X)	9	0	0	0	0	3/10	1/10
	9	0	0	0	0	3/4	1/4

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

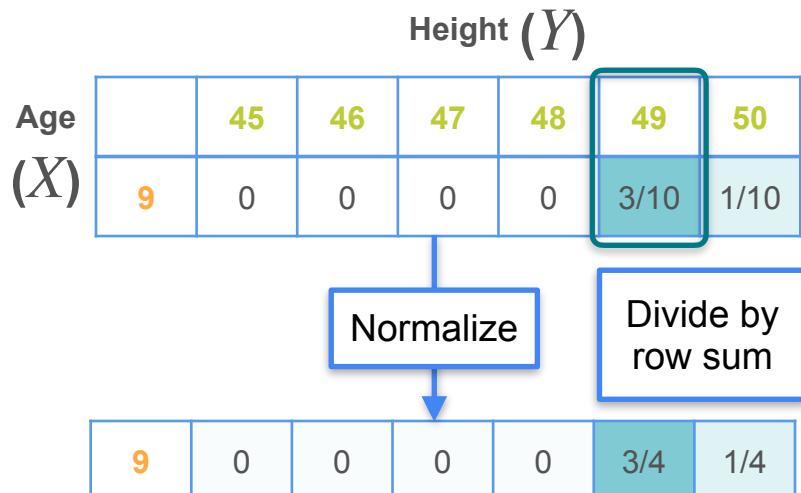
$$\mathbf{P}(X = 9, Y = 49) = \mathbf{P}(X = 9) \cdot \mathbf{P}(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{\mathbf{P}(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3/10}{4/10}$$

Conditional Distribution: Example 1



$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

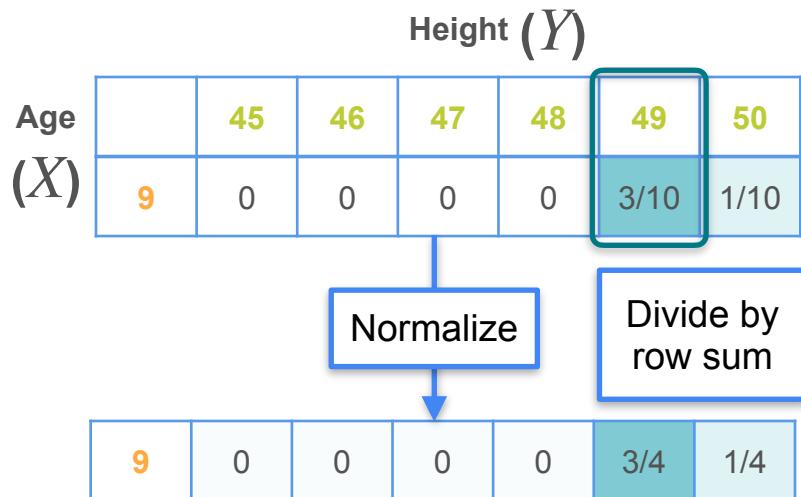
$$\mathbf{P}(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3/10}{4/10} = \frac{3}{4}$$

Conditional Distribution: Example 1



$$P(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = P(Y = y | X = 9)$$

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

$$P(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{P(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

Conditional Distribution: Example 1

Age (X)	Height (Y)					P(X = 9)
	45	46	47	48	49	
9	0	0	0	0	3/10	1/10
	Normalize	Divide by row sum				
9	0	0	0	0	3/4	1/4

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3}{4}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

Conditional Distribution: Example 1

(X)	Height (Y)					Sum
	45	46	47	48	P($X = 9$)	
9	0	0	0	0	3/10	1/10

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3/10}{4/10} = \frac{3}{4}$$

Conditional Distribution: Example 1

(X)	Height (Y)					P(X = 9)	Sum
	45	46	47	48	49		
9	0	0	0	0	3/10	1/10	

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y | X = 9)$$

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

$$\mathbf{P}(X = 9, Y = 49) = P(X = 9) \cdot P(Y = 49 | X = 9)$$

$$P(Y = 49 | X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 | X = 9) = \frac{3/10}{4/10} = \frac{3}{4}$$

Discrete Conditional Distribution: Formula

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Discrete Conditional Distribution: Formula

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Joint PDF of X and Y



Discrete Conditional Distribution: Formula

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Conditional PDF of Y

Joint PDF of X and Y

A diagram illustrating the formula for conditional probability. On the left, the term "Conditional PDF of Y " is written. A curved teal arrow originates from this text and points to the numerator $p_{XY}(x, y)$ in the formula. On the right, the term "Joint PDF of X and Y " is written. Another curved teal arrow originates from this text and points to the denominator $p_X(x)$ in the formula.

Discrete Conditional Distribution: Formula

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Conditional PDF of Y

Joint PDF of X and Y

Marginal distribution of X

The diagram illustrates the formula for the conditional probability of Y given $X=x$. It features a central equation $p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$. Three curved arrows originate from text labels to specific components of the equation: one arrow points from 'Conditional PDF of Y ' to the term $p_{Y|X=x}(y)$; another arrow points from 'Joint PDF of X and Y ' to the term $p_{XY}(x, y)$; and a third arrow points from 'Marginal distribution of X ' to the term $p_X(x)$.

Conditional Distributions: Example 2



Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

X

	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Conditional Distributions: Example 2



Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

$$p_{Y|X=4}(y = 1)$$

X

	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Conditional Distributions: Example 2



Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

$$p_{Y|X=4}(y = 1) = \frac{p_{XY}(x = 4, y = 1)}{p_X(x = 4)}$$

X

	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Conditional Distributions: Example 2



Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

$$p_{Y|X=4}(y = 1) = \frac{p_{XY}(x = 4, y = 1)}{p_X(x = 4)}$$

	Y							
	1	2	3	4	5	6	Sum	
X	1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	6	1/36	1/36	1/36	1/36	1/36	1/36	1/6

Conditional Distributions: Example 2



Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

$$p_{Y|X=4}(y=1) = \frac{p_{XY}(x=4, y=1)}{p_X(x=4)}$$
$$= \frac{1/36}{1/6}$$

	Y							
	1	2	3	4	5	6	Sum	
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6	
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6	
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6	
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6	
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6	
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6	

Conditional Distributions: Example 2

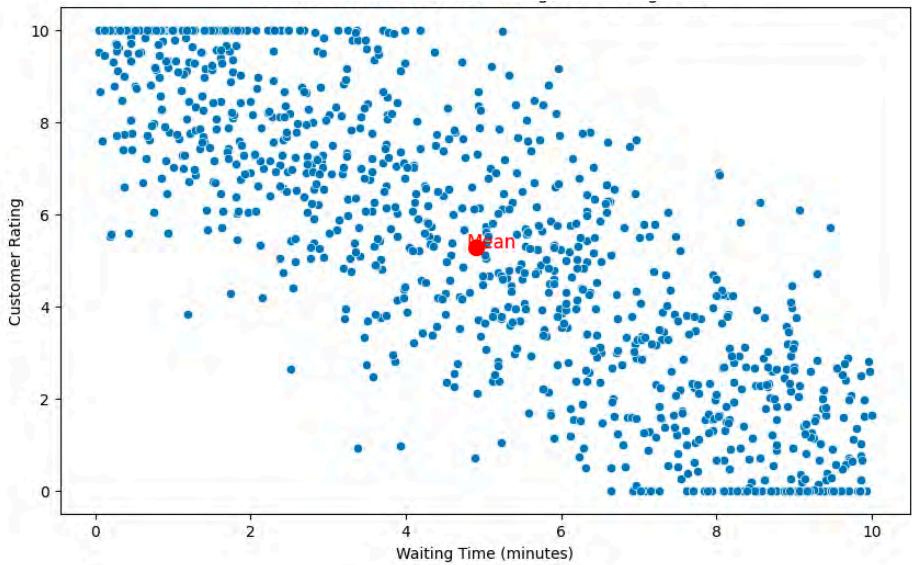


Die 1: 1/6 1/6 1/6 1/6 1/6 1/6
Die 2: 1/6 1/6 1/6 1/6 1/6 1/6

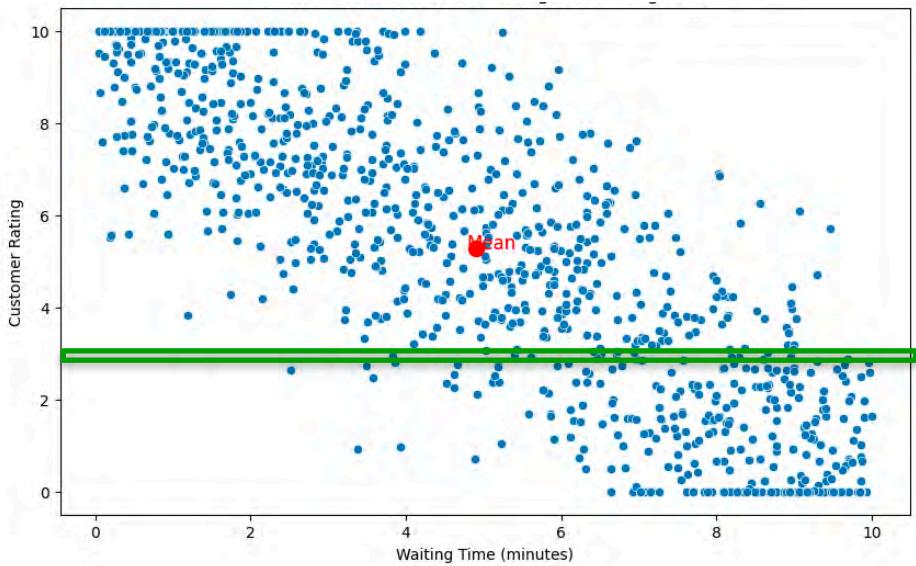
$$\begin{aligned} p_{Y|X=4}(y=1) &= \frac{p_{XY}(x=4, y=1)}{p_X(x=4)} \\ &= \frac{1/36}{1/6} \\ &= \frac{1}{6} \end{aligned}$$

	Y							
	1	2	3	4	5	6	Sum	
X	1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	6	1/36	1/36	1/36	1/36	1/36	1/36	1/6

Conditional Distributions: Example 4



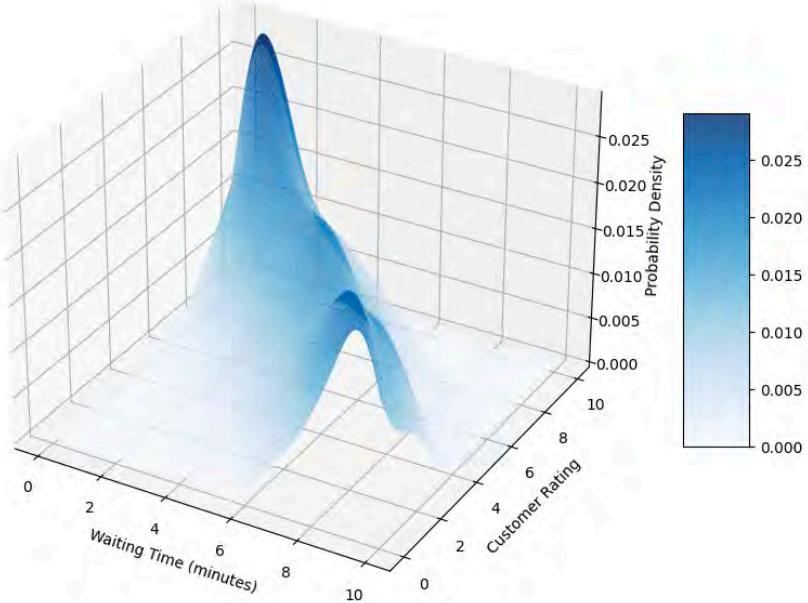
Conditional Distributions: Example 4



Continuous Conditional Distribution

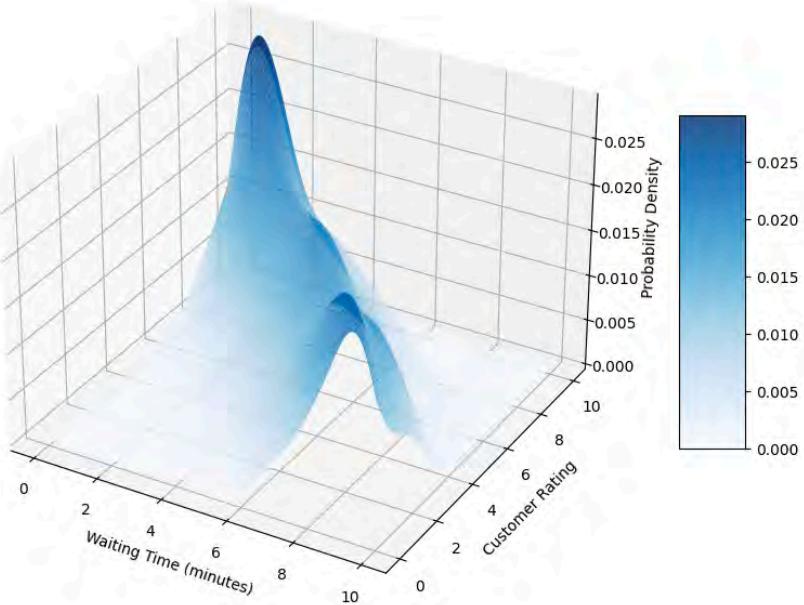
Continuous Conditional Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time



Continuous Conditional Distribution

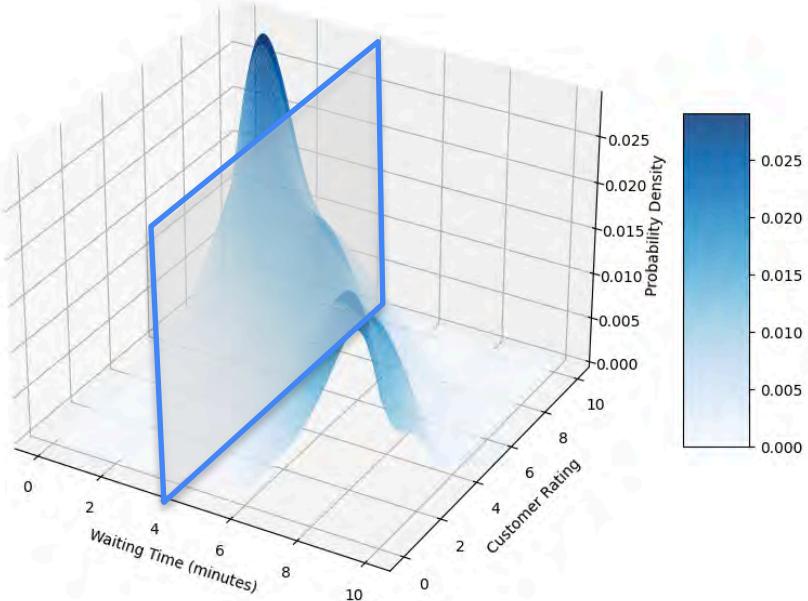
3D Probability Density Distribution for Customer Ratings vs Waiting Time



Probability distribution for rating given that waiting time was 4 minutes

Continuous Conditional Distribution

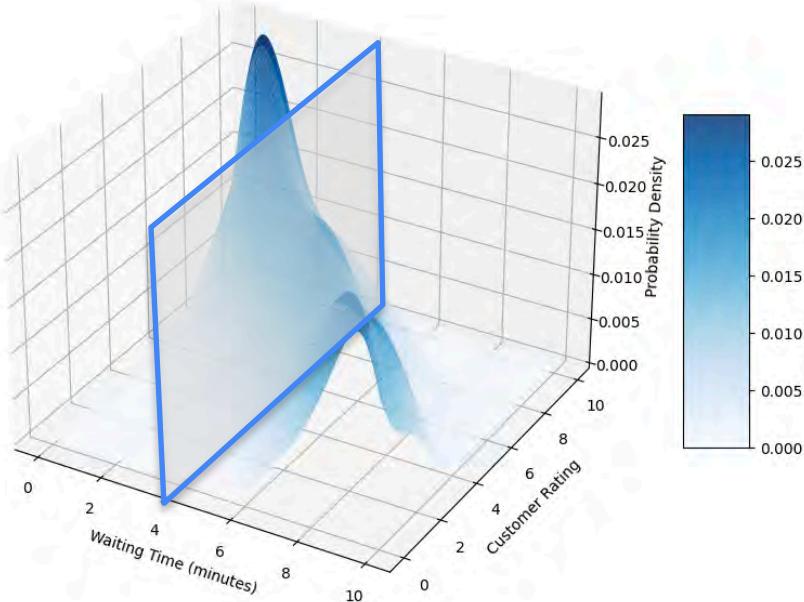
3D Probability Density Distribution for Customer Ratings vs Waiting Time



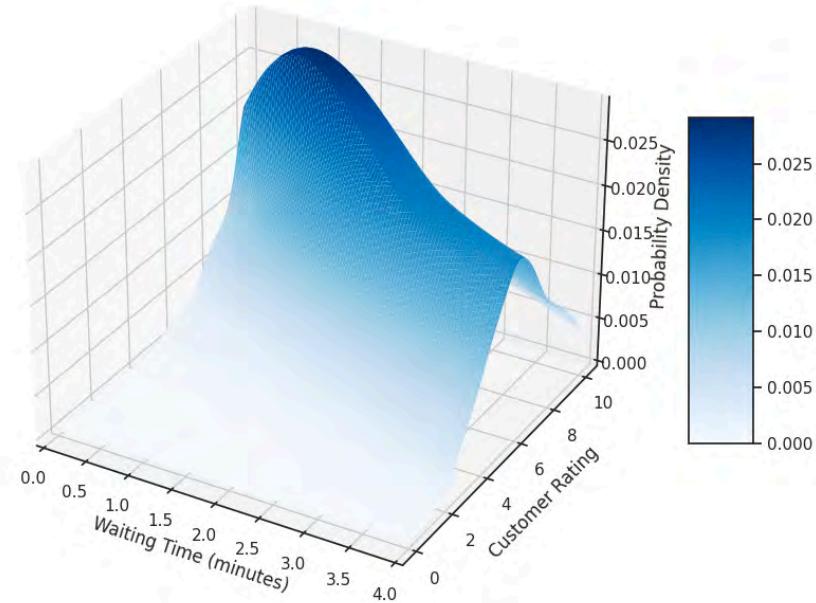
Probability distribution for rating given that waiting time was 4 minutes

Continuous Conditional Distribution

3D Probability Density Distribution for Customer Ratings vs Waiting Time

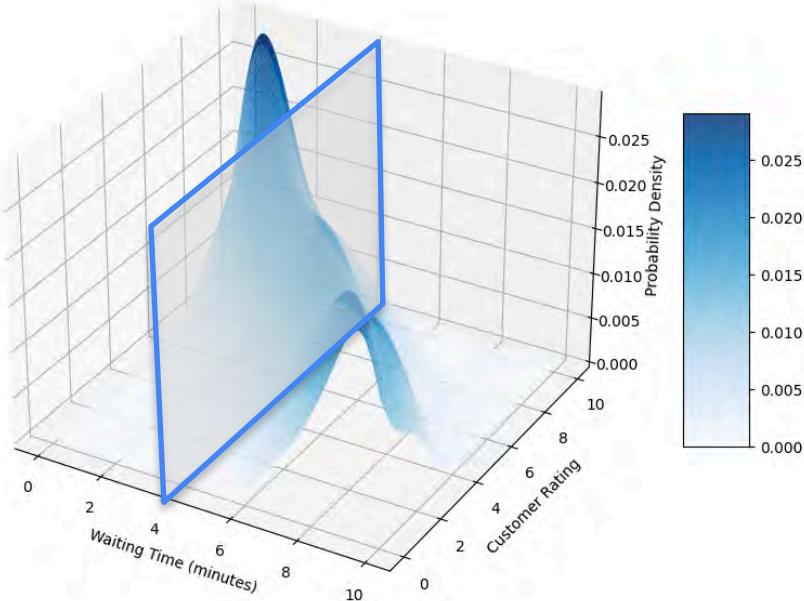


Probability distribution for rating given that waiting time was 4 minutes

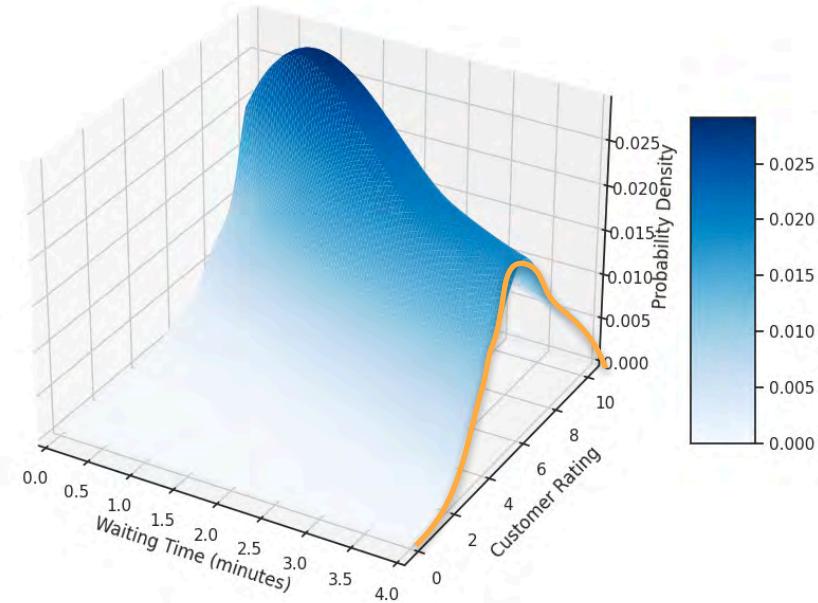


Continuous Conditional Distribution

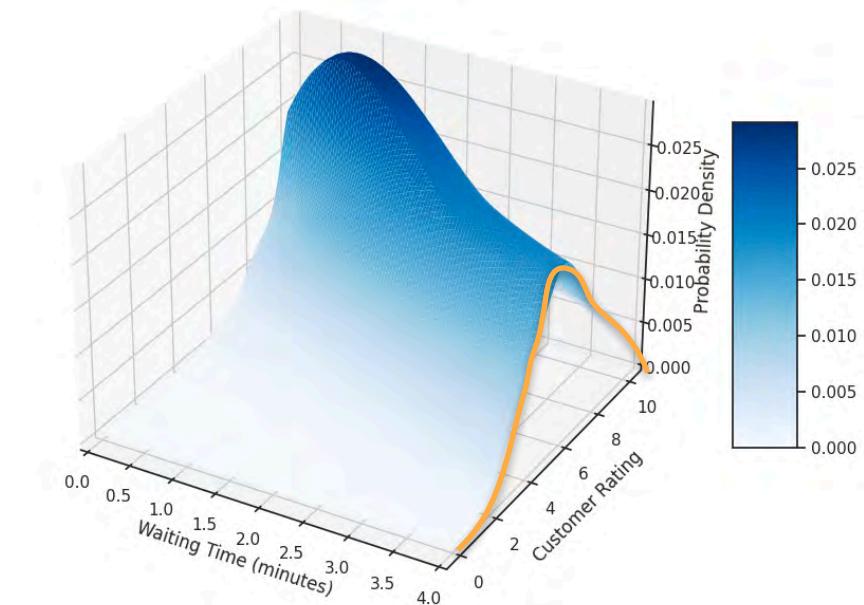
3D Probability Density Distribution for Customer Ratings vs Waiting Time



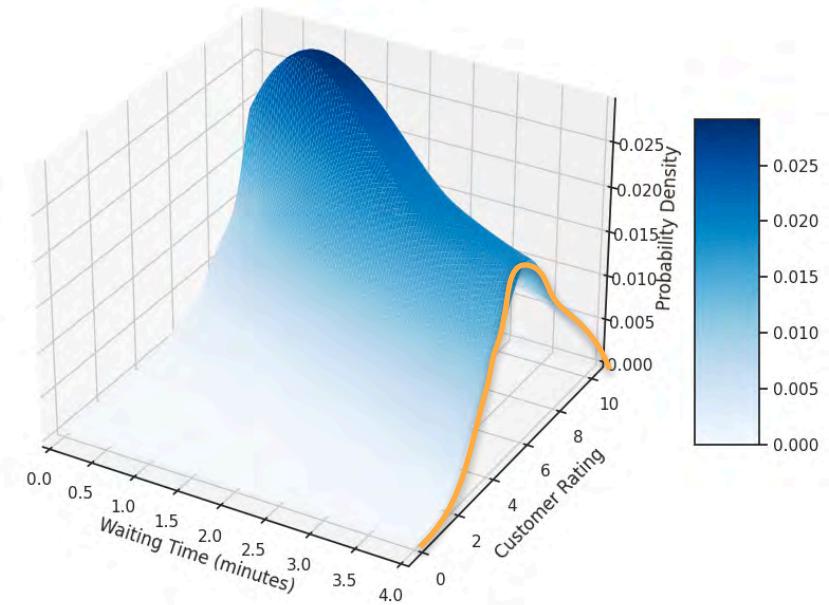
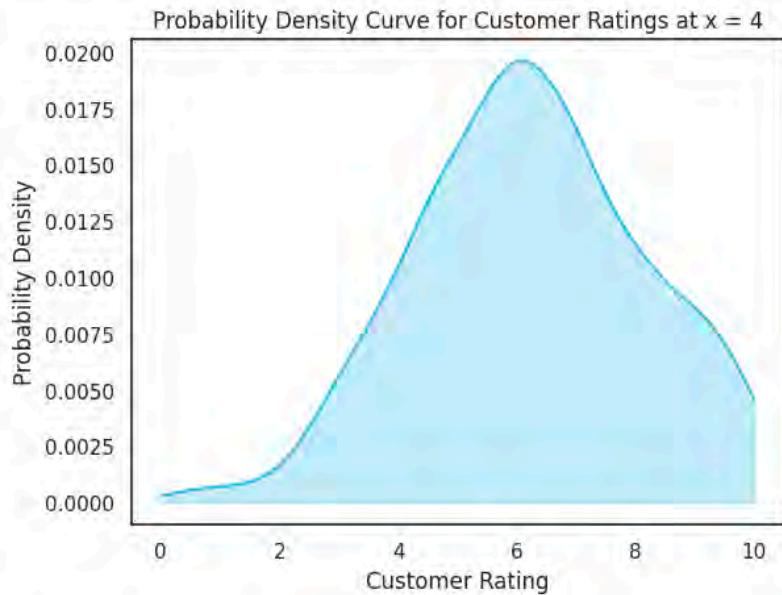
Probability distribution for rating given that waiting time was 4 minutes



Continuous Conditional Distribution



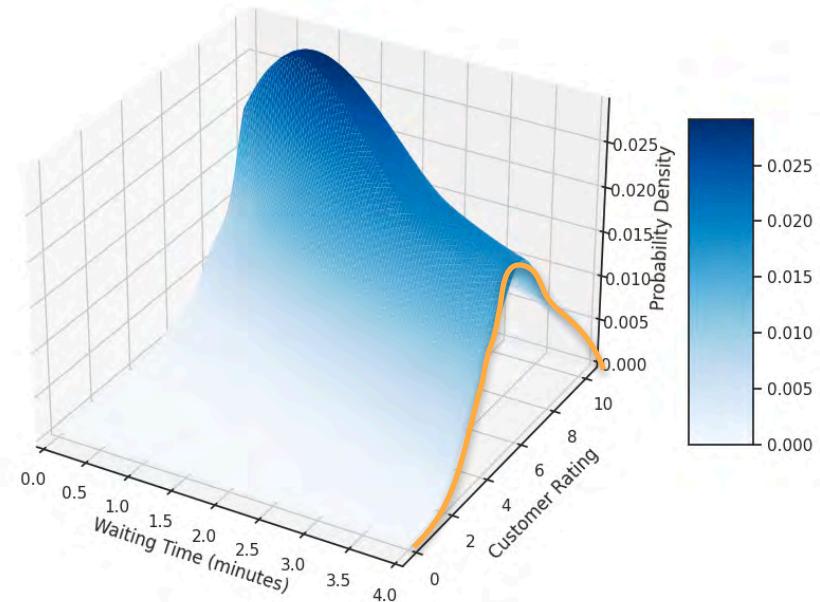
Continuous Conditional Distribution



Continuous Conditional Distribution



Conditional PDF of y given $x = 4$



Discrete Conditional Distribution

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Discrete Conditional Distribution

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Joint PDF of X and Y



Discrete Conditional Distribution

The diagram illustrates the formula for a discrete conditional distribution. On the left, the text "Conditional PDF of Y " is written next to a teal curved arrow pointing towards the term $p_{Y|X=x}(y)$. On the right, the text "Joint PDF of X and Y " is written next to another teal curved arrow pointing towards the term $p_{XY}(x,y)$. The central equation is:

$$p_{Y|X=x}(y) = \frac{p_{XY}(x,y)}{p_X(x)}$$

Discrete Conditional Distribution

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

Conditional PDF of Y

Joint PDF of X and Y

Marginal distribution of X

The diagram illustrates the derivation of the conditional probability formula. It shows three components: 'Conditional PDF of Y ' pointing to the numerator $p_{XY}(x, y)$; 'Marginal distribution of X ' pointing to the denominator $p_X(x)$; and 'Joint PDF of X and Y ' pointing to the overall fraction.

Continuous Conditional Distribution: Formula

$$f_{Y|X=x}(y) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Conditional PDF of Y

Joint PDF of X and Y

Marginal distribution of X



DeepLearning.AI

Probability Distributions with Multiple Variables

Covariance

Introduction to Covariance

Introduction to Covariance

X : age of a child

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Y_2 : grades in a test

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Y_2 : grades in a test

Y_3 : number of naps per day

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Y_2 : grades in a test

Y_3 : number of naps per day

Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Y_2 : grades in a test

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

Y_3 : number of naps per day

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Y_2 : grades in a test

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

Y_3 : number of naps per day

Age (X)	Naps per day (Y_3)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

Introduction to Covariance

X : age of a child



Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

Age (X)	Naps per day (Y_3)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

Introduction to Covariance

Y_1 : height of the child (in)

Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

X : age of a child

Y_2 : grades in a test

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

- How is variable X related to each of the Y variables?
- How do you compare these relations?

Y_3 : number of naps per day

Age (X)	Naps per day (Y_3)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)

Age (X)	Height (Y_1)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Y_2 : grades in a test

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

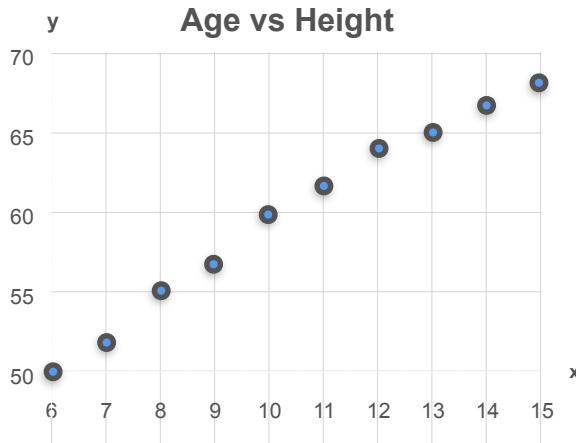
Y_3 : number of naps per day

Age (X)	Naps per day (Y_3)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)



Y_2 : grades in a test

Age (X)	Grades (Y_2)
6	5
7	7
8	8
9	3
10	1
11	1
12	6
13	10
14	2
15	7

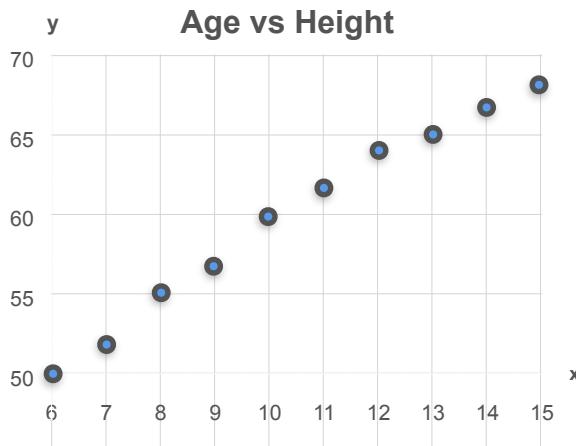
Y_3 : number of naps per day

Age (X)	Naps per day (Y_3)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

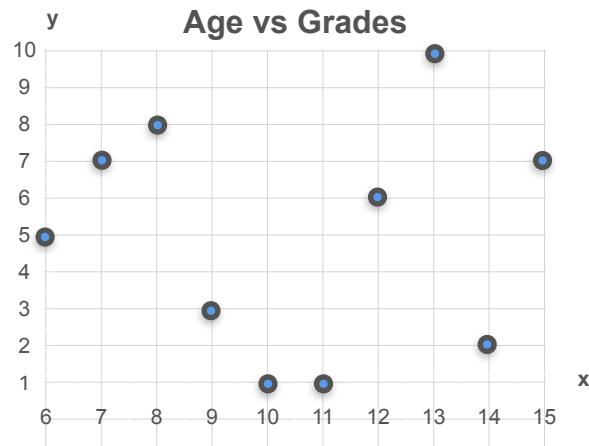
Introduction to Covariance

X : age of a child

Y_1 : height of the child (in)



Y_2 : grades in a test



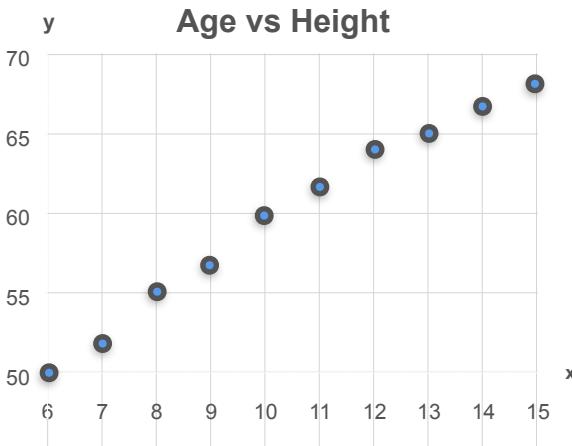
Y_3 : number of naps per day

Age (X)	Naps per day (Y ₃)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

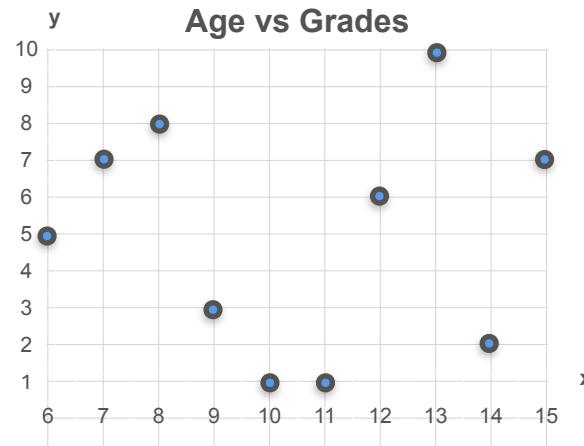
Introduction to Covariance

X : age of a child

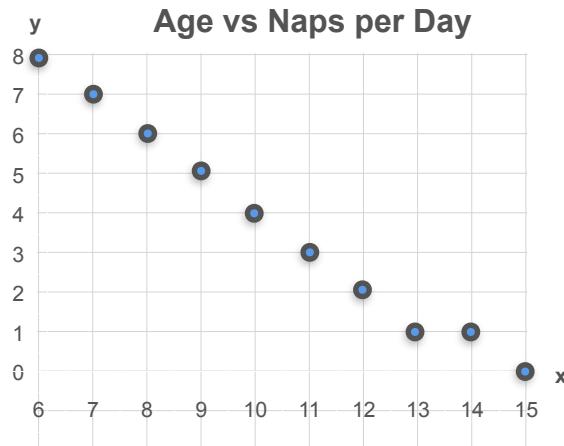
Y_1 : height of the child (in)



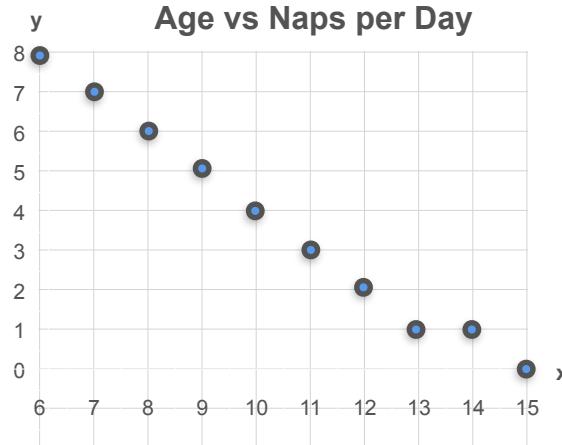
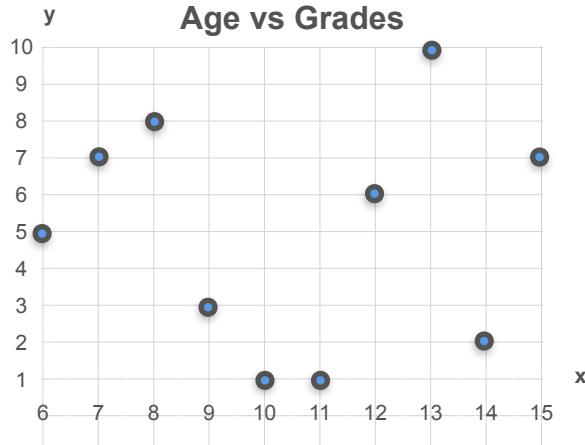
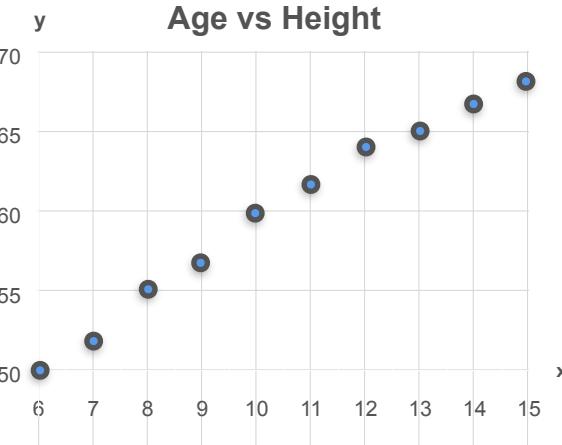
Y_2 : grades in a test



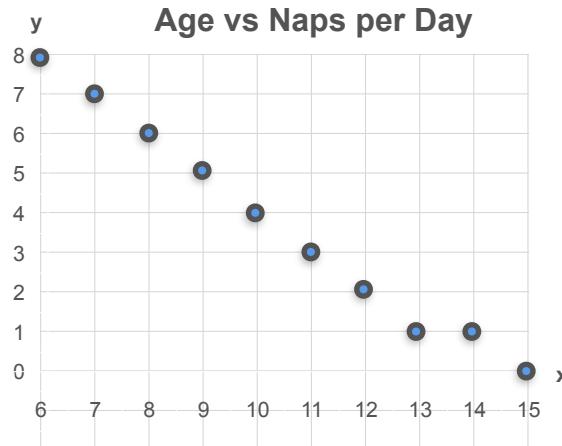
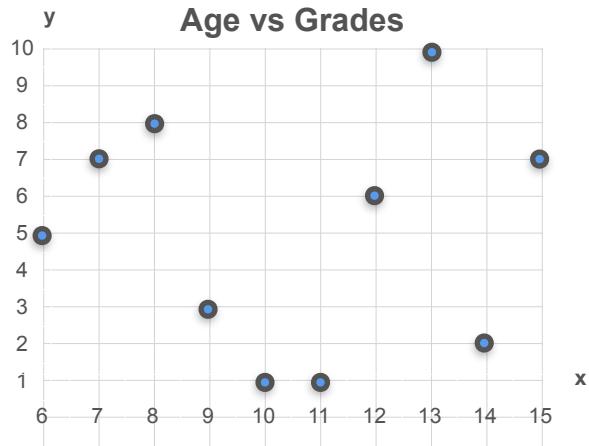
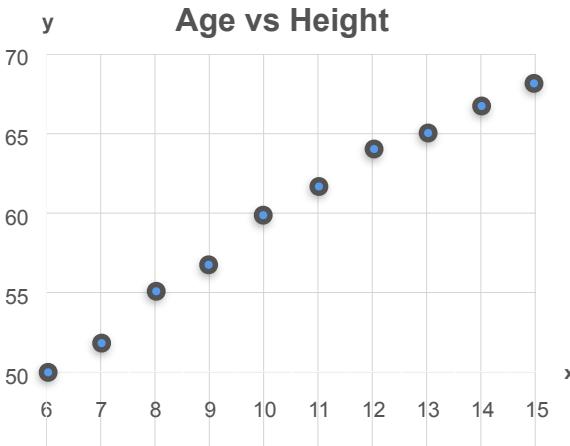
Y_3 : number of naps per day



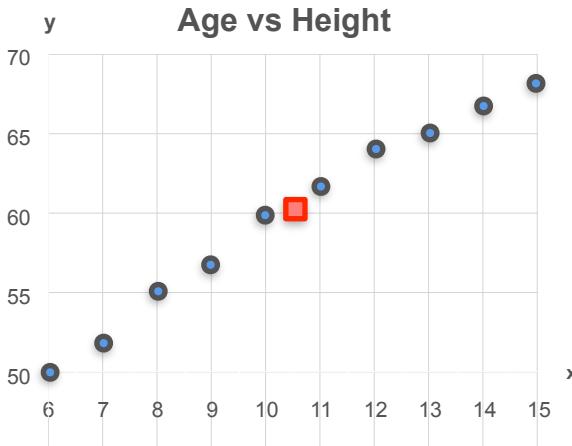
How To Compare These?



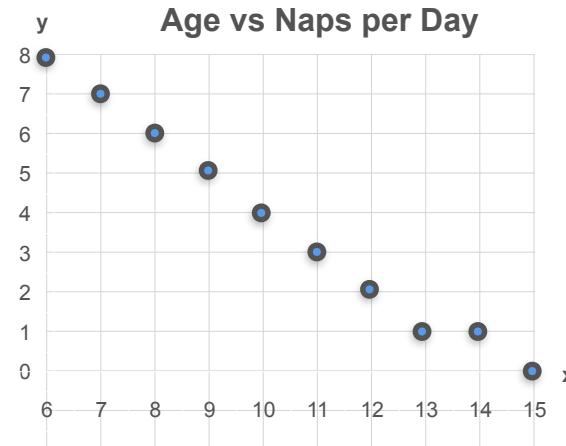
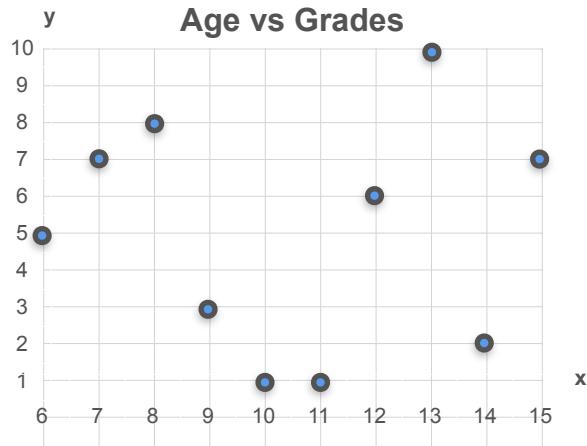
Mean?



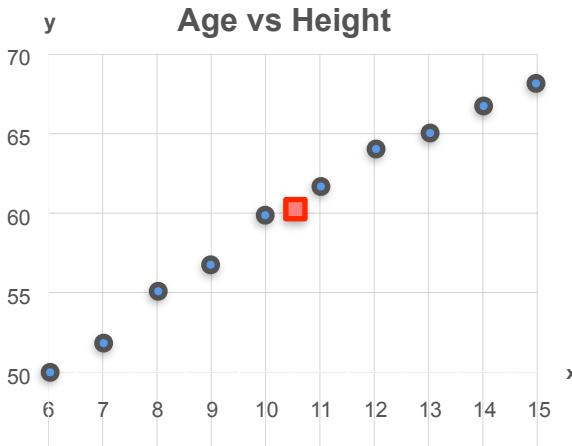
Mean?



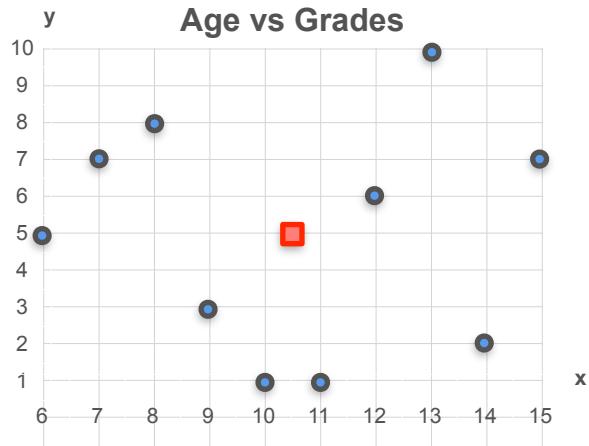
$$\mu_x = 10.5 \quad \mu_y = 60$$



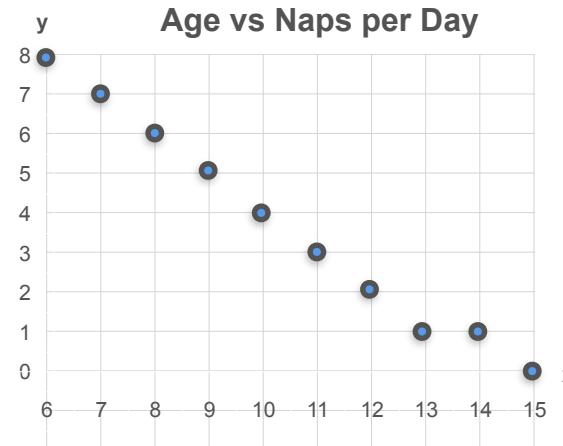
Mean?



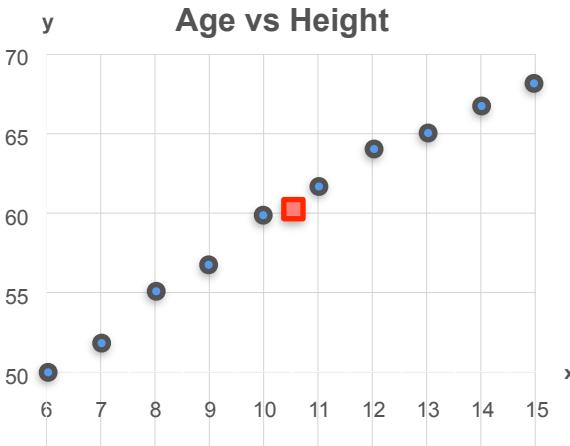
$$\mu_x = 10.5 \quad \mu_y = 60$$



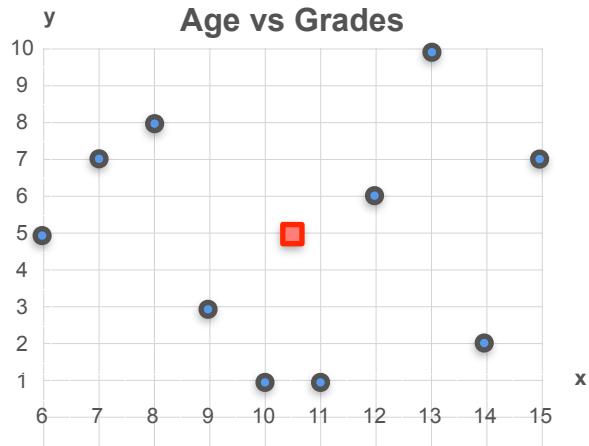
$$\mu_x = 10.5 \quad \mu_y = 5$$



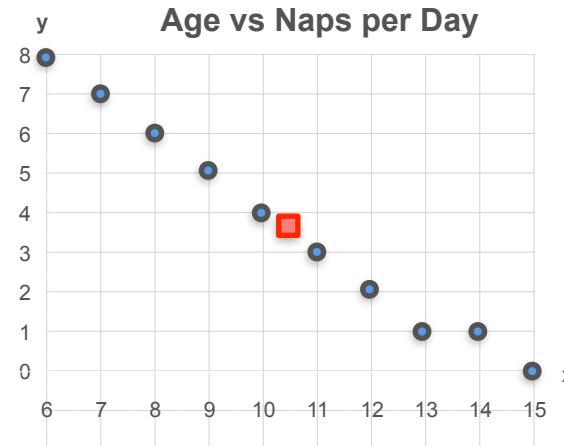
Mean?



$$\mu_x = 10.5 \quad \mu_y = 60$$

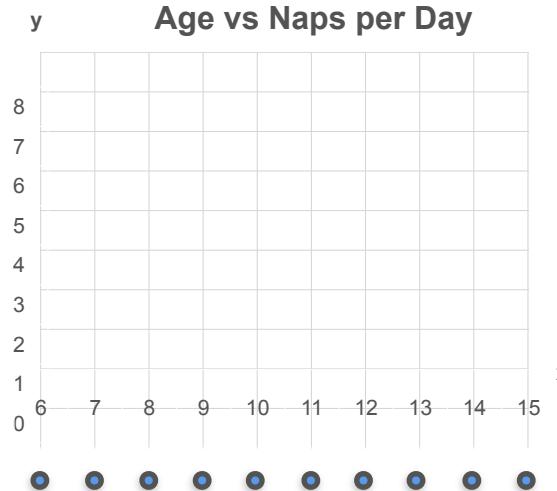
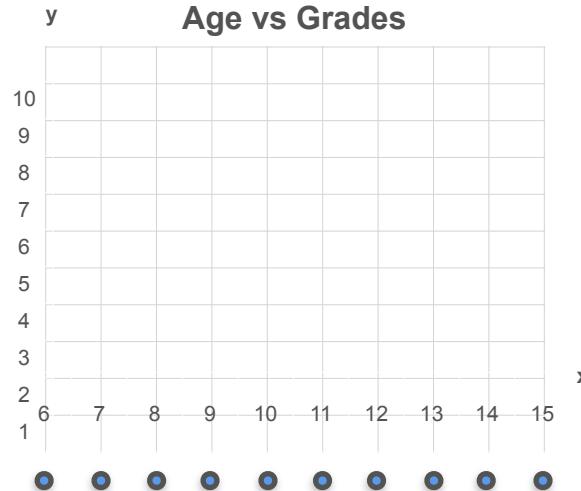
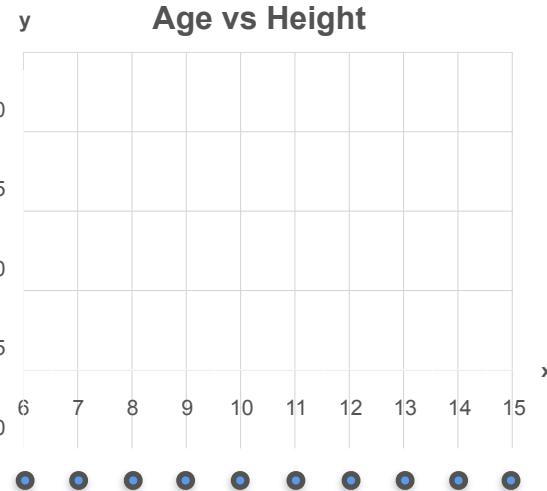


$$\mu_x = 10.5 \quad \mu_y = 5$$

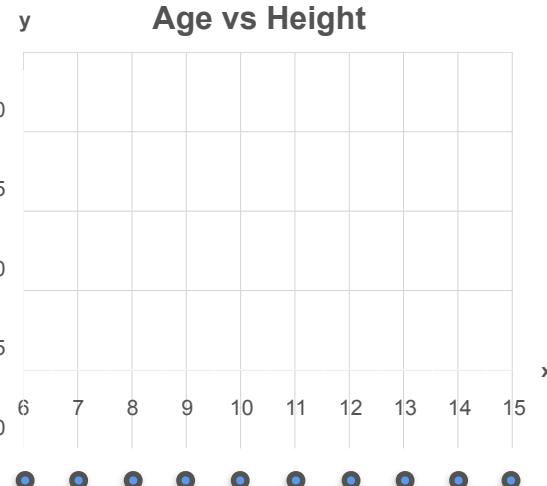


$$\mu_x = 10.5 \quad \mu_y = 3.7$$

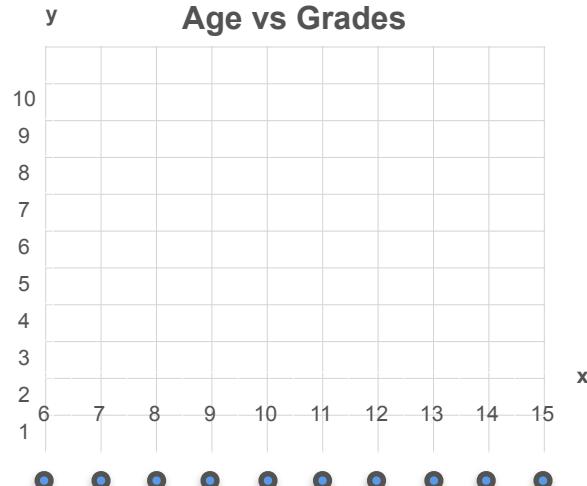
Horizontal (X) Variance



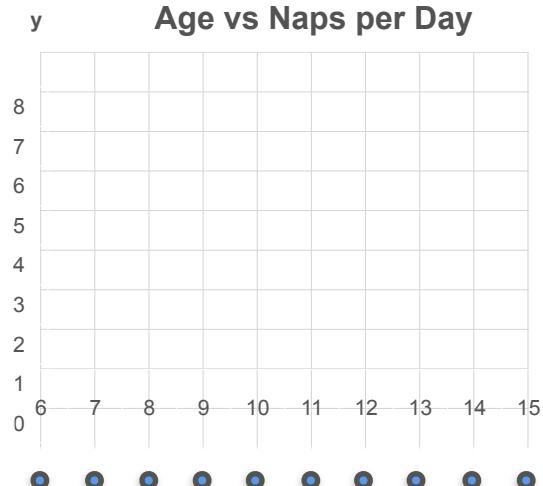
Horizontal (X) Variance



$$Var(X) = 9.17$$

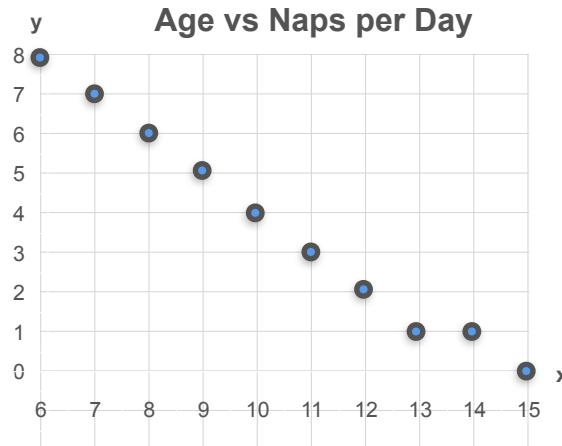
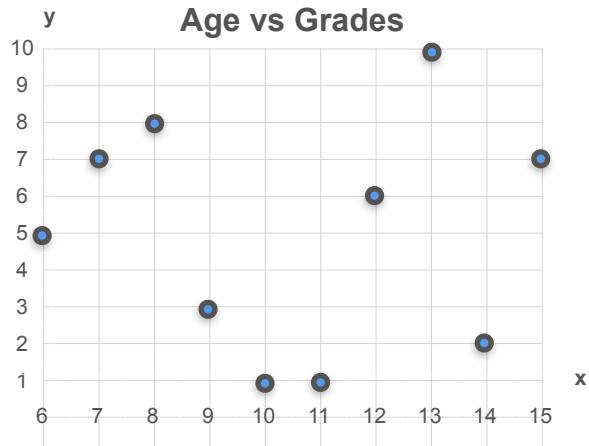
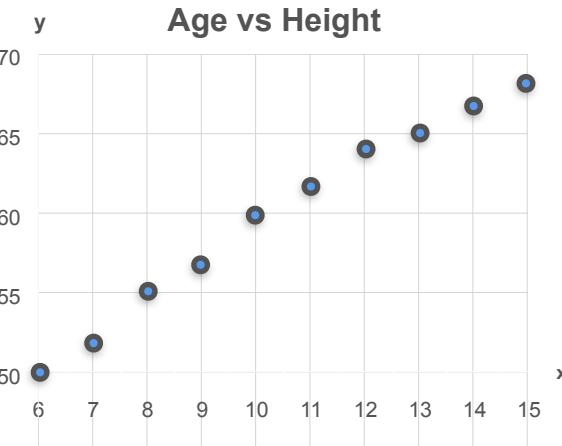


$$Var(X) = 9.17$$

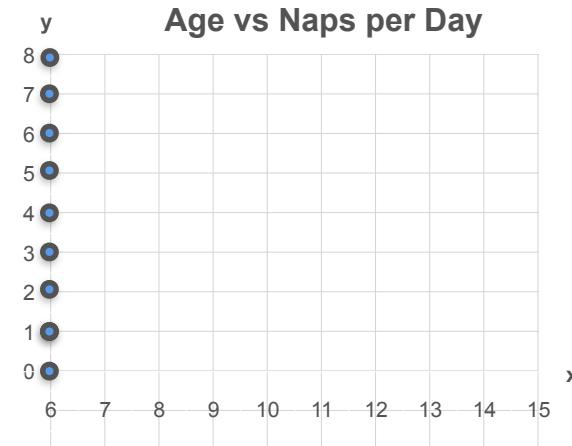
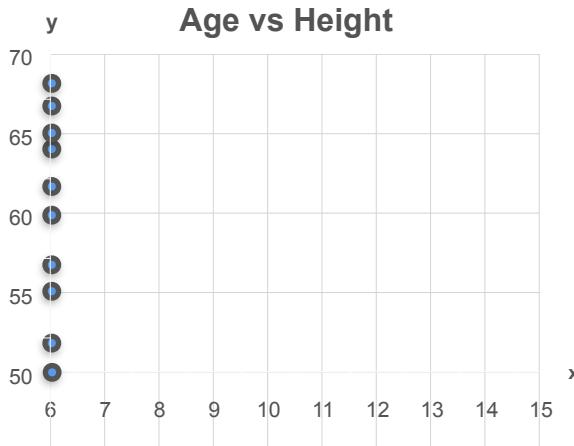


$$Var(X) = 9.17$$

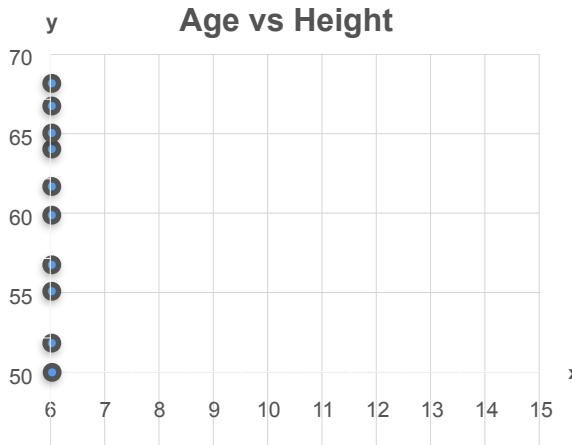
Anything Else?



Vertical (Y) Variance



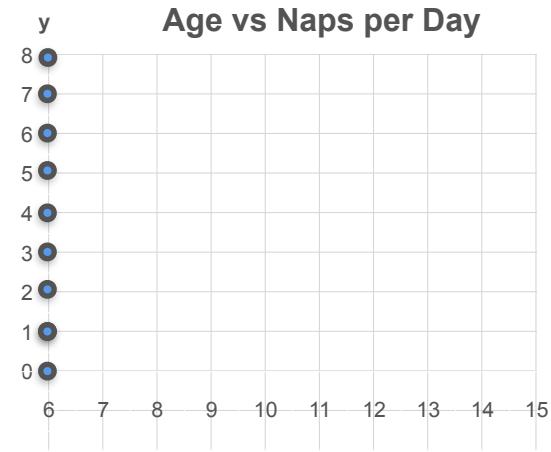
Vertical (Y) Variance



$$Var(Y) = 39.56$$

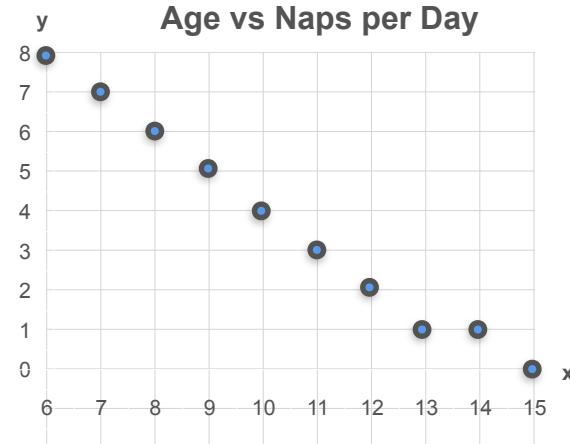
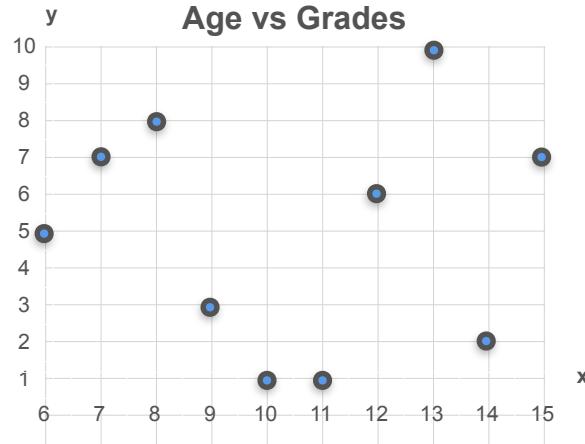
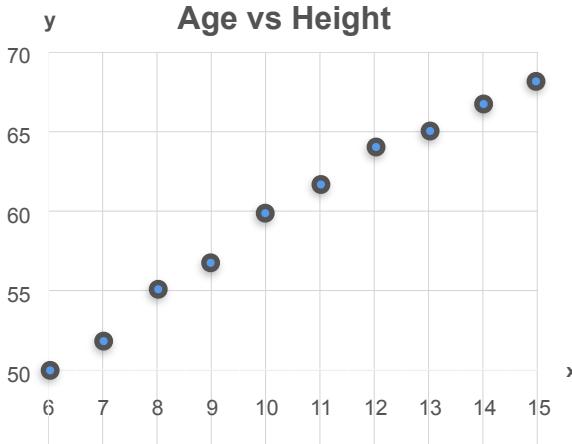


$$Var(Y) = 9.78$$

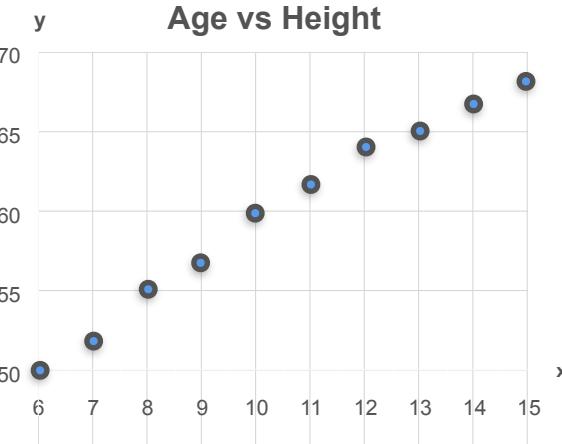


$$Var(Y) = 7.57$$

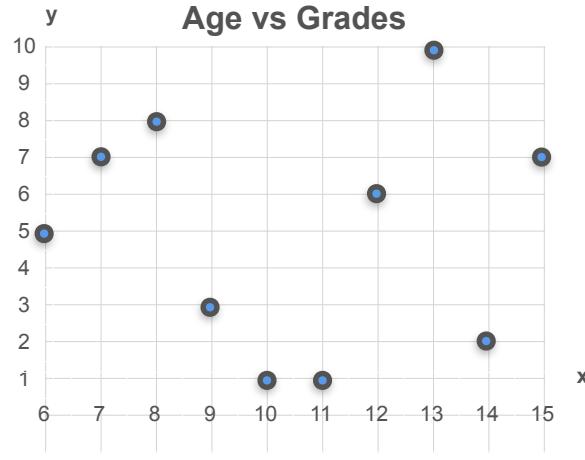
Still no Way To Compare Them



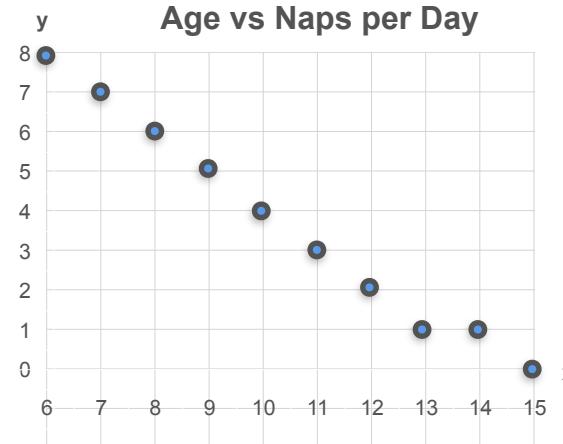
Still no Way To Compare Them



Covariance > 0



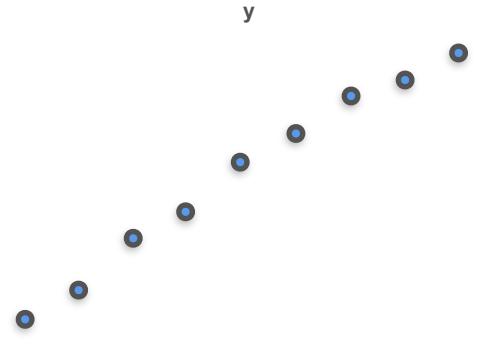
Covariance ≈ 0



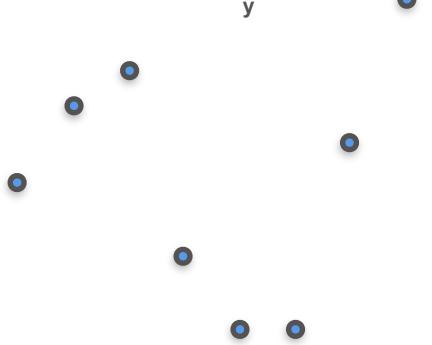
Covariance < 0

First Step: Center Them

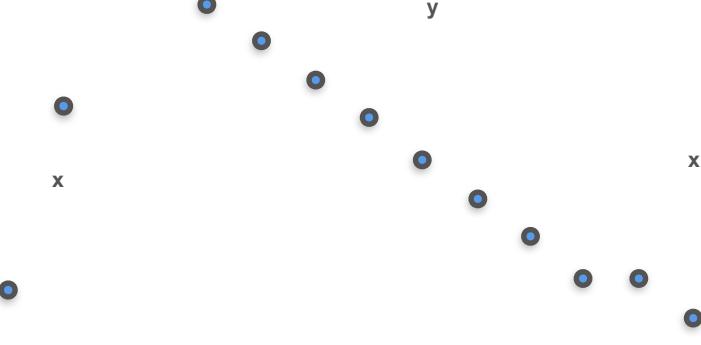
Age vs Height



Age vs Grades

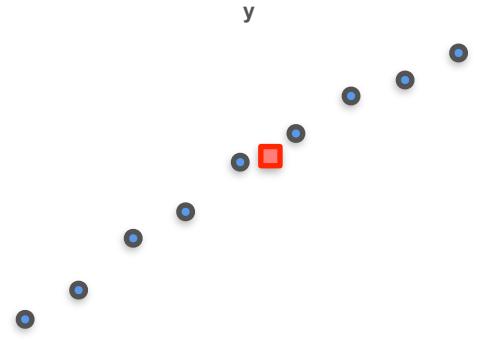


Age vs Naps per Day

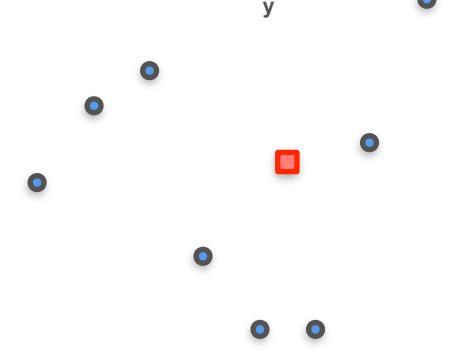


First Step: Center Them

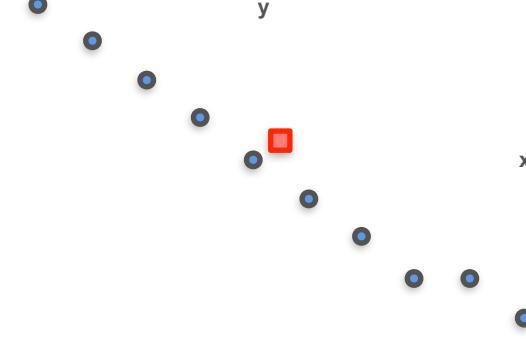
Age vs Height



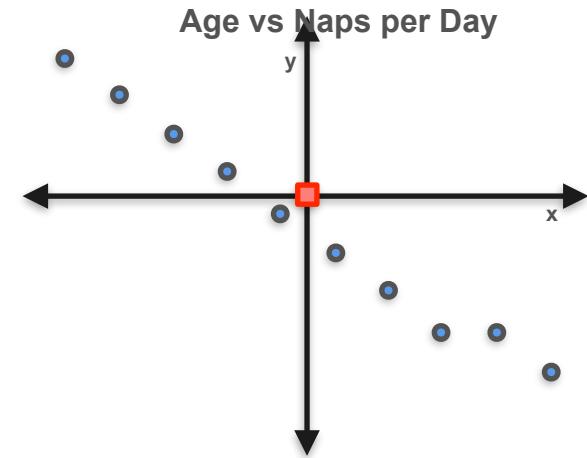
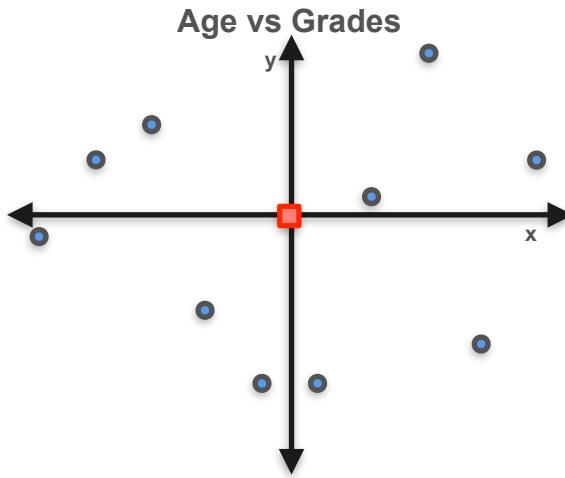
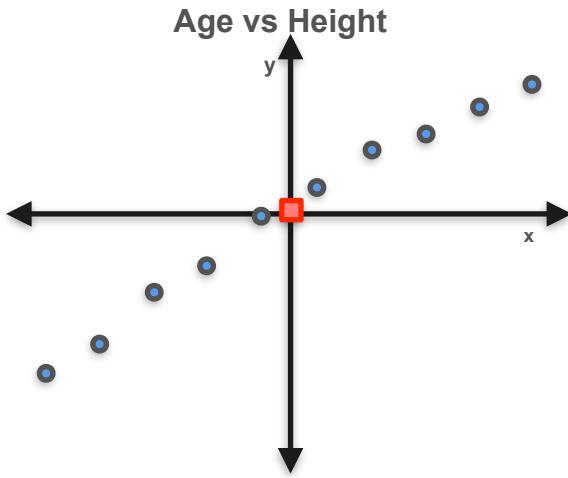
Age vs Grades



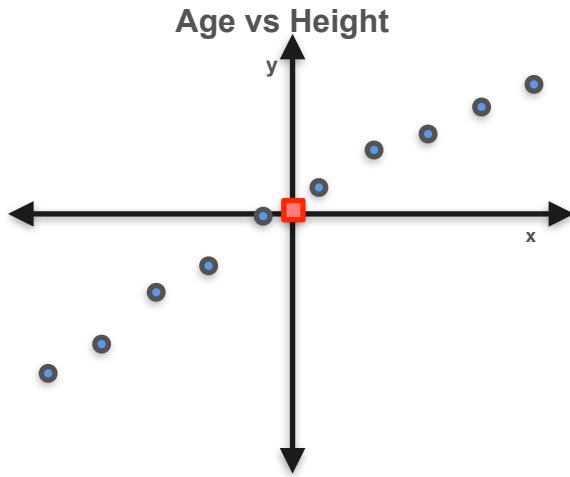
Age vs Naps per Day



First Step: Center Them



First Step: Center Them

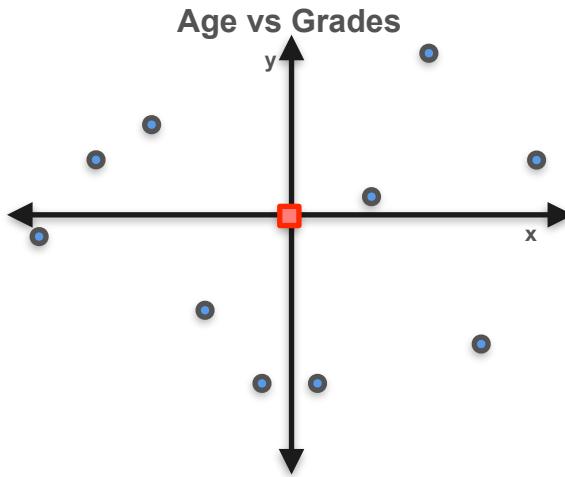


$$\mu_x = 0$$

$$\mu_y = 0$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

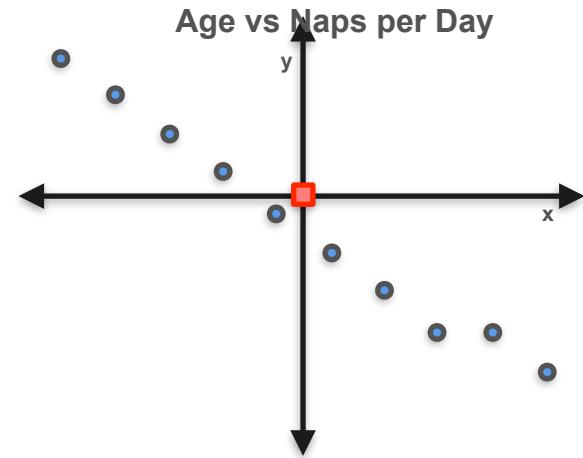


$$\mu_x = 0$$

$$\mu_y = 0$$

$$Var(X) = 1$$

$$Var(Y) = 1$$



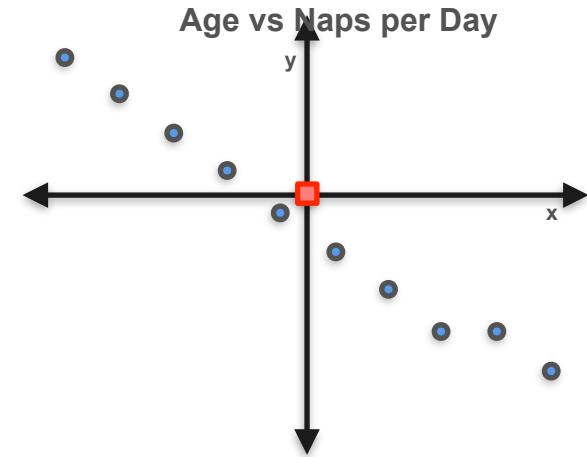
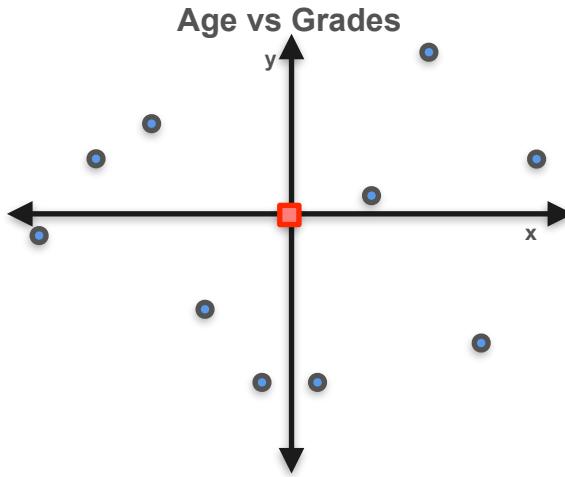
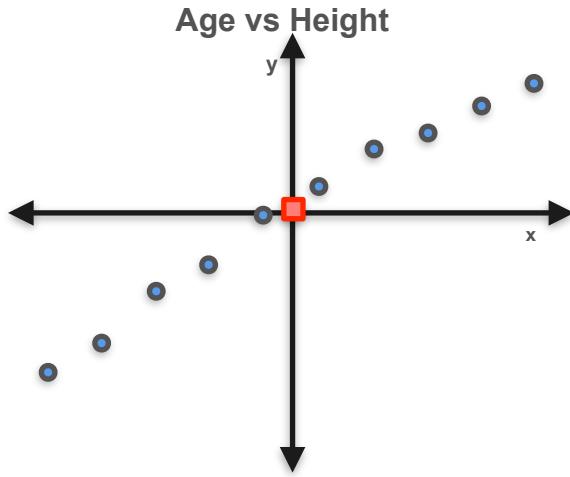
$$\mu_x = 0$$

$$\mu_y = 0$$

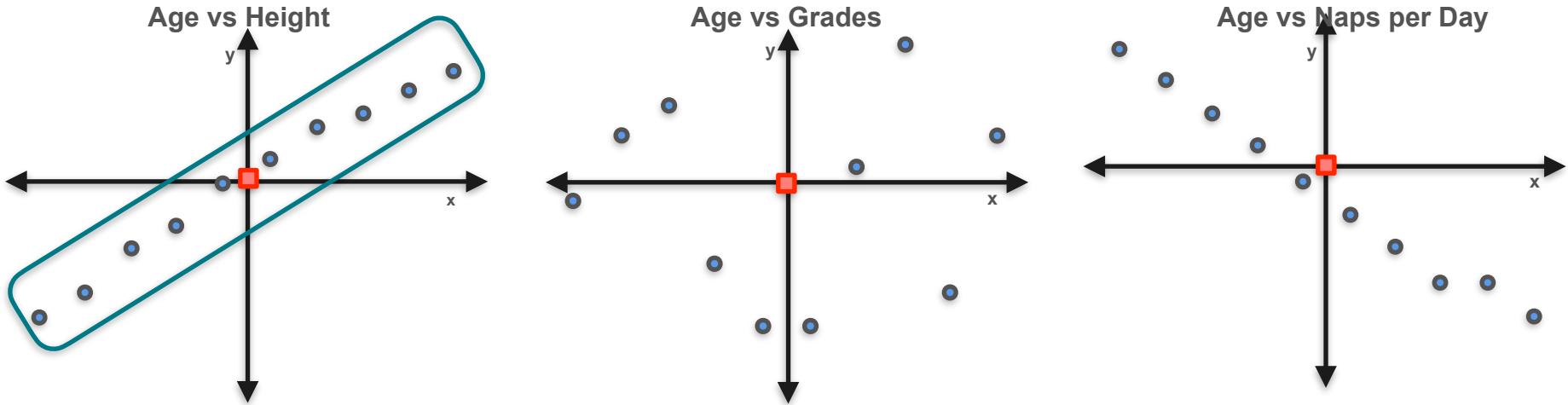
$$Var(X) = 1$$

$$Var(Y) = 1$$

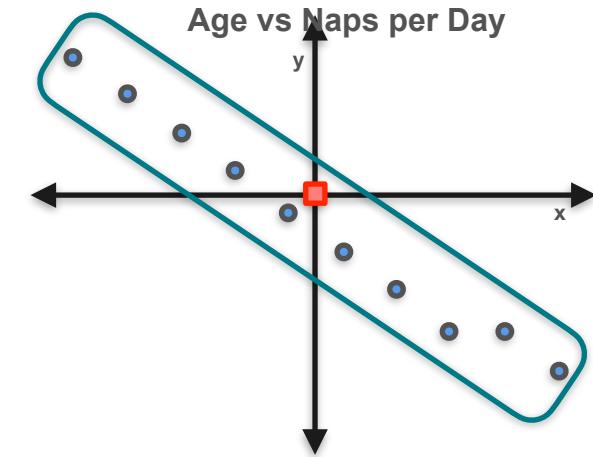
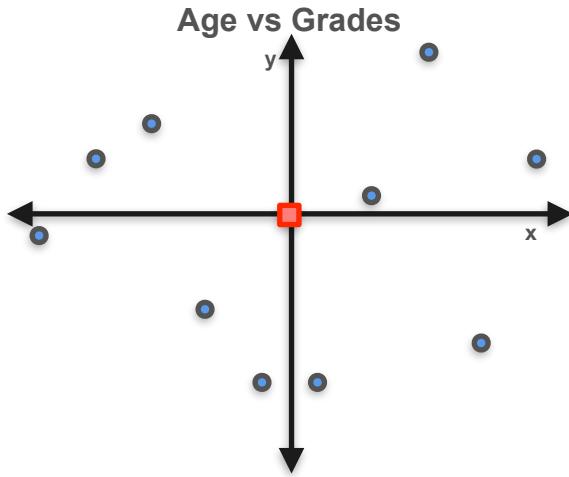
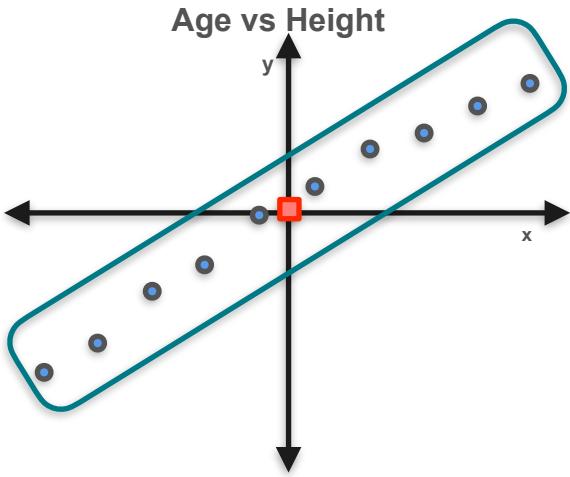
Second Step: Notice Trend



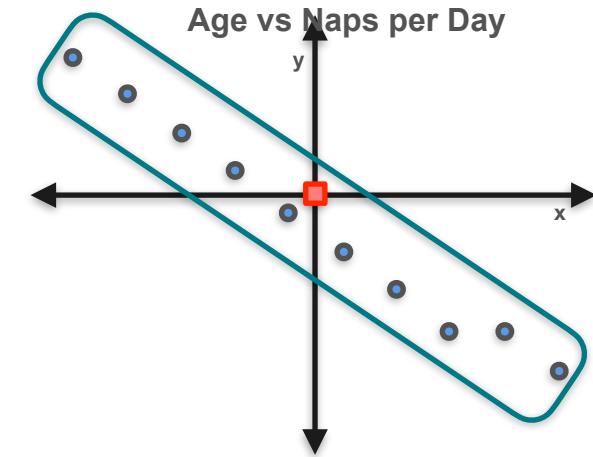
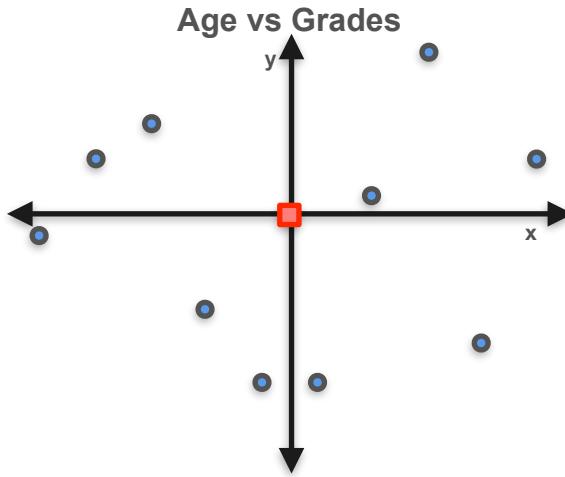
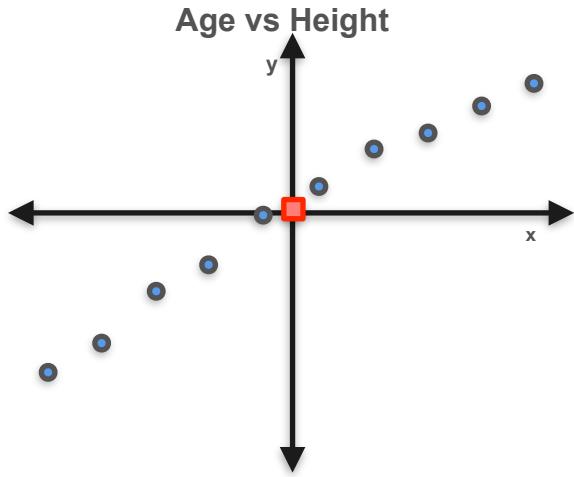
Second Step: Notice Trend



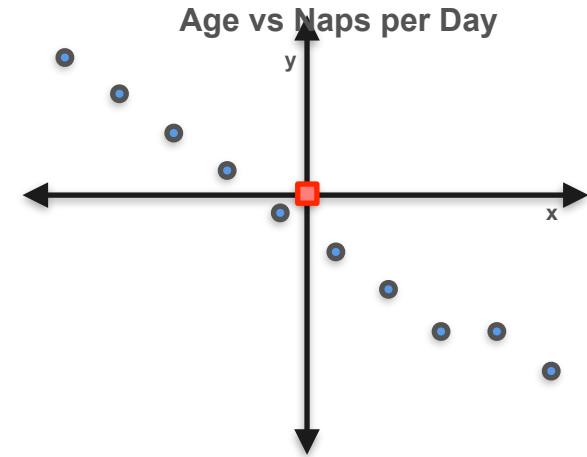
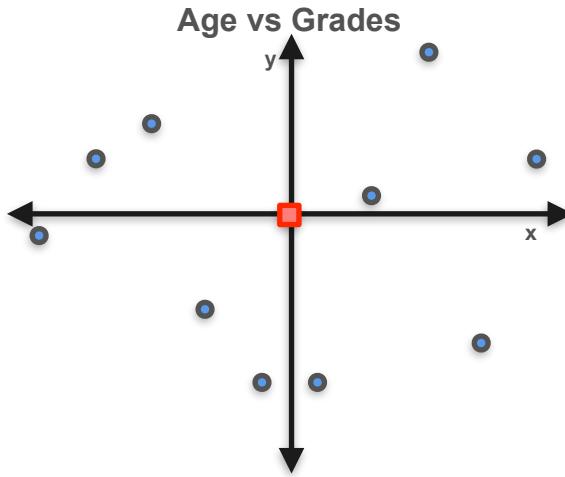
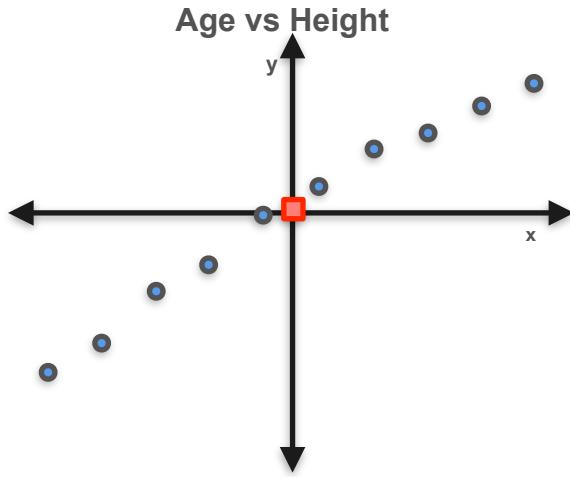
Second Step: Notice Trend



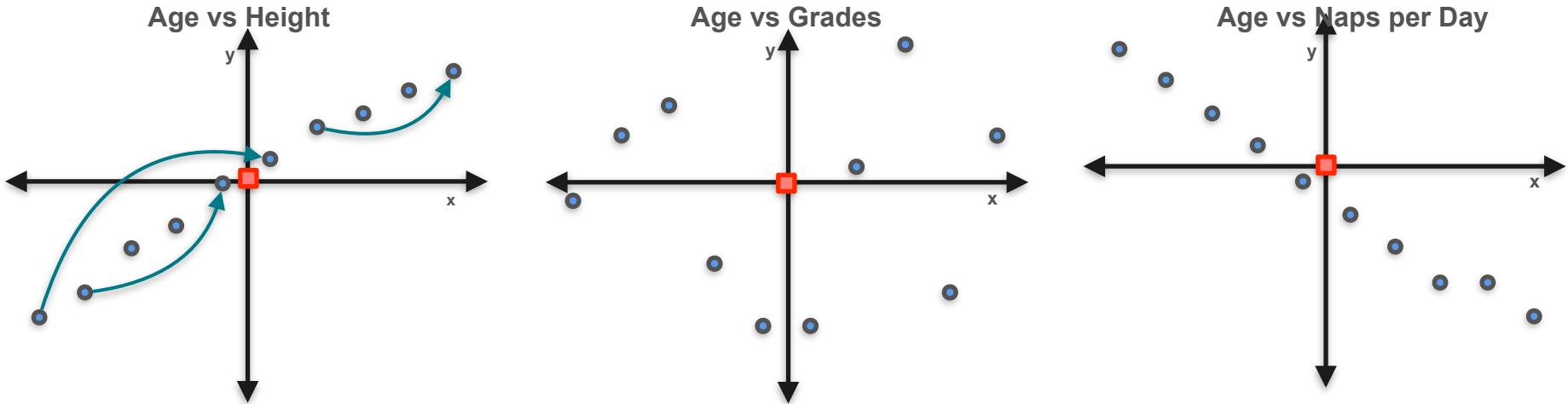
Second Step: Notice Trend



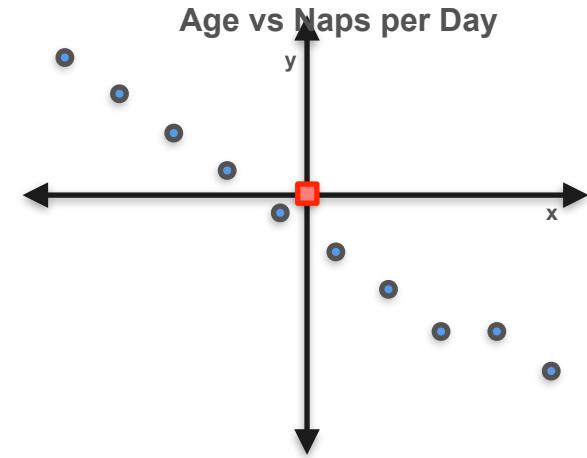
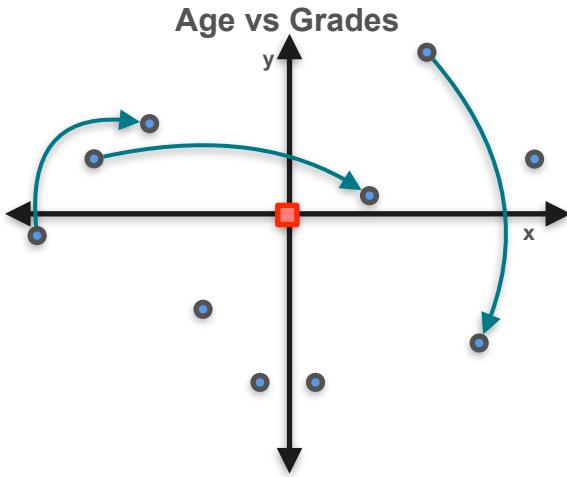
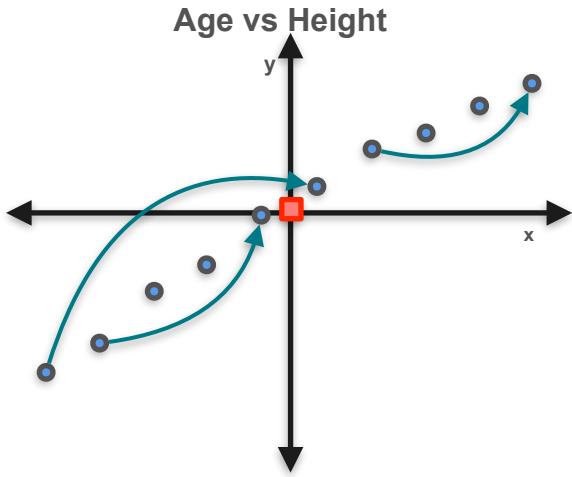
Second Step: Notice Trend



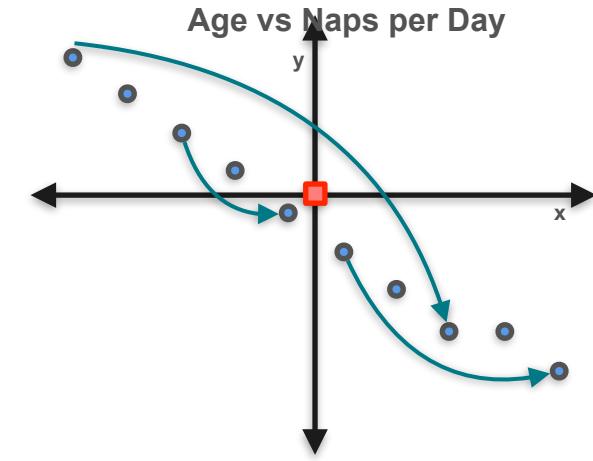
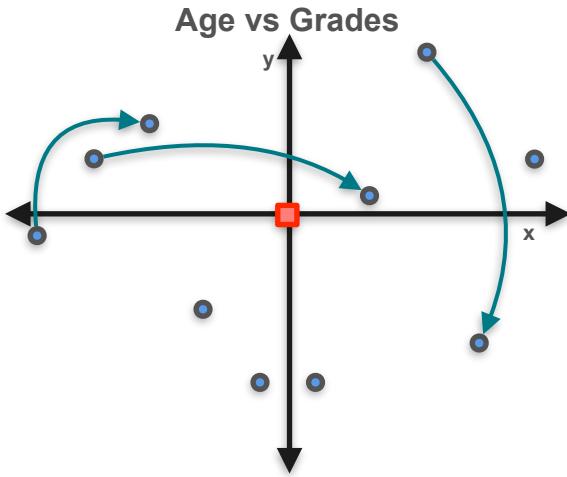
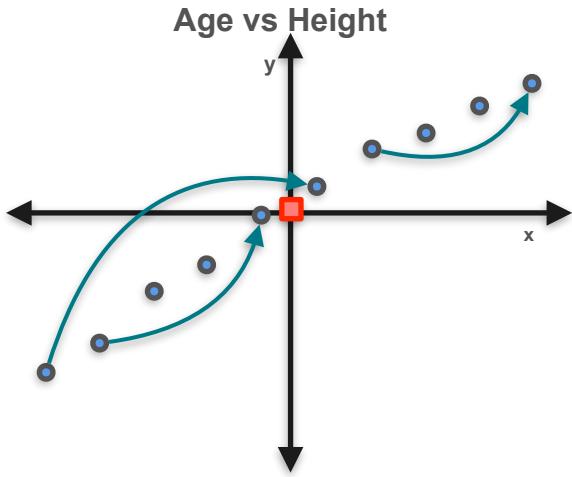
Second Step: Notice Trend



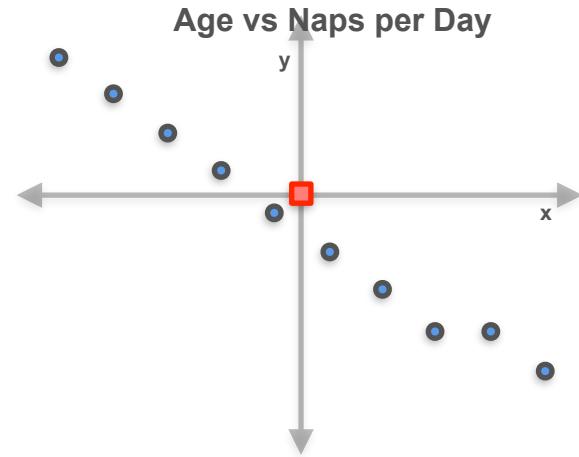
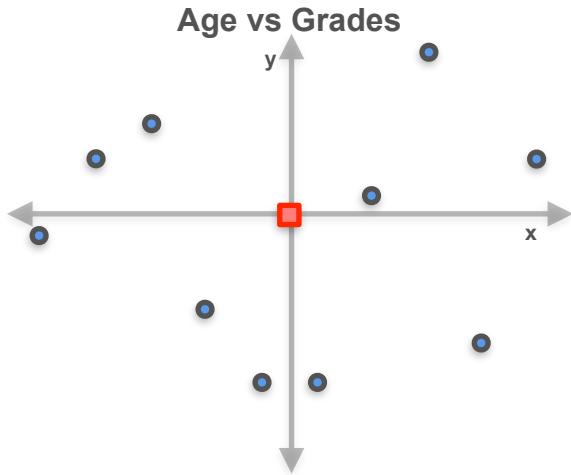
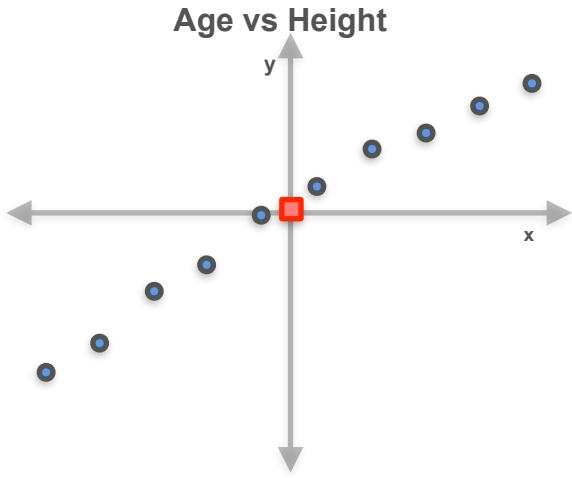
Second Step: Notice Trend



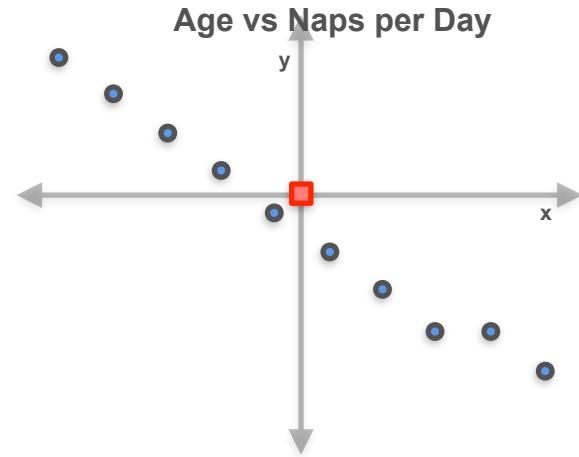
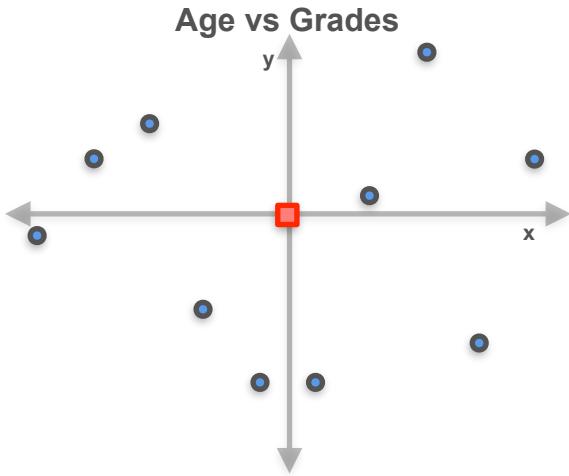
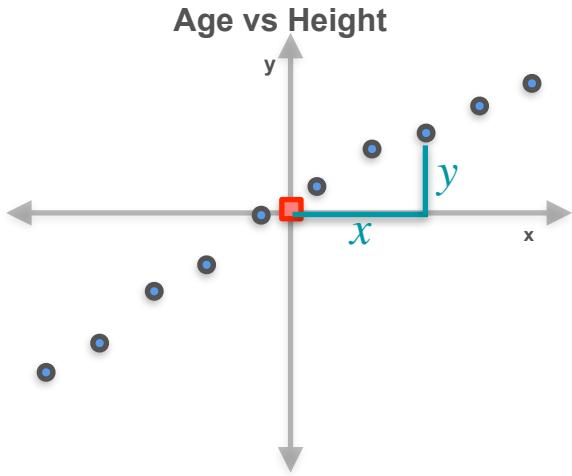
Second Step: Notice Trend



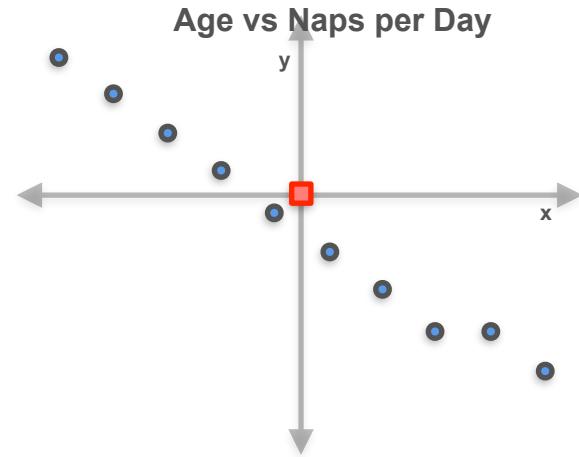
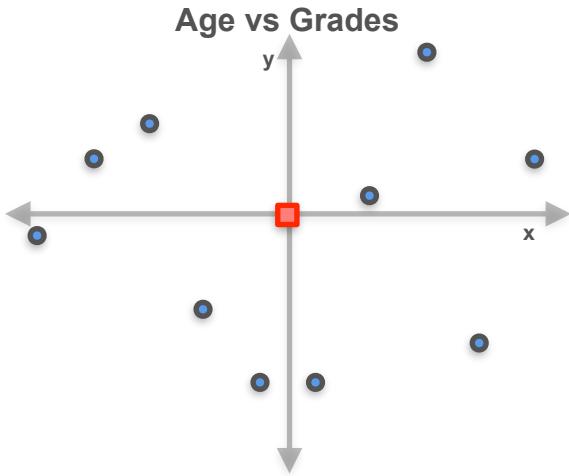
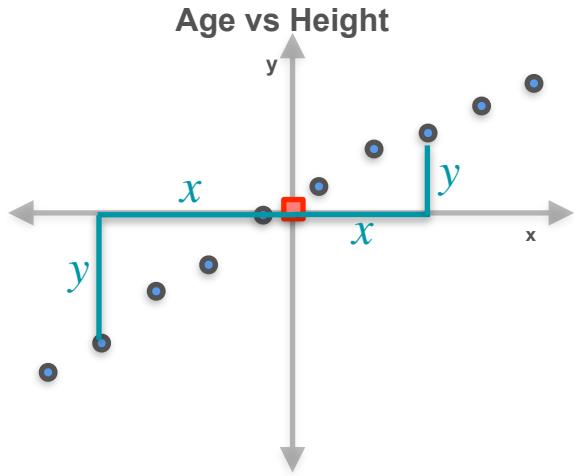
Positives and Negatives



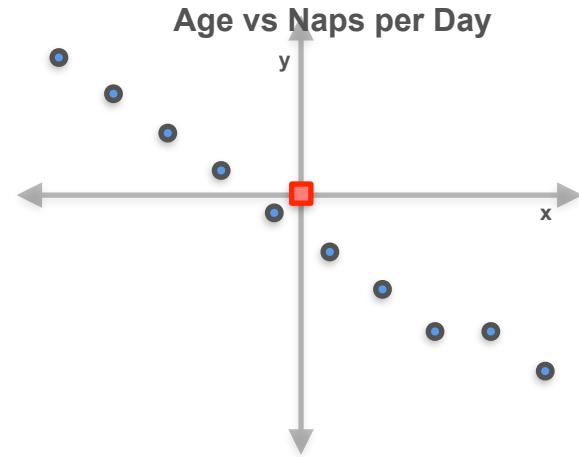
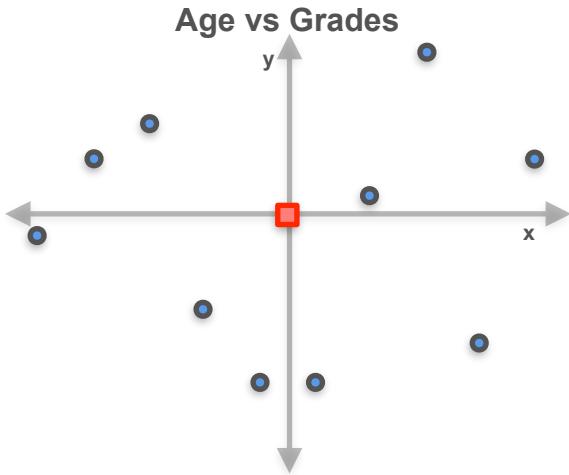
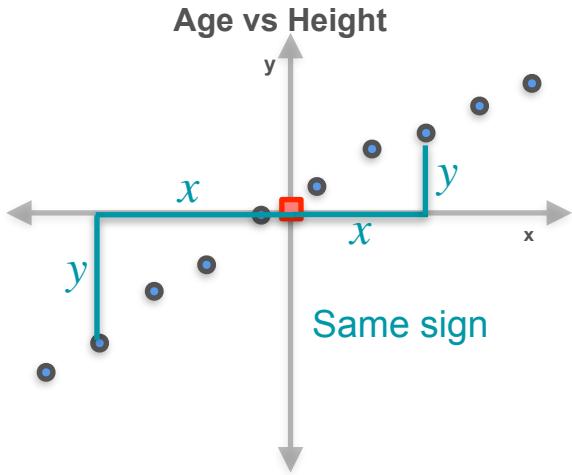
Positives and Negatives



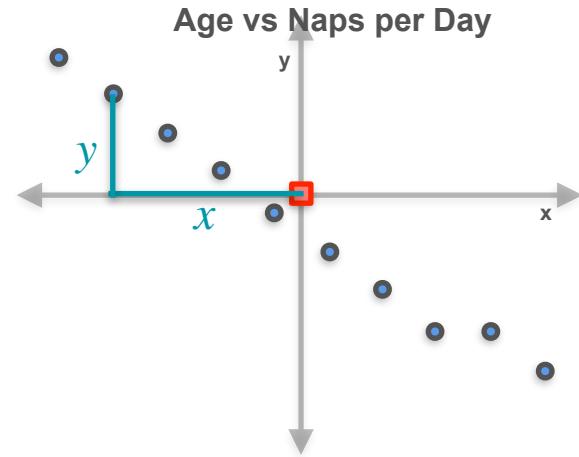
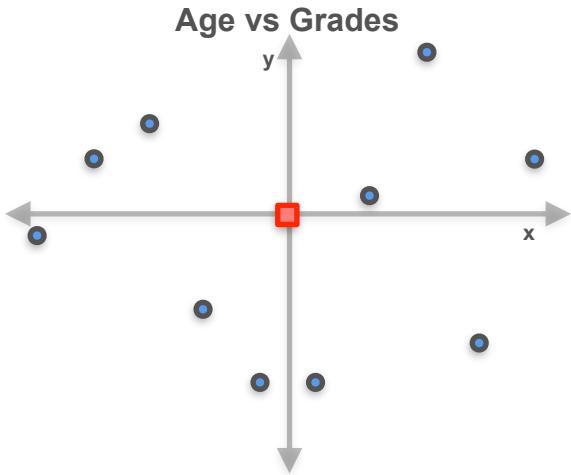
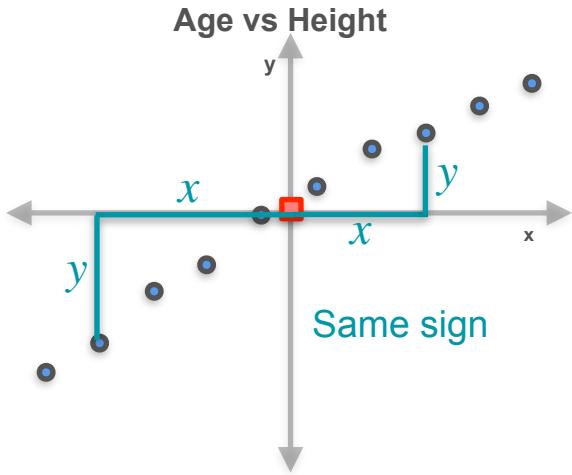
Positives and Negatives



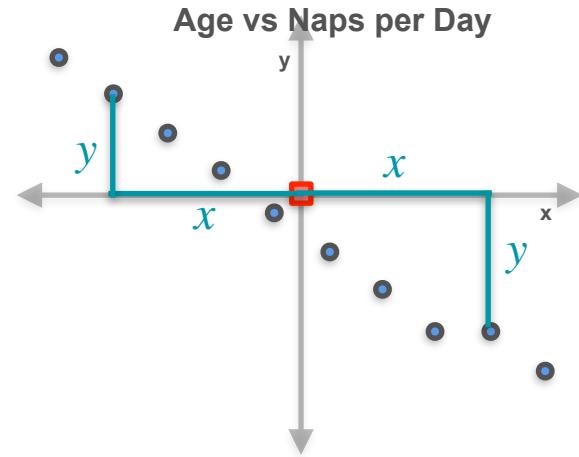
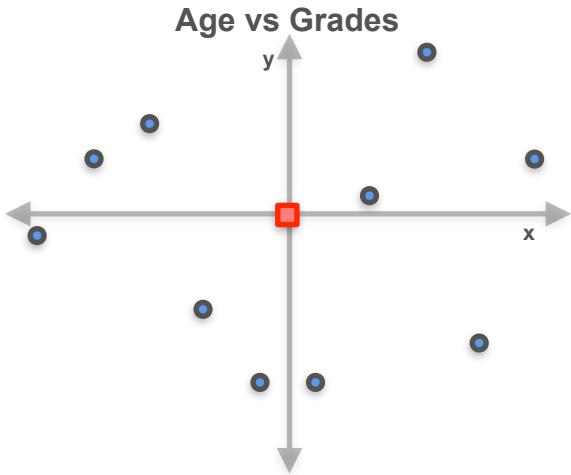
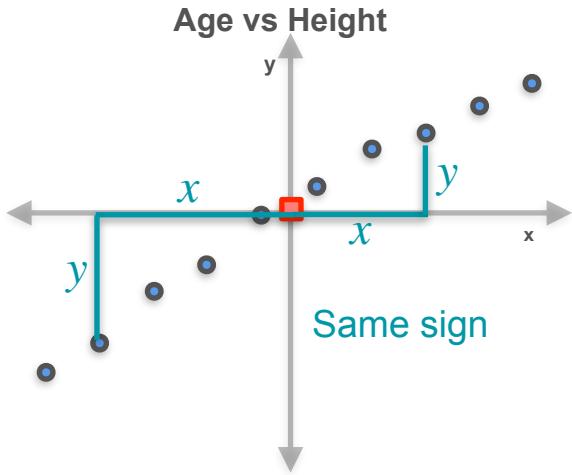
Positives and Negatives



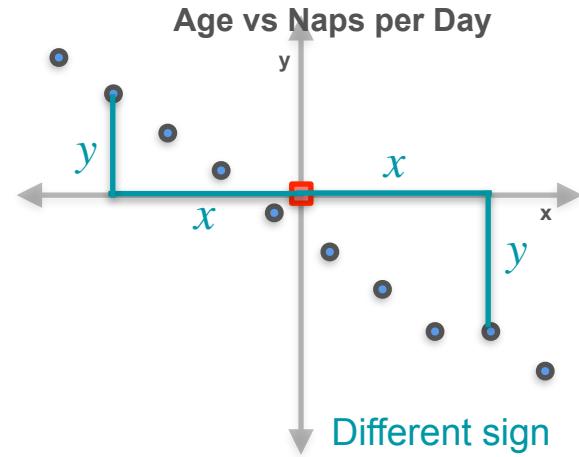
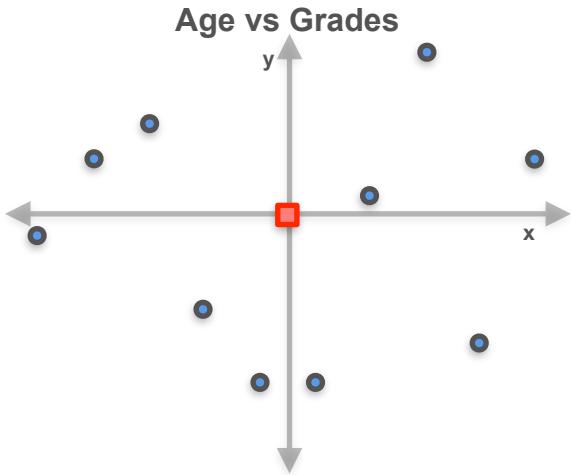
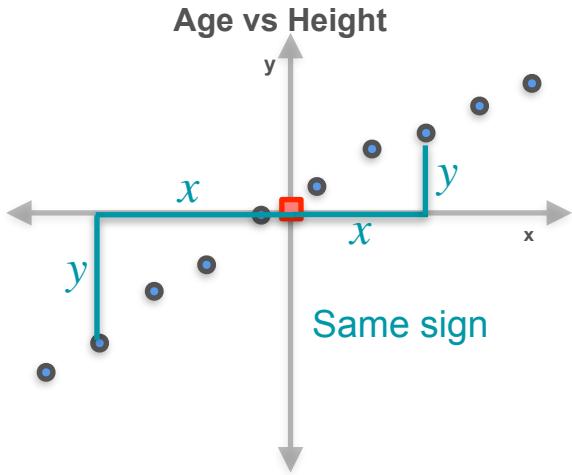
Positives and Negatives



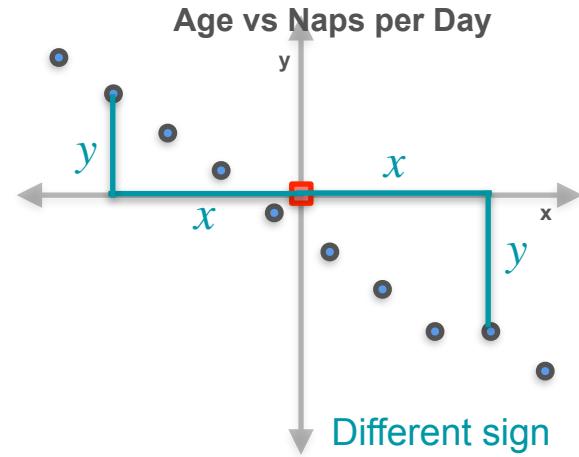
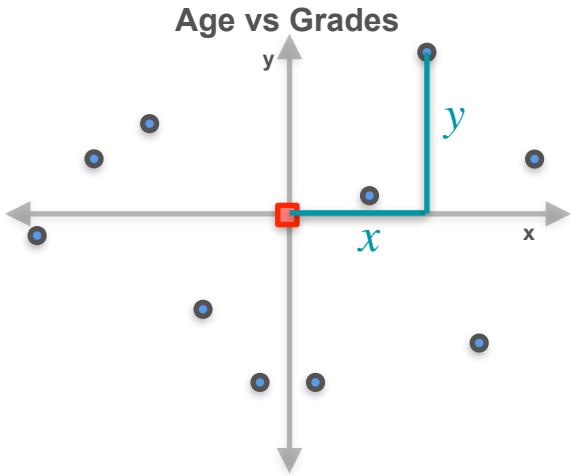
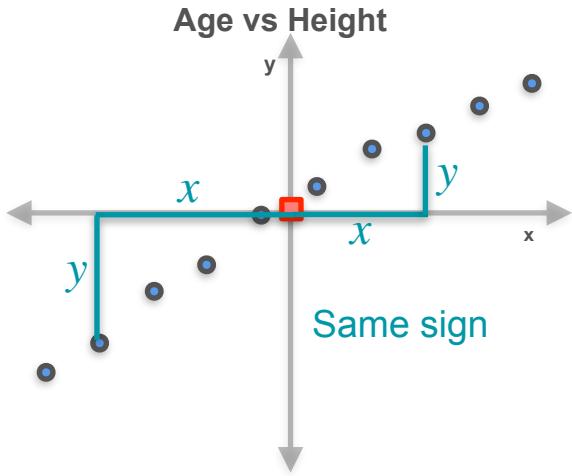
Positives and Negatives



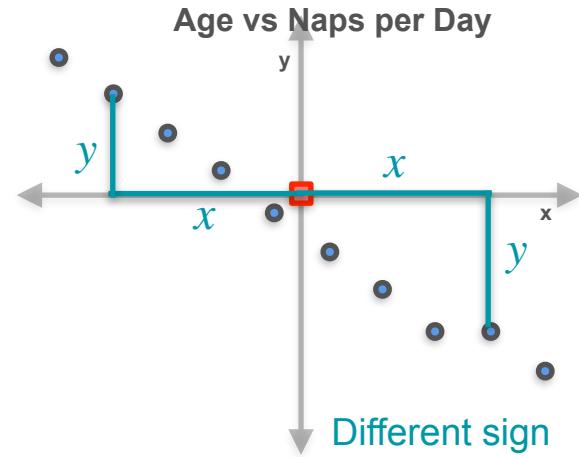
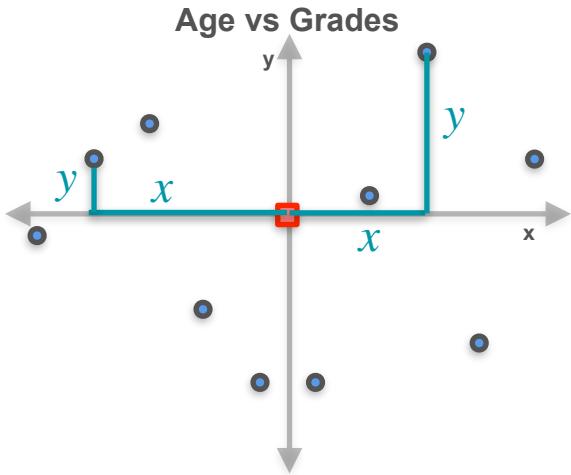
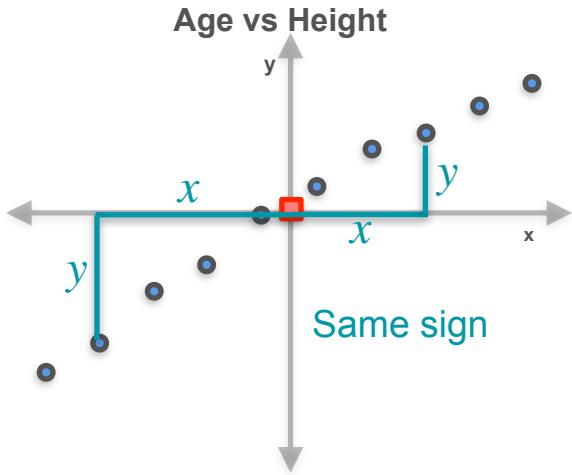
Positives and Negatives



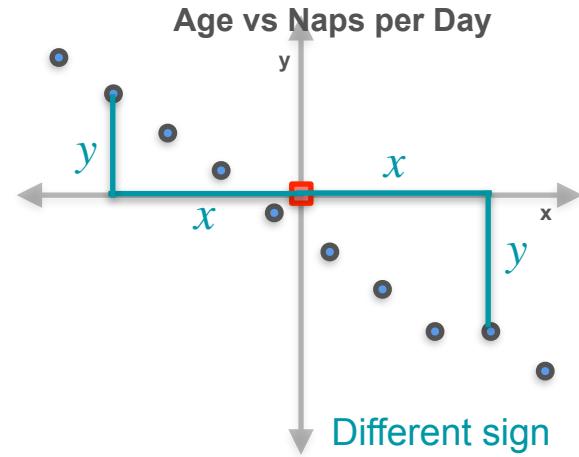
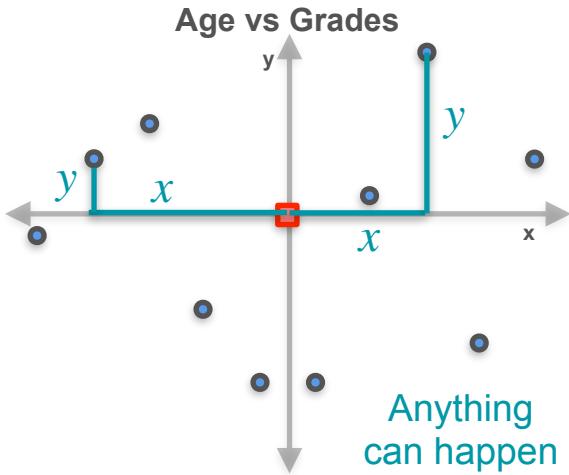
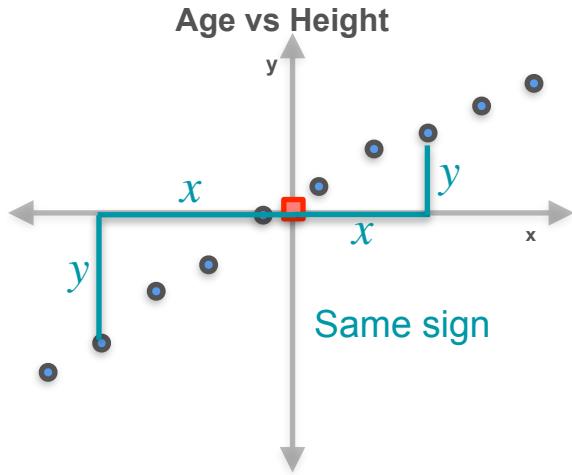
Positives and Negatives



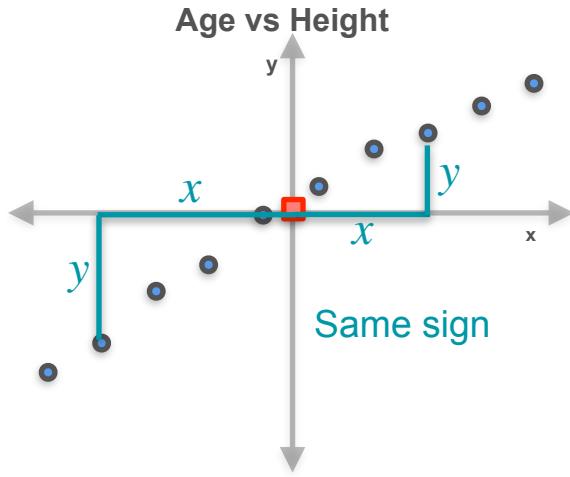
Positives and Negatives



Positives and Negatives

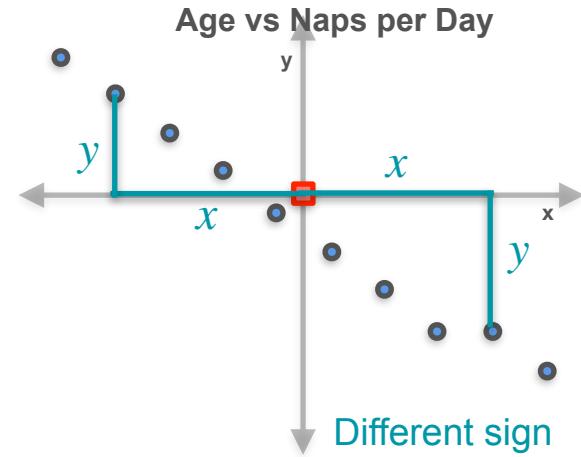
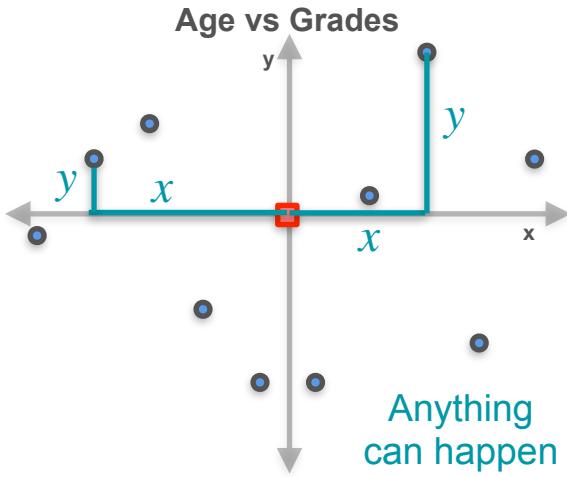


Positives and Negatives

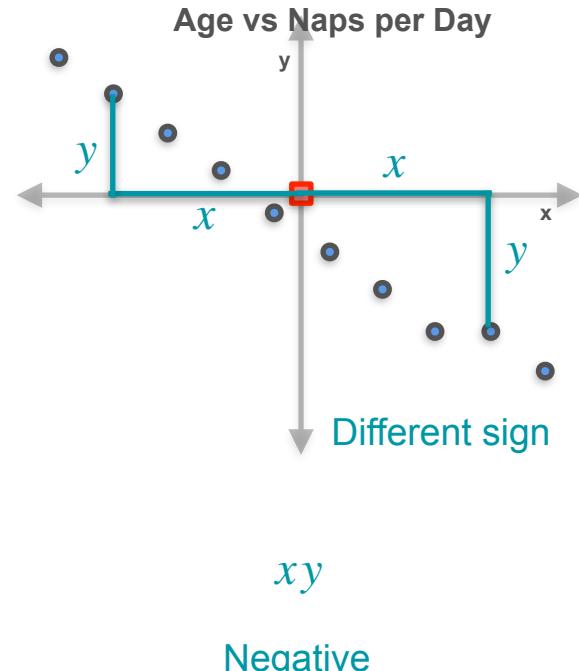
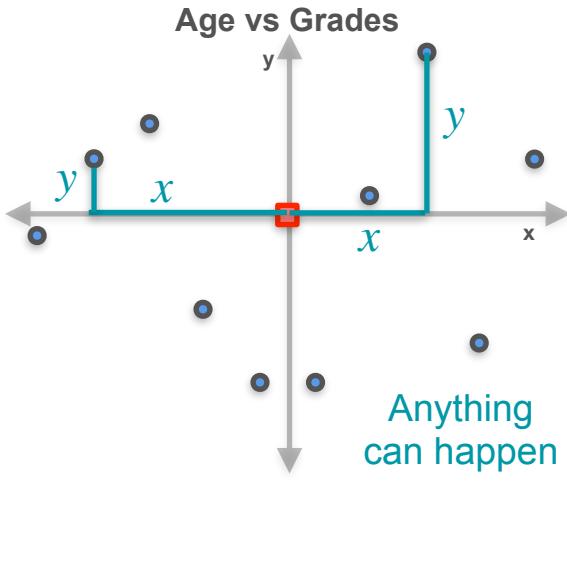
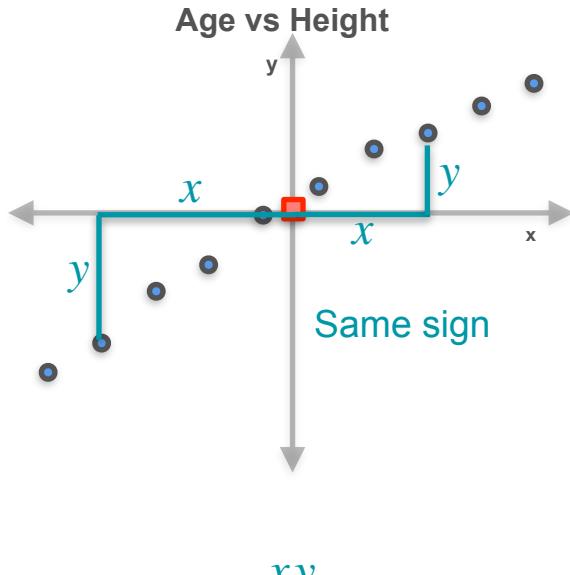


xy

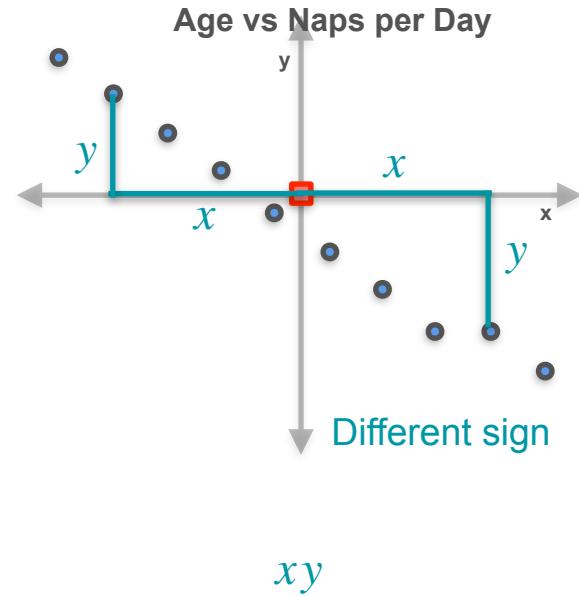
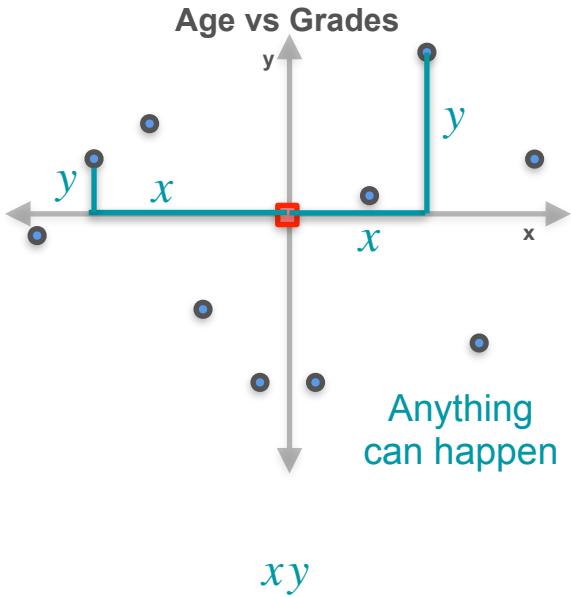
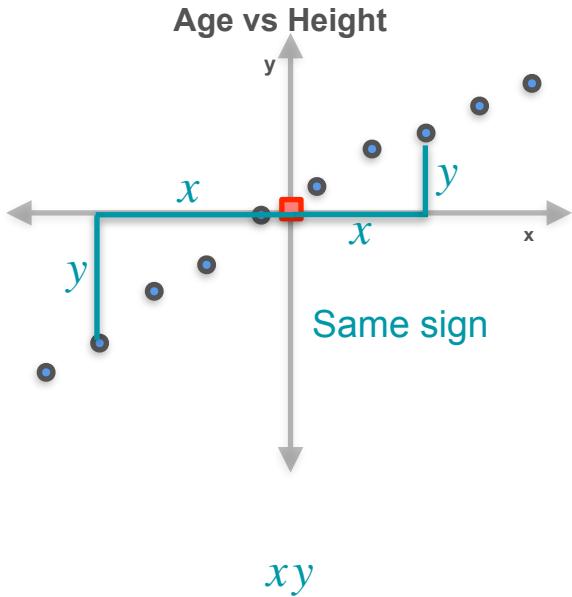
Positive



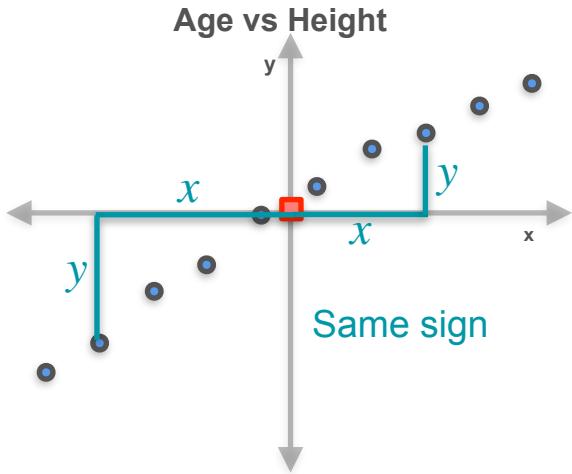
Positives and Negatives



Positives and Negatives

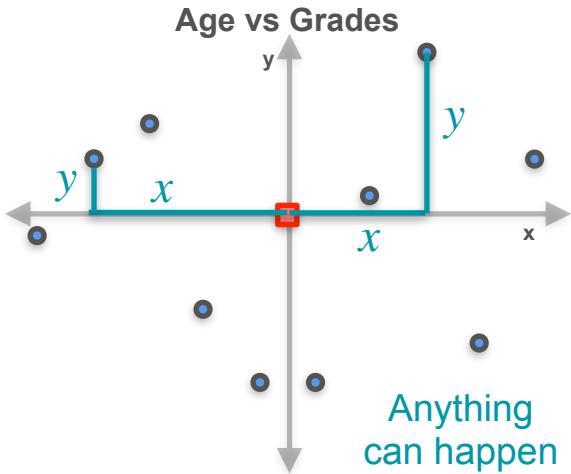


Positives and Negatives



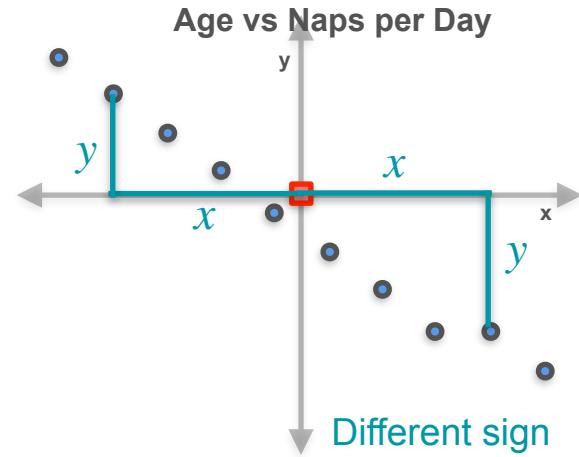
$$\sum xy$$

Positive



$$\sum xy$$

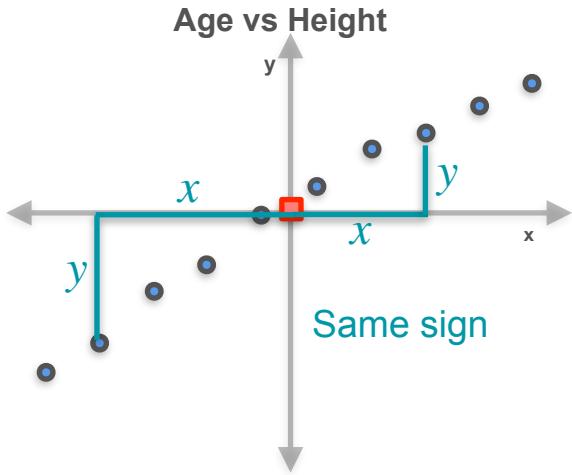
Both positive
and negative



$$\sum xy$$

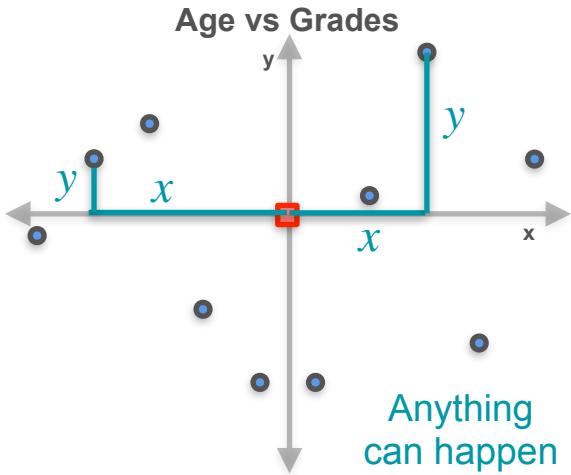
Negative

Positives and Negatives



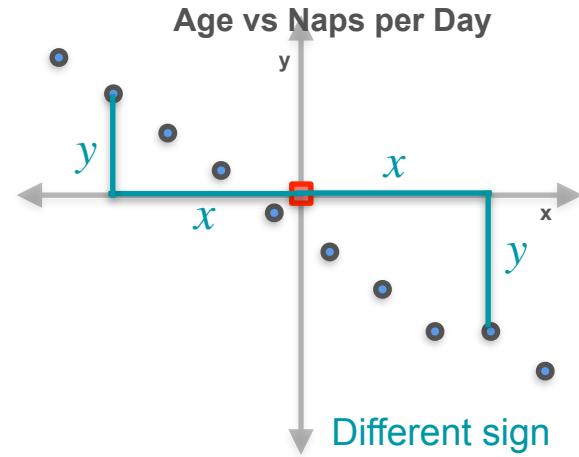
$$\sum xy > 0$$

Positive



$$\sum xy$$

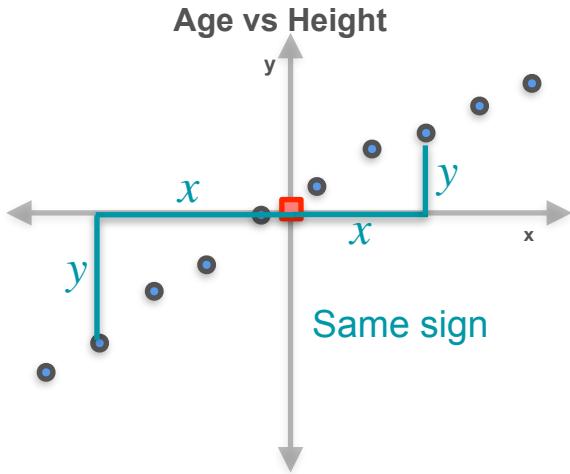
Both positive
and negative



$$\sum xy$$

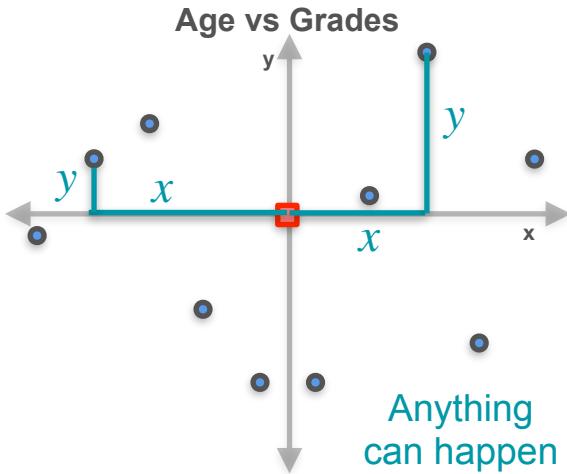
Negative

Positives and Negatives



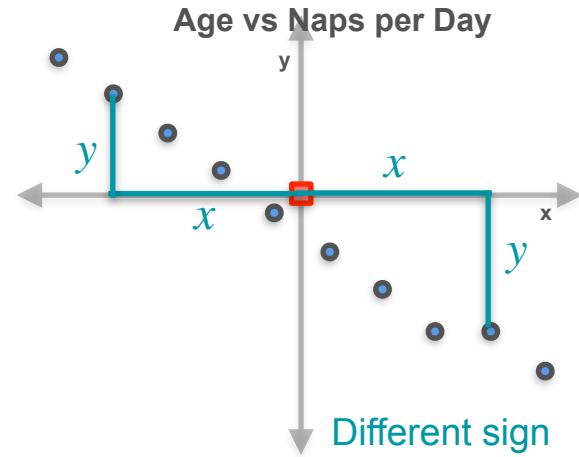
$$\sum xy > 0$$

Positive



$$\sum xy$$

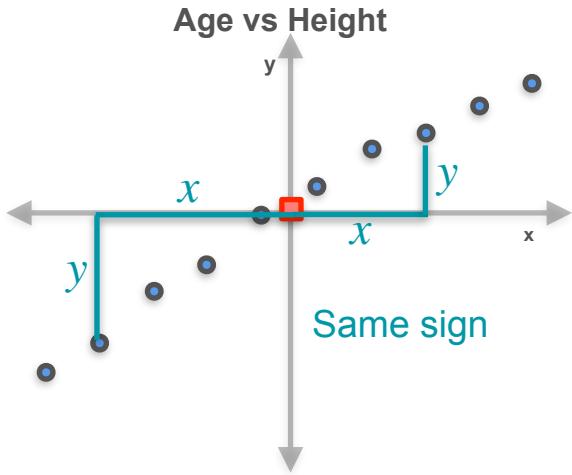
Both positive
and negative



$$\sum xy < 0$$

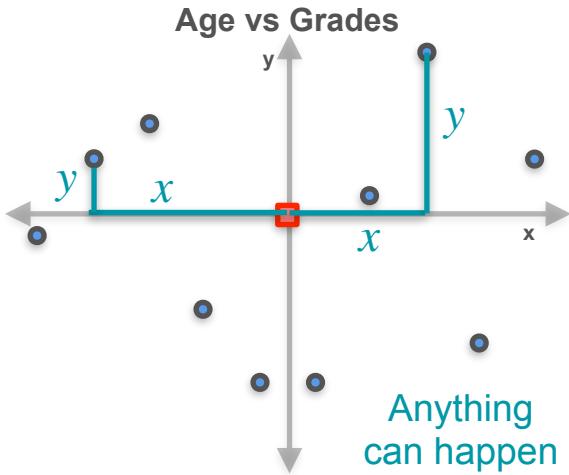
Negative

Positives and Negatives



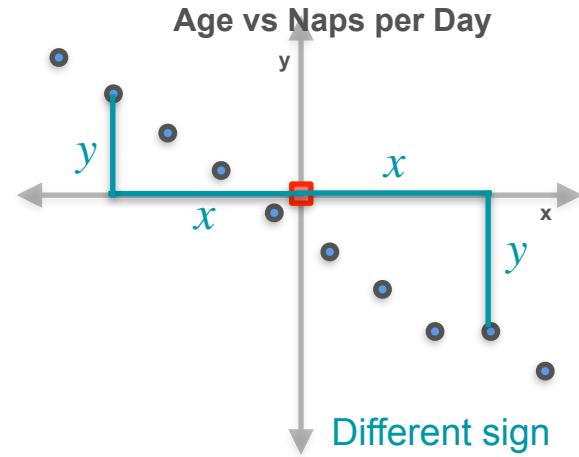
$$\sum xy > 0$$

Positive



$$\sum xy \approx 0$$

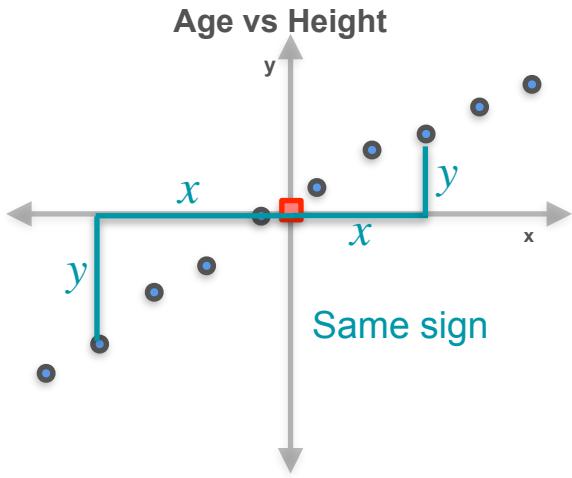
Both positive
and negative



$$\sum xy < 0$$

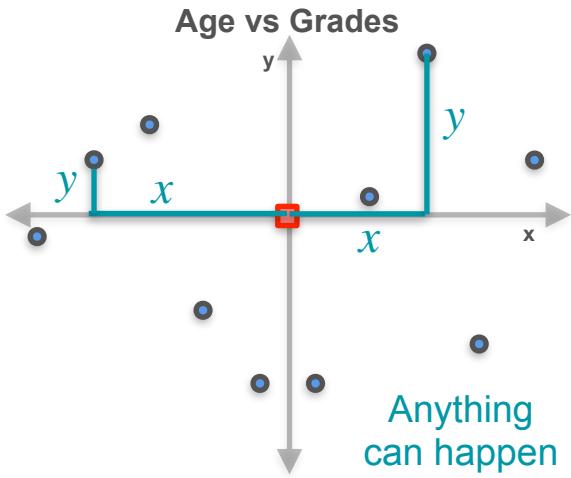
Negative

Covariance



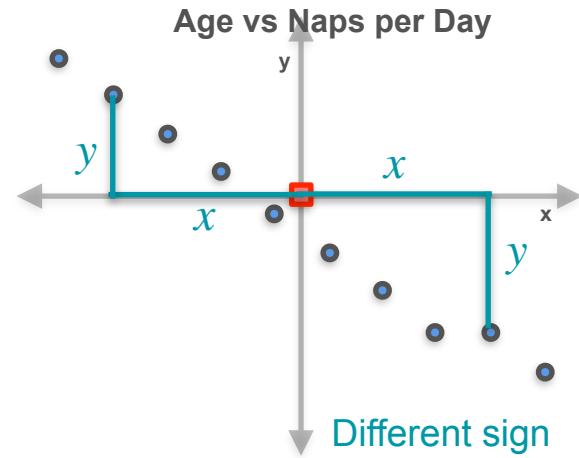
$$\sum xy > 0$$

Positive



$$\sum xy \approx 0$$

Both positive
and negative



$$\sum xy < 0$$

Negative

Covariance

Covariance

$$Cov(X, Y) = \sum xy$$

Covariance

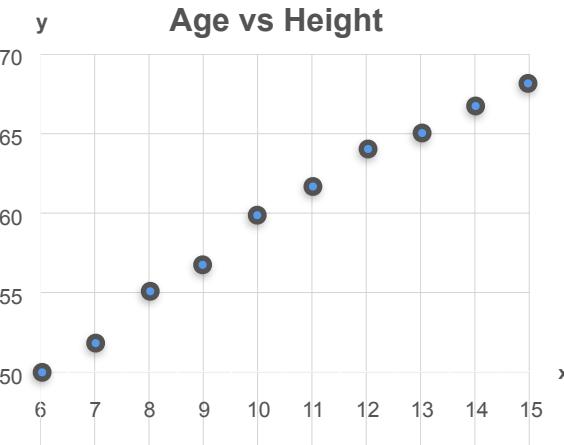
$$Cov(X, Y) = \sum xy \quad \text{Almost...}$$

Covariance

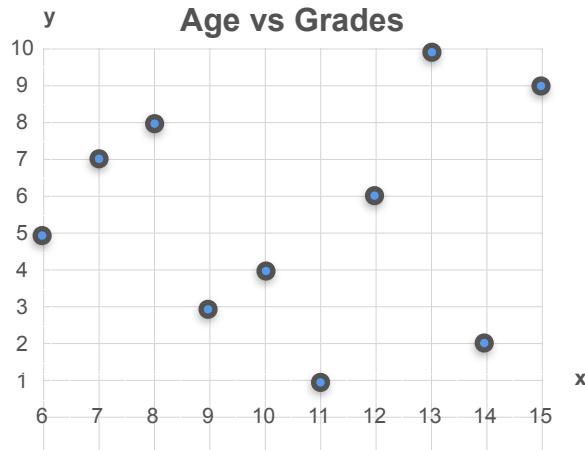
$$Cov(X, Y) = \sum xy \quad \text{Almost...}$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

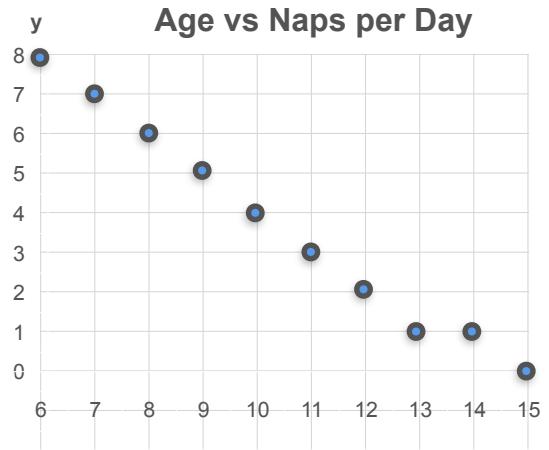
Covariance



$$\text{Cov}(X, Y) > 0$$



$$\text{Cov}(X, Y) \approx 0$$



$$\text{Cov}(X, Y) < 0$$

Covariance Formula

Covariance Formula

Age vs Height

Covariance Formula

Age vs Height

Covariance > 0

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

Age (x)	Height (y)
6	50
7	52
8	55
9	57
10	60
11	62
12	64
13	65
14	67
15	68

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	
7	52	-3.5	-8	
8	55	-2.5	-5	
9	57	-1.5	-3	
10	60	-0.5	0	
11	62	0.5	2	
12	64	1.5	4	
13	65	2.5	5	
14	67	3.5	7	
15	68	4.5	8	

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

$$\sum = 170$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

$$\sum = 170$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{170}{10}$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

$$\sum = 170$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{170}{10} = 17$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

$$\sum = 170$$

Covariance Formula

Age vs Height

Covariance > 0

$$\mu_x = 10.5 \quad \mu_y = 60$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{170}{10} = 17 > 0$$

Age (x)	Height (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	50	-4.5	-10	45
7	52	-3.5	-8	28
8	55	-2.5	-5	12.5
9	57	-1.5	-3	4.5
10	60	-0.5	0	0
11	62	0.5	2	1
12	64	1.5	4	6
13	65	2.5	5	12.5
14	67	3.5	7	24.5
15	68	4.5	8	36

$$\sum = 170$$

Covariance Formula

Covariance Formula

Age vs Naps per Day

Covariance Formula

Age vs Naps per Day

Covariance < 0

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5$$

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

Age (x)
6
7
8
9
10
11
12
13
14
15

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

Age (x)	Naps (y)
6	8
7	7
8	6
9	5
10	4
11	3
12	2
13	1
14	1
15	0

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

Age (x)	Naps (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	8	-4.5	4.3	-19.35
7	7	-3.5	3.3	-11.55
8	6	-2.5	2.3	-5.75
9	5	-1.5	1.3	-1.95
10	4	-0.5	0.3	-0.15
11	3	0.5	-0.7	-0.35
12	2	1.5	-1.7	-2.55
13	1	2.5	-2.7	-6.75
14	1	3.5	-2.7	-9.45
15	0	4.5	-3.7	-16.65

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Age (x)	Naps (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	8	-4.5	4.3	-19.35
7	7	-3.5	3.3	-11.55
8	6	-2.5	2.3	-5.75
9	5	-1.5	1.3	-1.95
10	4	-0.5	0.3	-0.15
11	3	0.5	-0.7	-0.35
12	2	1.5	-1.7	-2.55
13	1	2.5	-2.7	-6.75
14	1	3.5	-2.7	-9.45
15	0	4.5	-3.7	-16.65

$$\sum = -74.5$$

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{-74.5}{10}$$

Age (x)	Naps (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	8	-4.5	4.3	-19.35
7	7	-3.5	3.3	-11.55
8	6	-2.5	2.3	-5.75
9	5	-1.5	1.3	-1.95
10	4	-0.5	0.3	-0.15
11	3	0.5	-0.7	-0.35
12	2	1.5	-1.7	-2.55
13	1	2.5	-2.7	-6.75
14	1	3.5	-2.7	-9.45
15	0	4.5	-3.7	-16.65

$$\sum = -74.5$$

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{-74.5}{10} = -7.45$$

Age (x)	Naps (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	8	-4.5	4.3	-19.35
7	7	-3.5	3.3	-11.55
8	6	-2.5	2.3	-5.75
9	5	-1.5	1.3	-1.95
10	4	-0.5	0.3	-0.15
11	3	0.5	-0.7	-0.35
12	2	1.5	-1.7	-2.55
13	1	2.5	-2.7	-6.75
14	1	3.5	-2.7	-9.45
15	0	4.5	-3.7	-16.65

$$\sum = -74.5$$

Covariance Formula

Age vs Naps per Day

Covariance < 0

$$\mu_x = 10.5 \quad \mu_y = 3.7$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{-74.5}{10} = -7.45 < 0$$

Age (x)	Naps (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	8	-4.5	4.3	-19.35
7	7	-3.5	3.3	-11.55
8	6	-2.5	2.3	-5.75
9	5	-1.5	1.3	-1.95
10	4	-0.5	0.3	-0.15
11	3	0.5	-0.7	-0.35
12	2	1.5	-1.7	-2.55
13	1	2.5	-2.7	-6.75
14	1	3.5	-2.7	-9.45
15	0	4.5	-3.7	-16.65

$$\sum = -74.5$$

Covariance Formula

$$(x_i - \mu_x)$$

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

$$\sum = 1$$

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

$$\sum = 1$$

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{1}{10}$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

$$\sum = 1$$

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$Cov(X, Y) = \frac{1}{10} = 0.1$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

$$\sum = 1$$

Covariance Formula

Age vs Grades

Covariance ≈ 0

$$\mu_x = 10.5 \quad \mu_y = 5$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

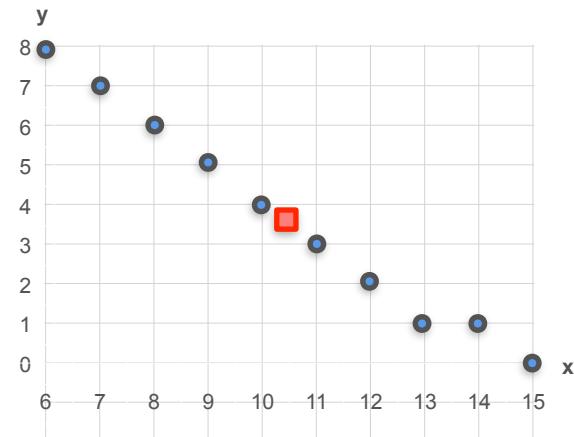
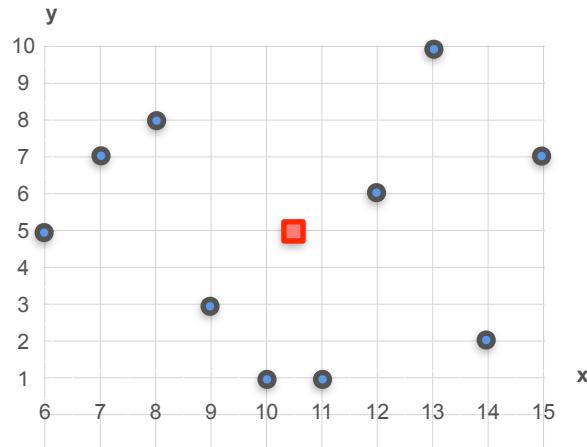
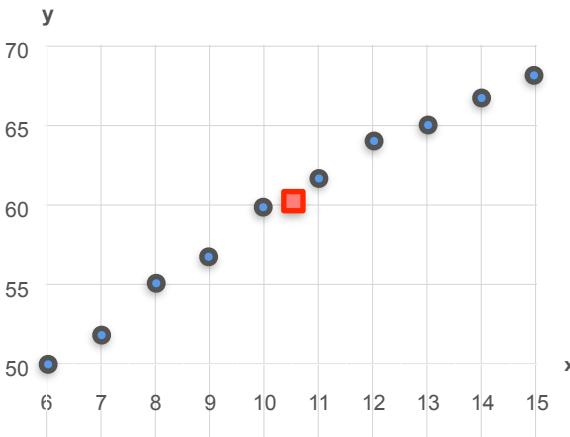
$$Cov(X, Y) = \frac{1}{10} = 0.1 \quad \approx 0$$

Age (x)	Grades (y)	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
6	5	-4.5	0	0
7	7	-3.5	2	-7
8	8	-2.5	3	-7.5
9	3	-1.5	-2	3
10	1	-0.5	-4	2
11	1	0.5	-4	-2
12	6	1.5	1	1.5
13	10	2.5	5	12.5
14	2	3.5	-3	-10.5
15	7	4.5	2	9

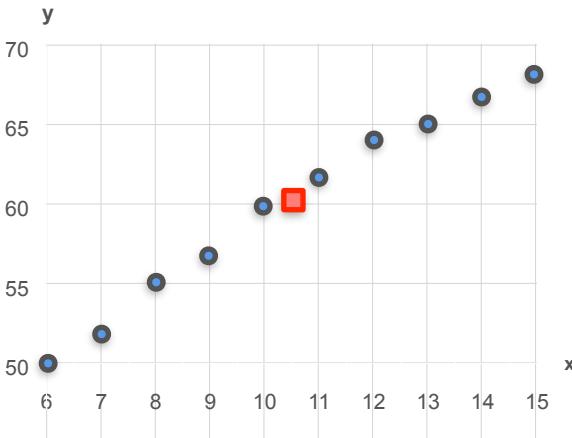
$$\sum = 1$$

Comparing Correlations

Comparing Correlations



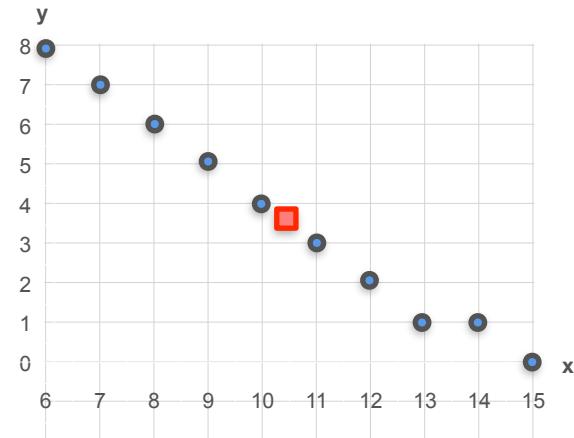
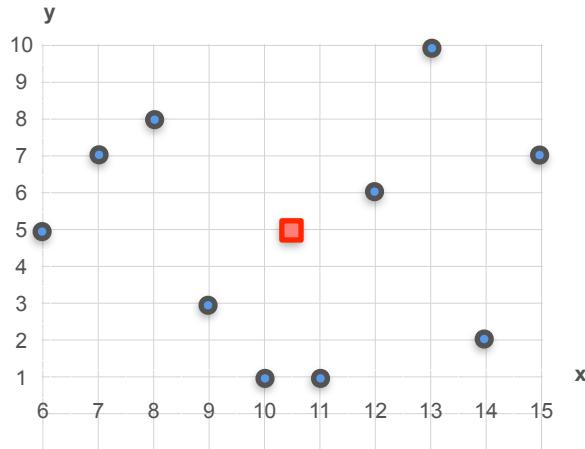
Comparing Correlations



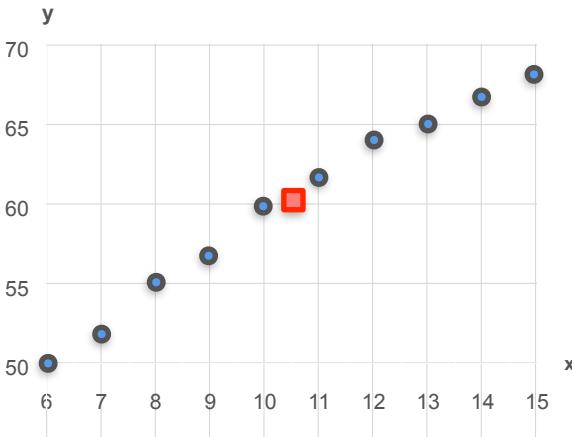
Age vs Height

Covariance > 0

$$Cov(x, y) = 17$$



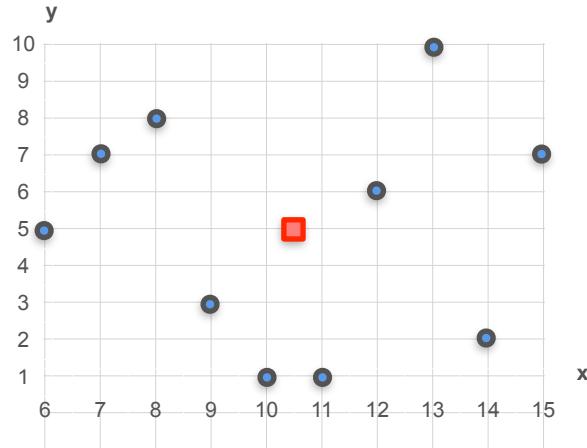
Comparing Correlations



Age vs Height

Covariance > 0

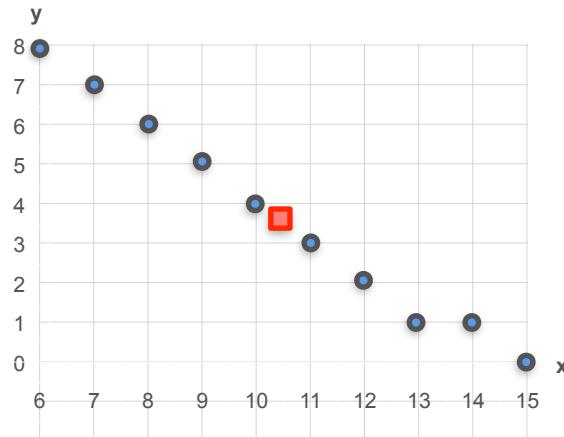
$$Cov(x, y) = 17$$



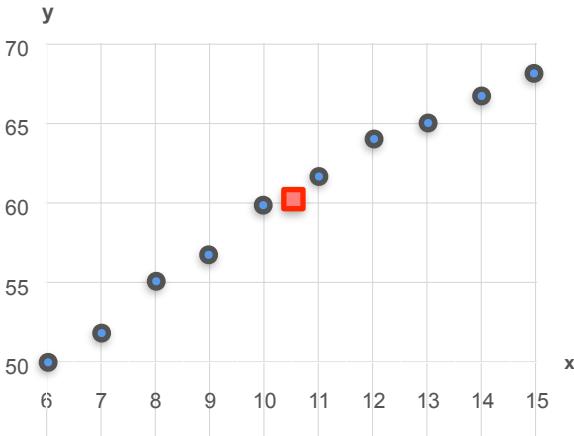
Age vs Grades

Covariance ≈ 0

$$Cov(x, y) = 0.1$$



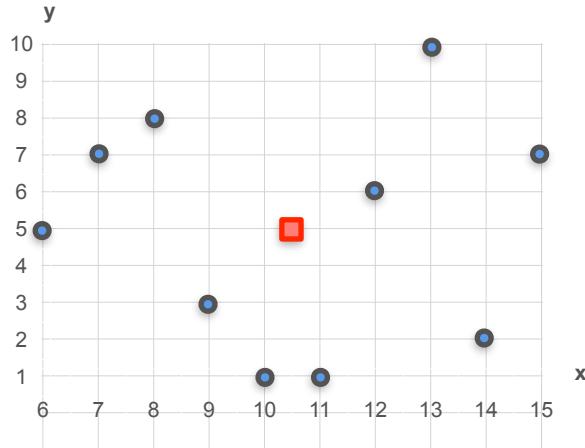
Comparing Correlations



Age vs Height

Covariance > 0

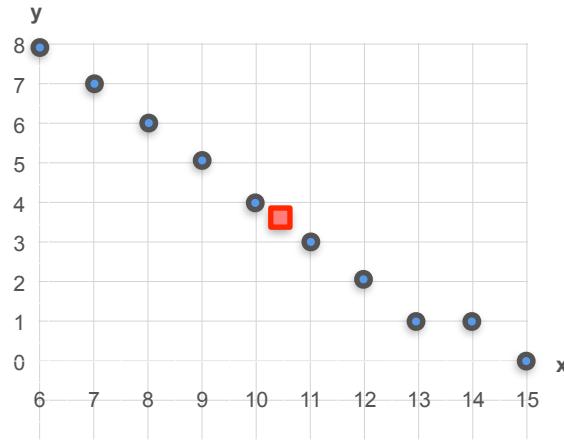
$$\text{Cov}(x, y) = 17$$



Age vs Grades

Covariance ≈ 0

$$\text{Cov}(x, y) = 0.1$$



Age vs Naps per Day

Covariance < 0

$$\text{Cov}(x, y) = -7.45$$

Covariance of a Probability Distribution: Motivation

Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

GAME 2

GAME 3

Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

GAME 2

GAME 3

a: Both players win \$1 each

Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

b: Both players lose \$1 each

GAME 2

GAME 3

Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

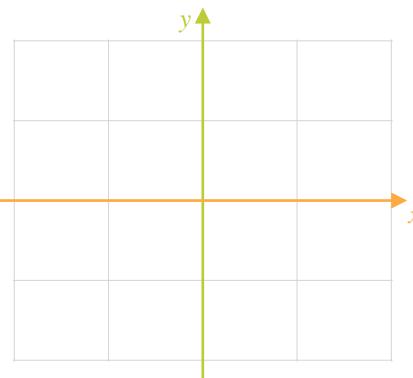
GAME 1

a: Both players win \$1 each

b: Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

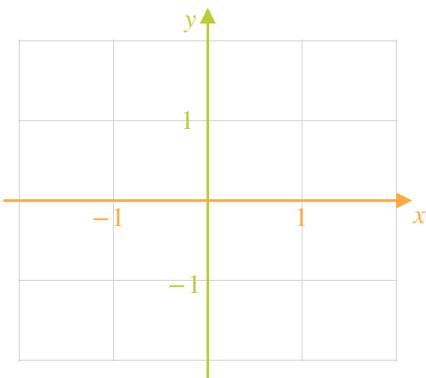
GAME 1

a : Both players win \$1 each

b : Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

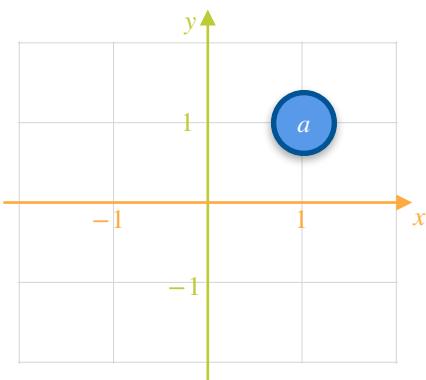
GAME 1

a: Both players win \$1 each

b: Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

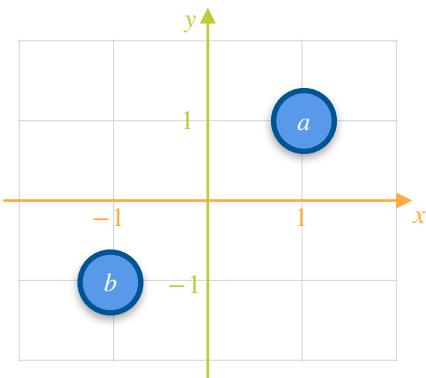
GAME 1

a : Both players win \$1 each

b : Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

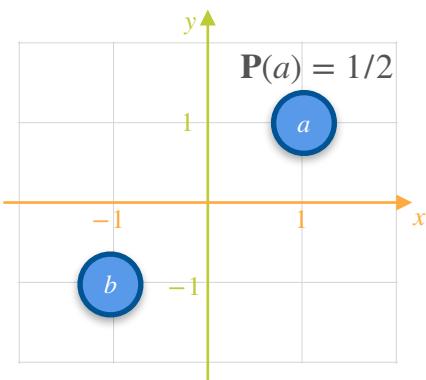
GAME 1

a : Both players win \$1 each

b : Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

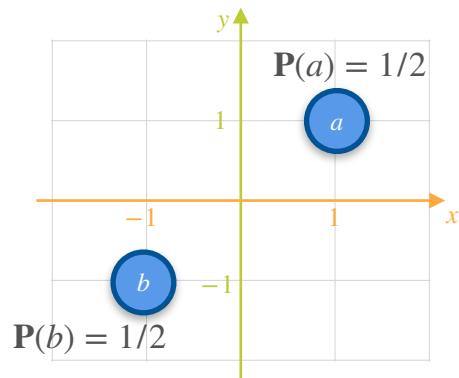
GAME 1

a : Both players win \$1 each

b : Both players lose \$1 each

GAME 2

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

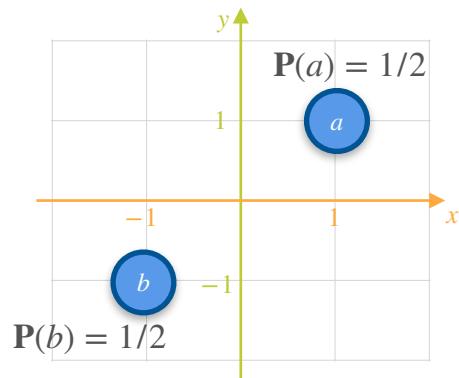
a: Both players win \$1 each

b: Both players lose \$1 each

GAME 2

c: X wins \$1 and Y loses \$1

GAME 3



Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

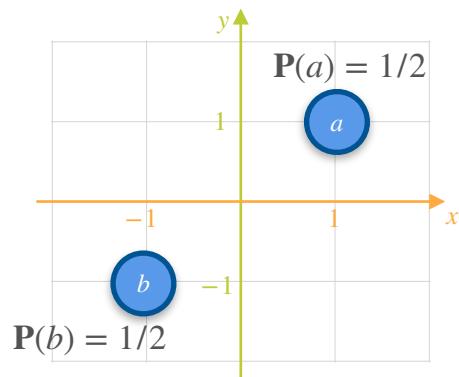
b: Both players lose \$1 each

GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1

GAME 3



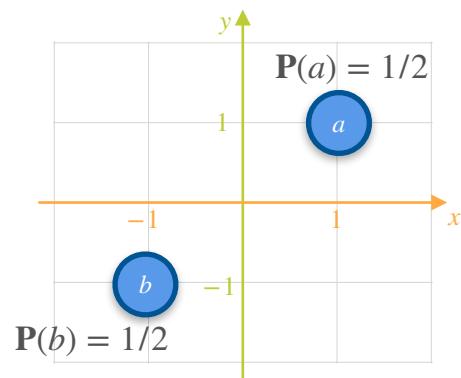
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

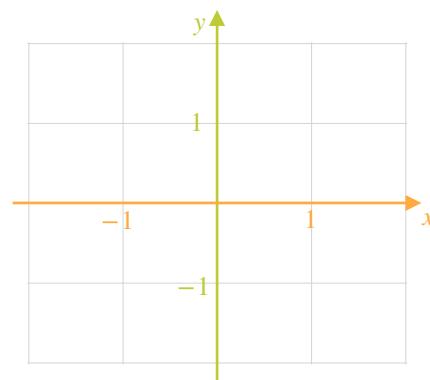
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

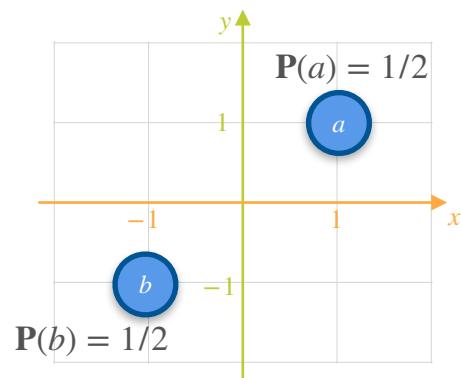
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

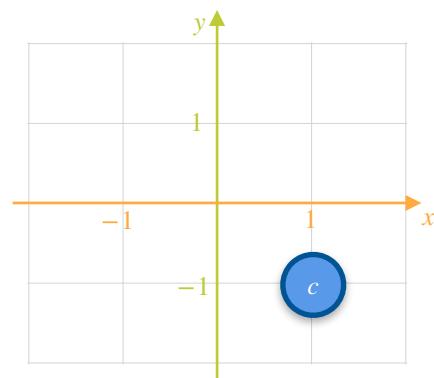
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

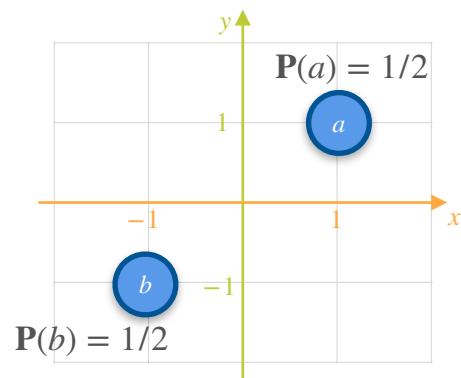
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

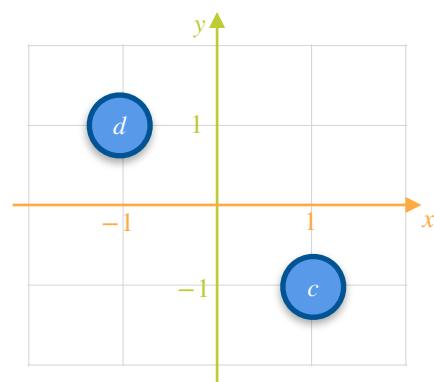
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

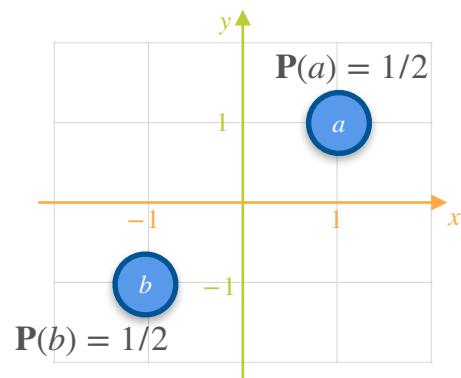
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

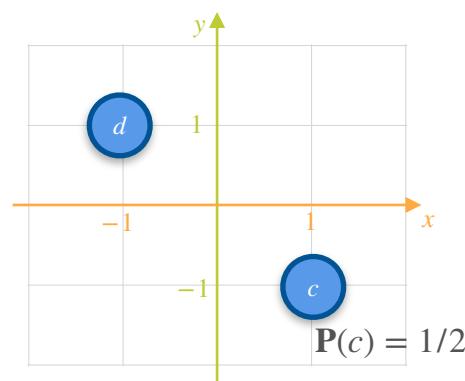
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

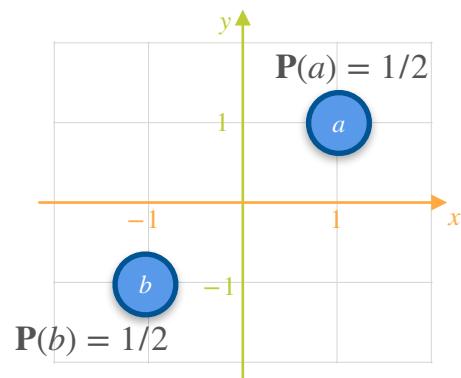
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

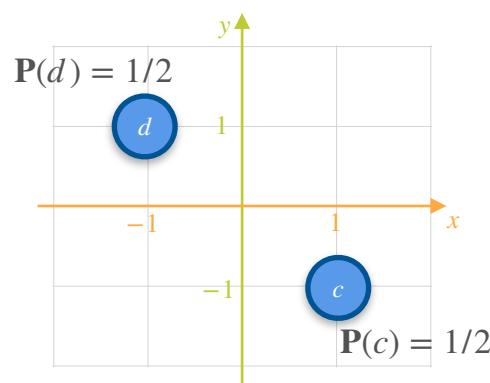
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

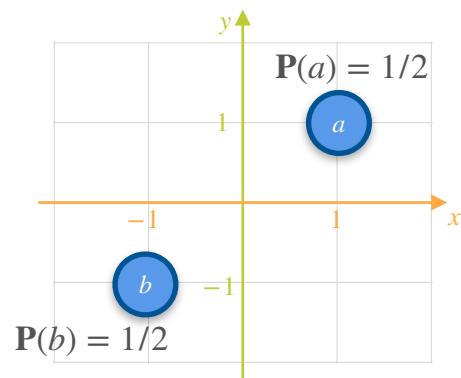
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

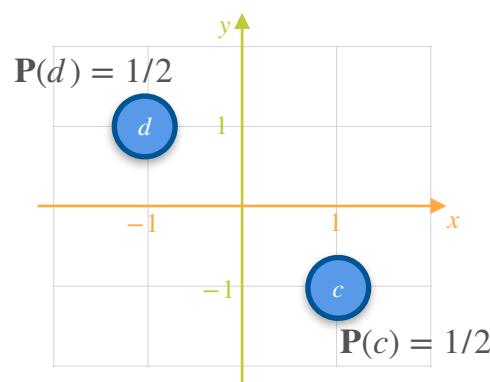
b: Both players lose \$1 each



GAME 2

c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

a, b, c or d

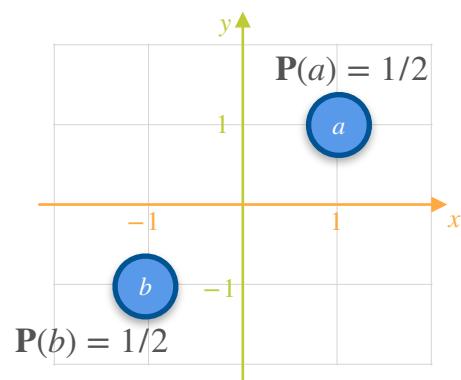
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

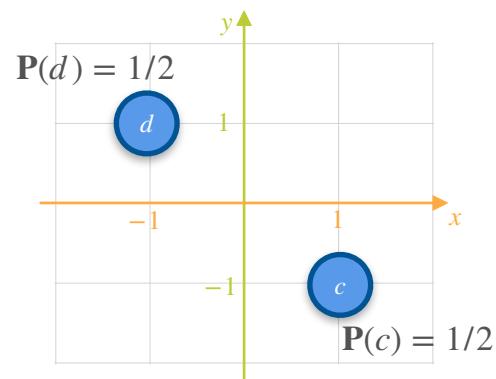
b: Both players lose \$1 each



GAME 2

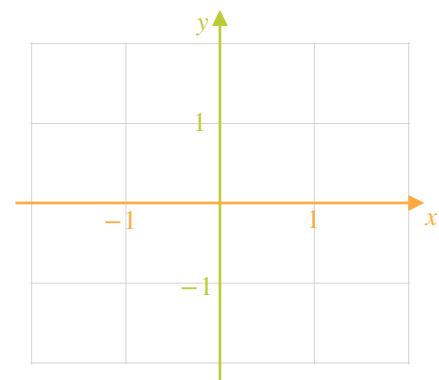
c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

a, b, c or d



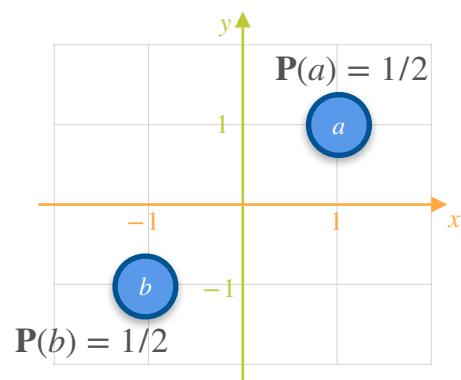
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

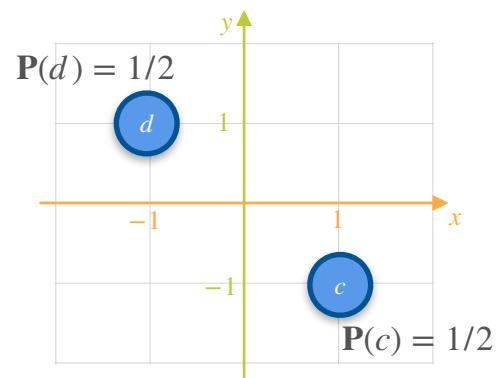
b: Both players lose \$1 each



GAME 2

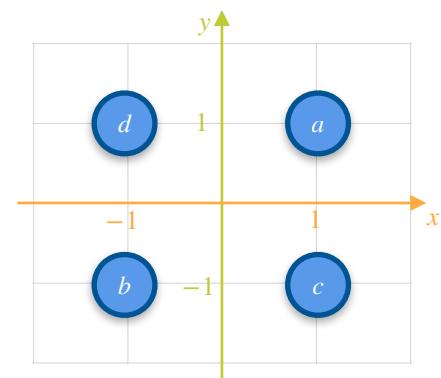
c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

a, b, c or d



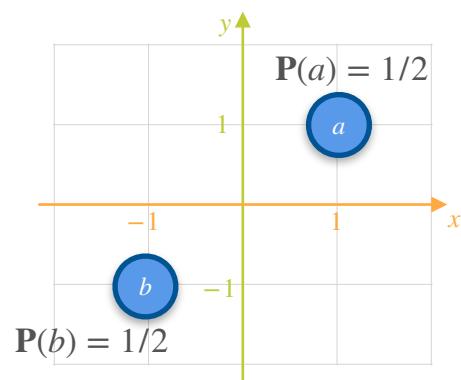
Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

GAME 1

a: Both players win \$1 each

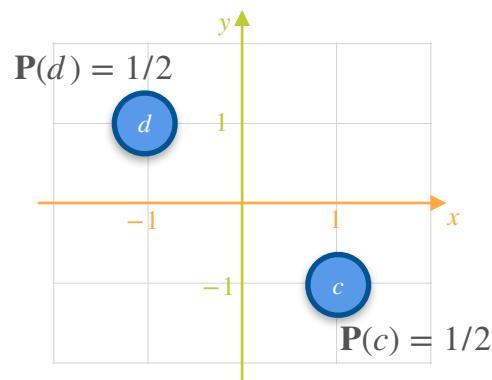
b: Both players lose \$1 each



GAME 2

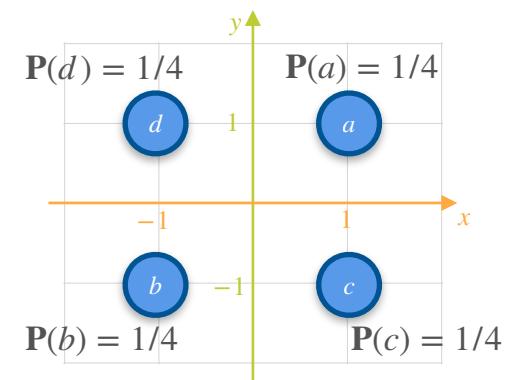
c: X wins \$1 and Y loses \$1

d: X loses \$1 and Y win \$1



GAME 3

a, b, c or d

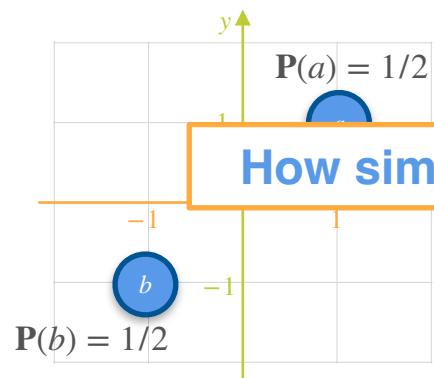


Covariance of a Probability Distribution: Motivation

X and Y are playing 3 games to either win or lose a dollar

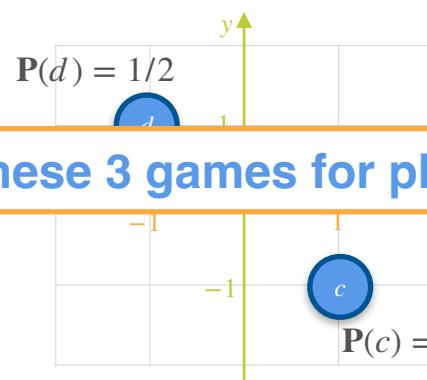
GAME 1

- a : Both players win \$1 each
 b : Both players lose \$1 each



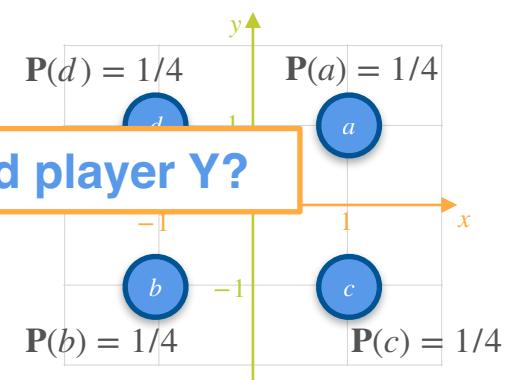
GAME 2

- c : X wins \$1 and Y loses \$1
 d : X loses \$1 and Y win \$1



GAME 3

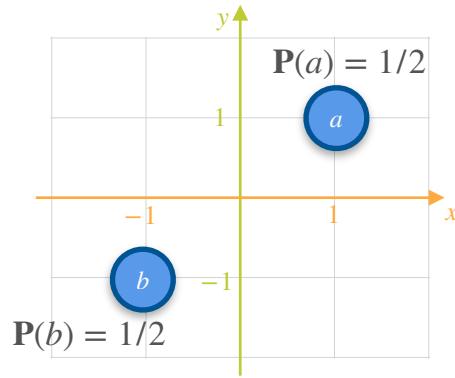
- a, b, c or d



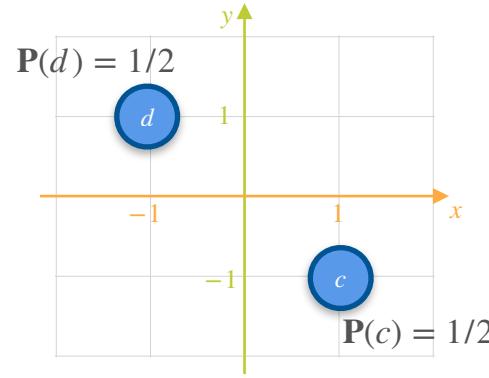
How similar are these 3 games for player X and player Y?

Covariance of a Probability Distribution: Motivation

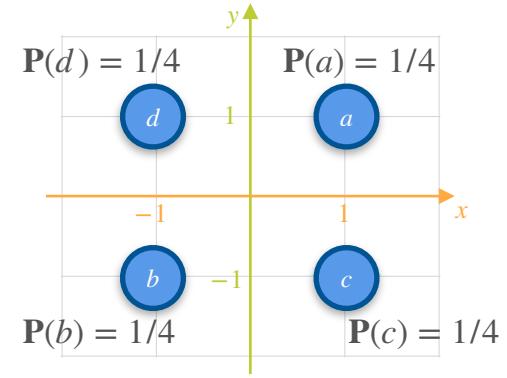
GAME 1



GAME 2



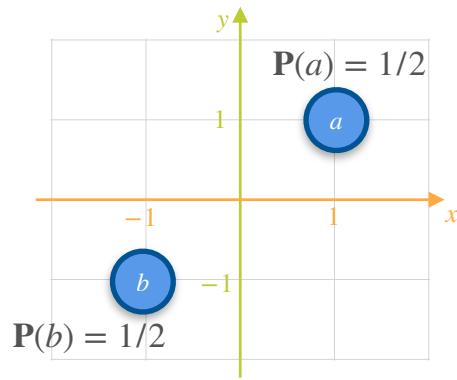
GAME 3



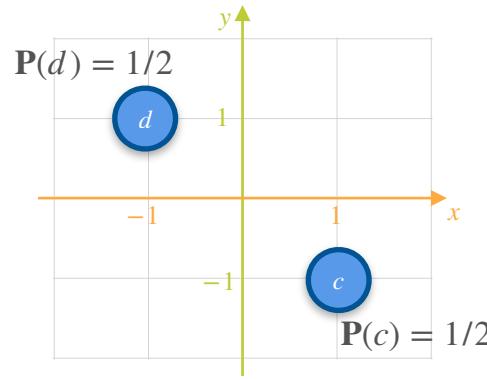
How similar are these 3 games for player X and player Y?

Covariance of a Probability Distribution: Motivation

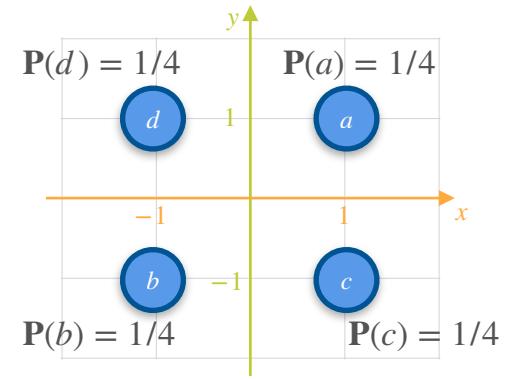
GAME 1



GAME 2



GAME 3

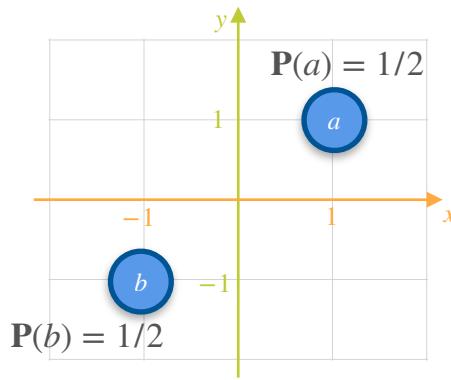


How similar are these 3 games for player X and player Y?

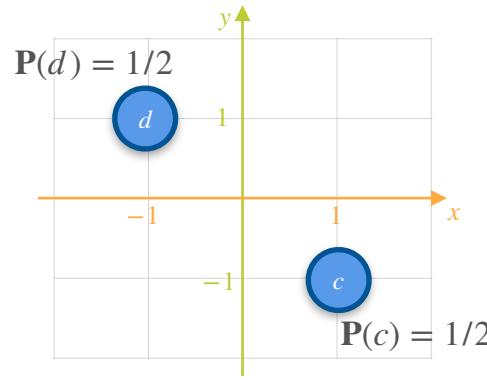
X : how much money in dollars player X wins

Covariance of a Probability Distribution: Motivation

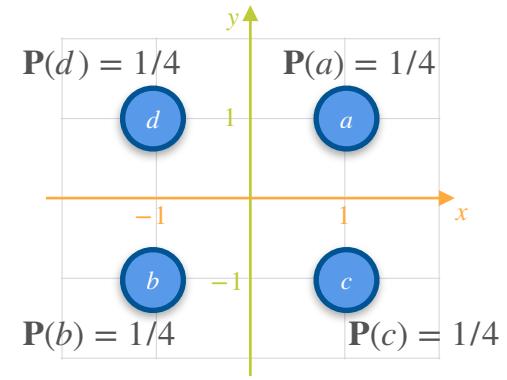
GAME 1



GAME 2



GAME 3



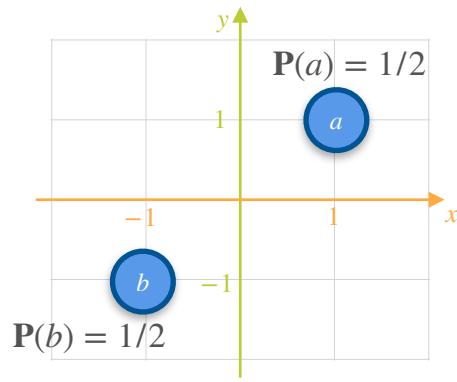
How similar are these 3 games for player X and player Y?

X : how much money in dollars player X wins

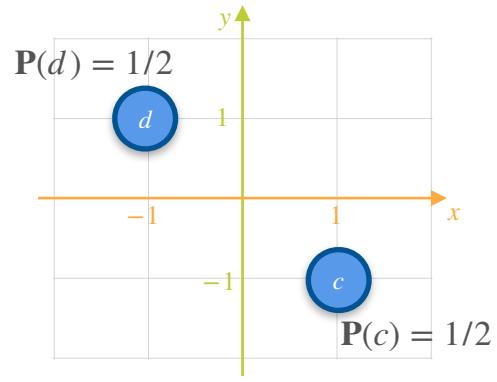
Y : how much money in dollars player Y wins

Covariance of a Probability Distribution: Motivation

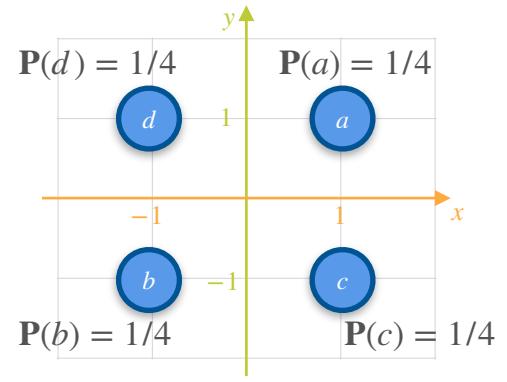
GAME 1



GAME 2

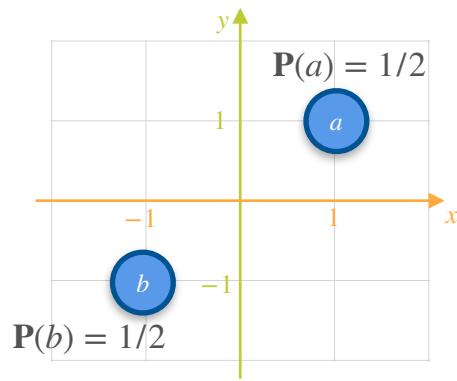


GAME 3

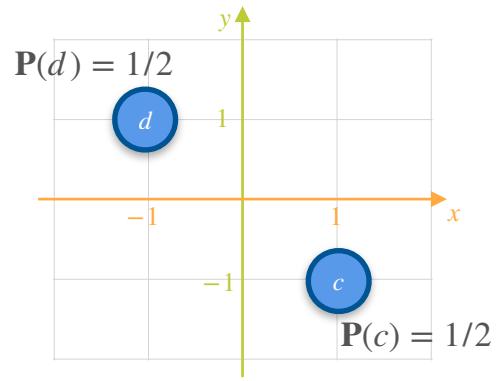


Covariance of a Probability Distribution: Motivation

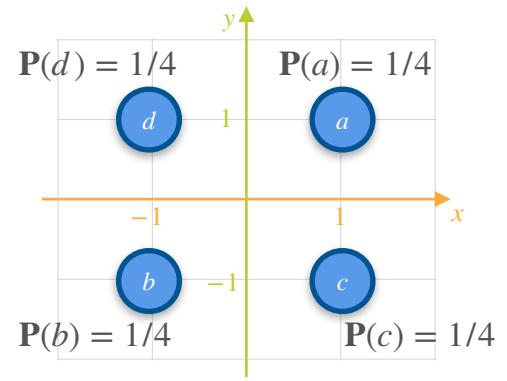
GAME 1



GAME 2



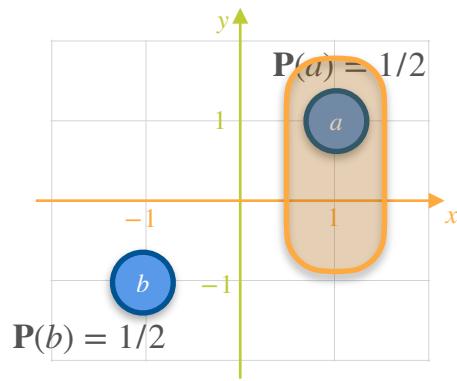
GAME 3



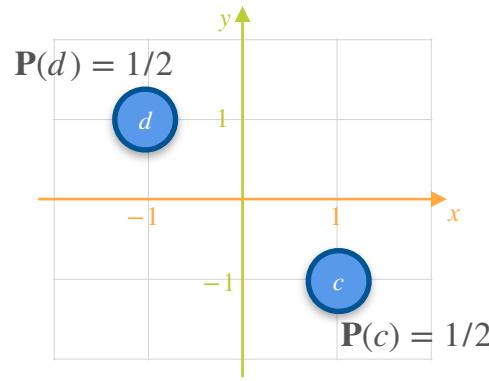
$$\mathbb{E}[X_1] =$$

Covariance of a Probability Distribution: Motivation

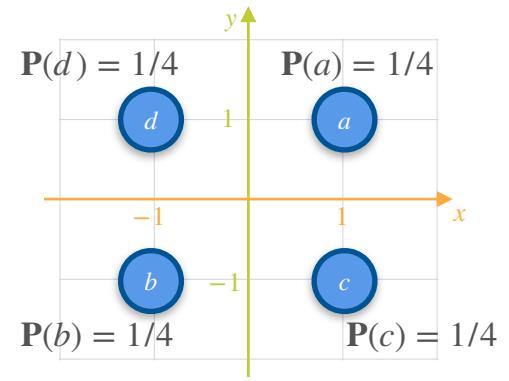
GAME 1



GAME 2



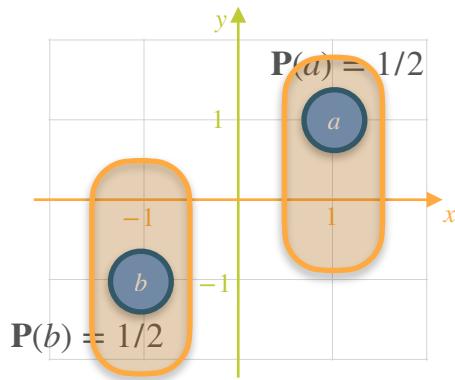
GAME 3



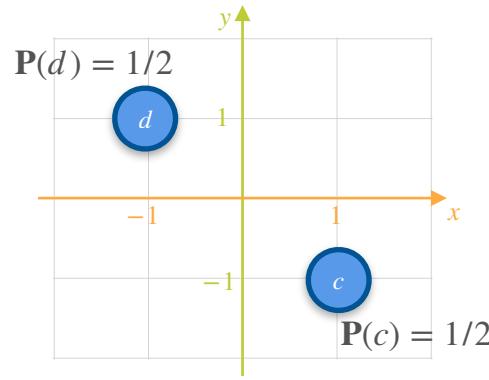
$$\mathbb{E}[X_1] = \frac{1}{2}(1)$$

Covariance of a Probability Distribution: Motivation

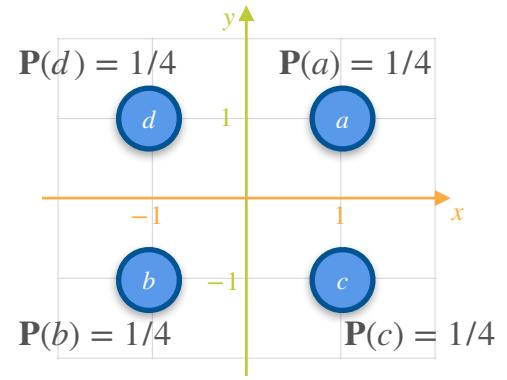
GAME 1



GAME 2



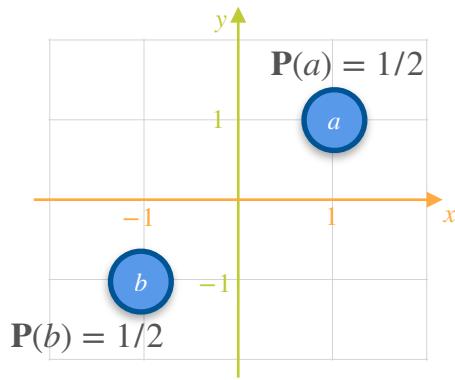
GAME 3



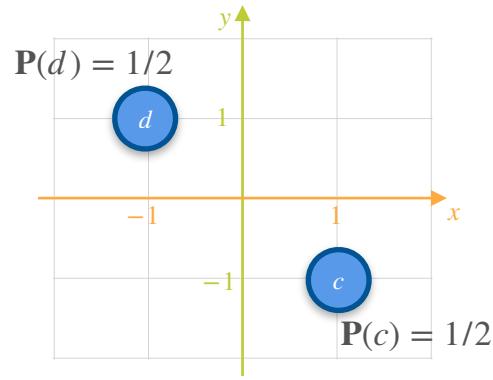
$$\mathbb{E}[X_1] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

Covariance of a Probability Distribution: Motivation

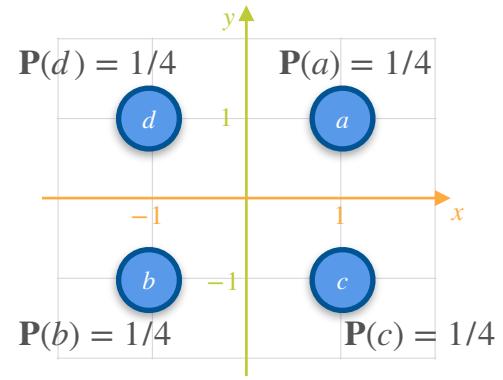
GAME 1



GAME 2



GAME 3

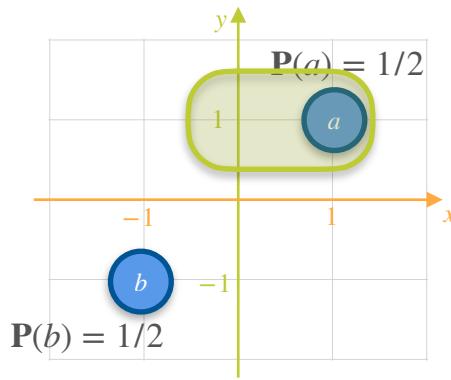


$$\mathbb{E}[X_1] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

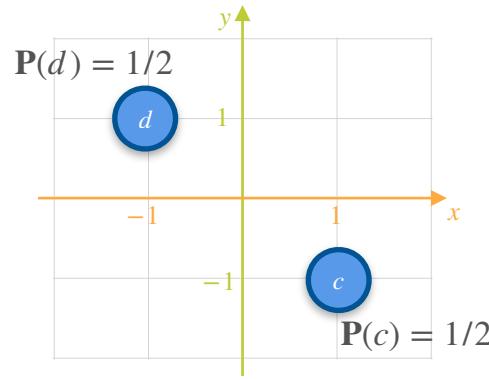
$$\mathbb{E}[Y_1] =$$

Covariance of a Probability Distribution: Motivation

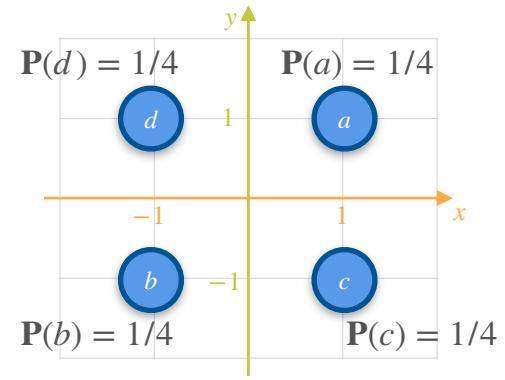
GAME 1



GAME 2



GAME 3

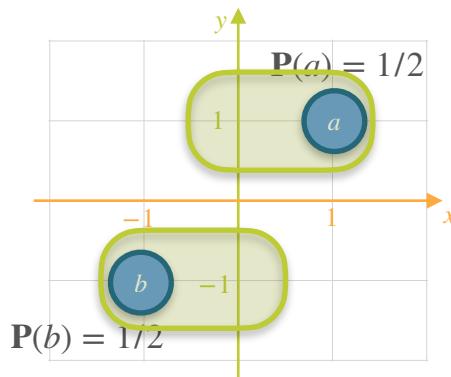


$$\mathbb{E}[X_1] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

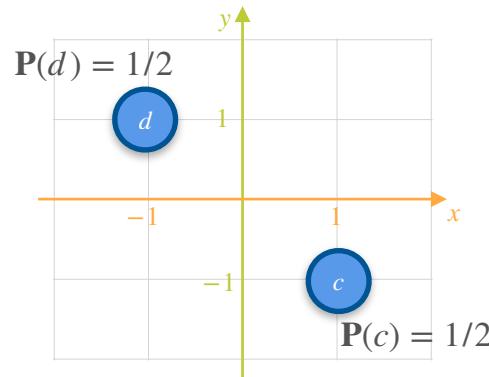
$$\mathbb{E}[Y_1] = \frac{1}{2}(1)$$

Covariance of a Probability Distribution: Motivation

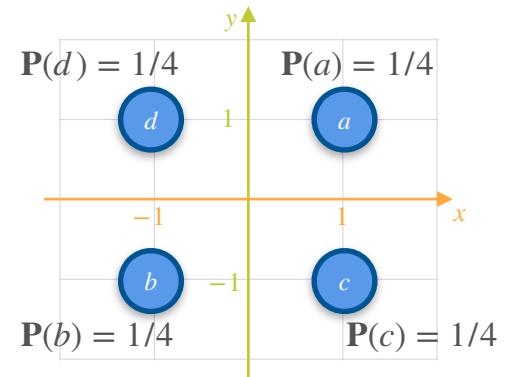
GAME 1



GAME 2



GAME 3

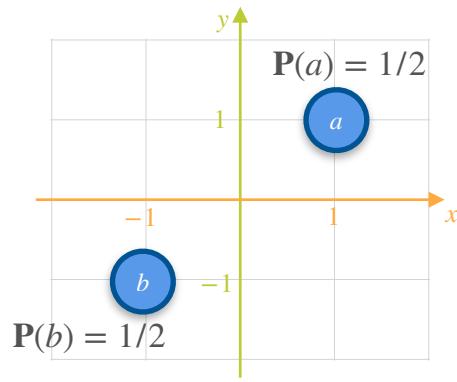


$$\mathbb{E}[X_1] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

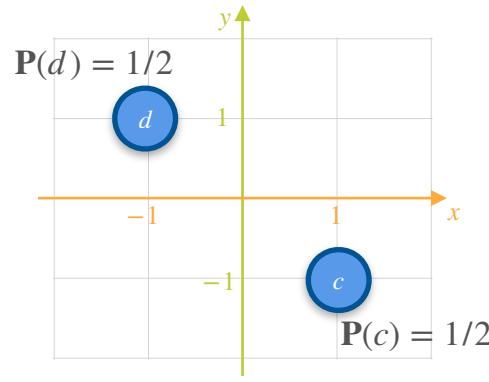
$$\mathbb{E}[Y_1] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

Covariance of a Probability Distribution: Motivation

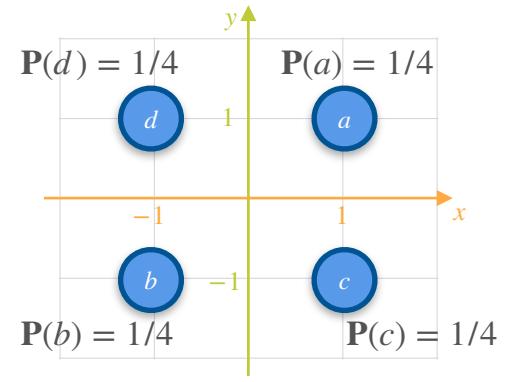
GAME 1



GAME 2



GAME 3

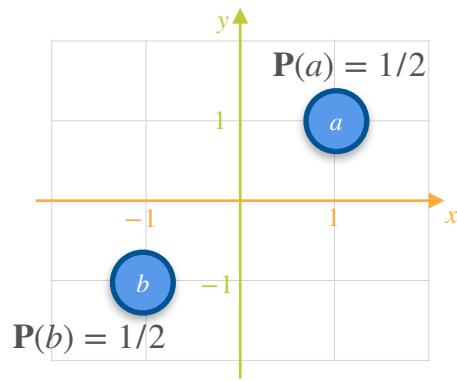


$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

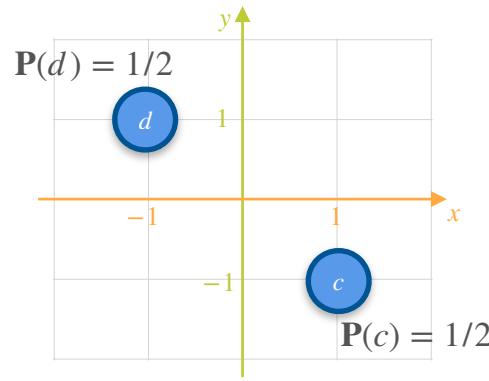
GAME 1



$$\mathbb{E}[X_1] = 0$$

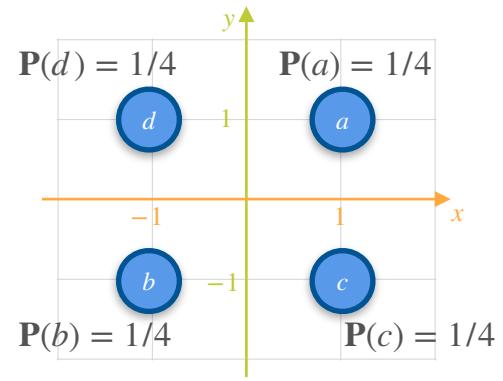
$$\mathbb{E}[Y_1] = 0$$

GAME 2



$$\mathbb{E}[X_2] =$$

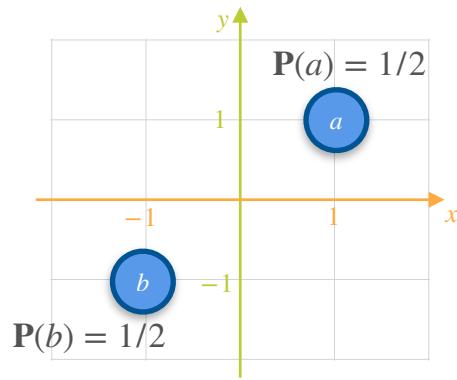
GAME 3



$$\mathbb{E}[X_3] = 0$$

Covariance of a Probability Distribution: Motivation

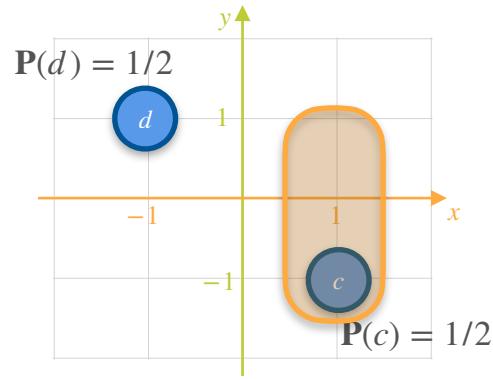
GAME 1



$$\mathbb{E}[X_1] = 0$$

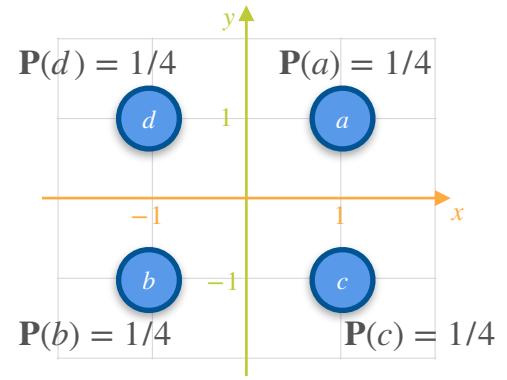
$$\mathbb{E}[Y_1] = 0$$

GAME 2



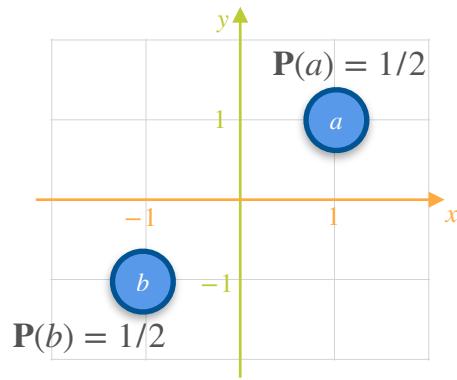
$$\mathbb{E}[X_2] = \frac{1}{2}(1)$$

GAME 3



Covariance of a Probability Distribution: Motivation

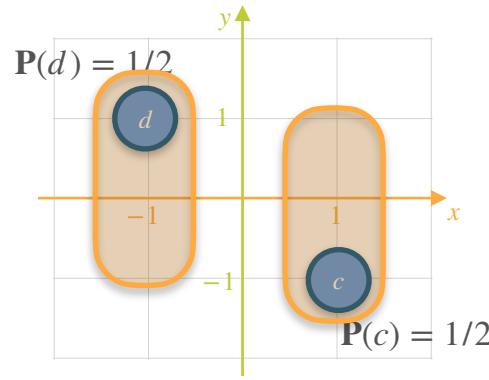
GAME 1



$$\mathbb{E}[X_1] = 0$$

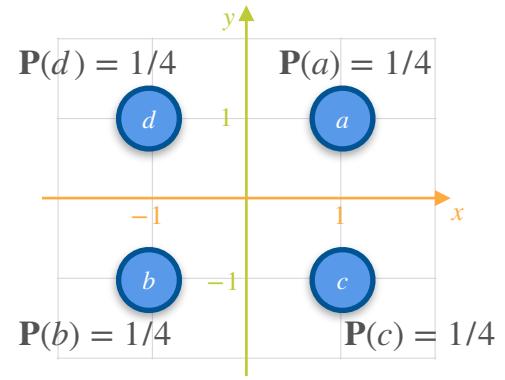
$$\mathbb{E}[Y_1] = 0$$

GAME 2



$$\mathbb{E}[X_2] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

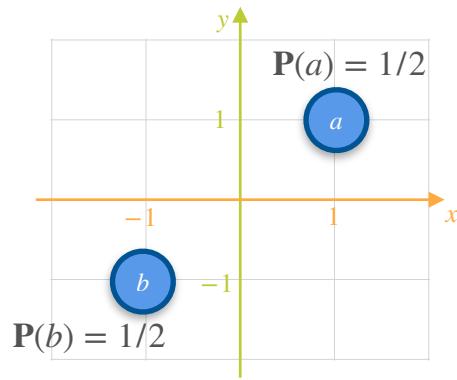
GAME 3



$$\mathbb{E}[b] = 0$$

Covariance of a Probability Distribution: Motivation

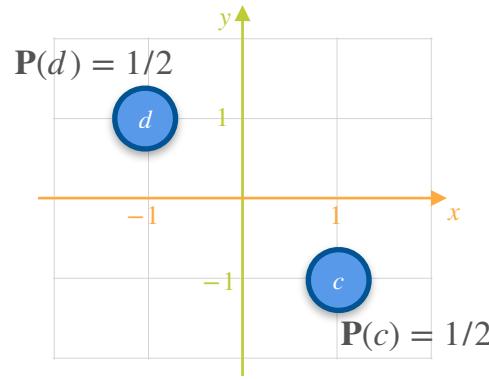
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

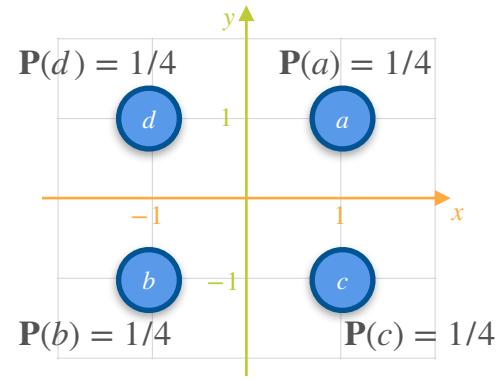
GAME 2



$$\mathbb{E}[X_2] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

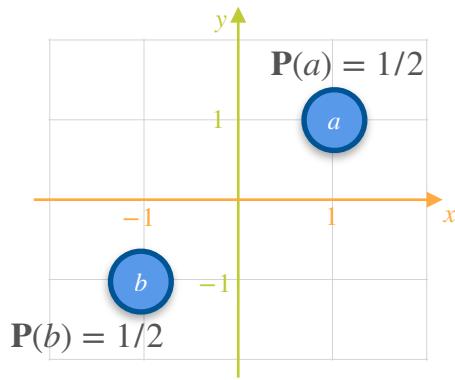
$$\mathbb{E}[Y_2] =$$

GAME 3



Covariance of a Probability Distribution: Motivation

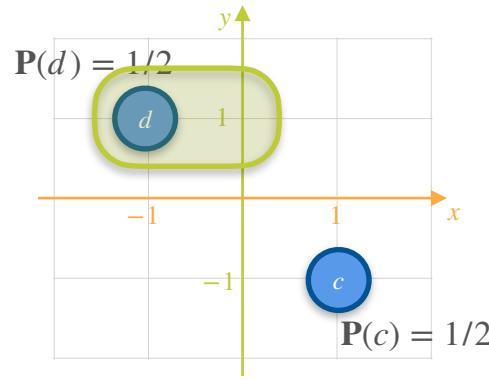
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

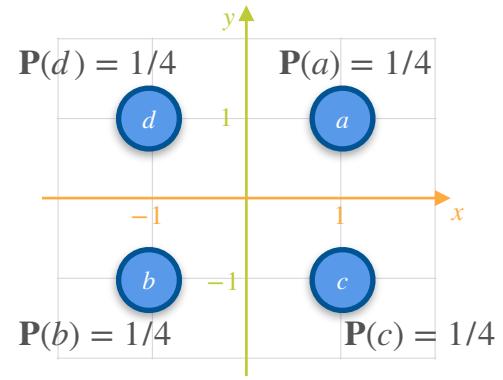
GAME 2



$$\mathbb{E}[X_2] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

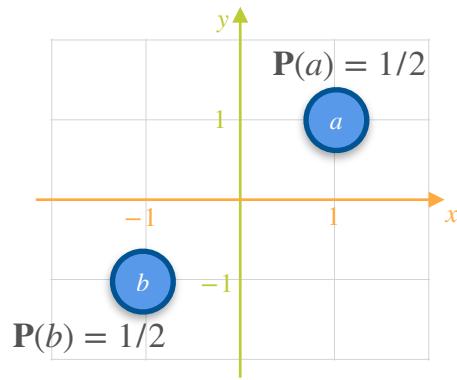
$$\mathbb{E}[Y_2] = \frac{1}{2}(1)$$

GAME 3



Covariance of a Probability Distribution: Motivation

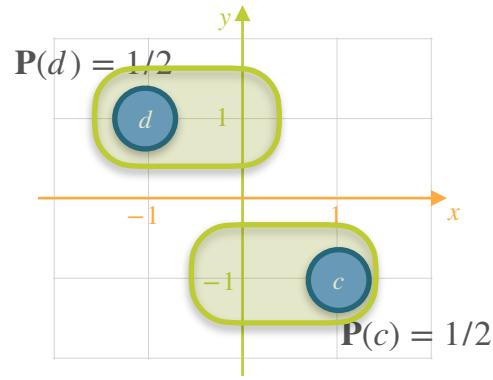
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

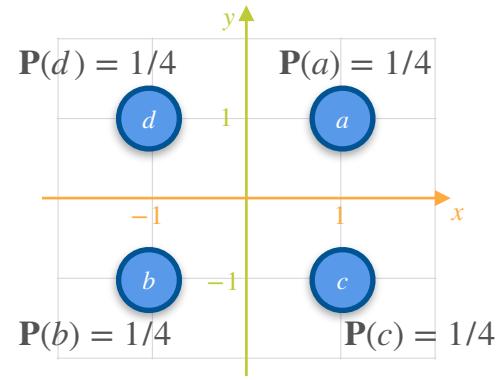
GAME 2



$$\mathbb{E}[X_2] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

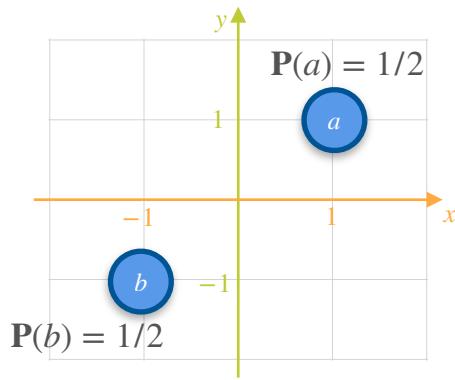
$$\mathbb{E}[Y_2] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

GAME 3

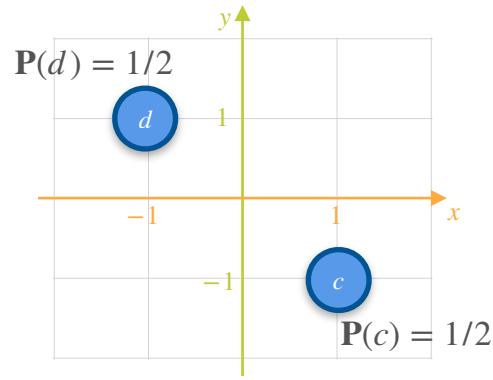


Covariance of a Probability Distribution: Motivation

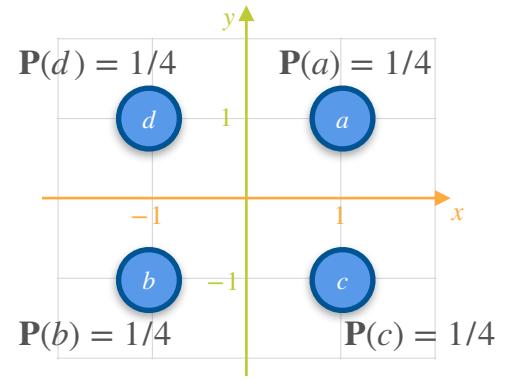
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0$$

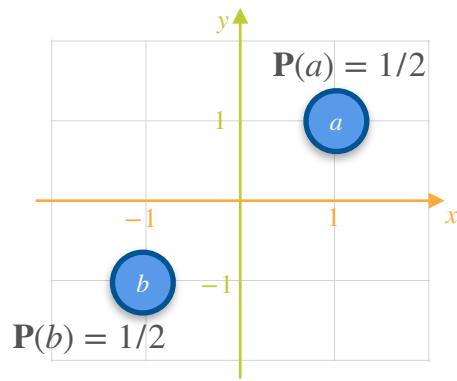
$$\mathbb{E}[Y_1] = 0$$

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation

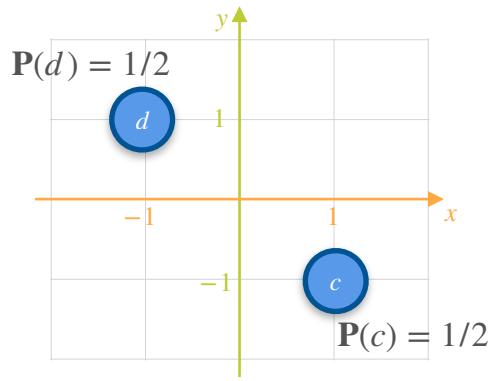
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

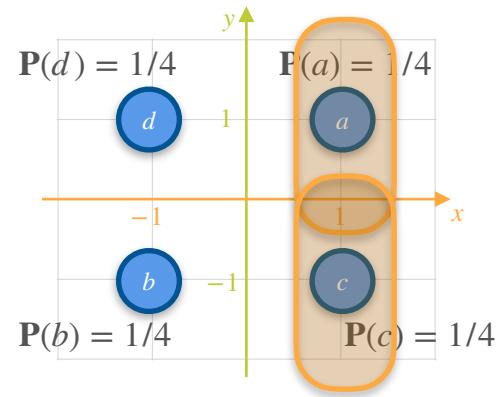
GAME 2



$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

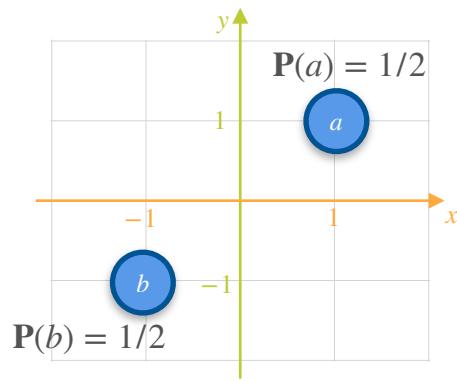
GAME 3



$$\mathbb{E}[X_3] = 2\left(\frac{1}{4}(1)\right)$$

Covariance of a Probability Distribution: Motivation

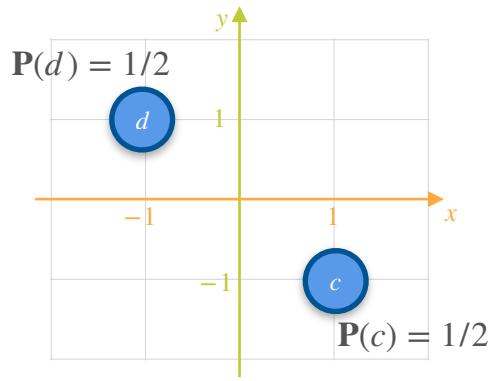
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

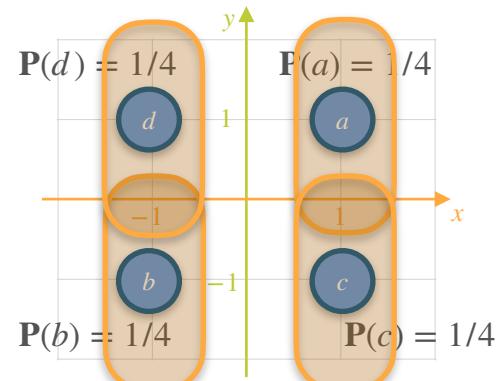
GAME 2



$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

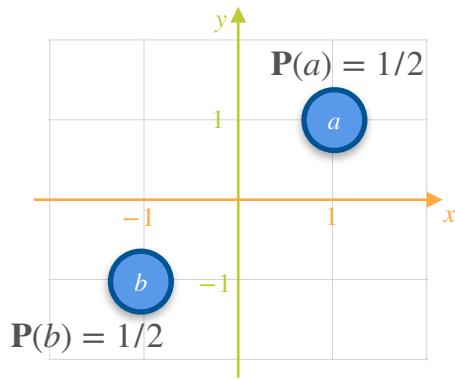
GAME 3



$$\mathbb{E}[X_3] = 2\left(\frac{1}{4}(1)\right) + 2\left(\frac{1}{4}(-1)\right) = 0$$

Covariance of a Probability Distribution: Motivation

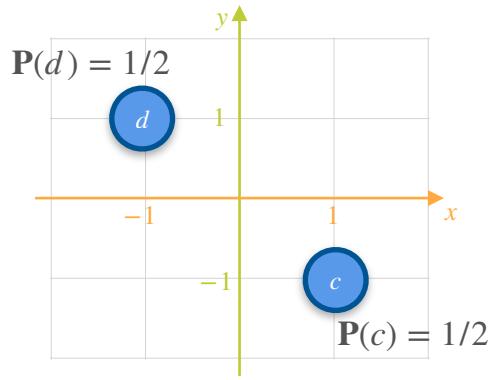
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

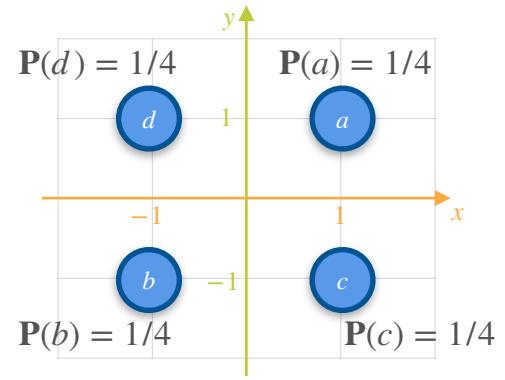
GAME 2



$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

GAME 3

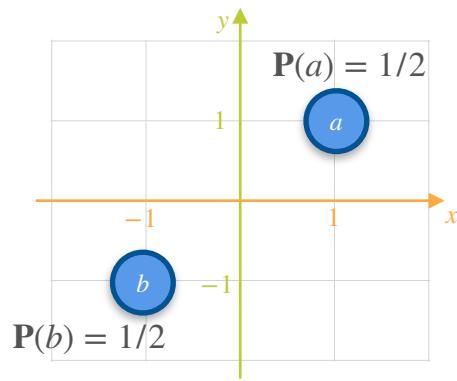


$$\mathbb{E}[X_3] = 2\left(\frac{1}{4}(1)\right) + 2\left(\frac{1}{4}(-1)\right) = 0$$

$$\mathbb{E}[Y_3] =$$

Covariance of a Probability Distribution: Motivation

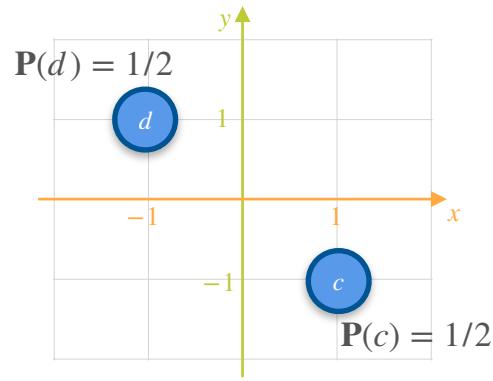
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

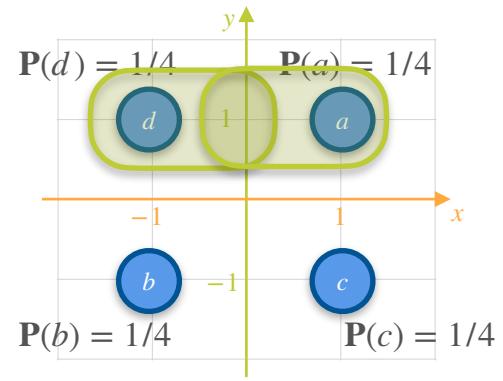
GAME 2



$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

GAME 3

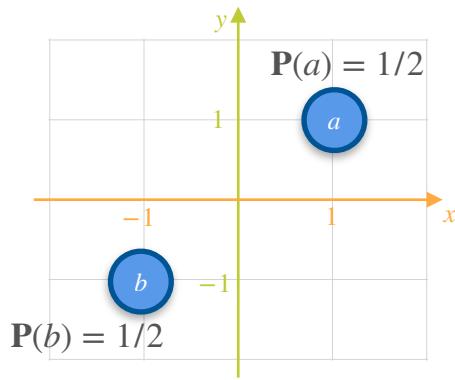


$$\mathbb{E}[X_3] = 2\left(\frac{1}{4}(1)\right) + 2\left(\frac{1}{4}(-1)\right) = 0$$

$$\mathbb{E}[Y_3] = 2\left(\frac{1}{4}(1)\right)$$

Covariance of a Probability Distribution: Motivation

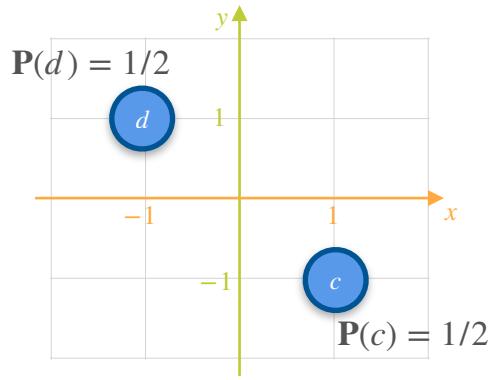
GAME 1



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

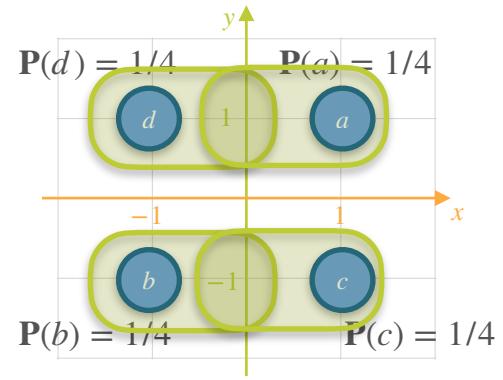
GAME 2



$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[Y_2] = 0$$

GAME 3

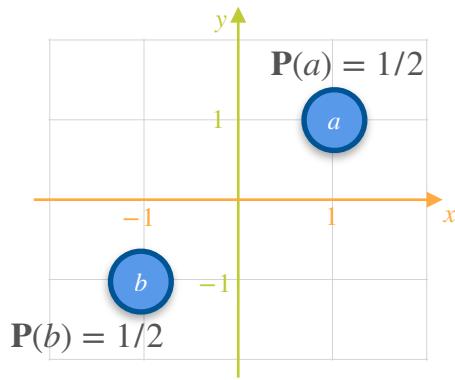


$$\mathbb{E}[X_3] = 2\left(\frac{1}{4}(1)\right) + 2\left(\frac{1}{4}(-1)\right) = 0$$

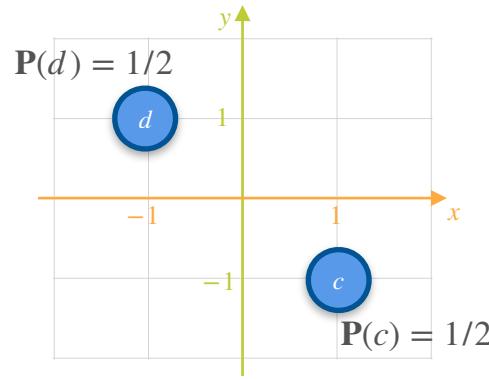
$$\mathbb{E}[Y_3] = 2\left(\frac{1}{4}(1)\right) + 2\left(\frac{1}{4}(-1)\right) = 0$$

Covariance of a Probability Distribution: Motivation

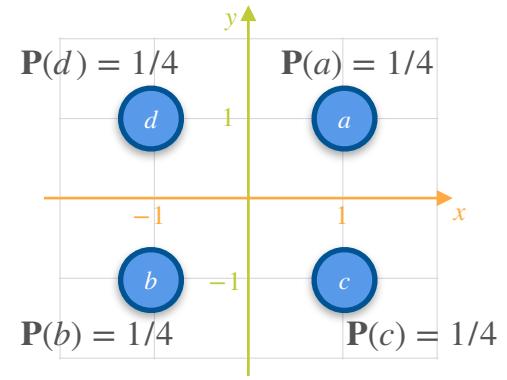
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[Y_1] = 0$$

$$\mathbb{E}[X_2] = 0$$

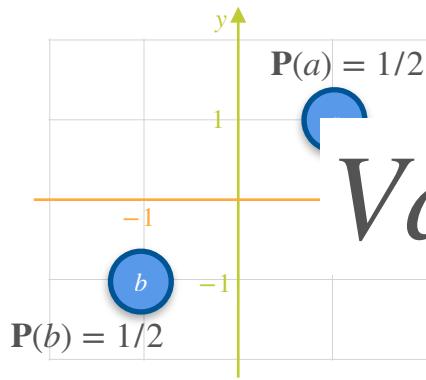
$$\mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0$$

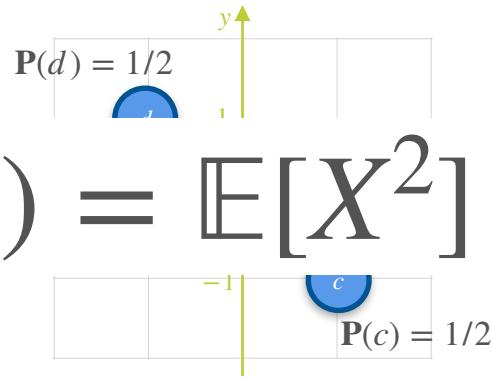
$$\mathbb{E}[Y_3] = 0$$

Covariance of a Probability Distribution: Motivation

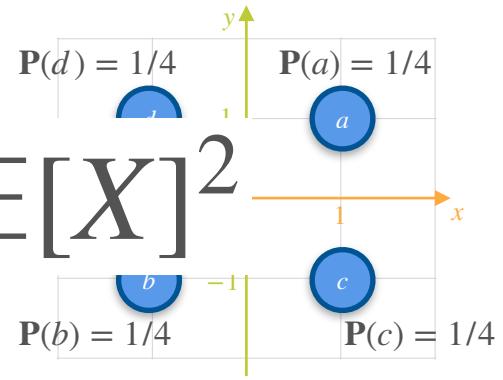
GAME 1



GAME 2



GAME 3



$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$\mathbb{E}[X_1] = 0$$

$$\mathbb{E}[X_2] = 0$$

$$\mathbb{E}[X_3] = 0$$

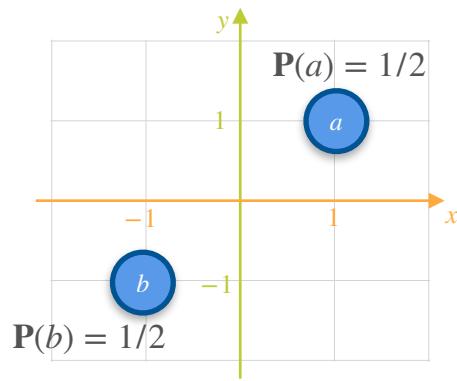
$$\mathbb{E}[Y_1] = 0$$

$$\mathbb{E}[Y_2] = 0$$

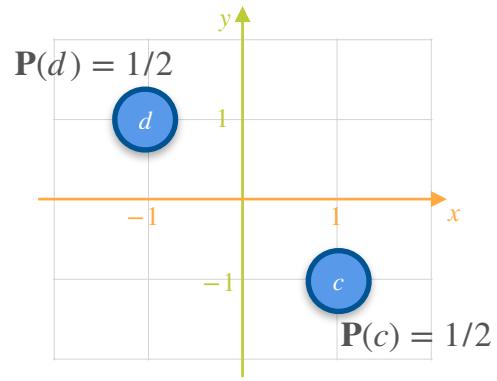
$$\mathbb{E}[Y_3] = 0$$

Covariance of a Probability Distribution: Motivation

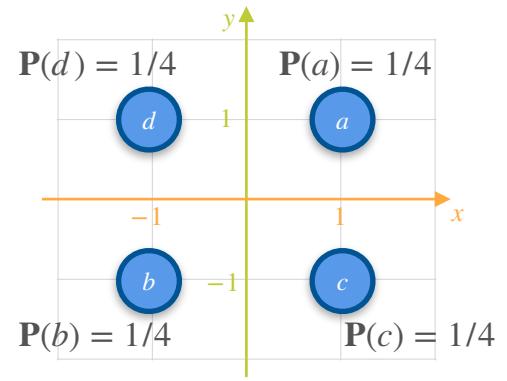
GAME 1



GAME 2



GAME 3



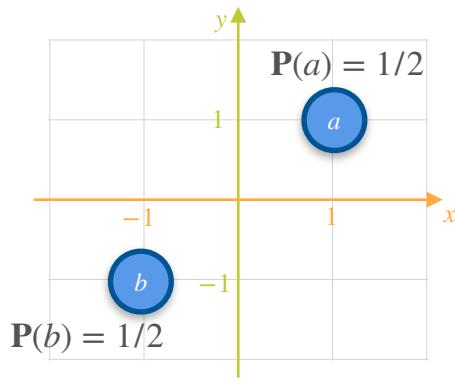
$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

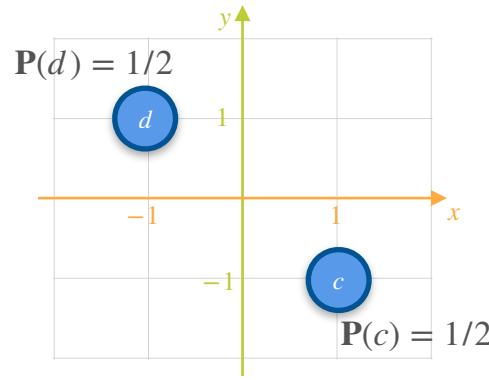
$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

Covariance of a Probability Distribution: Motivation

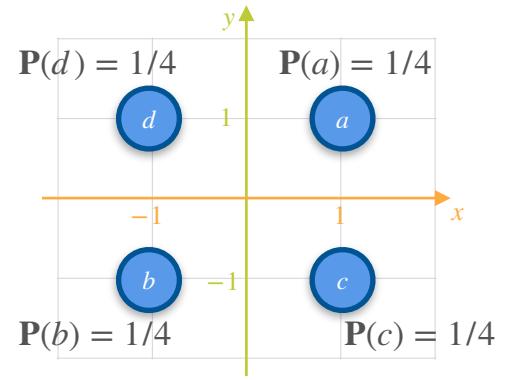
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

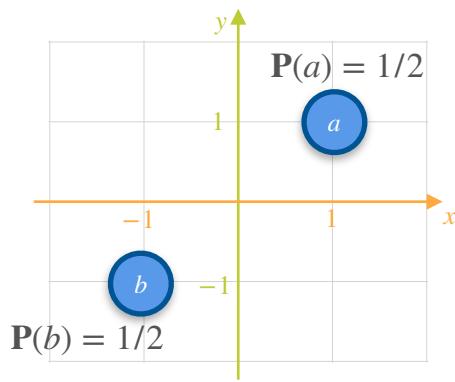
$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

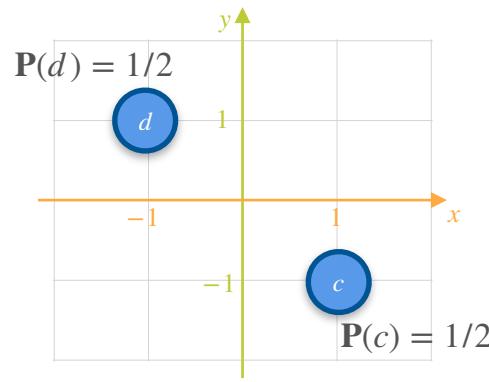
$$Var(X_1) =$$

Covariance of a Probability Distribution: Motivation

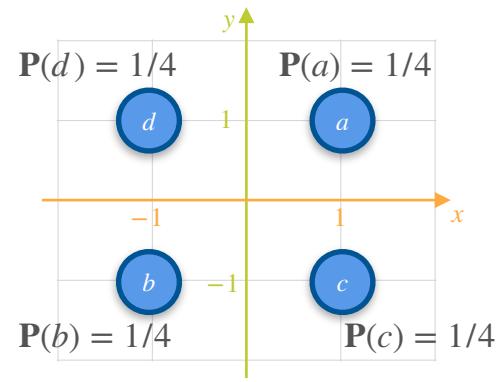
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

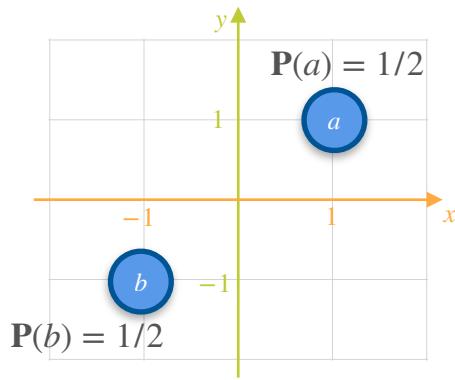
$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

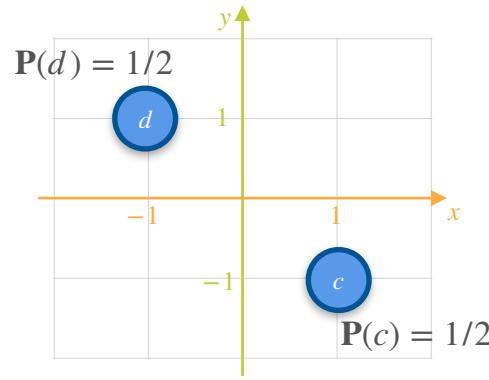
$$Var(X_1) = \mathbb{E}[X_1^2] - E[X_1]^2$$

Covariance of a Probability Distribution: Motivation

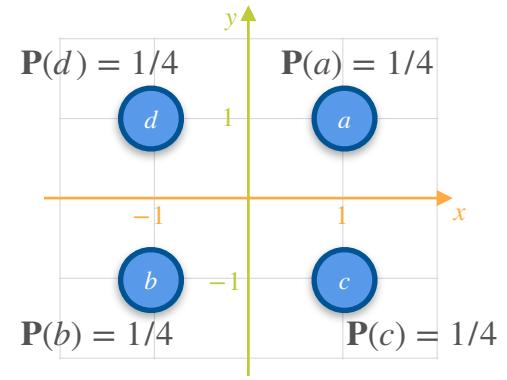
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

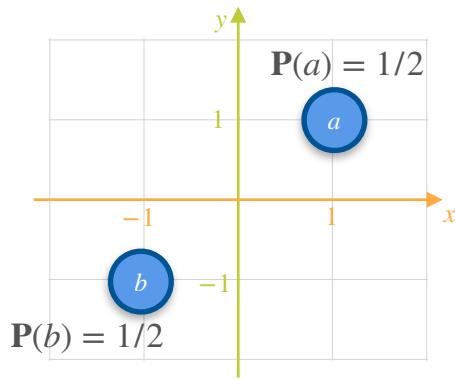
$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

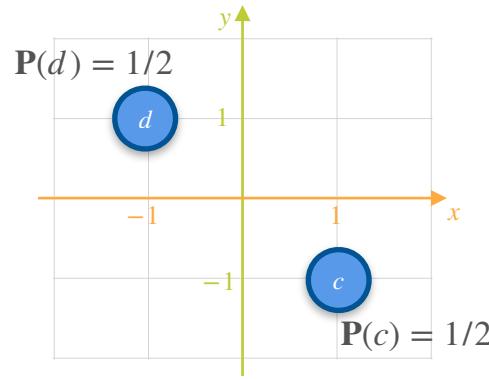
$$Var(X_1) = \mathbb{E}[X_1^2] - E[X_1]^2 = \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 - 0^2 =$$

Covariance of a Probability Distribution: Motivation

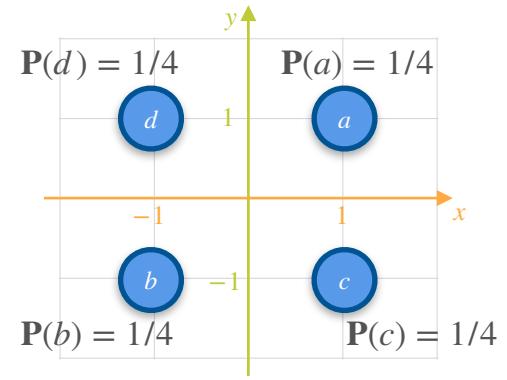
GAME 1



GAME 2



GAME 3



$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

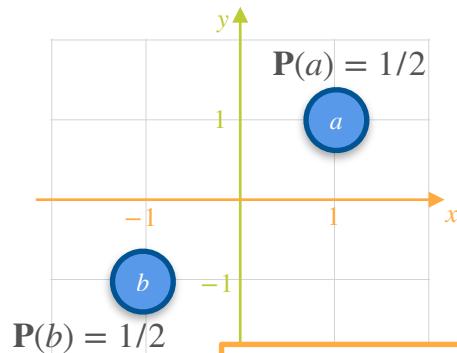
$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

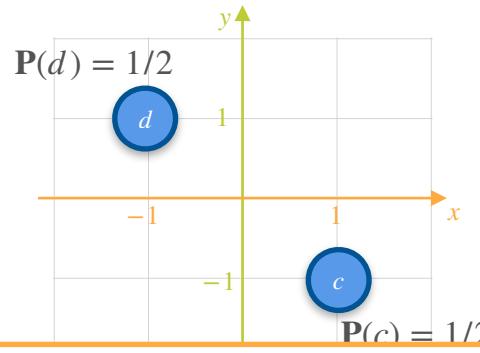
$$Var(X_1) = \mathbb{E}[X_1^2] - E[X_1]^2 = \frac{1}{2}(-1)^2 + \frac{1}{2}(1)^2 - 0^2 = 1$$

Covariance of a Probability Distribution: Motivation

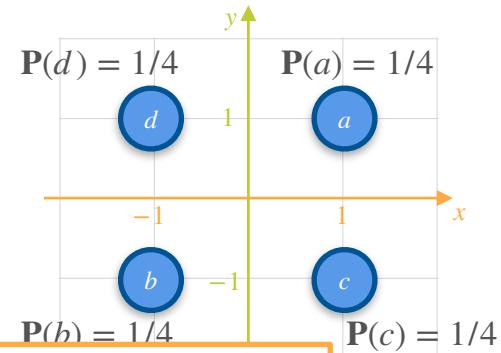
GAME 1



GAME 2



GAME 3



How similar are these 3 games for player X and player Y?

$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

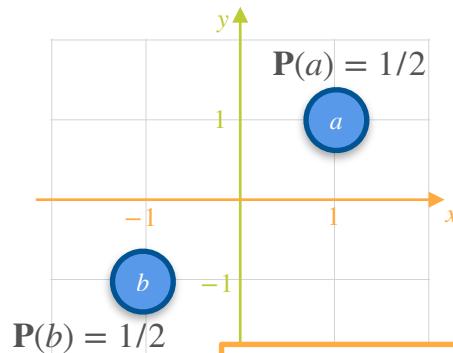
$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

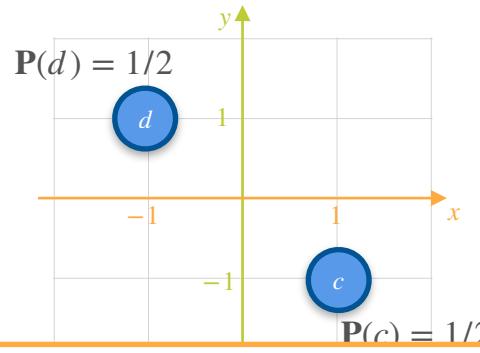
$$Var(X_1) = 1$$

Covariance of a Probability Distribution: Motivation

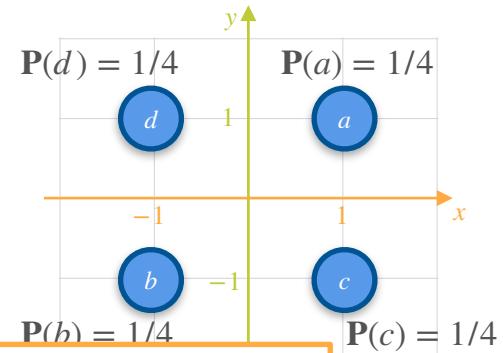
GAME 1



GAME 2



GAME 3



How similar are these 3 games for player X and player Y?

$$\mathbb{E}[X_1] = 0 \quad \mathbb{E}[Y_1] = 0$$

$$Var(X_1) = 1$$

$$Var(Y_1) = 1$$

$$\mathbb{E}[X_2] = 0 \quad \mathbb{E}[Y_2] = 0$$

$$Var(X_2) = 1$$

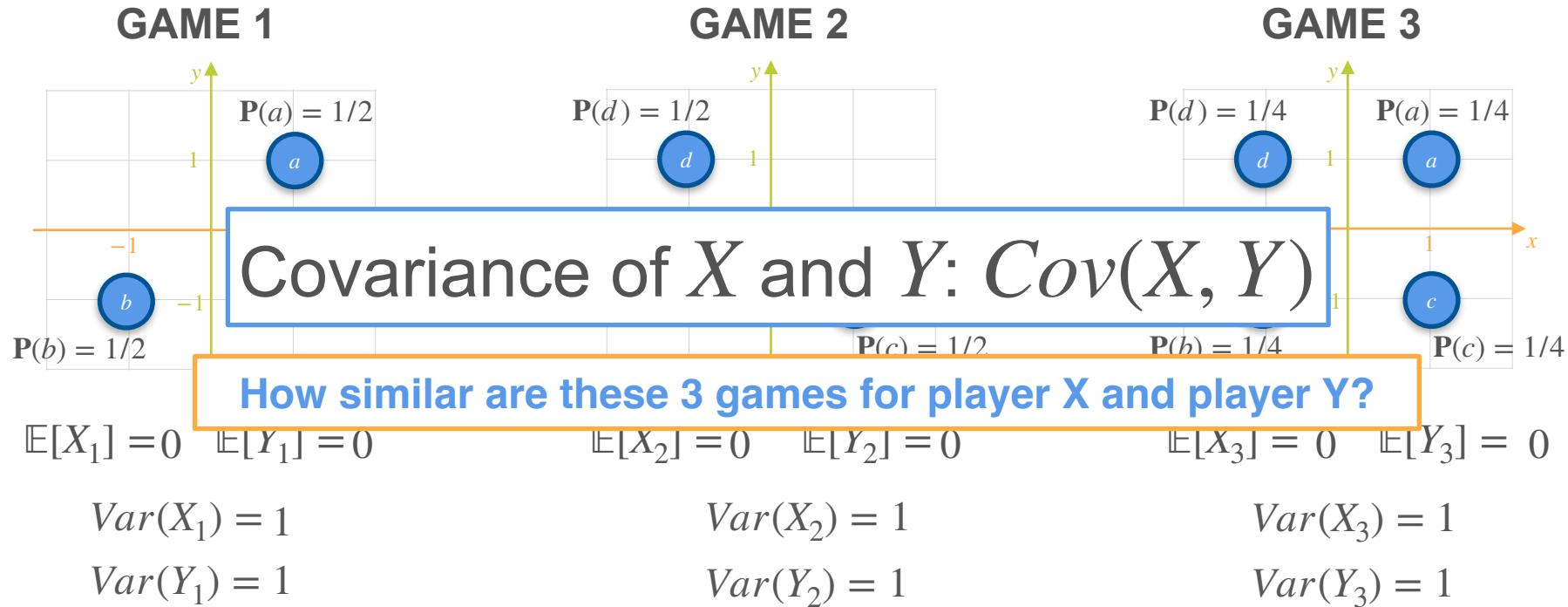
$$Var(Y_2) = 1$$

$$\mathbb{E}[X_3] = 0 \quad \mathbb{E}[Y_3] = 0$$

$$Var(X_3) = 1$$

$$Var(Y_3) = 1$$

Covariance of a Probability Distribution: Motivation



Covariance of a Probability Distribution: Motivation

Covariance of a Probability Distribution: Motivation

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

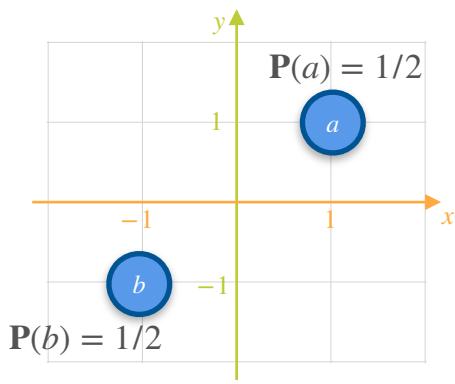
Covariance of a Probability Distribution: Motivation

Covariance of X and Y : $\text{Cov}(X, Y)$

$$\text{Cov}(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

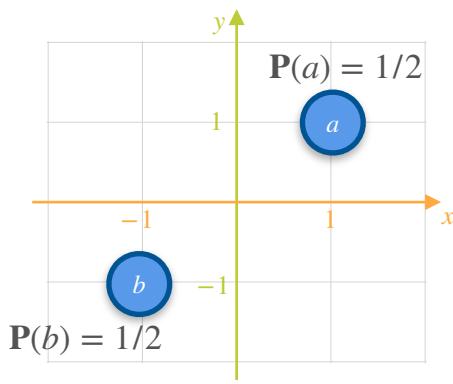
GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

GAME 1

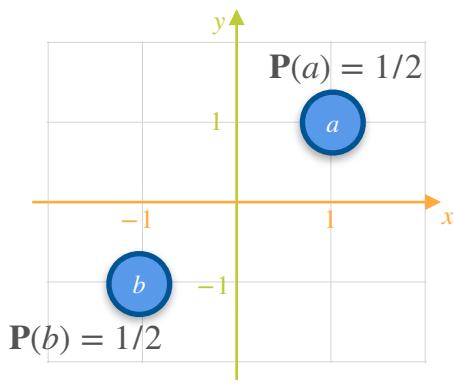


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
-----	-----	-----------------	-----------------	------------------------------

Covariance of a Probability Distribution: Motivation

GAME 1

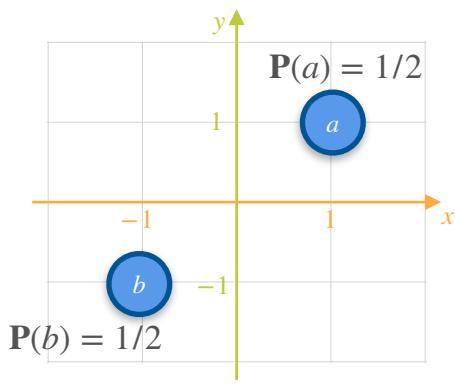


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1				
-1				

Covariance of a Probability Distribution: Motivation

GAME 1

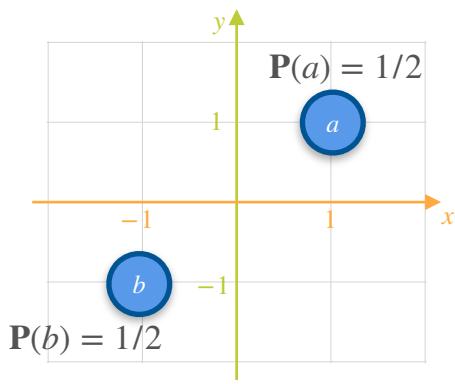


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1			
-1	-1			

Covariance of a Probability Distribution: Motivation

GAME 1

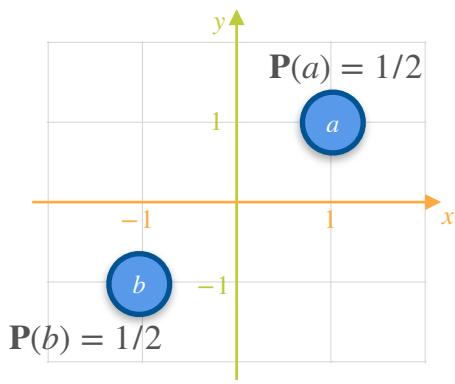


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	-1

Covariance of a Probability Distribution: Motivation

GAME 1

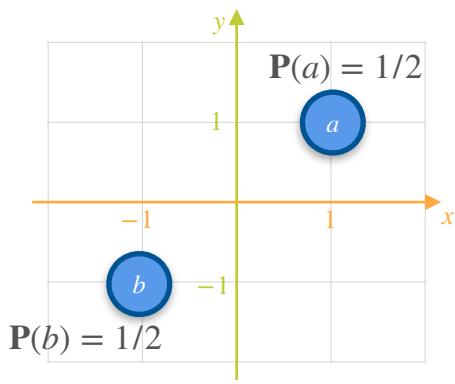


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	-1

Covariance of a Probability Distribution: Motivation

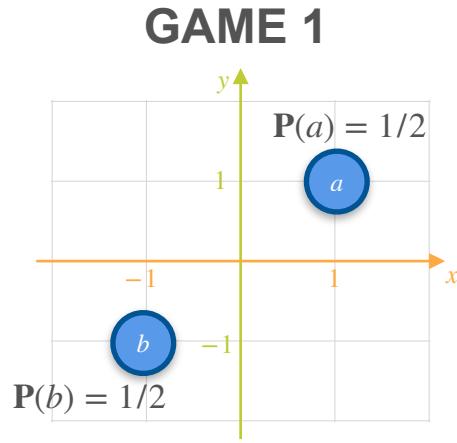
GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

Covariance of a Probability Distribution: Motivation

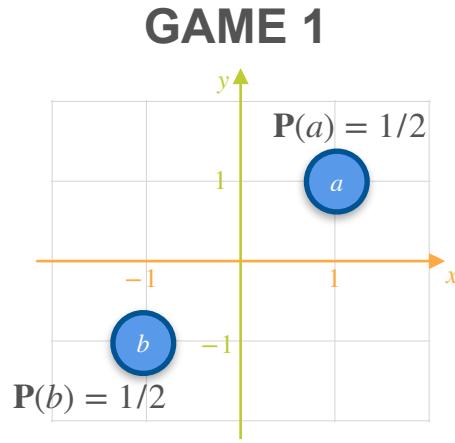


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

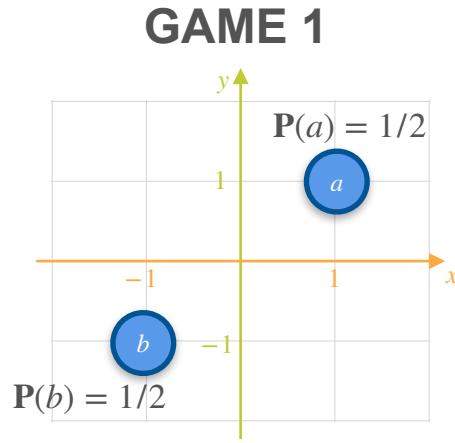


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

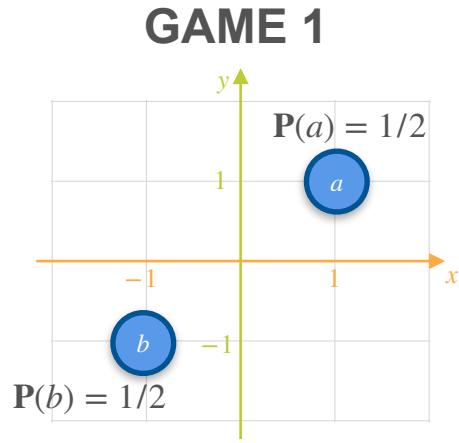


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) = 2$$

Covariance of a Probability Distribution: Motivation

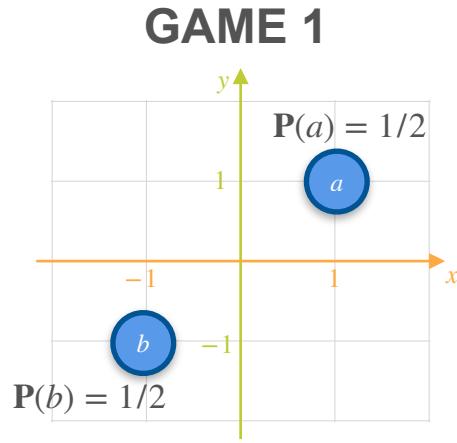


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \sum (x_i - \mu_x)(y_i - \mu_y) = 2$$

Covariance of a Probability Distribution: Motivation

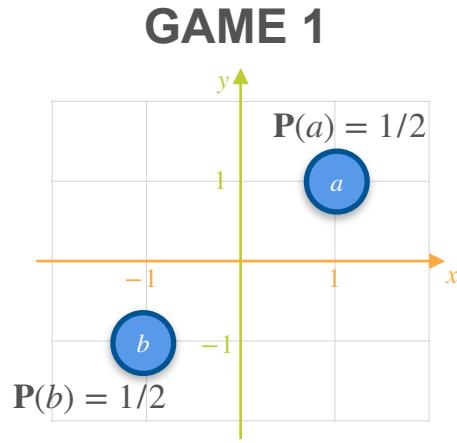


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = 2$$

Covariance of a Probability Distribution: Motivation

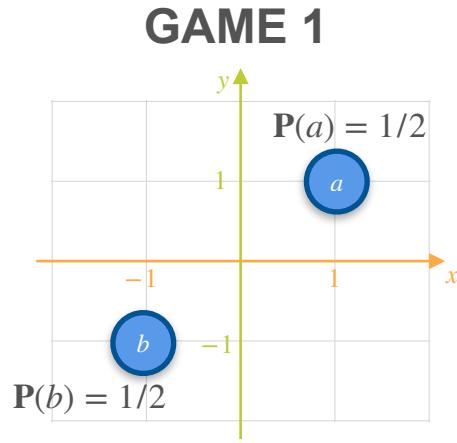


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{2}{2}$$

Covariance of a Probability Distribution: Motivation



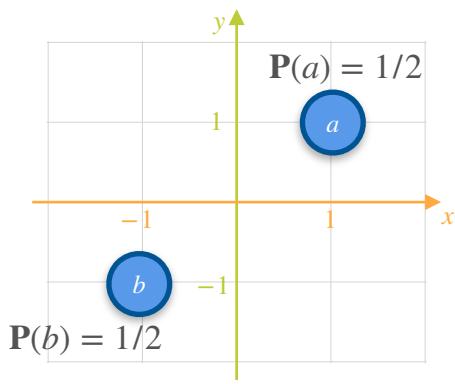
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{2}{2} = 1$$

Covariance of a Probability Distribution: Motivation

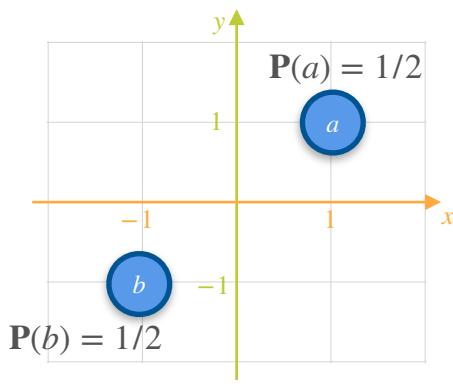
GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

GAME 1

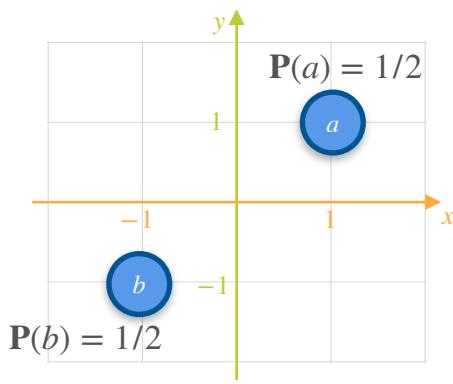


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



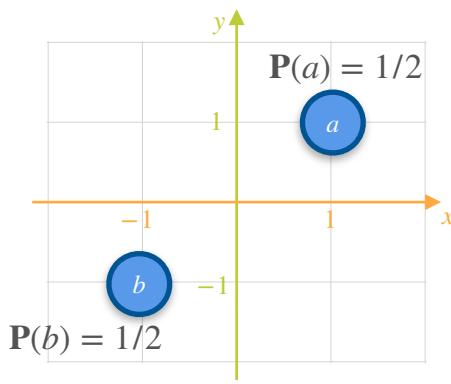
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

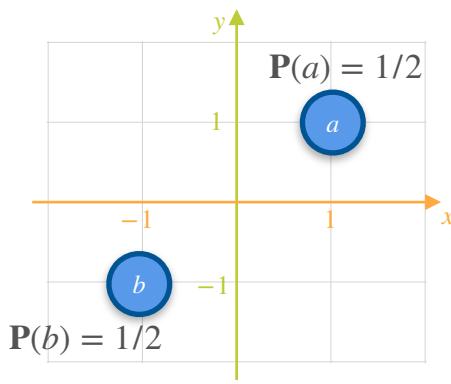
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
-----	-----	-----------------	-----------------	------------------------------

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

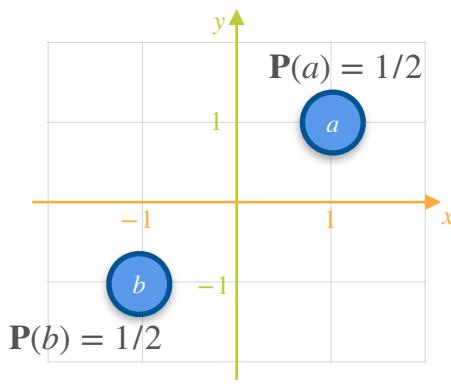
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1				
-1				

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

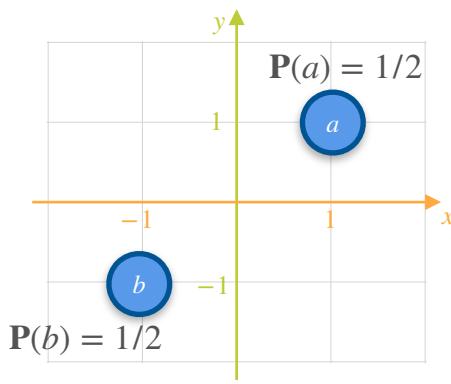
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	0	0	0
-1	-1	-2	-2	4

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

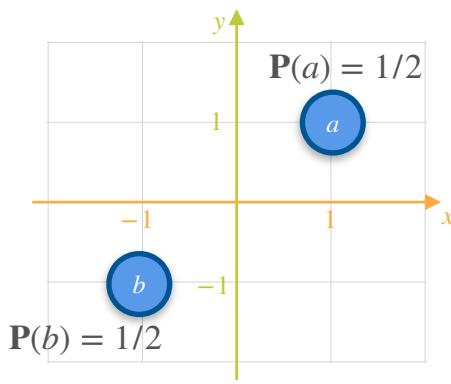
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	-1

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

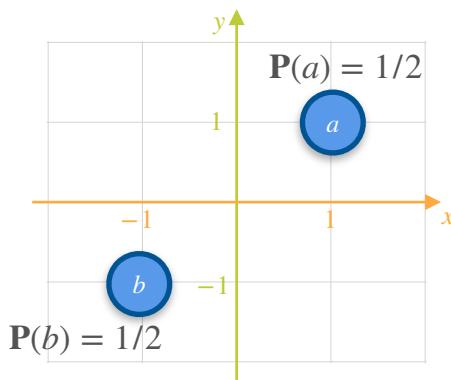
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	-1

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

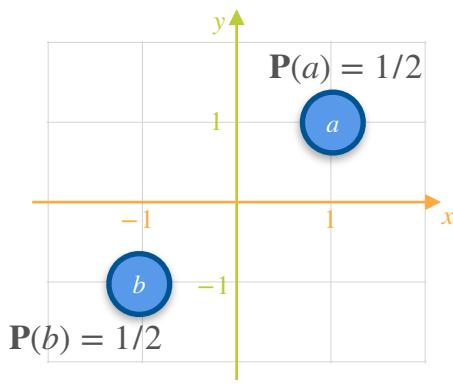
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

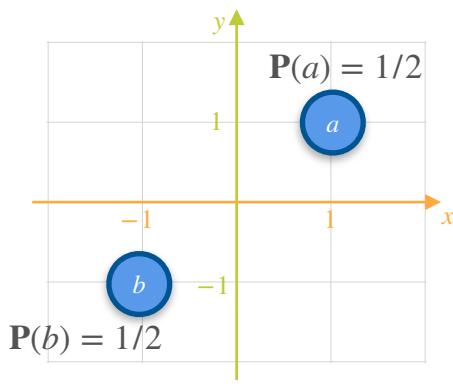
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

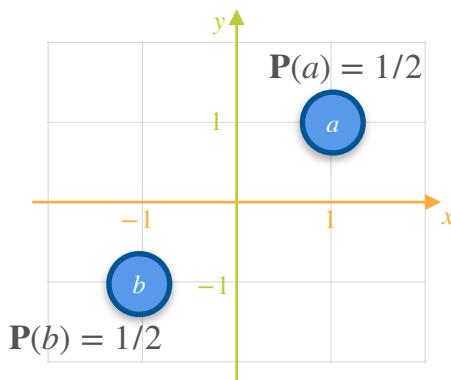
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

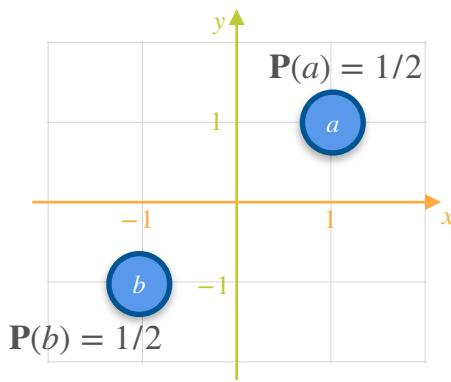
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) = 2$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$\mu_x = \mathbb{E}[X_1] = 0$$

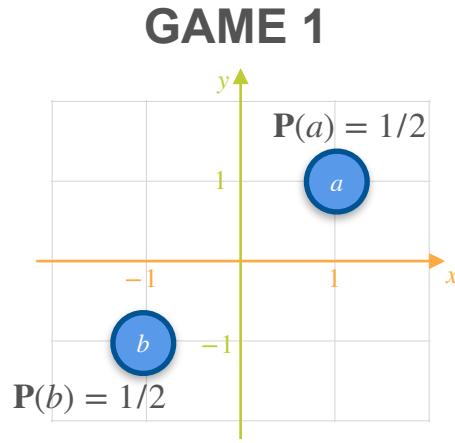
$$\mu_y = \mathbb{E}[Y_1] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \sum (x_i - \mu_x)(y_i - \mu_y) = 2$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_1] = 0$$

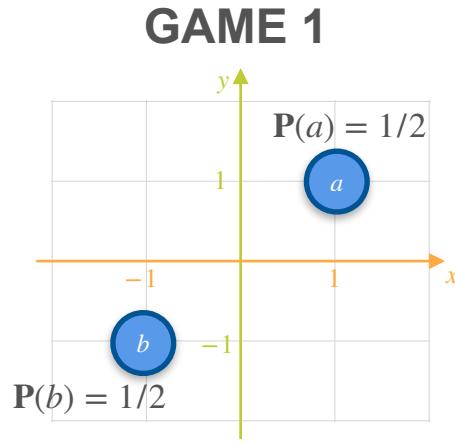
$$\mu_y = \mathbb{E}[Y_1] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = 2$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

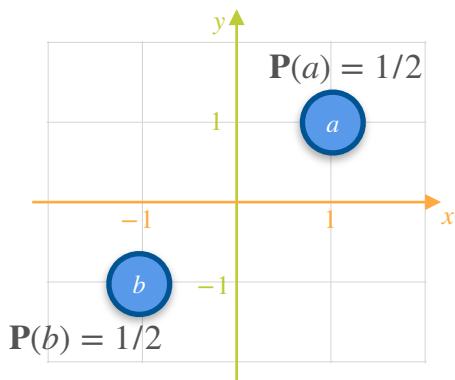
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{2}{2}$$

Covariance of a Probability Distribution: Motivation

GAME 1



$$\mu_x = \mathbb{E}[X_1] = 0$$

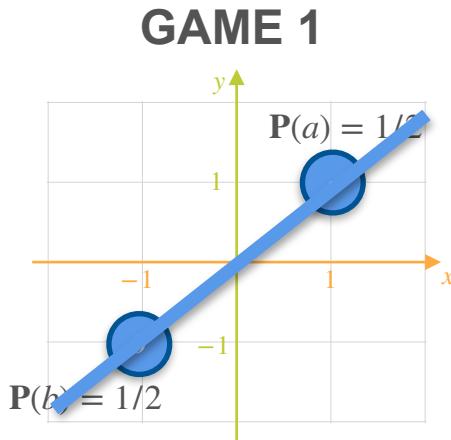
$$\mu_y = \mathbb{E}[Y_1] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{2}{2} = 1$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_1] = 0$$

$$\mu_y = \mathbb{E}[Y_1] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{2}{2} = 1$$

Covariance of a Probability Distribution: Motivation

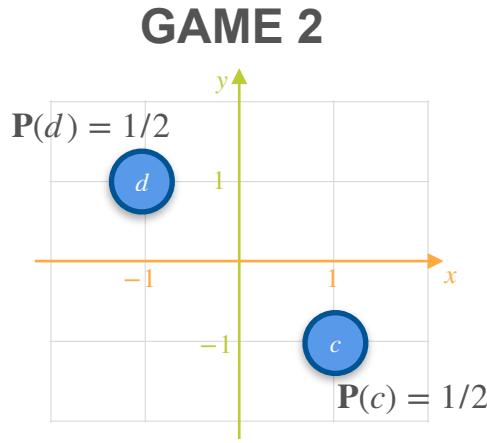
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

GAME 2

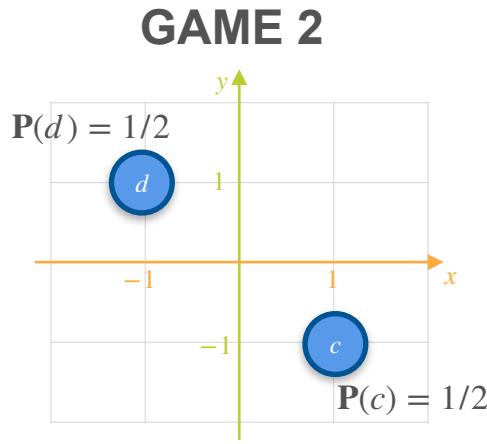
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

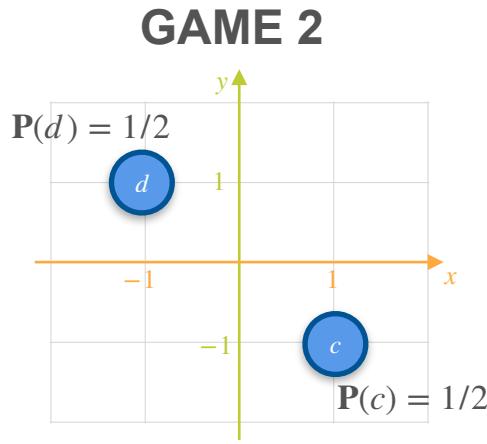
Covariance of a Probability Distribution: Motivation



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_2] = 0$$

Covariance of a Probability Distribution: Motivation

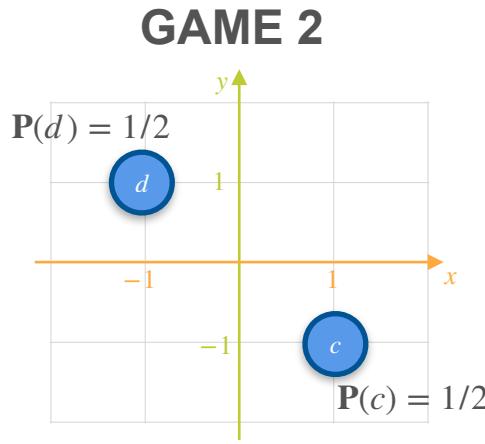


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



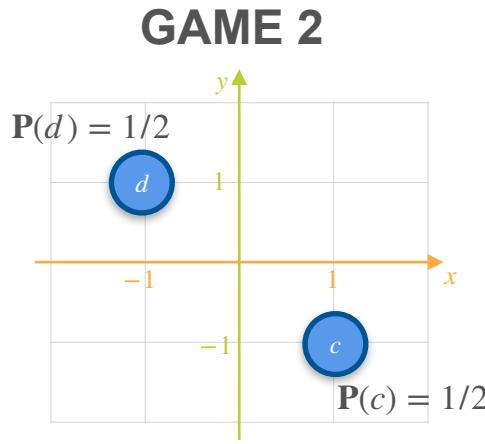
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
-----	-----	-----------------	-----------------	------------------------------

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

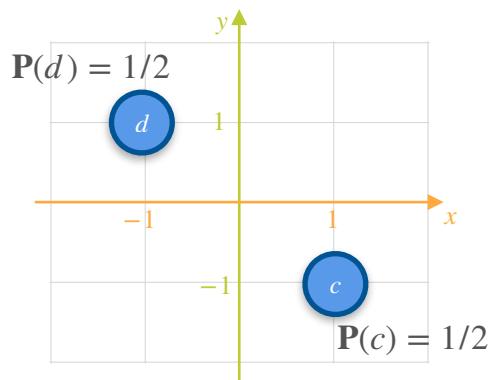
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1				
-1				

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 2



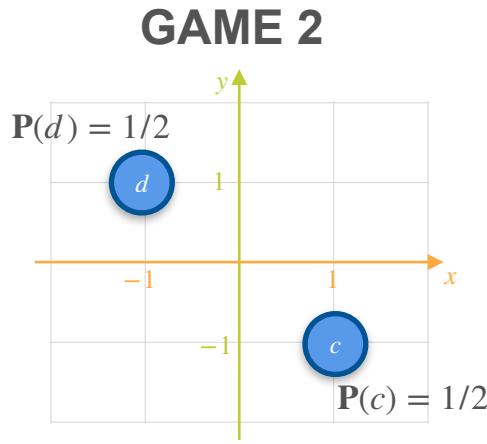
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	0	-2	0
-1	1	-2	0	-2

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



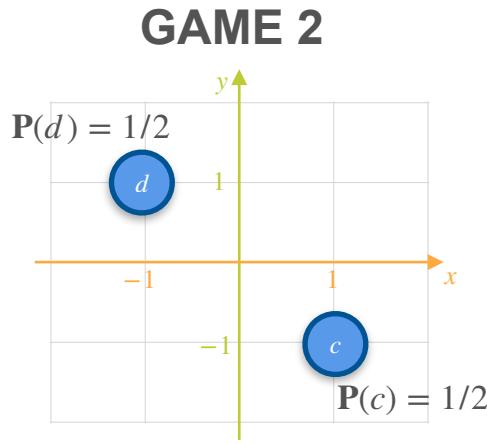
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	1

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



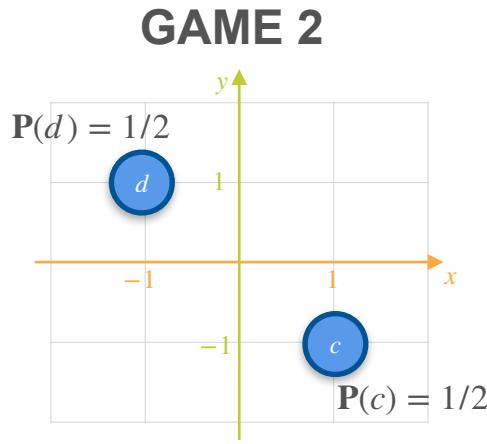
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	1

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



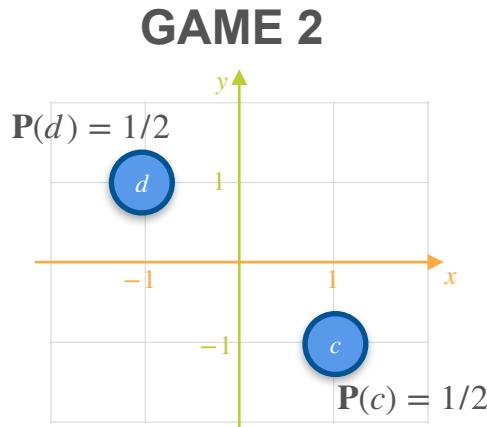
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

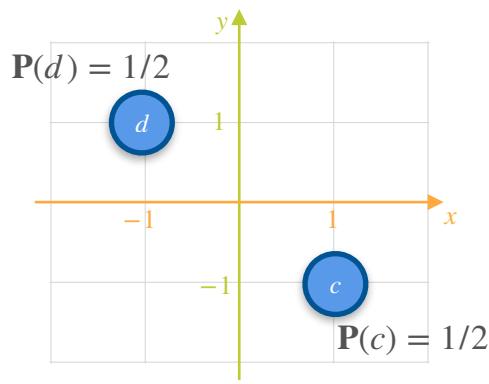
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

GAME 2



$$\mu_x = \mathbb{E}[X_2] = 0$$

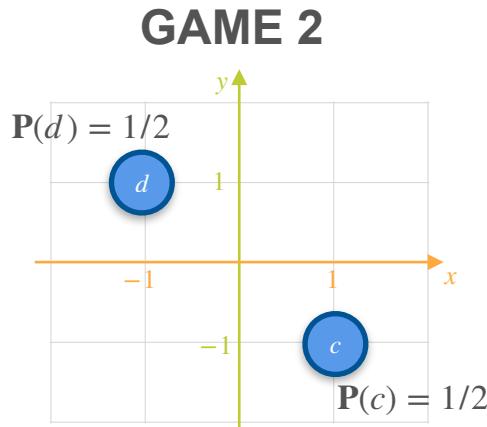
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

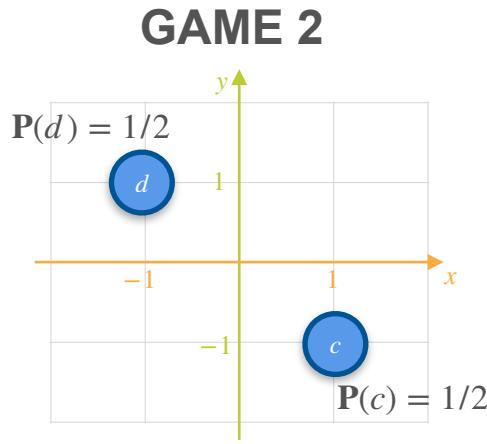
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$\sum (x_i - \mu_x)(y_i - \mu_y) = -2$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

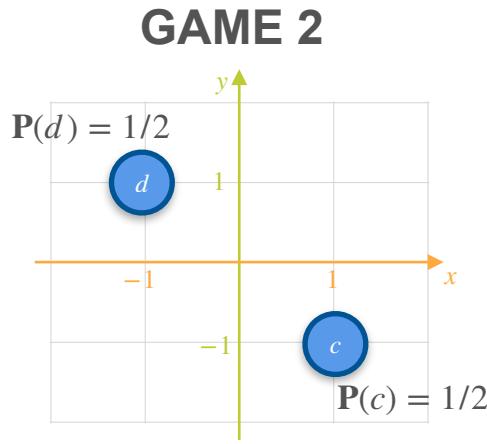
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$Cov(X, Y) = \sum (x_i - \mu_x)(y_i - \mu_y) = -2$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

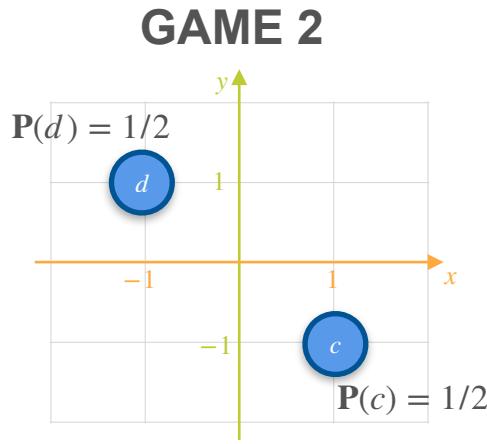
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = -2$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

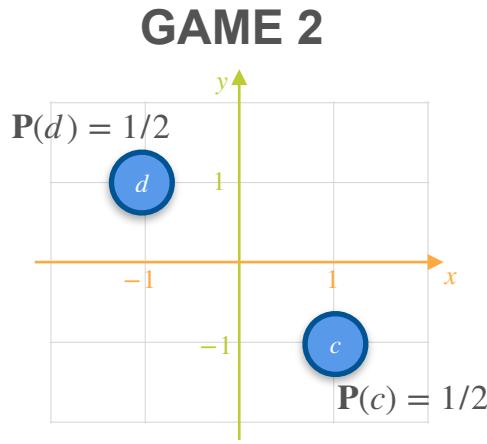
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{-2}{2}$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

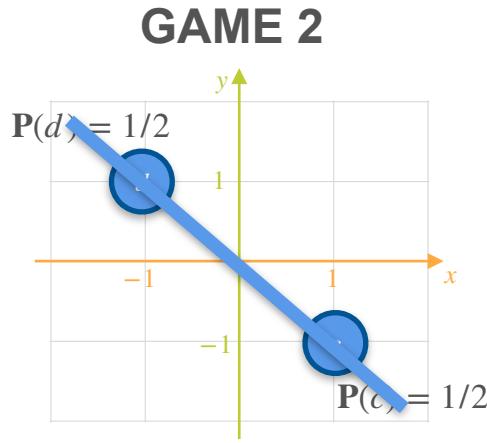
$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{-2}{2} = -1$$

Covariance of a Probability Distribution: Motivation



$$\mu_x = \mathbb{E}[X_2] = 0$$

$$\mu_y = \mathbb{E}[Y_2] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	-1	1	-1	-1
-1	1	-1	1	-1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{-2}{2} = -1$$

Covariance of a Probability Distribution: Motivation

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

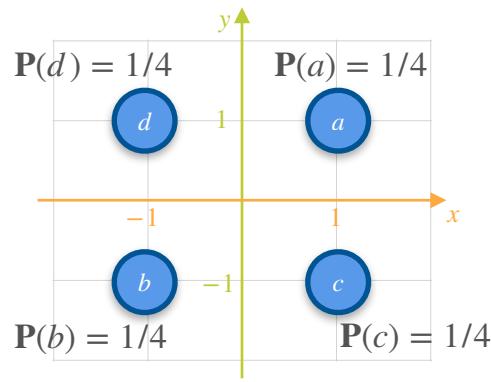
Covariance of a Probability Distribution: Motivation

GAME 3

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

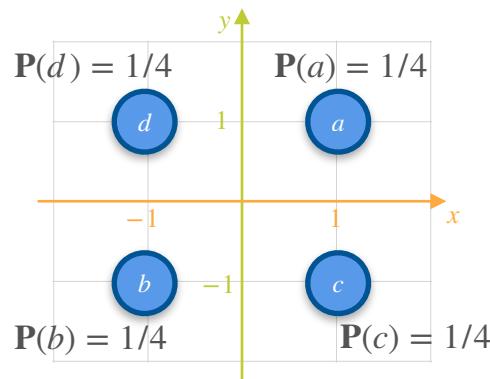
GAME 3



$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of a Probability Distribution: Motivation

GAME 3

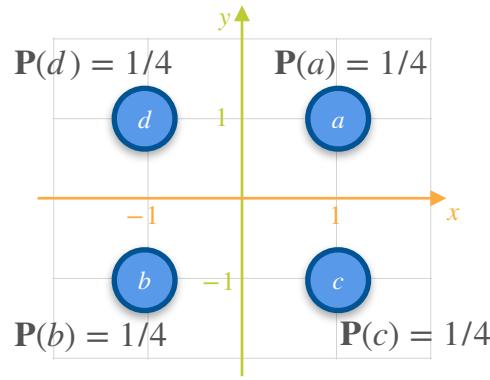


$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_3] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



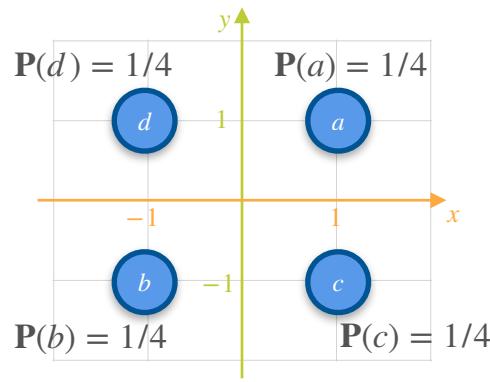
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



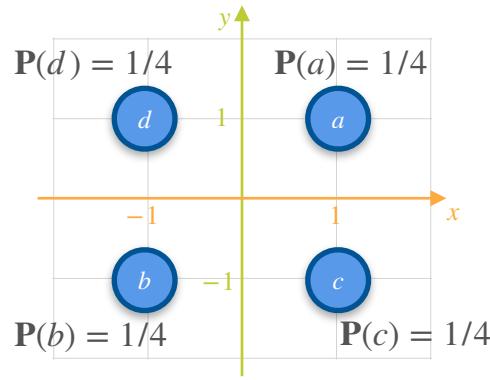
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

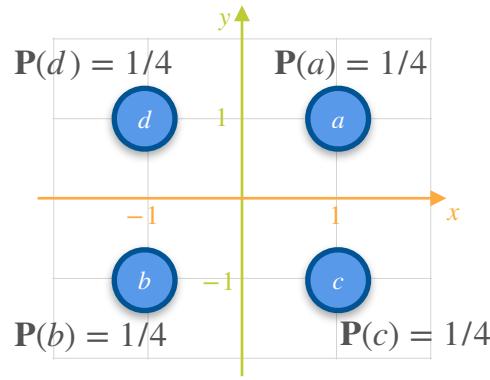
$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1				
1				
-1				
-1				

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

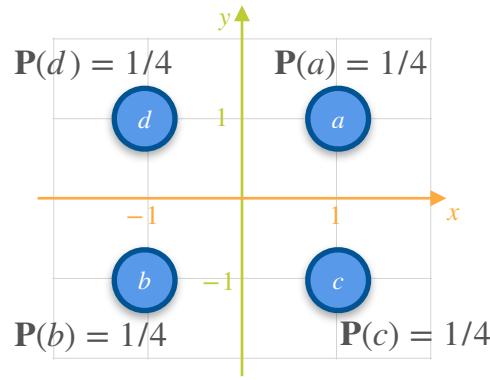
$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	0	0	0
1	-1	0	-2	0
-1	1	-2	0	-2
-1	-1	-2	-2	4

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

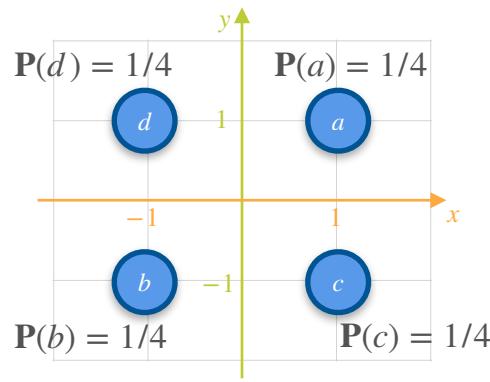
$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	1	-1	-1
-1	1	-1	1	-1
-1	-1	-1	-1	1

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

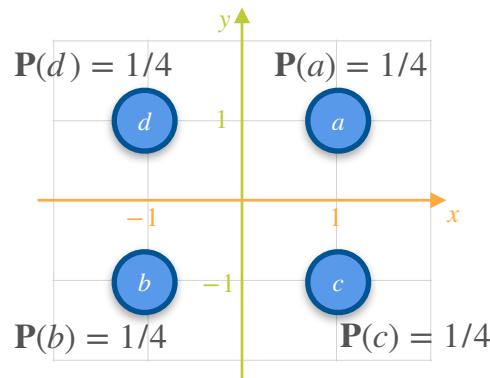
$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	-1	1
-1	1	1	-1	-1
-1	-1	-1	-1	-1

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

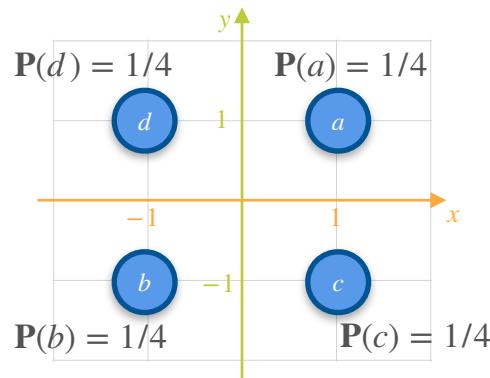
$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

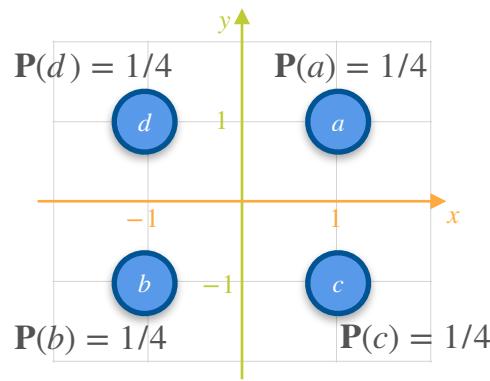
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

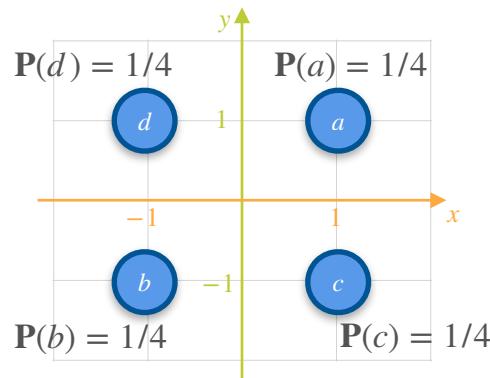
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) =$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

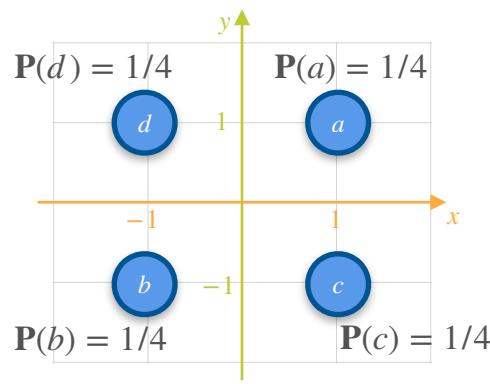
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$\sum (x_i - \mu_x)(y_i - \mu_y) = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

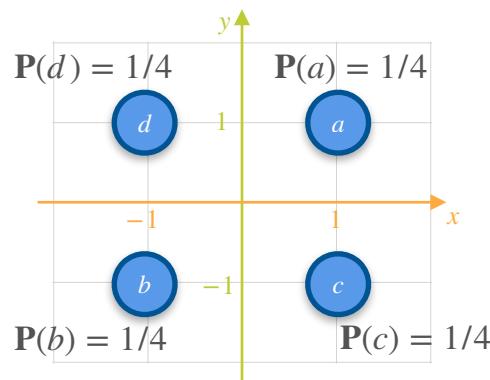
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$Cov(X, Y) = \sum (x_i - \mu_x)(y_i - \mu_y) = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

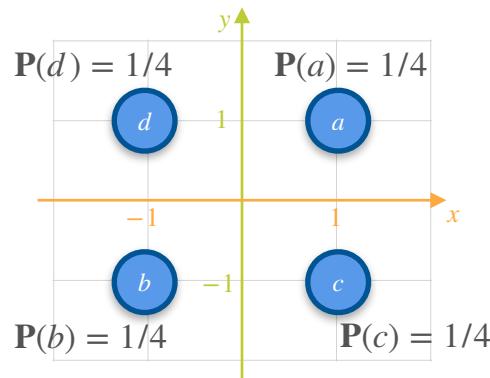
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

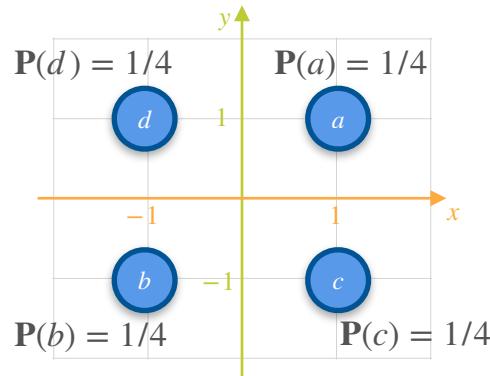
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{0}{4}$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

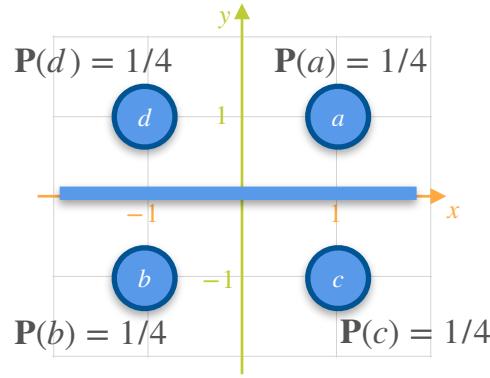
$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{0}{4} = 0$$

Covariance of a Probability Distribution: Motivation

GAME 3



$$\mu_x = \mathbb{E}[X_3] = 0$$

$$\mu_y = \mathbb{E}[Y_3] = 0$$

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

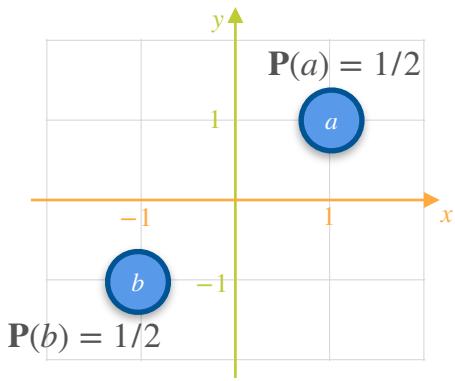
x	y	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$
1	1	1	1	1
1	-1	-1	1	-1
-1	1	1	-1	-1
-1	-1	-1	-1	1

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{0}{4} = 0$$

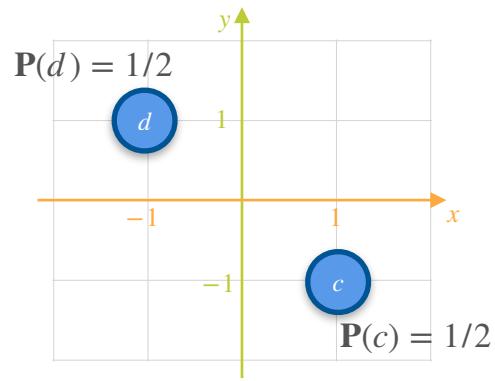
Covariance of a Probability Distribution: Motivation

Covariance of a Probability Distribution: Motivation

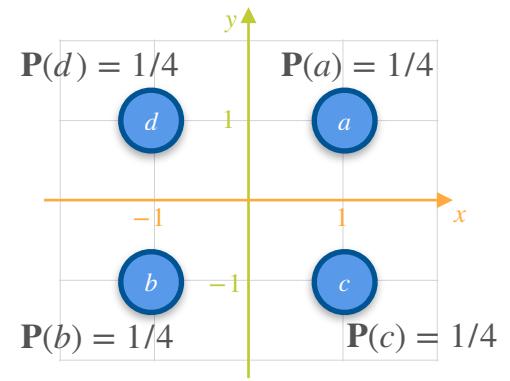
GAME 1



GAME 2

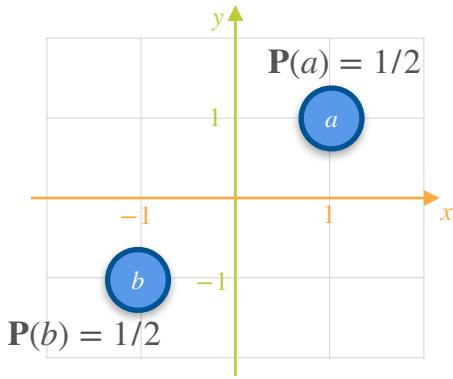


GAME 3



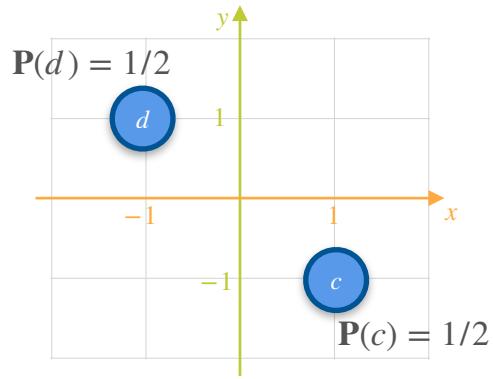
Covariance of a Probability Distribution: Motivation

GAME 1



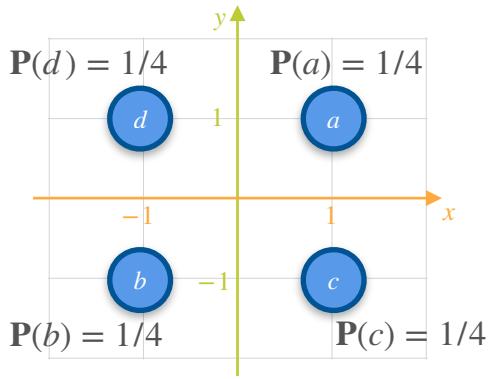
$$Cov(X, Y) = 1$$

GAME 2



$$Cov(X, Y) = -1$$

GAME 3



$$Cov(X, Y) = 0$$

Covariance of a Probability Distribution: Motivation

Covariance of a Probability Distribution: Motivation

GAME 4

Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each

Covariance of a Probability Distribution: Motivation

GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each

b: Both players lose \$1 each

c: Neither players wins nor lose anything

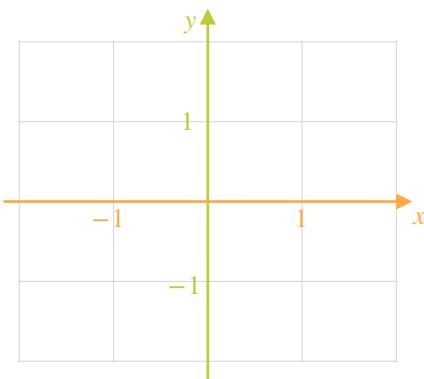
Covariance of a Probability Distribution: Motivation

GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

c : Neither players wins nor lose anything



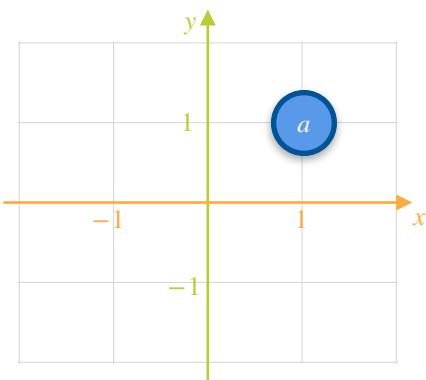
Covariance of a Probability Distribution: Motivation

GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

c : Neither players wins nor lose anything



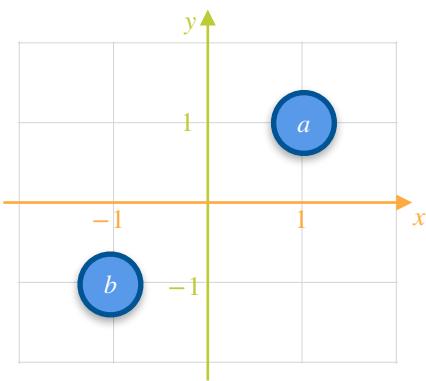
Covariance of a Probability Distribution: Motivation

GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

c : Neither players wins nor lose anything



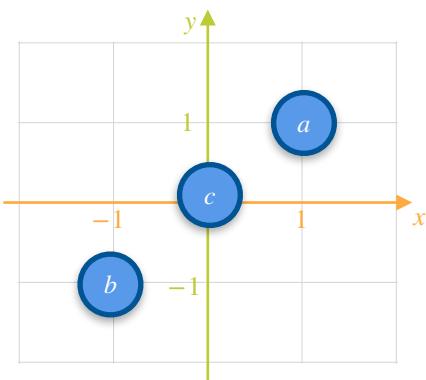
Covariance of a Probability Distribution: Motivation

GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

c : Neither players wins nor lose anything



Covariance of a Probability Distribution: Motivation

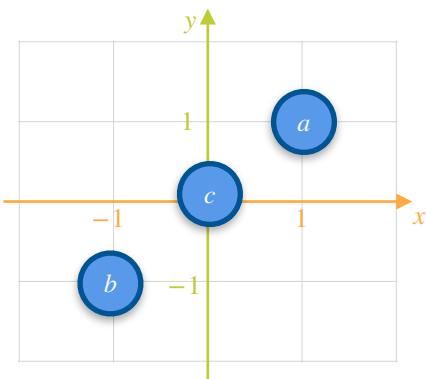
GAME 4

a : Both players win \$1 each

b : Both players lose \$1 each

c : Neither players wins nor lose anything

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

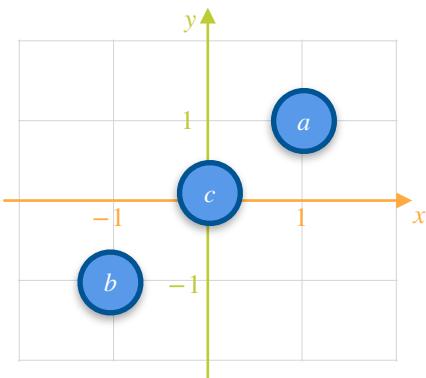
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each

c : Neither players wins nor lose anything

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

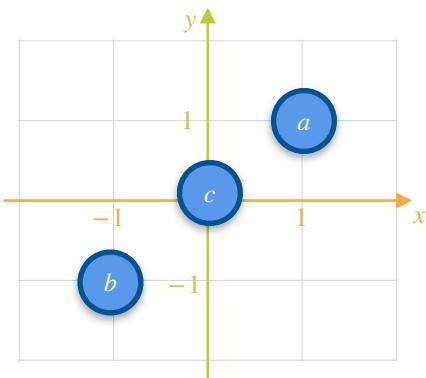
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each $P(b) = 1/3$

c : Neither players wins nor lose anything

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

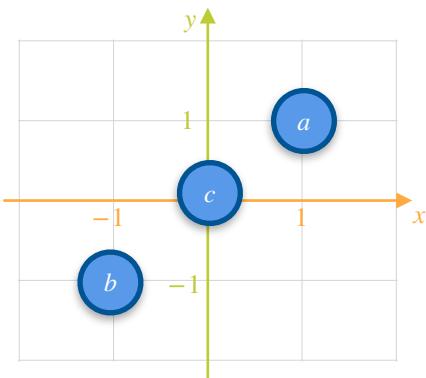
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each $P(b) = 1/3$

c : Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

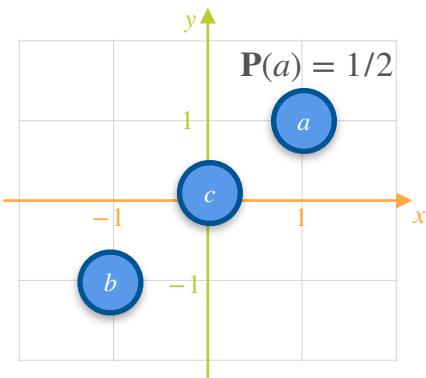
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each $P(b) = 1/3$

c : Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

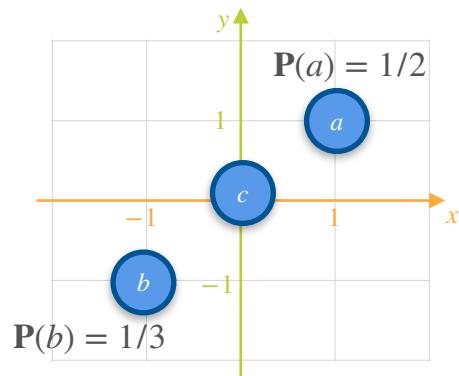
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each $P(b) = 1/3$

c : Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

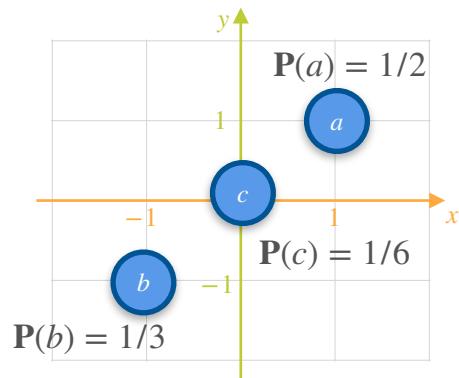
GAME 4

a : Both players win \$1 each $P(a) = 1/2$

b : Both players lose \$1 each $P(b) = 1/3$

c : Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



Covariance of a Probability Distribution: Motivation

GAME 4

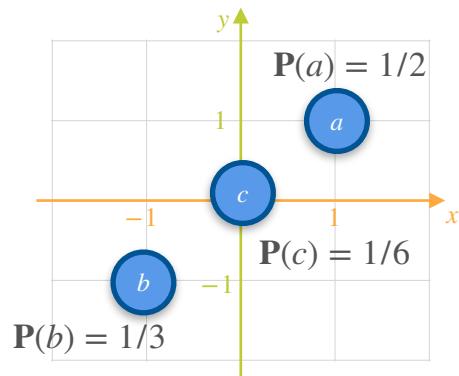
a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities

$$\mathbb{E}[X_4] =$$



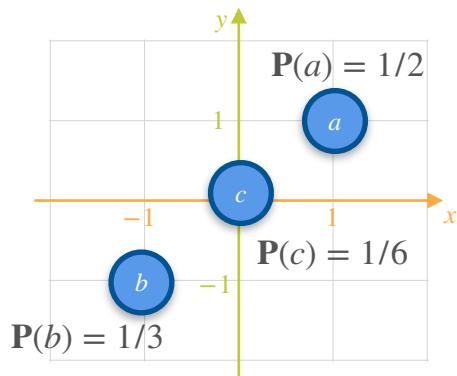
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$



Unequal Probabilities

$$\mathbb{E}[X_4] = \frac{1}{2}(1) + \frac{1}{6}(0) + \frac{1}{3}(-1) =$$

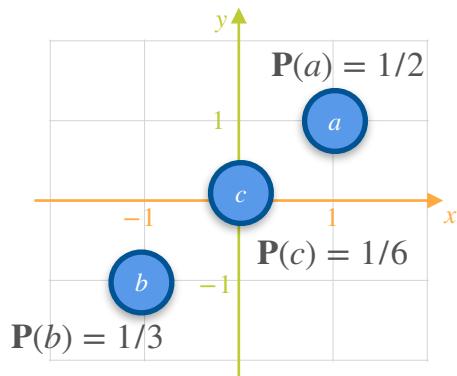
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$



Unequal Probabilities

$$\mathbb{E}[X_4] = \frac{1}{2}(1) + \frac{1}{6}(0) + \frac{1}{3}(-1) = \frac{1}{6}$$

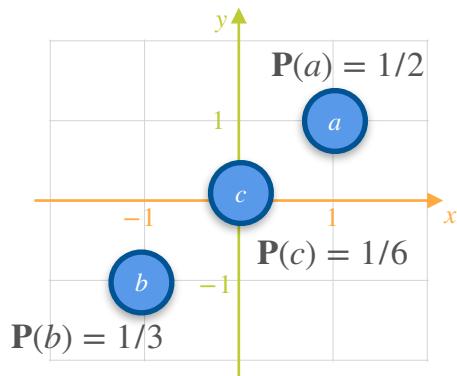
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$



Unequal Probabilities

$$\mathbb{E}[X_4] = \frac{1}{2}(1) + \frac{1}{6}(0) + \frac{1}{3}(-1) = \frac{1}{6}$$

$$\mathbb{E}[Y_4] =$$

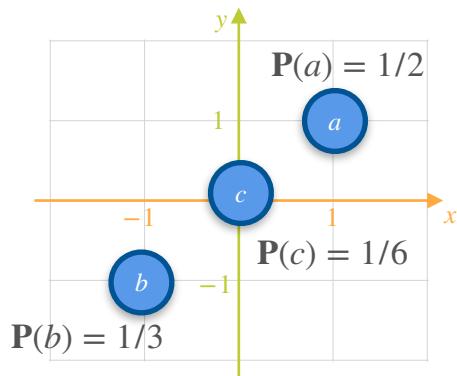
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$



Unequal Probabilities

$$\mathbb{E}[X_4] = \frac{1}{2}(1) + \frac{1}{6}(0) + \frac{1}{3}(-1) = \frac{1}{6}$$

$$\mathbb{E}[Y_4] = \frac{1}{2}(1) + \frac{1}{6}(0) + \frac{1}{3}(-1) = \frac{1}{6}$$

Covariance of a Probability Distribution: Motivation

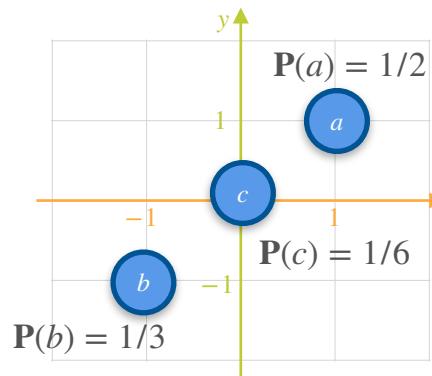
GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



$$\mathbb{E}[X_4] = \frac{1}{6}$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

Covariance of a Probability Distribution: Motivation

GAME 4

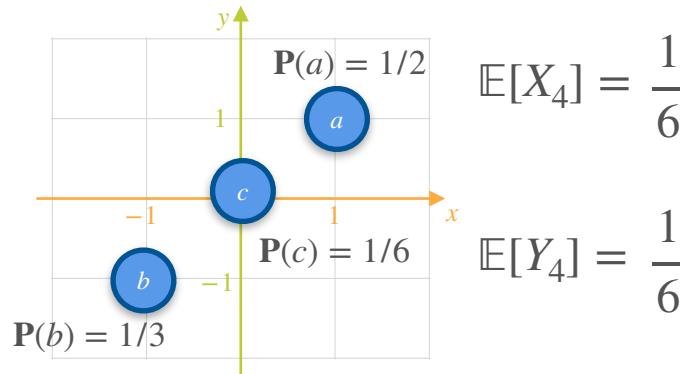
a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities

$$Var(X_4) =$$



$$\mathbb{E}[X_4] = \frac{1}{6}$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

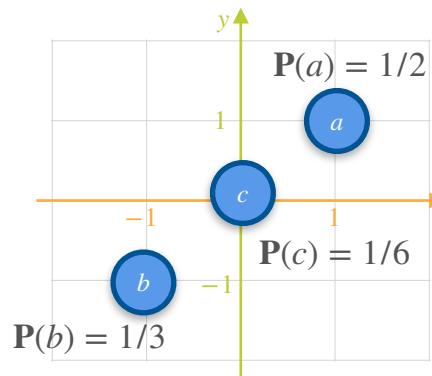
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $\mathbf{P}(a) = 1/2$

b: Both players lose \$1 each $\mathbf{P}(b) = 1/3$

c: Neither players wins nor lose anything $\mathbf{P}(c) = 1/6$



$$\mathbb{E}[X_4] = \frac{1}{6}$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

Unequal Probabilities

$$Var(X_4) = \sum_{n=1}^N (\mathbb{E}[X_4] - \mu_x)^2 \cdot \mathbf{P}(x_i)$$

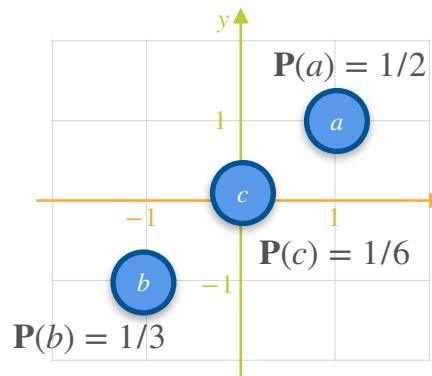
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $\mathbf{P}(a) = 1/2$

b: Both players lose \$1 each $\mathbf{P}(b) = 1/3$

c: Neither players wins nor lose anything $\mathbf{P}(c) = 1/6$



Unequal Probabilities

$$\text{Var}(X_4) = \sum_{n=1}^N (\mathbb{E}[X_4] - \mu_x)^2 \cdot \mathbf{P}(x_i)$$

$$\mathbb{E}[X_4] = \frac{1}{6} \quad \text{Var}(X_4) =$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

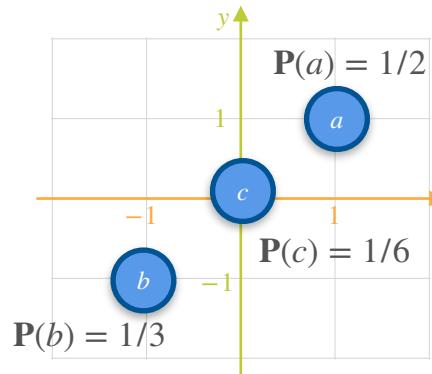
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $\mathbf{P}(a) = 1/2$

b: Both players lose \$1 each $\mathbf{P}(b) = 1/3$

c: Neither players wins nor lose anything $\mathbf{P}(c) = 1/6$



Unequal Probabilities

$$\text{Var}(X_4) = \sum_{n=1}^N (\mathbb{E}[X_4] - \mu_x)^2 \cdot \mathbf{P}(x_i)$$

$$\mathbb{E}[X_4] = \frac{1}{6}$$

$$\text{Var}(X_4) = \frac{1}{2}\left(1 - \frac{1}{6}\right)^2 + \frac{1}{6}\left(0 - \frac{1}{6}\right)^2 + \frac{1}{3}\left(-1 - \frac{1}{6}\right)^2$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

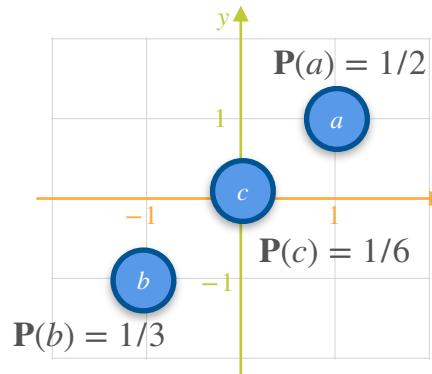
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $\mathbf{P}(a) = 1/2$

b: Both players lose \$1 each $\mathbf{P}(b) = 1/3$

c: Neither players wins nor lose anything $\mathbf{P}(c) = 1/6$



Unequal Probabilities

$$\text{Var}(X_4) = \sum_{n=1}^N (\mathbb{E}[X_4] - \mu_x)^2 \cdot \mathbf{P}(x_i)$$

$$\mathbb{E}[X_4] = \frac{1}{6}$$

$$\text{Var}(X_4) = \frac{1}{2}\left(1 - \frac{1}{6}\right)^2 + \frac{1}{6}\left(0 - \frac{1}{6}\right)^2 + \frac{1}{3}\left(-1 - \frac{1}{6}\right)^2$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

$$= 0.806$$

Covariance of a Probability Distribution: Motivation

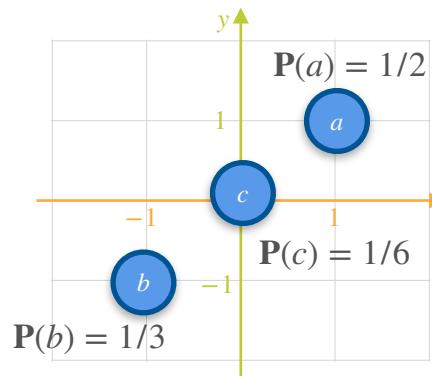
GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



$$\mathbb{E}[X_4] = \frac{1}{6} \quad \text{Var}(X_4) = 0.806$$

$$\mathbb{E}[Y_4] = \frac{1}{6}$$

Covariance of a Probability Distribution: Motivation

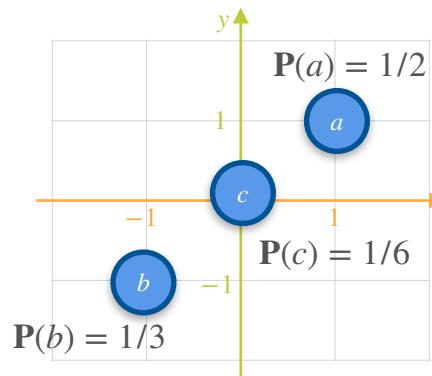
GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$

Unequal Probabilities



$$\mathbb{E}[X_4] = \frac{1}{6} \quad \text{Var}(X_4) = 0.806$$

$$\mathbb{E}[Y_4] = \frac{1}{6} \quad \text{Var}(Y_4) = 0.806$$

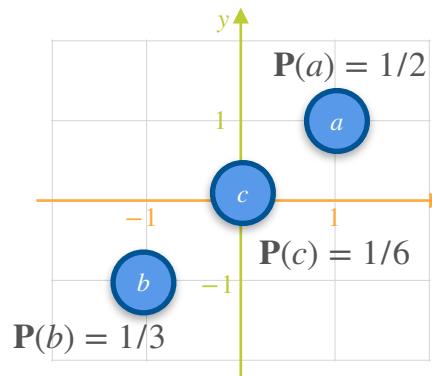
Covariance of a Probability Distribution: Motivation

GAME 4

a: Both players win \$1 each $P(a) = 1/2$

b: Both players lose \$1 each $P(b) = 1/3$

c: Neither players wins nor lose anything $P(c) = 1/6$



Unequal Probabilities

$$\mathbb{E}[X_4] = \frac{1}{6} \quad \text{Var}(X_4) = 0.806$$

$$\mathbb{E}[Y_4] = \frac{1}{6} \quad \text{Var}(Y_4) = 0.806$$

$\text{Cov}(X, Y) = ?$

Covariance of Probability Distributions

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n}$$

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$


equal probabilities

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$


equal probabilities

$$Cov(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y)$$

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$

equal probabilities

$$Cov(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y)$$

unequal probabilities

Covariance of Probability Distributions

$$Cov(X, Y) = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n} = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$

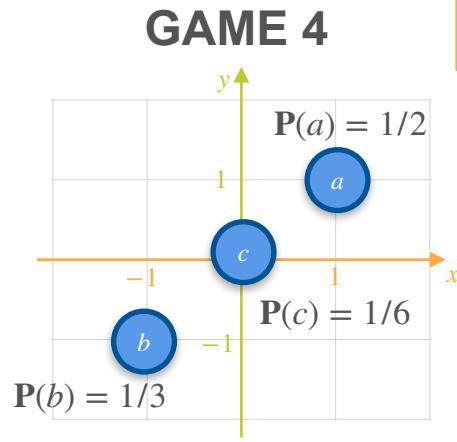
equal probabilities

$$Cov(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y)$$

unequal probabilities

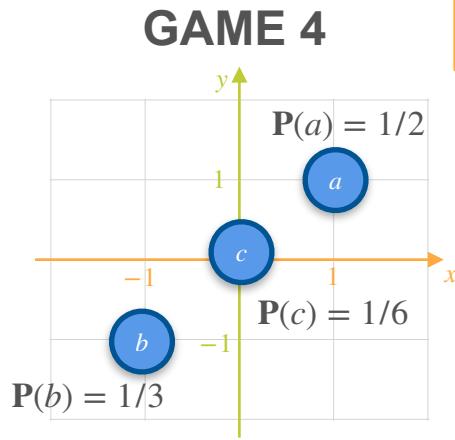
$$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Covariance of a Probability Distribution: Motivation



Unequal Probabilities

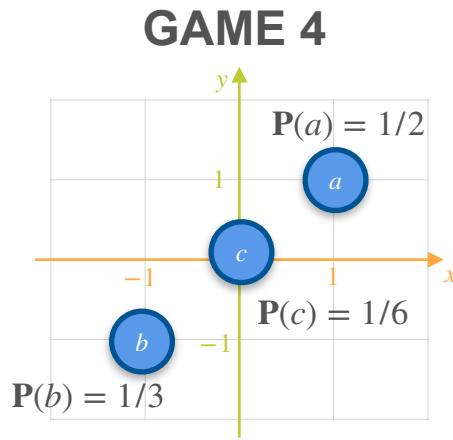
Covariance of a Probability Distribution: Motivation



Unequal Probabilities

$$Cov(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Covariance of a Probability Distribution: Motivation

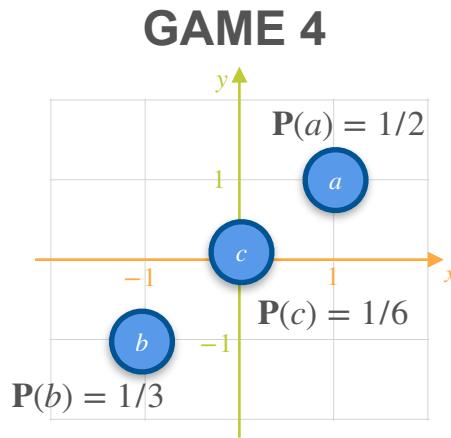


Unequal Probabilities

$$\text{Cov}(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\mu_x = \mu_y = \mathbb{E}[X_4] = \mathbb{E}[Y_4] = 1/6$$

Covariance of a Probability Distribution: Motivation



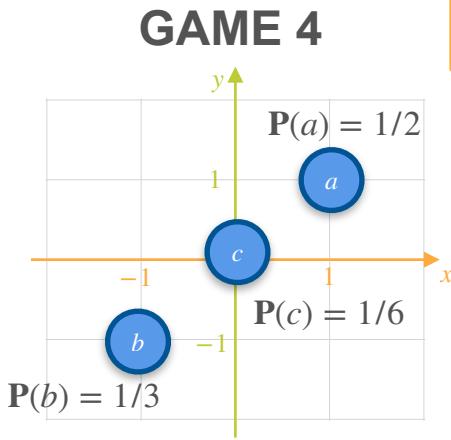
Unequal Probabilities

$$Cov(X, Y) = \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\mu_x = \mu_y = \mathbb{E}[X_4] = \mathbb{E}[Y_4] = 1/6$$

$$Var(X_4) = Var(Y_4) = 0.806$$

Covariance of a Probability Distribution: Motivation



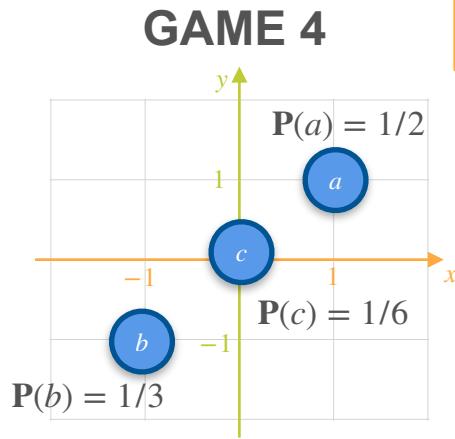
Unequal Probabilities

$$\begin{aligned} \text{Cov}(X, Y) &= \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \frac{1}{2}\left(1 - \frac{1}{6}\right)^2 + \frac{1}{6}\left(0 - \frac{1}{6}\right)^2 + \frac{1}{3}\left(-1 - \frac{1}{6}\right)^2 \end{aligned}$$

$$\mu_x = \mu_y = \mathbb{E}[X_4] = \mathbb{E}[Y_4] = 1/6$$

$$\text{Var}(X_4) = \text{Var}(Y_4) = 0.806$$

Covariance of a Probability Distribution: Motivation



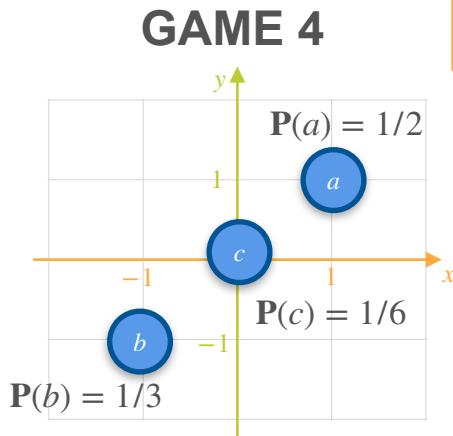
Unequal Probabilities

$$\begin{aligned} \text{Cov}(X, Y) &= \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \frac{1}{2}\left(1 - \frac{1}{6}\right)^2 + \frac{1}{6}\left(0 - \frac{1}{6}\right)^2 + \frac{1}{3}\left(-1 - \frac{1}{6}\right)^2 \end{aligned}$$

$$\mu_x = \mu_y = \mathbb{E}[X_4] = \mathbb{E}[Y_4] = 1/6$$

$$\text{Var}(X_4) = \text{Var}(Y_4) = 0.806$$

Covariance of a Probability Distribution: Motivation



Unequal Probabilities

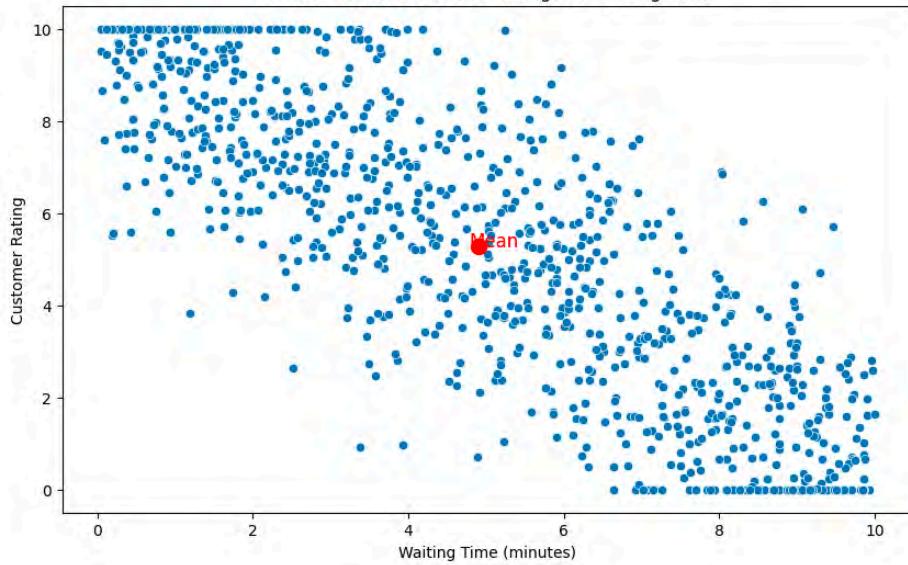
$$\begin{aligned} \text{Cov}(X, Y) &= \sum p_{XY}(x_i, y_i)(x_i - \mu_x)(y_i - \mu_y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \frac{1}{2}\left(1 - \frac{1}{6}\right)^2 + \frac{1}{6}\left(0 - \frac{1}{6}\right)^2 + \frac{1}{3}\left(-1 - \frac{1}{6}\right)^2 \end{aligned}$$

$$\mu_x = \mu_y = \mathbb{E}[X_4] = \mathbb{E}[Y_4] = 1/6$$

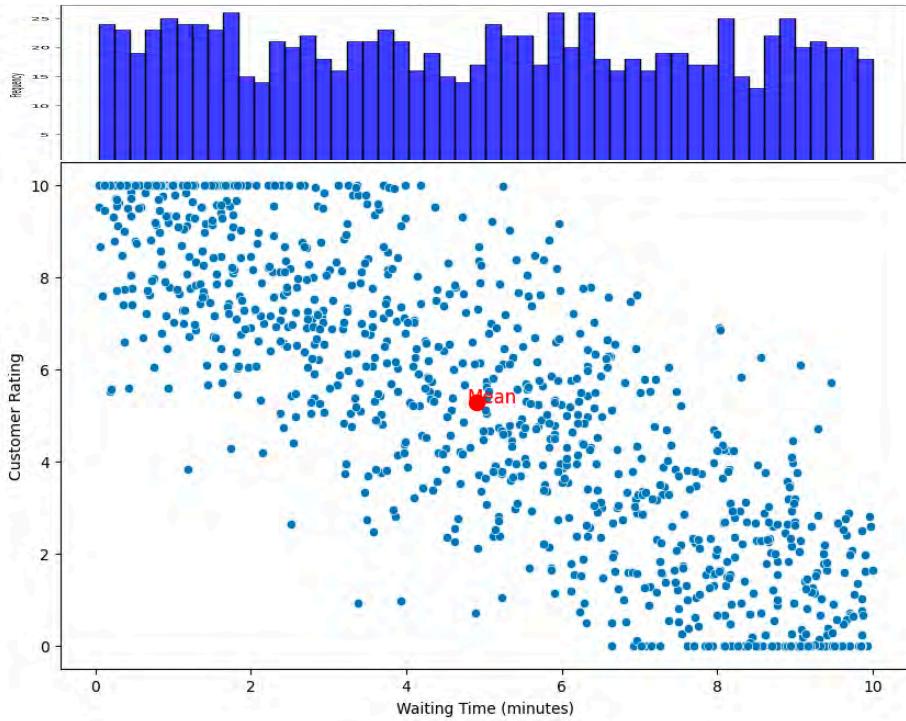
$$\text{Var}(X_4) = \text{Var}(Y_4) = 0.806$$

$$\text{Cov}(X, Y) = 0.806$$

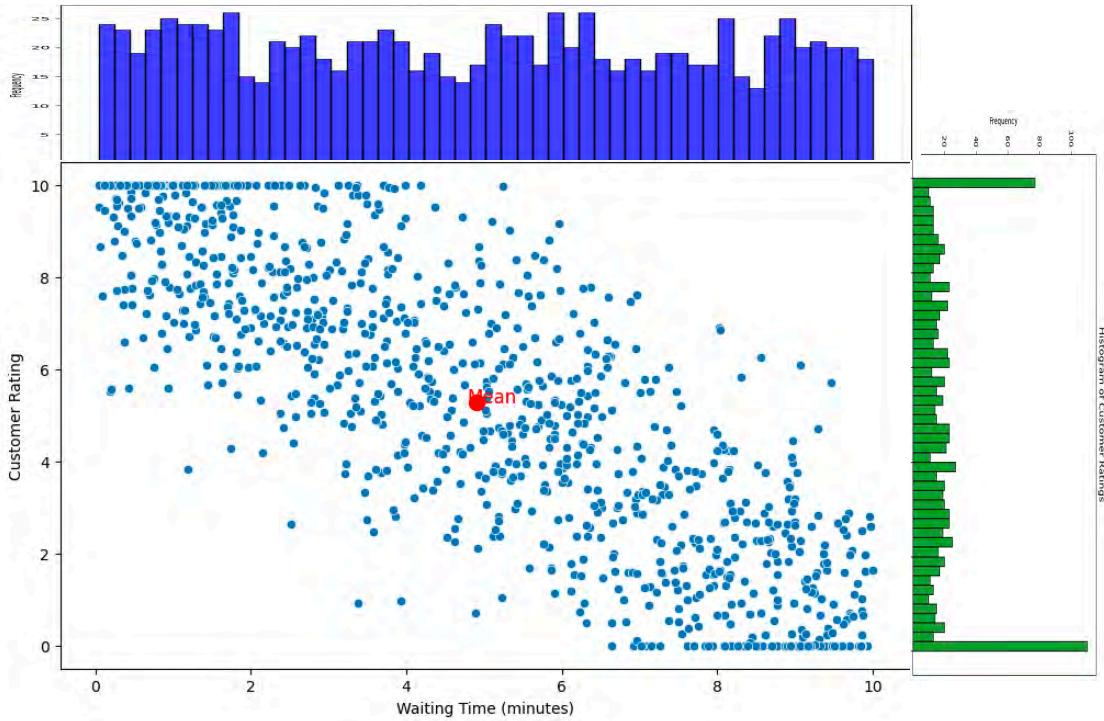
Covariance?



Covariance?

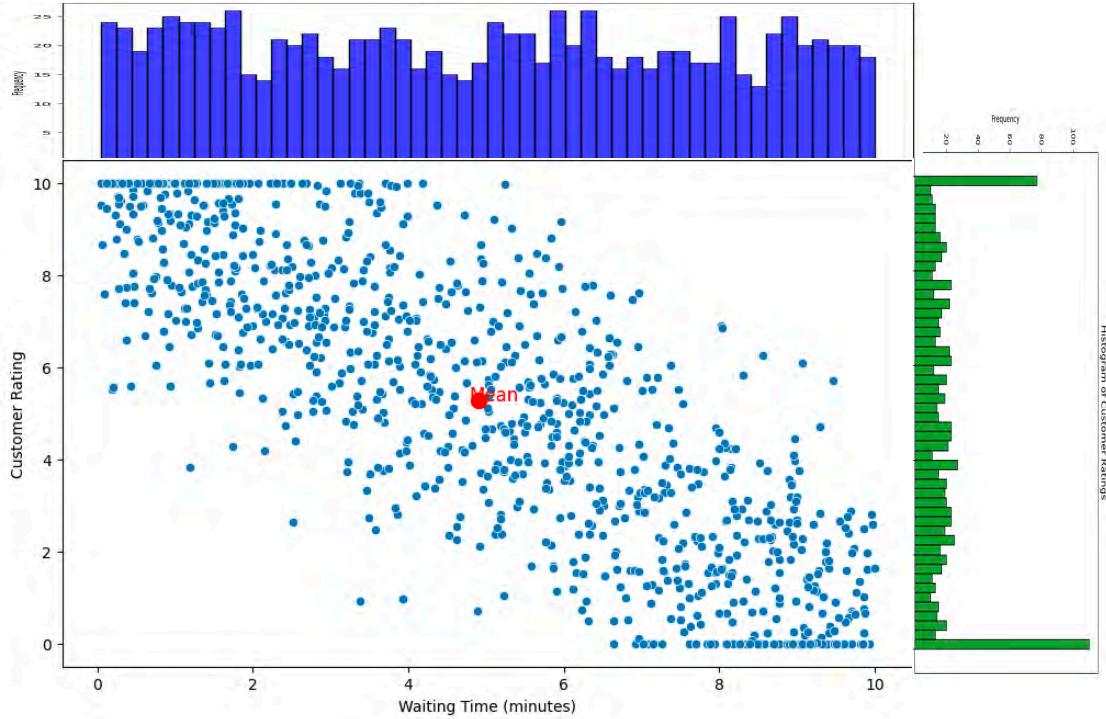


Covariance?



Covariance?

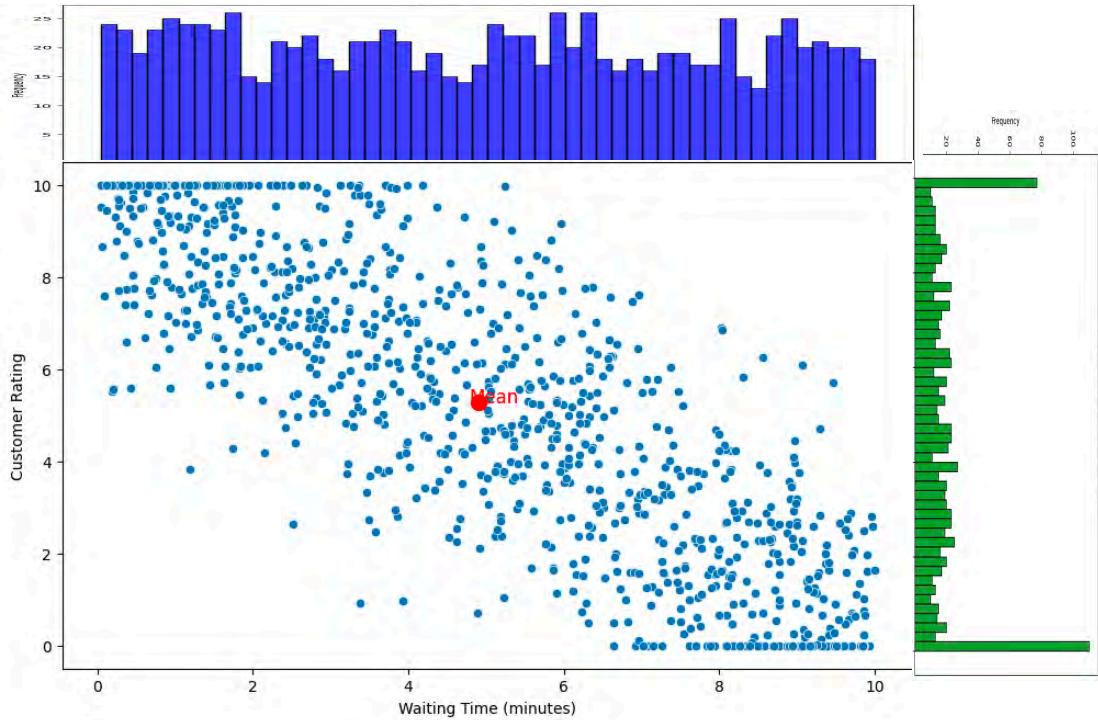
$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$



Covariance?

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\text{Cov}(X, Y) = 18.014 - (4.903)(5.280)$$

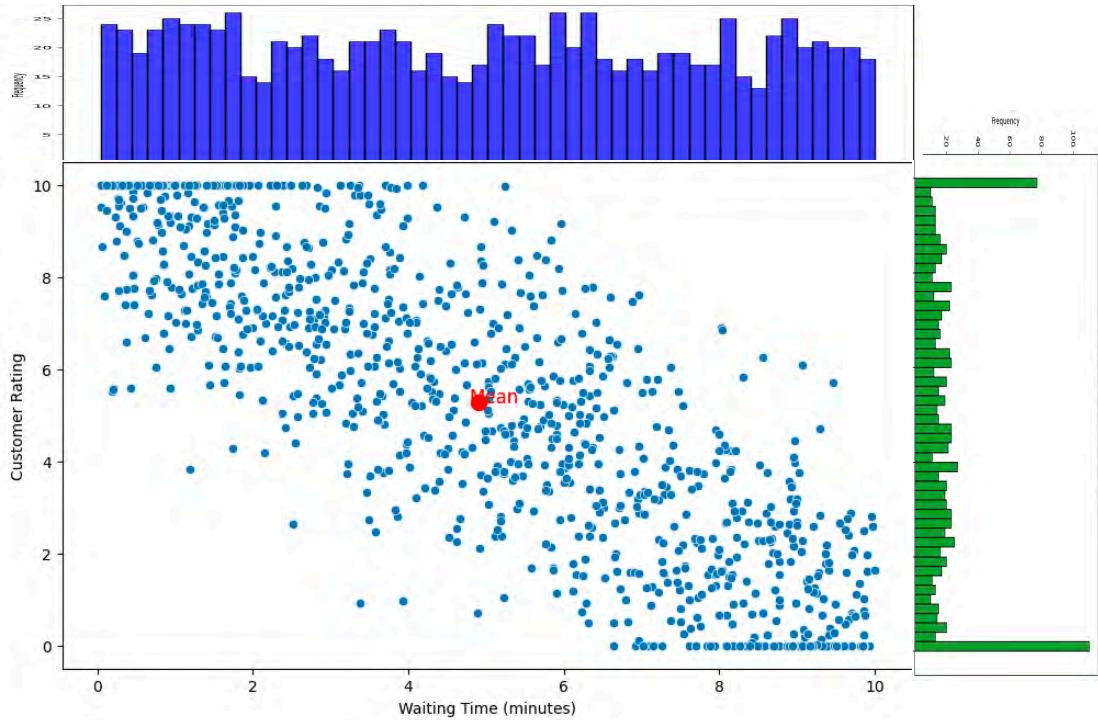


Covariance?

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\text{Cov}(X, Y) = 18.014 - (4.903)(5.280)$$

$$\text{Cov}(X, Y) = 18.014 - (4.903)(5.280)$$



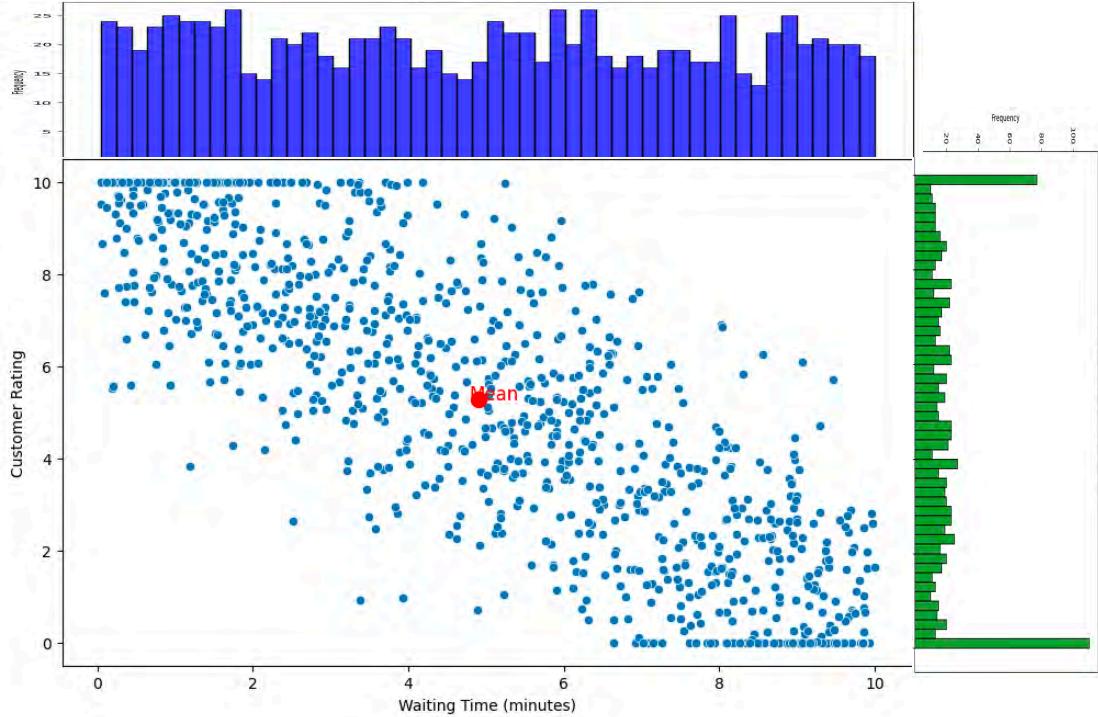
Covariance?

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\text{Cov}(X, Y) = 18.014 - (4.903)(5.280)$$

$$\text{Cov}(X, Y) = 18.014 - (4.903)(5.280)$$

$$\text{Cov}(X, Y) = -7.878$$



Covariance?

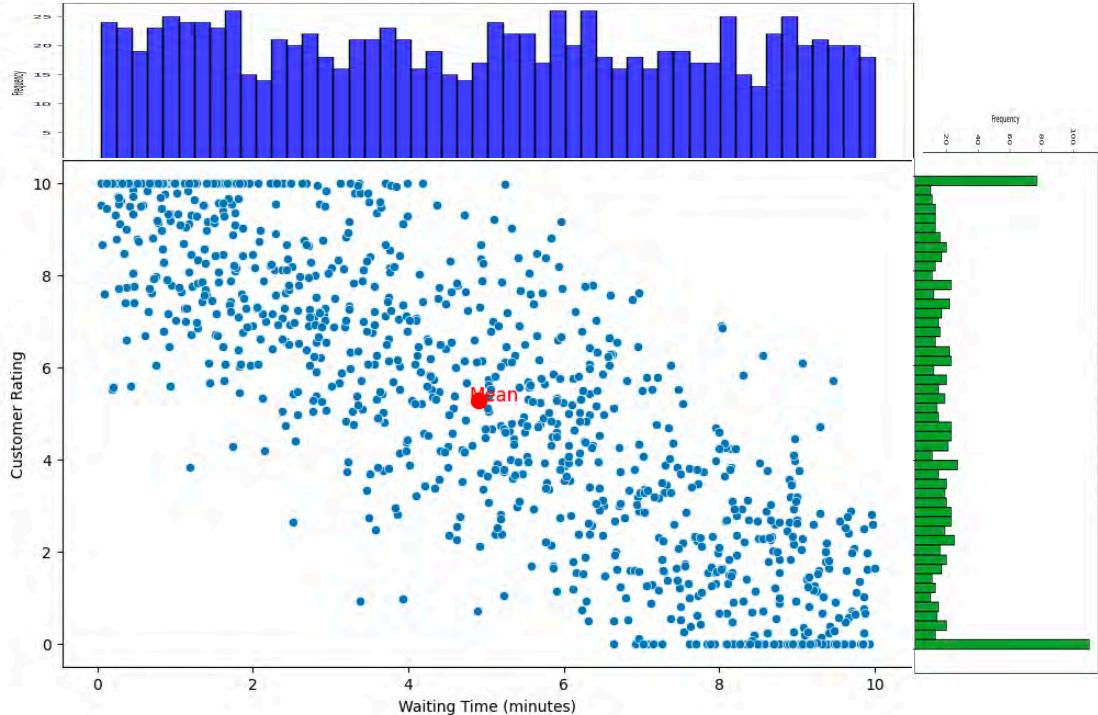
$$\mathbb{E}(X) = 4.903$$

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = -7.878$$



Covariance?

$$\mathbb{E}(X) = 4.903$$

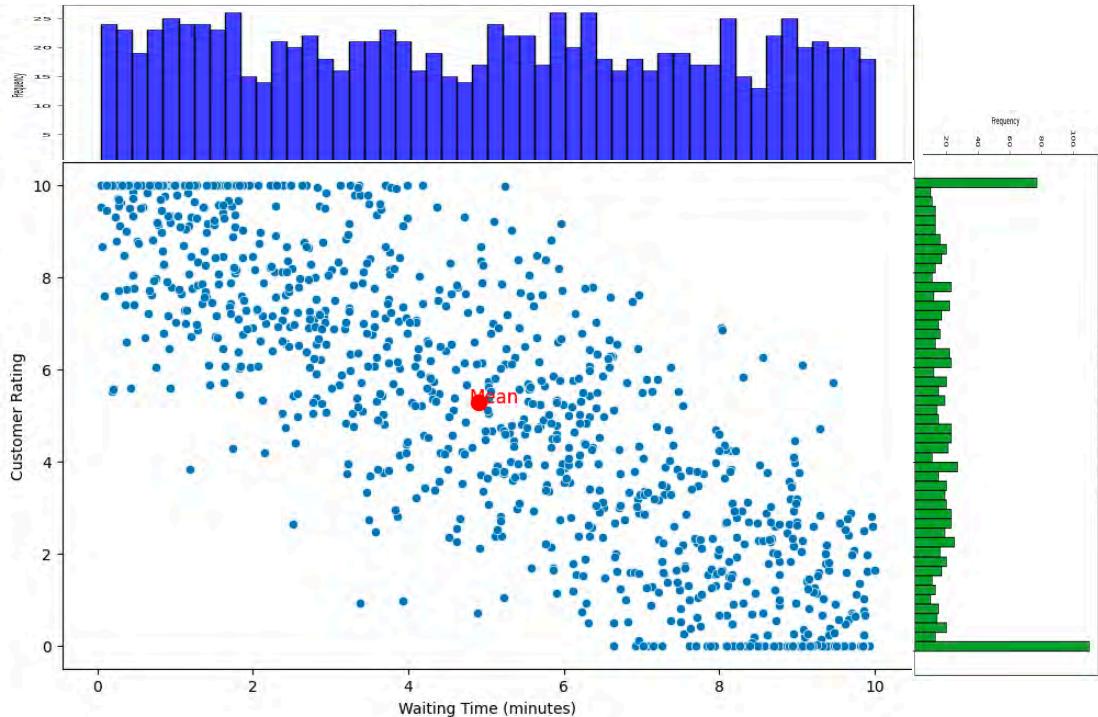
$$\mathbb{E}(Y) = 5.280$$

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = -7.878$$



Covariance?

$$\mathbb{E}(X) = 4.903$$

$$\mathbb{E}(Y) = 5.280$$

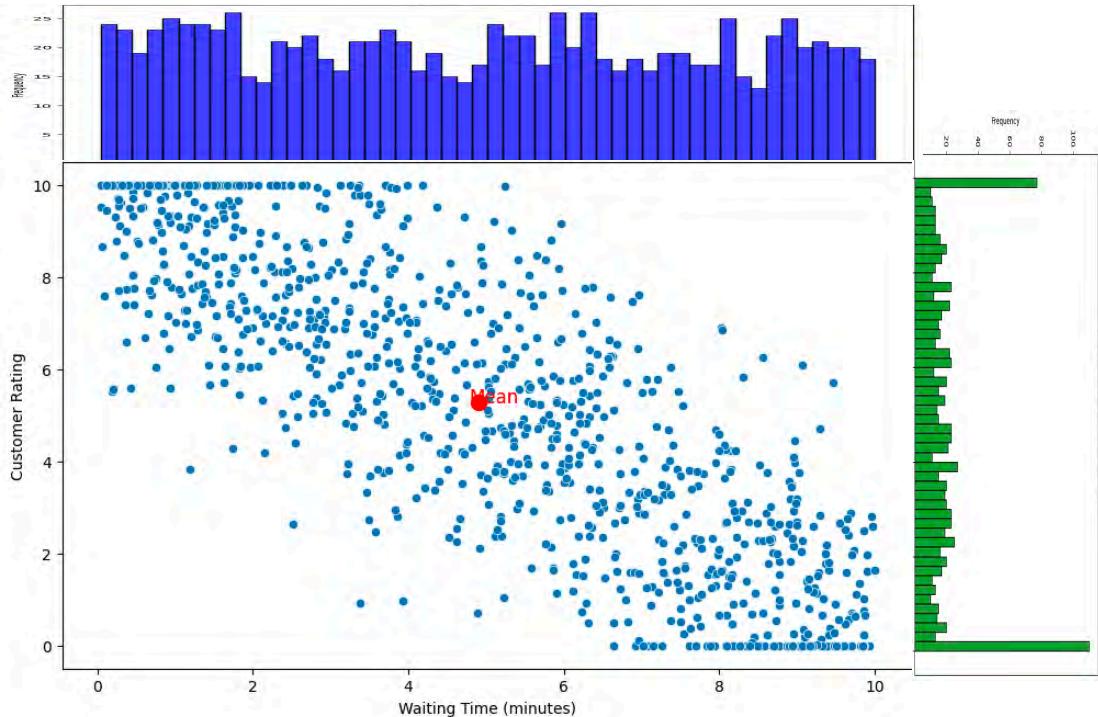
$$\mathbb{E}(XY) = 18.014$$

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$Cov(X, Y) = -7.878$$



Covariance of a Joint Continuous Distribution

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

$$\mathbb{E}(Y) = 5.280$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

$$\mathbb{E}(Y) = 5.280$$

$$\mathbb{E}(XY) = 18.014$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$\mathbb{E}(Y) = 5.280$$

$$\mathbb{E}(XY) = 18.014$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$\mathbb{E}(Y) = 5.280$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$\mathbb{E}(XY) = 18.014$$

Covariance of a Joint Continuous Distribution

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathbb{E}(X) = 4.903$$

$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$\mathbb{E}(Y) = 5.280$$

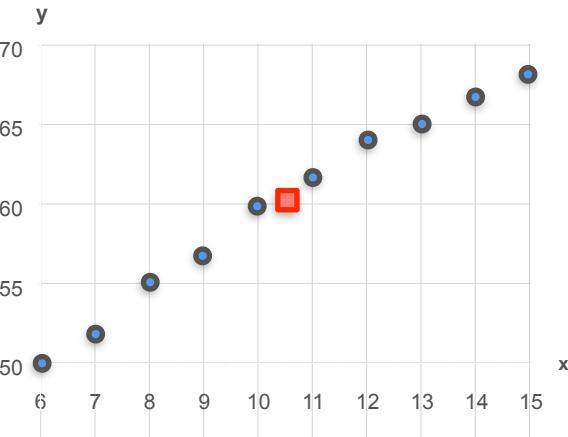
$$Cov(X, Y) = 18.014 - (4.903)(5.280)$$

$$\mathbb{E}(XY) = 18.014$$

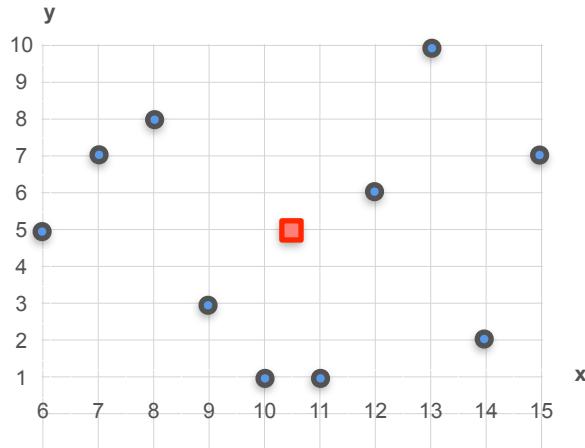
$$Cov(X, Y) = -7.878$$

Covariance Matrix

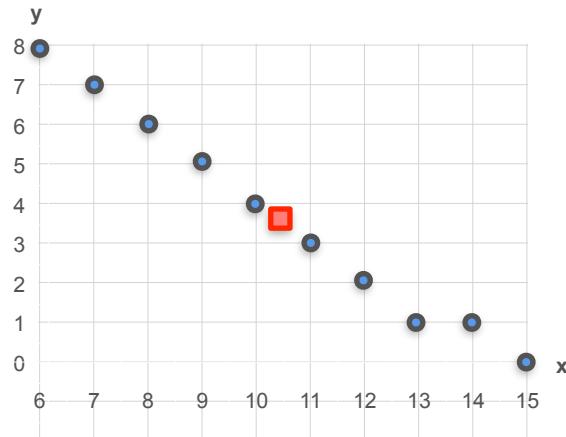
Covariance Matrix



Age vs Height

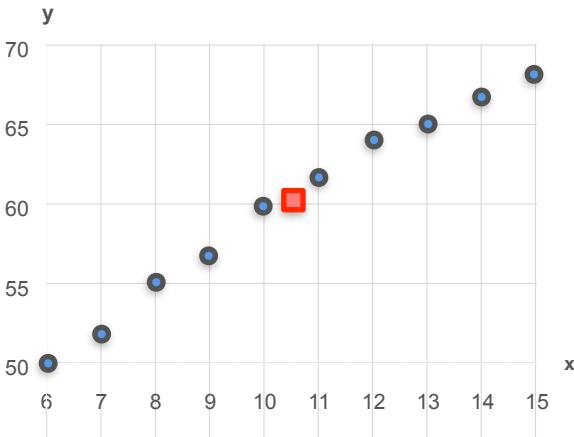


Age vs Grades



Age vs Naps per Day

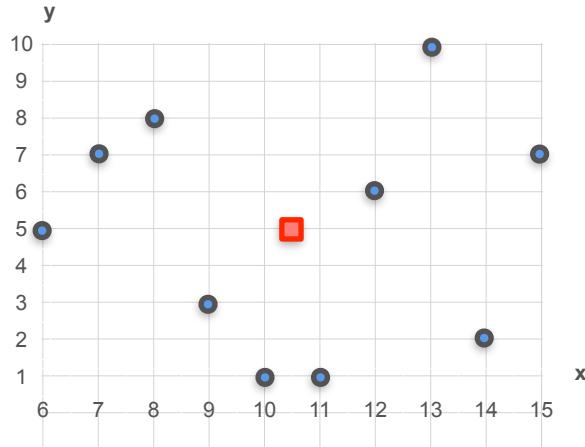
Covariance Matrix



Age vs Height

$$\text{Var}(X) = 9.17$$

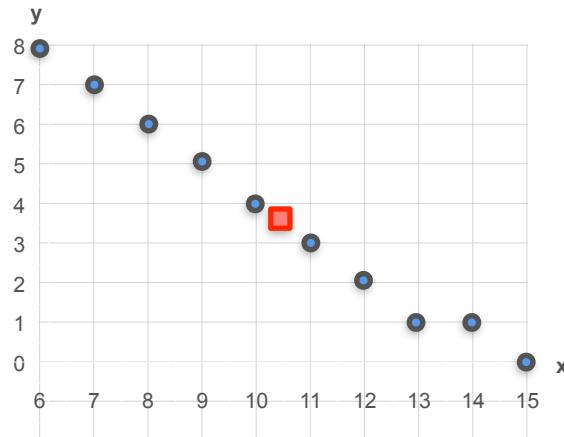
$$\text{Var}(Y) = 39.56$$



Age vs Grades

$$\text{Var}(X) = 9.17$$

$$\text{Var}(Y) = 9.78$$

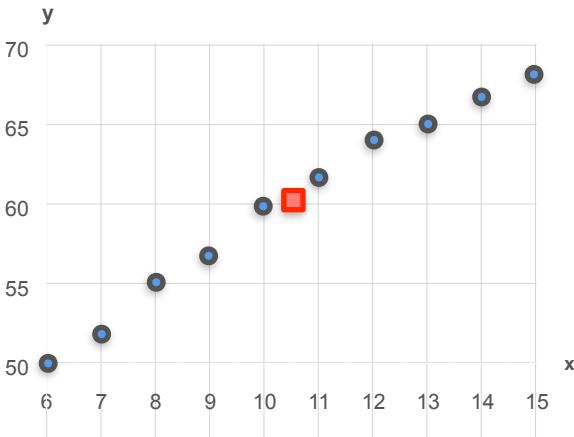


Age vs Naps per Day

$$\text{Var}(X) = 9.17$$

$$\text{Var}(Y) = 7.57$$

Covariance Matrix

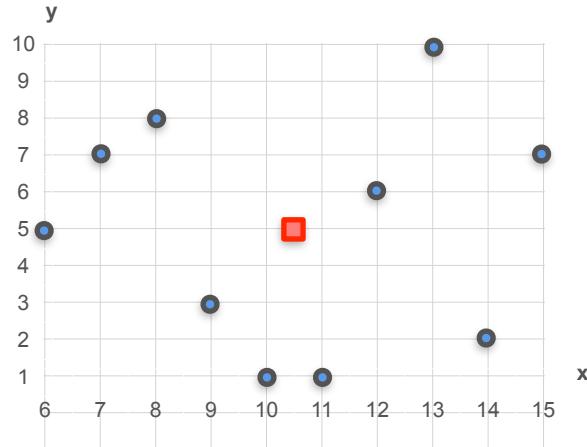


Age vs Height

$$\text{Var}(X) = 9.17$$

$$\text{Var}(Y) = 39.56$$

$$\text{Cov}(X, Y) = 17$$

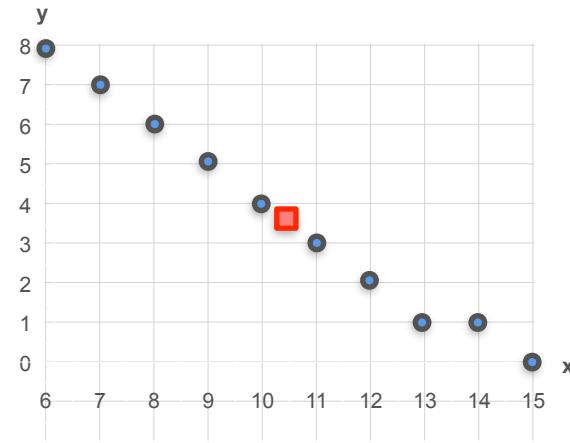


Age vs Grades

$$\text{Var}(X) = 9.17$$

$$\text{Var}(Y) = 9.78$$

$$\text{Cov}(X, Y) = 0.1$$



Age vs Naps per Day

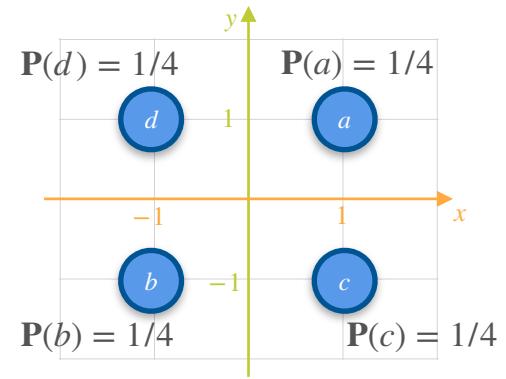
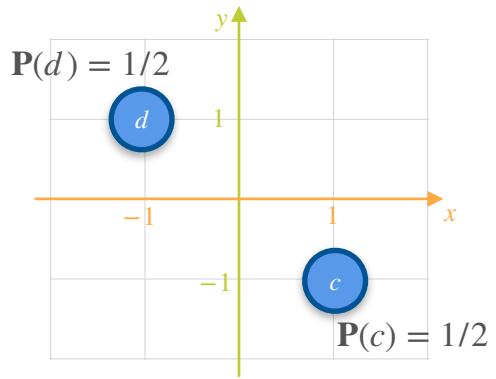
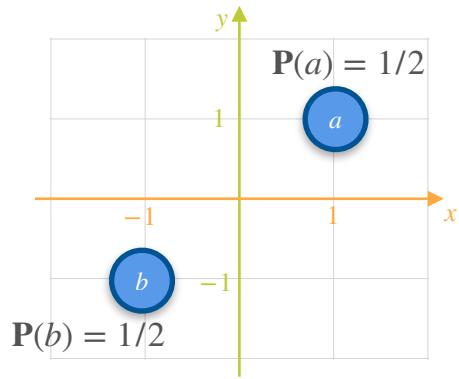
$$\text{Var}(X) = 9.17$$

$$\text{Var}(Y) = 7.57$$

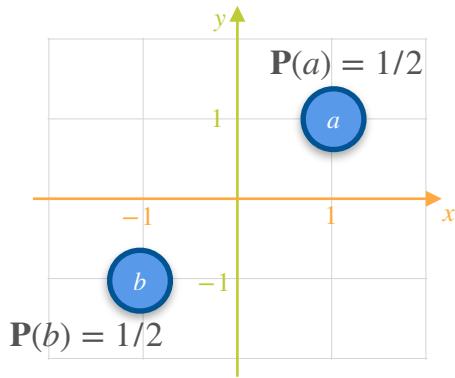
$$\text{Cov}(X, Y) = -7.45$$

Covariance Matrix

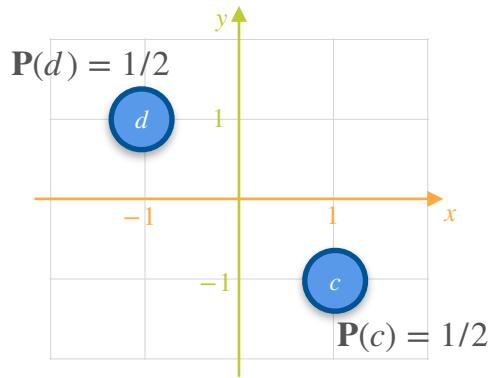
Covariance Matrix



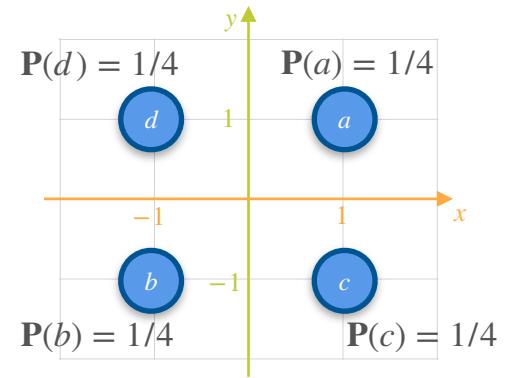
Covariance Matrix



$$\text{Var}(X) = \text{Var}(Y) = 1$$

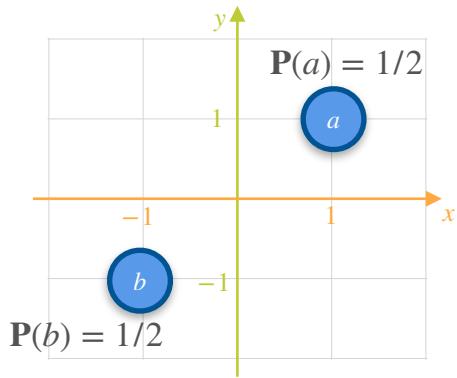


$$\text{Var}(X) = \text{Var}(Y) = 1$$



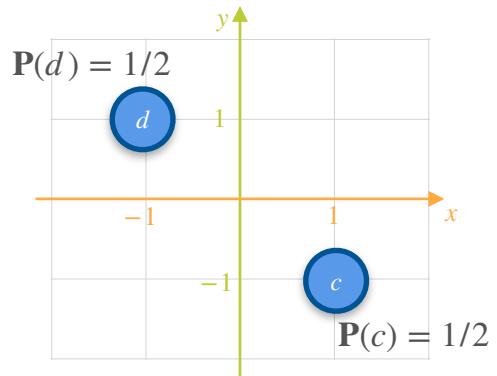
$$\text{Var}(X) = \text{Var}(Y) = 1$$

Covariance Matrix



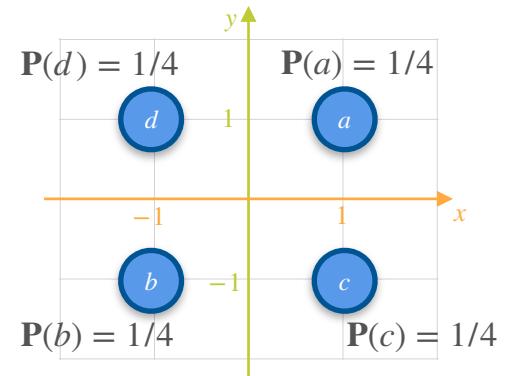
$$\text{Var}(X) = \text{Var}(Y) = 1$$

$$\text{Cov}(X, Y) = 1$$



$$\text{Var}(X) = \text{Var}(Y) = 1$$

$$\text{Cov}(X, Y) = -1$$

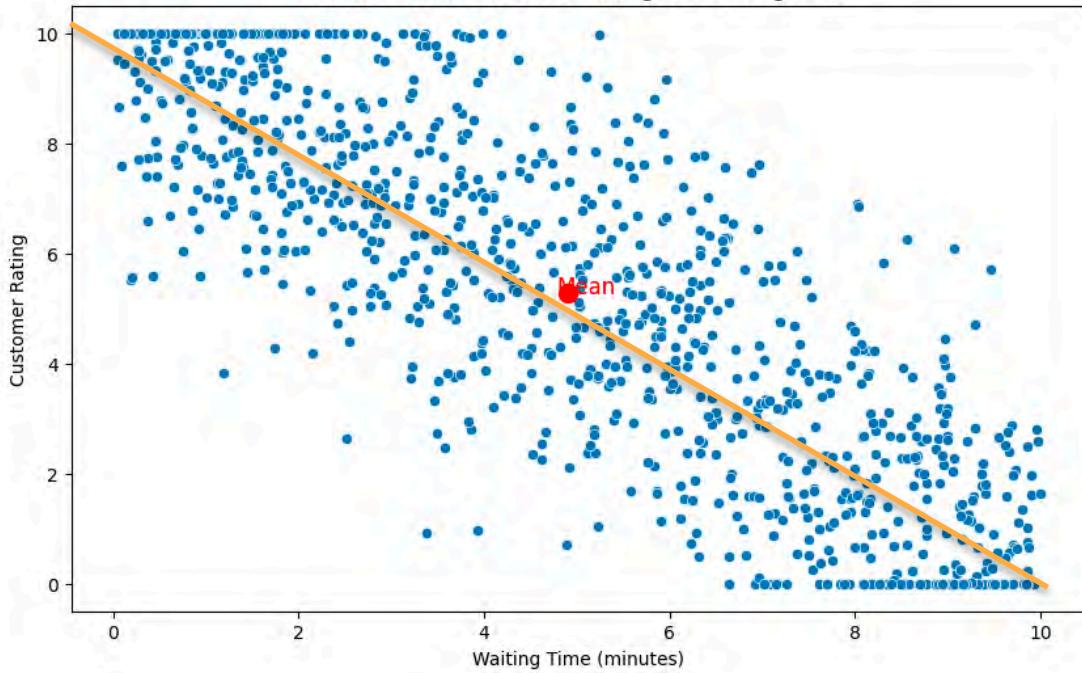


$$\text{Var}(X) = \text{Var}(Y) = 1$$

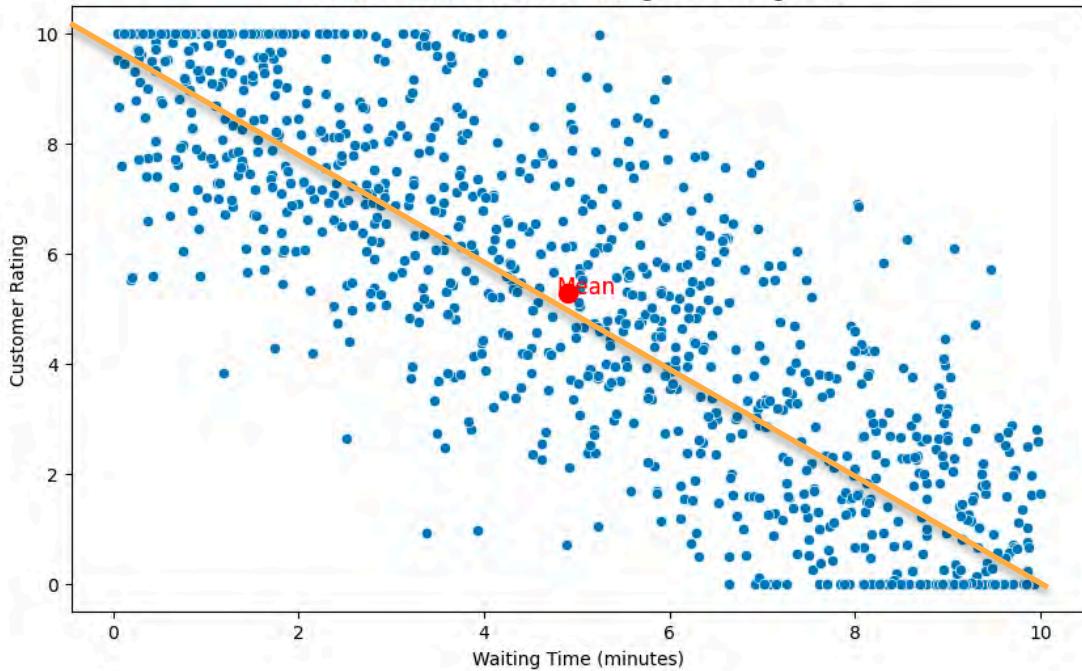
$$\text{Cov}(X, Y) = 0$$

Covariance Matrix

Covariance Matrix



Covariance Matrix

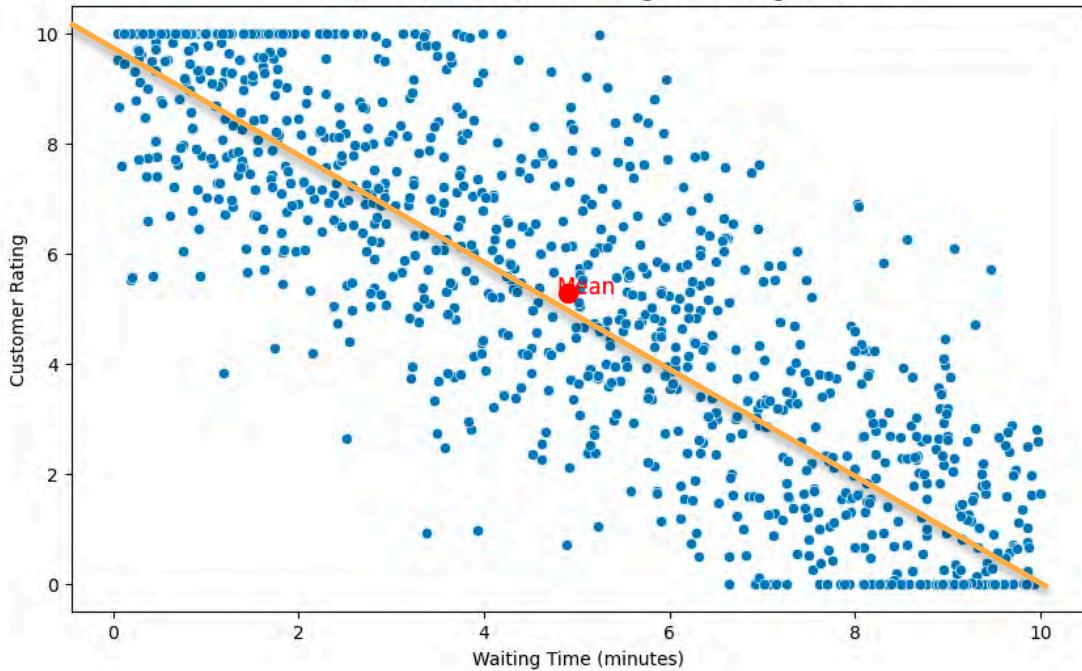


$$\text{Var}(X) = 8.526$$

$$\text{Var}(Y) = 10.163$$

$$\text{Cov}(X, Y) = -7.878$$

Covariance Matrix

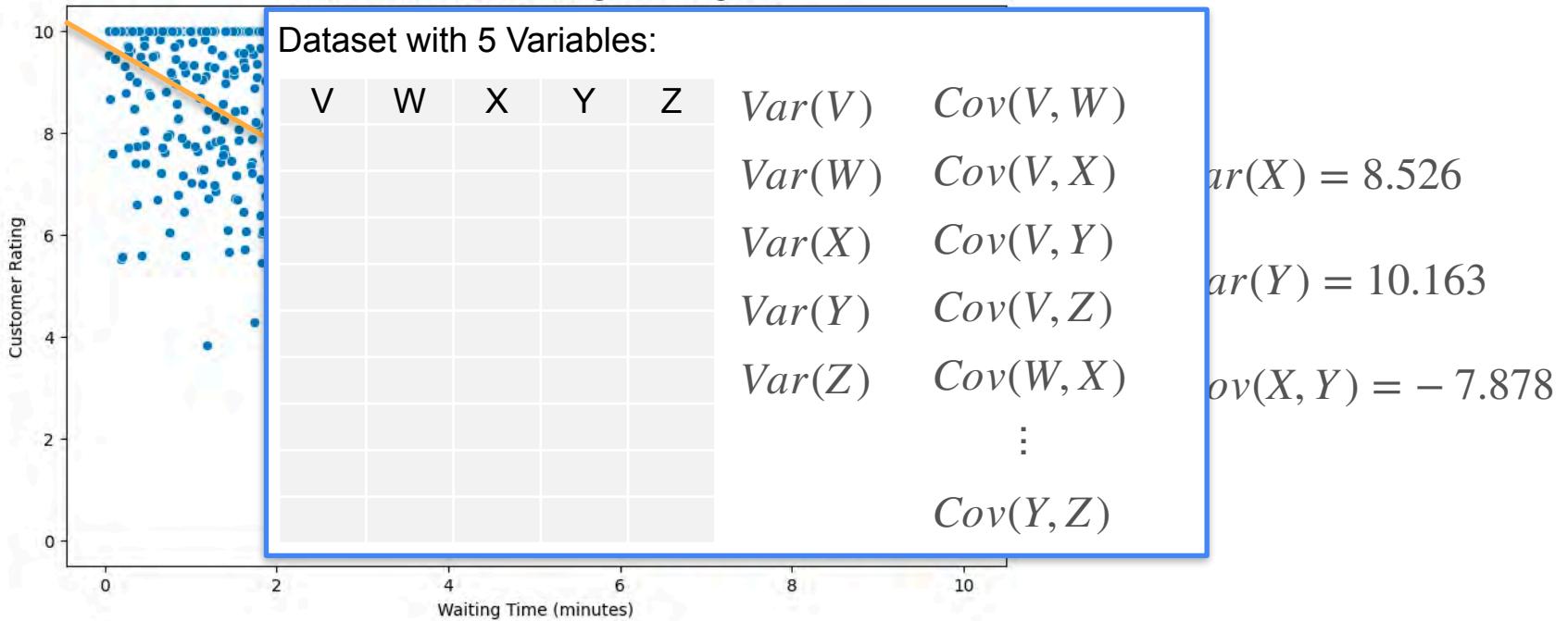


$$\text{Var}(X) = 8.526$$

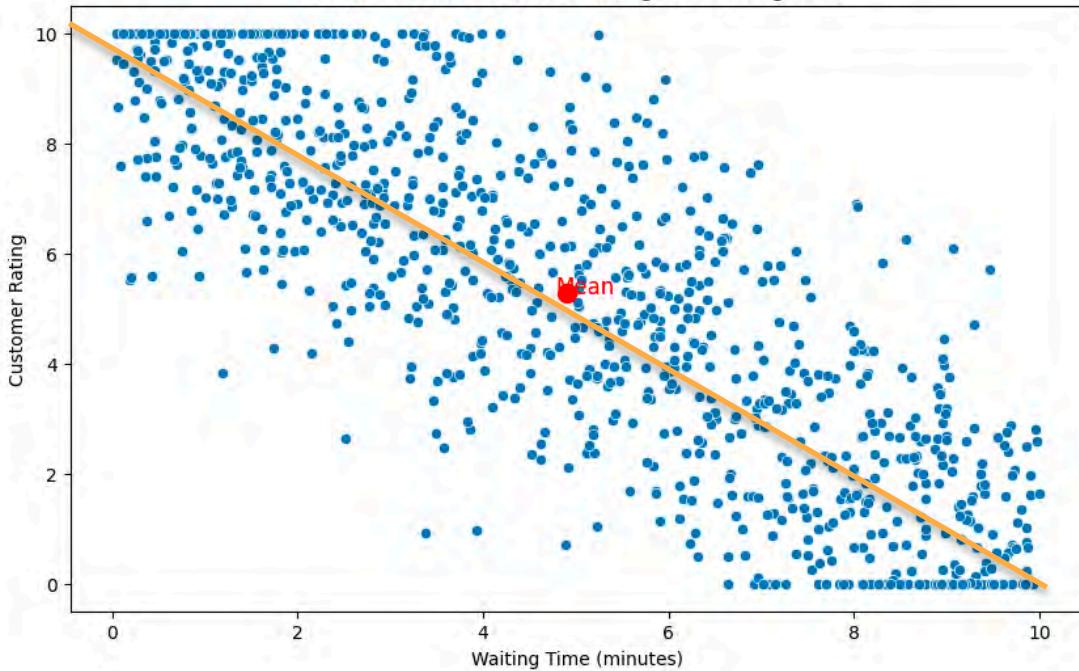
$$\text{Var}(Y) = 10.163$$

$$\text{Cov}(X, Y) = -7.878$$

Covariance Matrix



Covariance Matrix

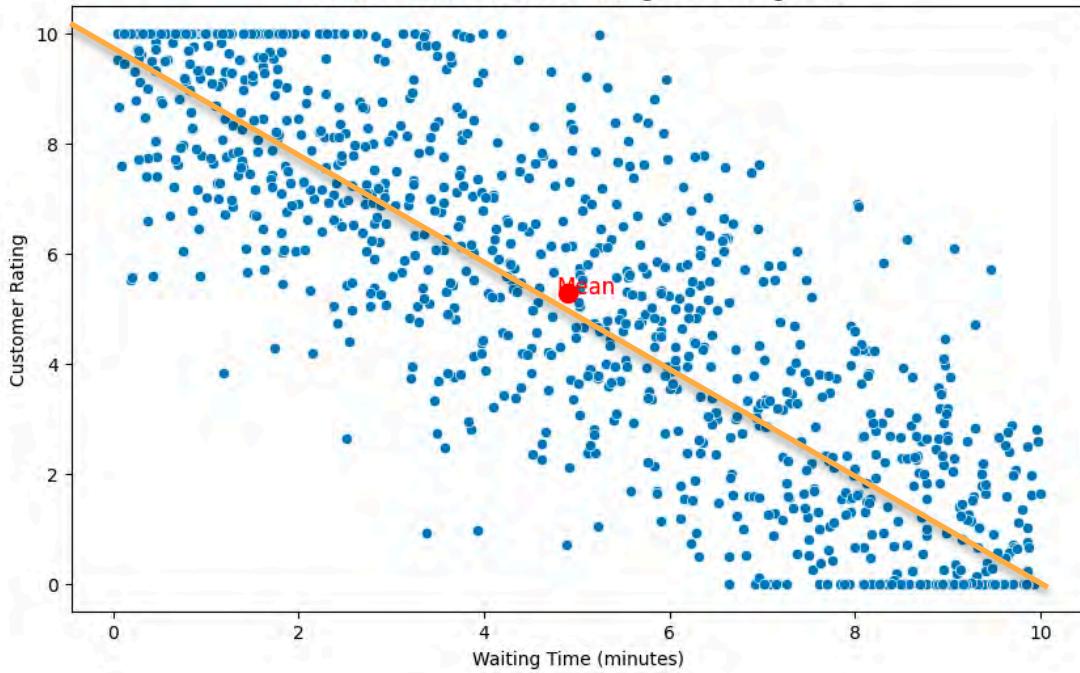


$$\text{Var}(X) = 8.526$$

$$\text{Var}(Y) = 10.163$$

$$\text{Cov}(X, Y) = -7.878$$

Covariance Matrix



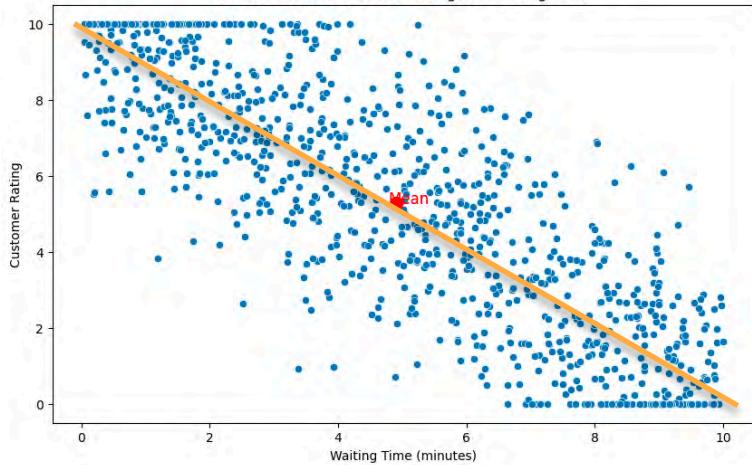
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

Covariance Matrix

Covariance Matrix

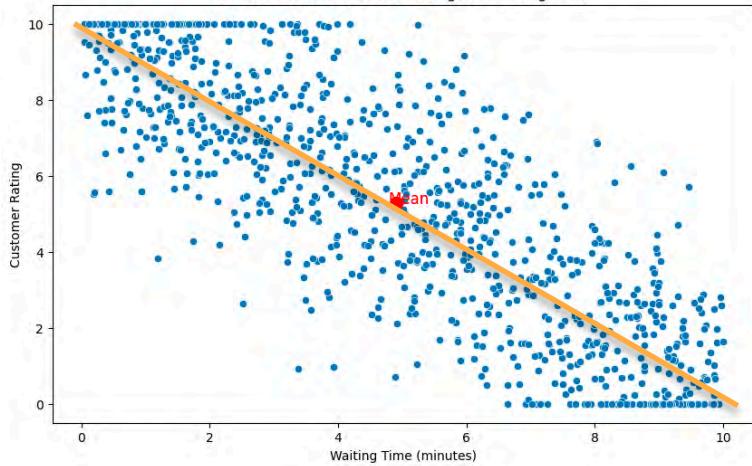


$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

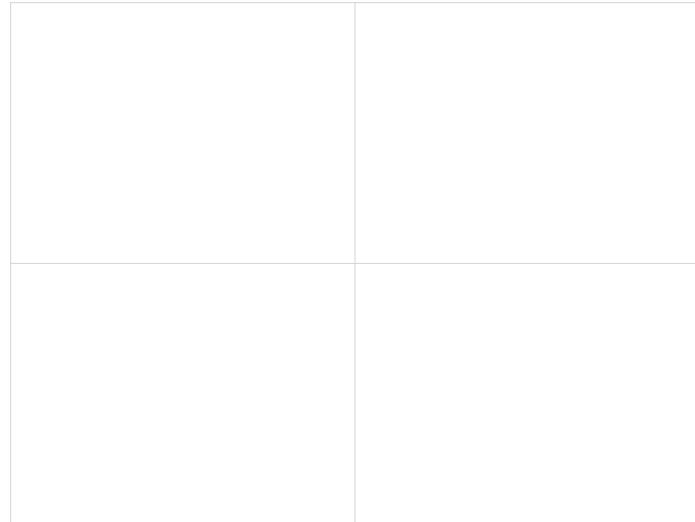
Covariance Matrix



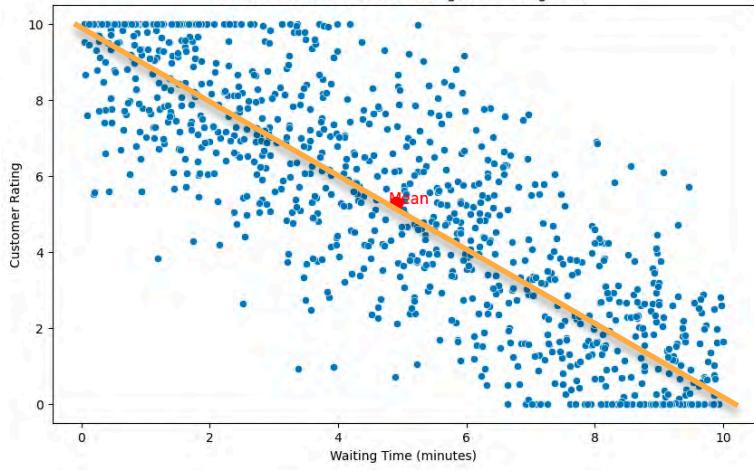
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$



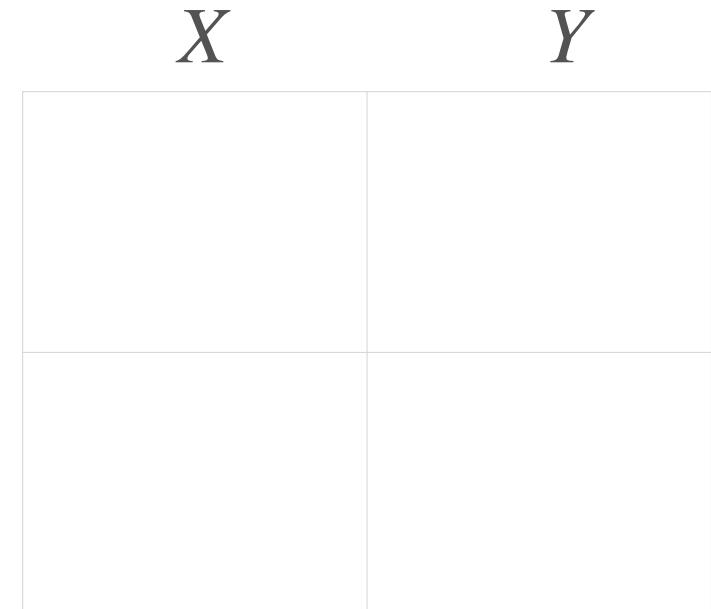
Covariance Matrix



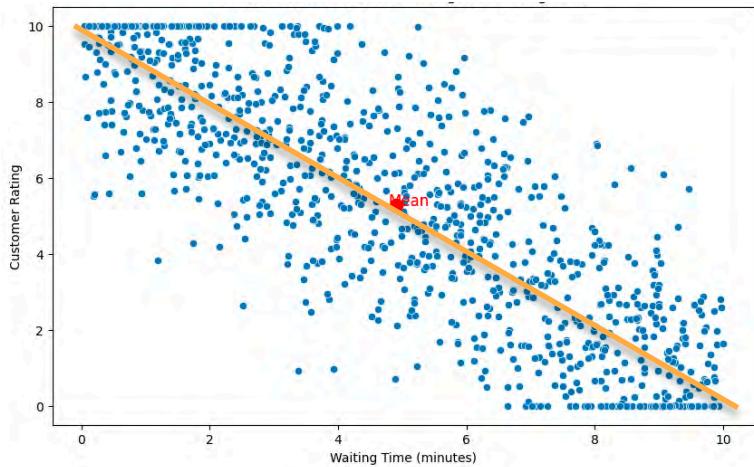
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$



Covariance Matrix



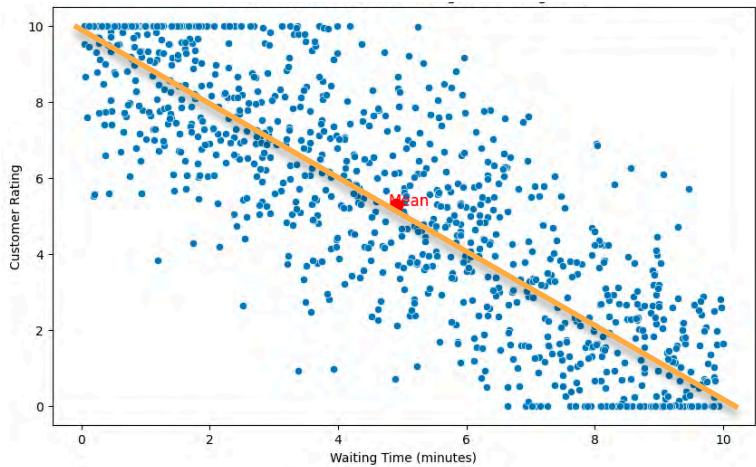
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

X		Y
	$Var(X)$	
Y		$Var(Y)$

Covariance Matrix



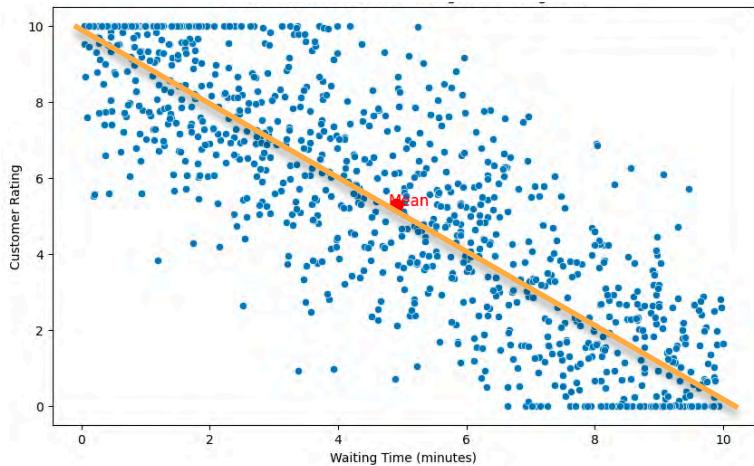
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

X	X	Y
$Var(X)$	$Cov(X, Y)$	$Var(Y)$
Y	$Cov(X, Y)$	$Var(Y)$

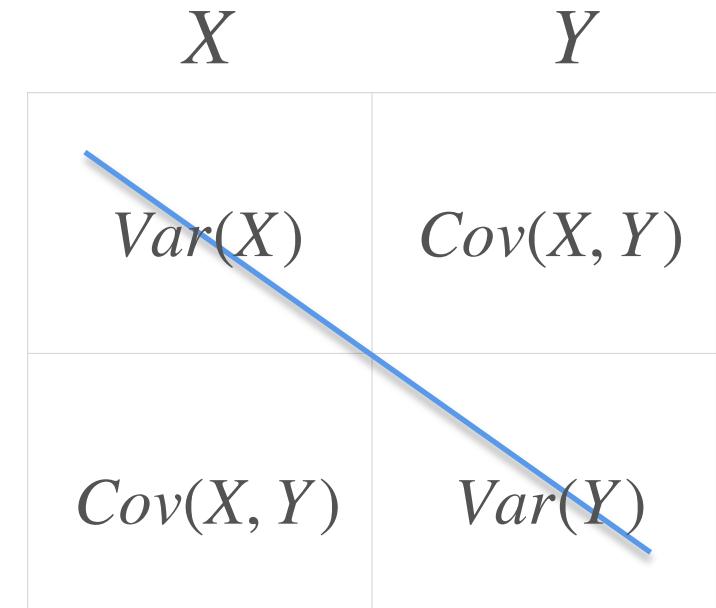
Covariance Matrix



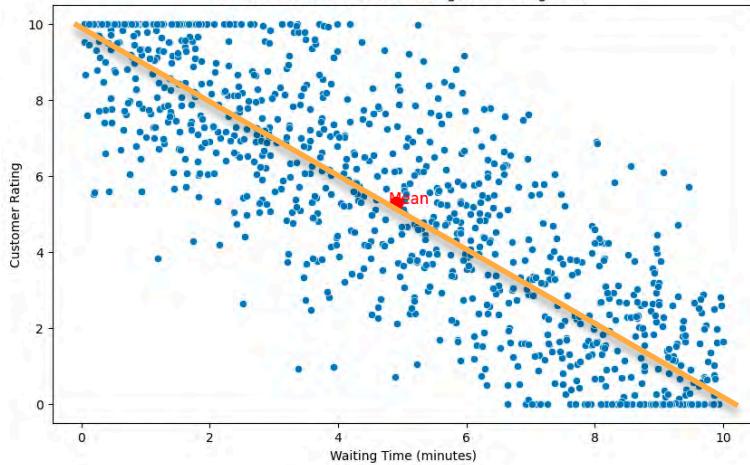
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$



Covariance Matrix



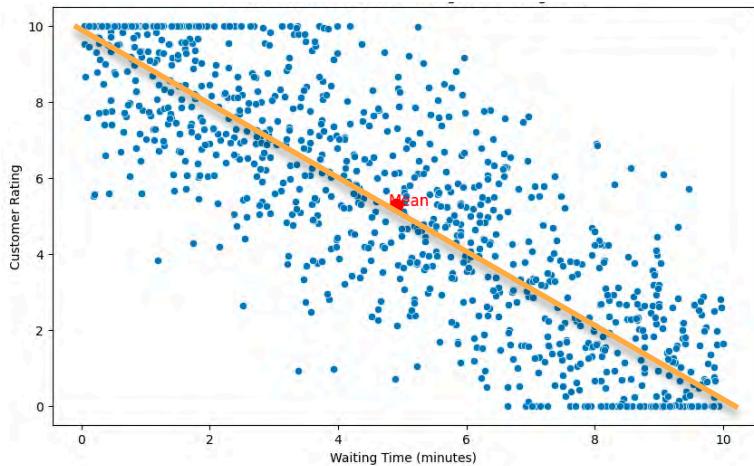
$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

	X	Y
X	$Var(X)$	$Cov(X, Y)$
Y	$Cov(X, Y)$	$Var(Y)$

Covariance Matrix



$$Var(X) = 8.526$$

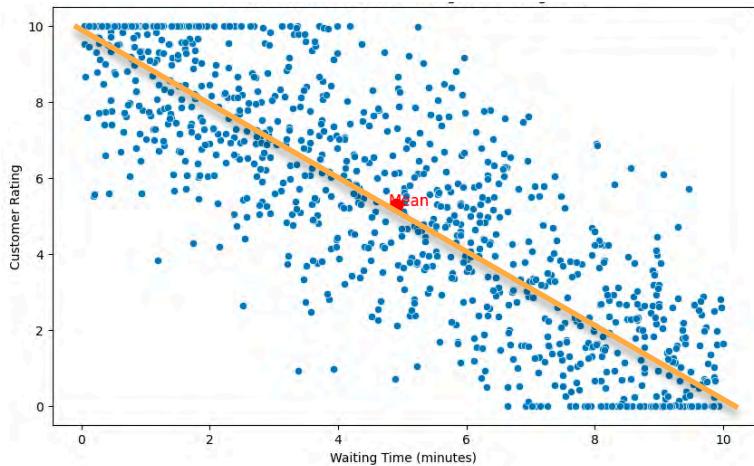
$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

	X	Y
X	$Var(X)$	$Cov(X, Y)$
Y	$Cov(X, Y)$	$Var(Y)$

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

Covariance Matrix



$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

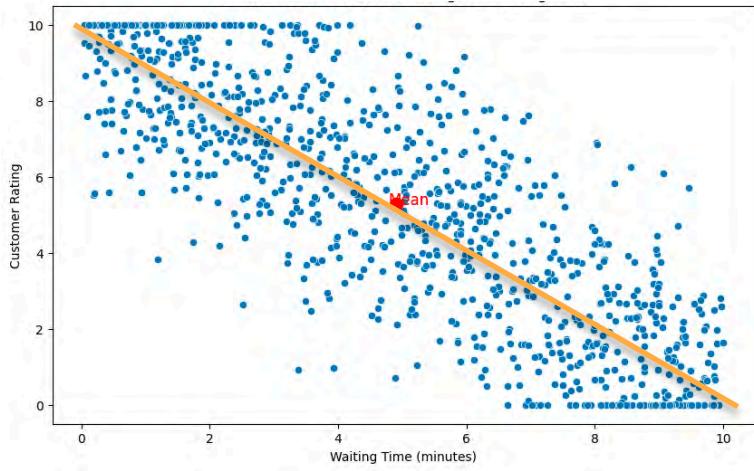
$$Cov(X, Y) = -7.878$$

	X	Y
X	$Var(X)$	$Cov(X, Y)$
Y	$Cov(X, Y)$	$Var(Y)$

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

8.526	-7.878
-7.878	10.163

Covariance Matrix



$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

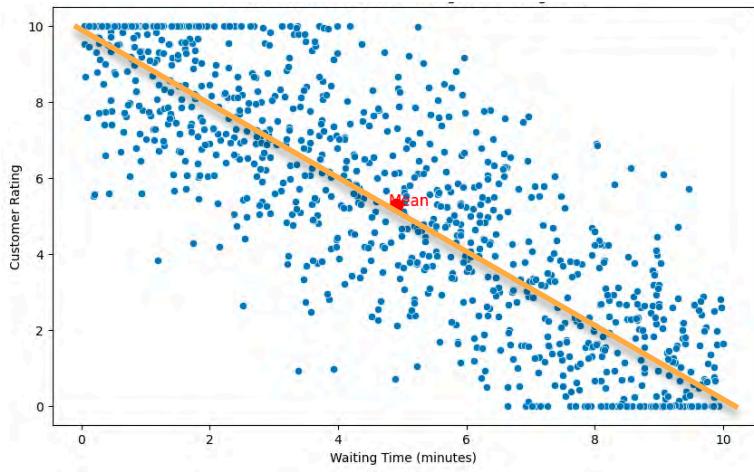
	X	Y
X	$Var(X)$	$Cov(X, Y)$
Y	$Cov(X, Y)$	$Var(Y)$

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

8.526	-7.878
-7.878	10.163

$$\begin{bmatrix} 8.534 & -7.878 \\ -7.878 & 10.173 \end{bmatrix}$$

Covariance Matrix



$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

	X	Y
X	$Var(X)$	$Cov(X, Y)$
Y	$Cov(X, Y)$	$Var(Y)$

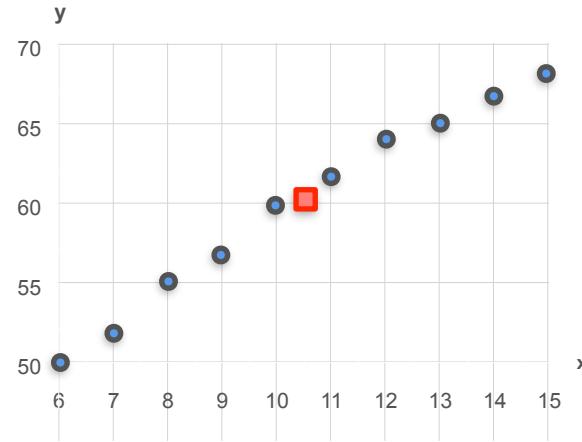
$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

8.526	-7.878
-7.878	10.163

$$\begin{bmatrix} 8.534 & -7.878 \\ -7.878 & 10.173 \end{bmatrix}$$

Covariance Matrix

Covariance Matrix



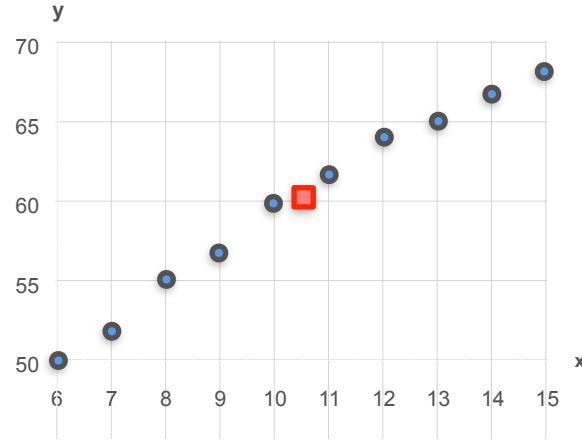
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Covariance Matrix



Age vs Height

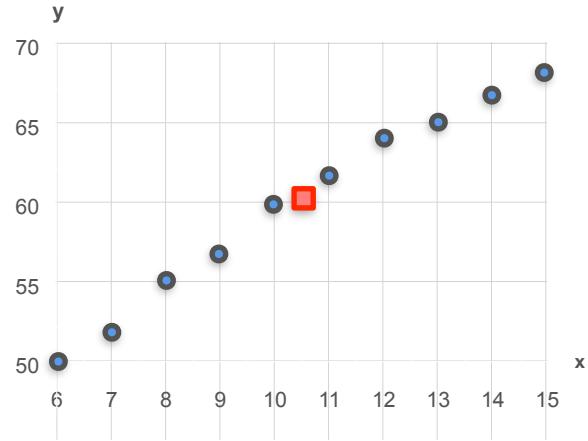
$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

Covariance Matrix



Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$$\begin{bmatrix} 9.17 & 17 \\ 17 & 39.56 \end{bmatrix}$$

Covariance Matrix

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = 1$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = -1$$

Covariance Matrix

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$$Var(X) = 1$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = 1$$

$$Cov(X, Y) = -1$$

Covariance Matrix

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$$\begin{aligned}Var(X) &= 1 \\Var(Y) &= 1 \\Cov(X, Y) &= 1\end{aligned}\quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{aligned}Var(X) &= 1 \\Var(Y) &= 1 \\Cov(X, Y) &= -1\end{aligned}$$

Covariance Matrix

$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = 1$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

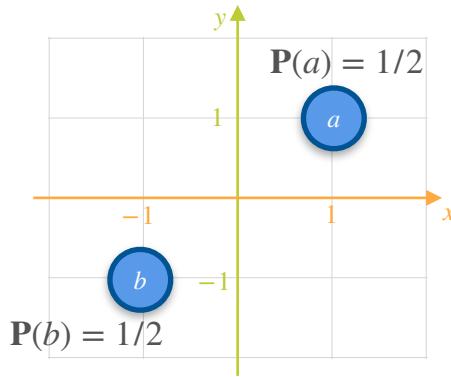
$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = -1$$

Covariance Matrix



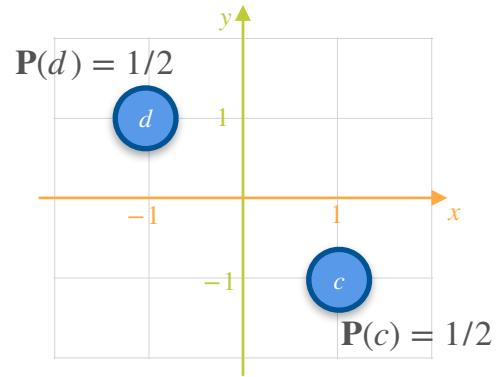
$$\begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = 1$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$



$$Var(X) = 1$$

$$Var(Y) = 1$$

$$Cov(X, Y) = -1$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Covariance of a Joint Continuous Distribution

Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$

Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$

X

Y

Z

X

Y

Z

Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$

X Y Z

X

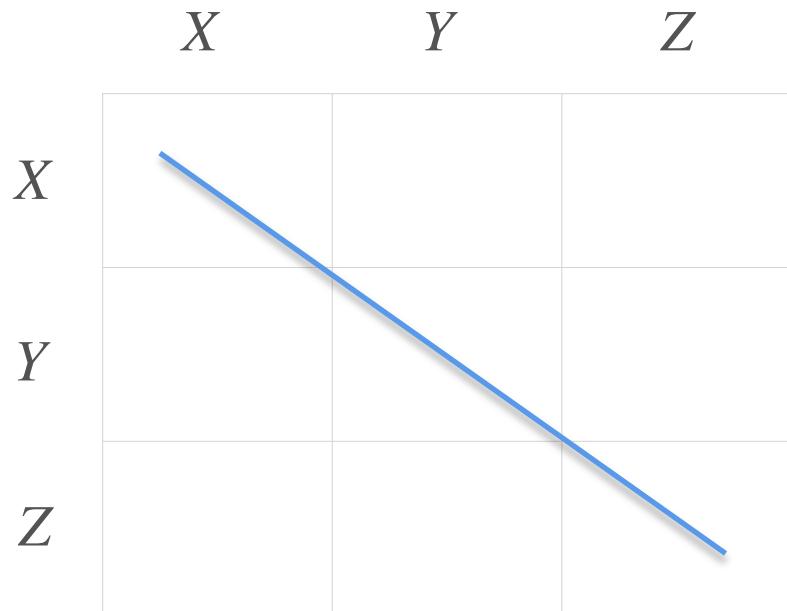
Y

Z

Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

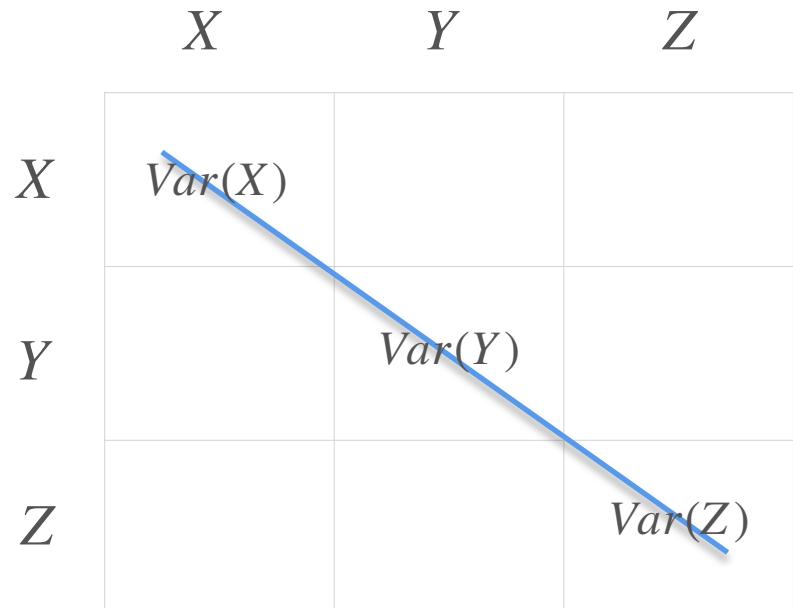
X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$



Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$



Covariance of a Joint Continuous Distribution

Dataset with 3 Variables:

X	Y	Z	$Var(X)$	$Cov(X, Y)$
			$Var(Y)$	$Cov(X, Z)$
			$Var(Z)$	$Cov(Y, Z)$

X	$Var(X)$	$Cov(X, Y)$	$Cov(X, Z)$
Y	$Cov(X, Y)$	$Var(Y)$	$Cov(Y, Z)$
Z	$Cov(X, Z)$	$Cov(Y, Z)$	$Var(Z)$

Covariance of a Joint Continuous Distribution

Covariance of a Joint Continuous Distribution

Dataset with 5 Variables:

V	W	X	Y	Z	$Var(V)$	$Cov(V, W)$
					$Var(W)$	$Cov(V, X)$
					$Var(X)$	$Cov(V, Y)$
					$Var(Y)$	$Cov(V, Z)$
					$Var(Z)$	$Cov(W, X)$
						\vdots
						$Cov(Y, Z)$

Covariance of a Joint Continuous Distribution

Dataset with 5 Variables:

$$Var(V) \quad Cov(V, W)$$

$$Var(W) \quad Cov(V, X)$$

$$Var(X) \quad Cov(V, Y)$$

$$Var(Y) \quad Cov(V, Z)$$

$$Var(Z) \quad Cov(W, X)$$

三

$$Cov(Y, Z)$$

V W X Y Z

V

W

X

Y

7

Covariance of a Joint Continuous Distribution

Dataset with 5 Variables:

$Var(V)$	$Cov(V, W)$
$Var(W)$	$Cov(V, X)$
$Var(X)$	$Cov(V, Y)$
$Var(Y)$	$Cov(V, Z)$
$Var(Z)$	$Cov(W, X)$
	⋮
	$Cov(Y, Z)$

Covariance of a Joint Continuous Distribution

Dataset with 5 Variables:

$Var(V)$	$Cov(V, W)$
$Var(W)$	$Cov(V, X)$
$Var(X)$	$Cov(V, Y)$
$Var(Y)$	$Cov(V, Z)$
$Var(Z)$	$Cov(W, X)$
	\vdots
	$Cov(Y, Z)$

V	W	X	Y	Z
V	$Var(V)$			
W		$Var(W)$		
X			$Var(X)$	
Y				$Var(Y)$
Z				$Var(Z)$

Covariance of a Joint Continuous Distribution

Dataset with 5 Variables:

V	W	X	Y	Z
			$Var(V)$	$Cov(V, W)$
			$Var(W)$	$Cov(V, X)$
			$Var(X)$	$Cov(V, Y)$
			$Var(Y)$	$Cov(V, Z)$
			$Var(Z)$	$Cov(W, X)$
				\vdots
				$Cov(Y, Z)$

$Var(V) \quad Cov(V, W)$
 $Var(W) \quad Cov(V, X)$
 $Var(X) \quad Cov(V, Y)$
 $Var(Y) \quad Cov(V, Z)$
 $Var(Z) \quad Cov(W, X)$
 \vdots
 $Cov(Y, Z)$

	V	W	X	Y	Z
V	$Var(V)$	$Cov(V, W)$	$Cov(V, X)$	$Cov(V, Y)$	$Cov(V, Z)$
W	$Cov(V, W)$	$Var(W)$	$Cov(W, X)$	$Cov(W, Y)$	$Cov(W, Z)$
X	$Cov(V, X)$	$Cov(W, X)$	$Var(X)$	$Cov(X, Y)$	$Cov(X, Z)$
Y	$Cov(V, Y)$	$Cov(W, Y)$	$Cov(X, Y)$	$Var(Y)$	$Cov(Y, Z)$
Z	$Cov(V, Z)$	$Cov(W, Z)$	$Cov(X, Z)$	$Cov(Y, Z)$	$Var(Z)$

Covariance of a Joint Continuous Distribution

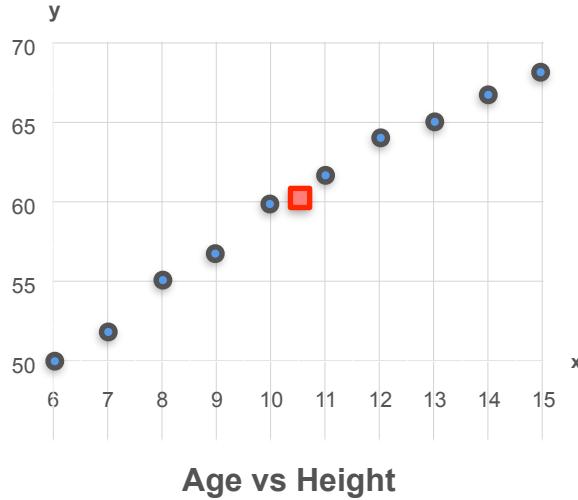
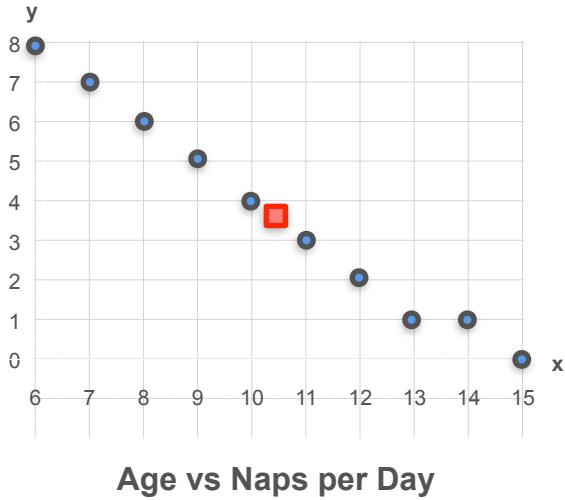
$\sum =$

Covariance Matrix

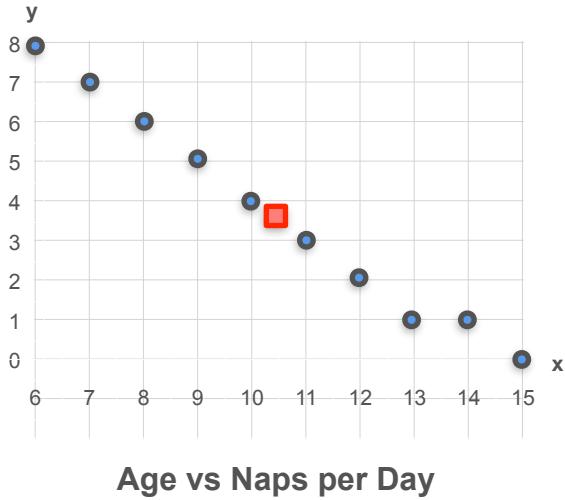
	V	W	X	Y	Z
V	$Var(V)$	$Cov(V, W)$	$Cov(V, X)$	$Cov(V, Y)$	$Cov(V, Z)$
W	$Cov(V, W)$	$Var(W)$	$Cov(W, X)$	$Cov(W, Y)$	$Cov(W, Z)$
X	$Cov(V, X)$	$Cov(W, X)$	$Var(X)$	$Cov(X, Y)$	$Cov(X, Z)$
Y	$Cov(V, Y)$	$Cov(W, Y)$	$Cov(X, Y)$	$Var(Y)$	$Cov(Y, Z)$
Z	$Cov(V, Z)$	$Cov(W, Z)$	$Cov(X, Z)$	$Cov(Y, Z)$	$Var(Z)$

Correlation Coefficient

Correlation Coefficient



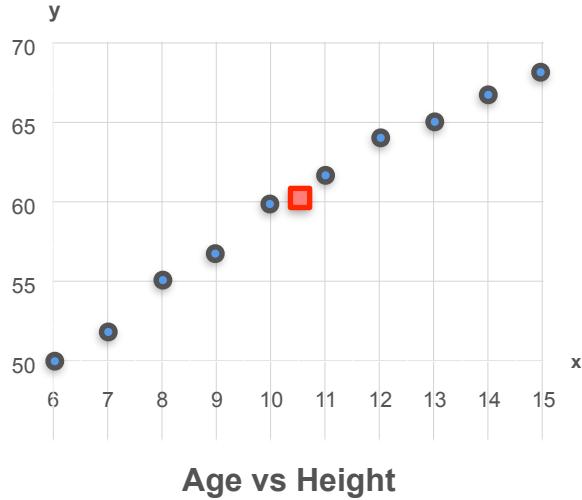
Correlation Coefficient



Age vs Naps per Day

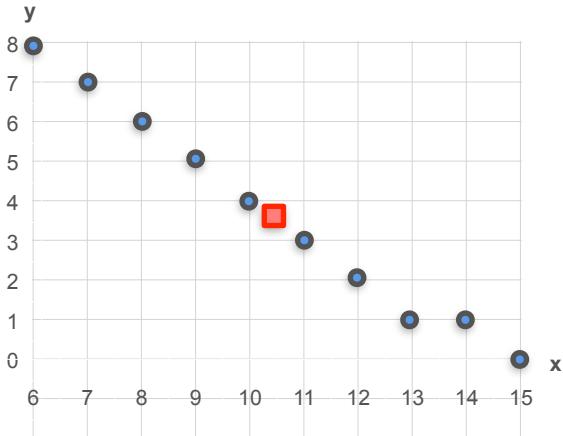
$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$



Age vs Height

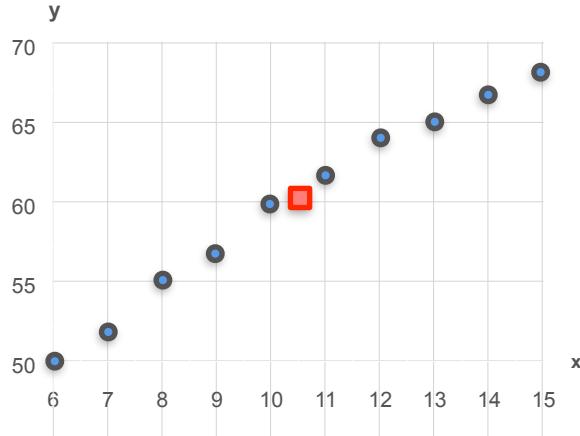
Correlation Coefficient



Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

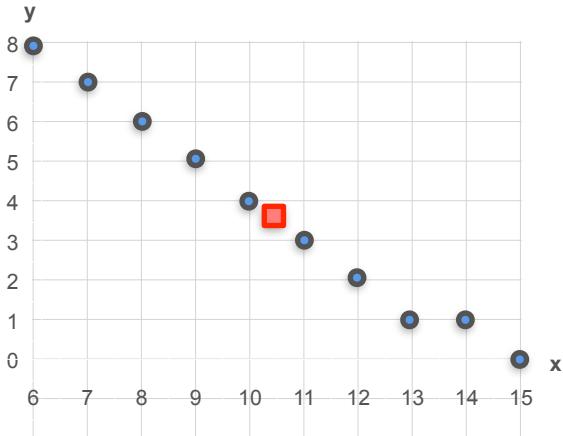


Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

Correlation Coefficient

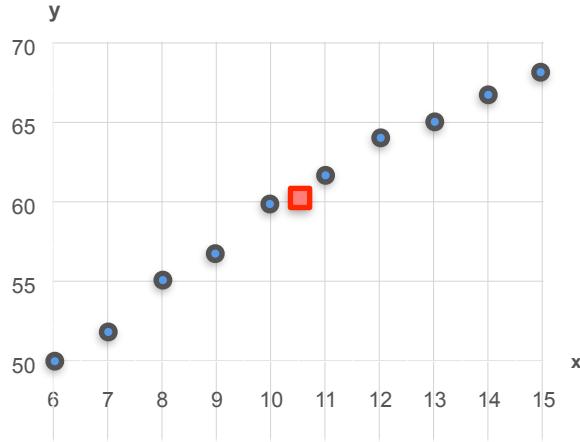


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



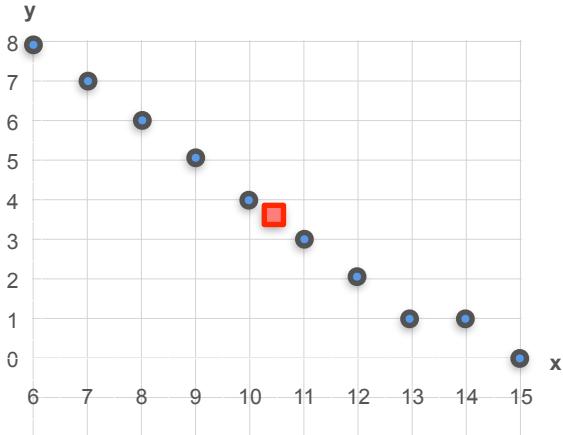
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

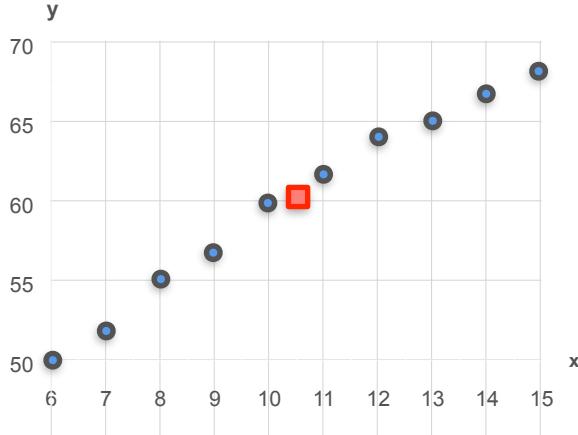


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



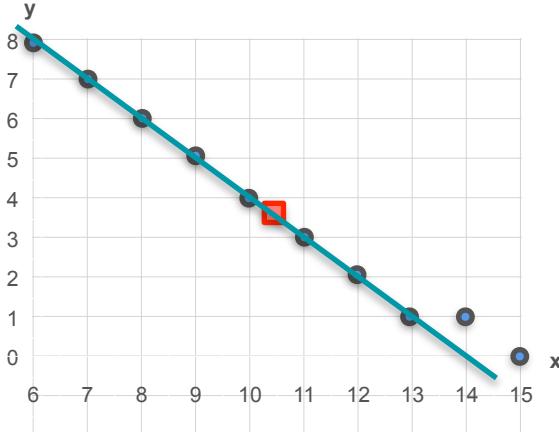
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

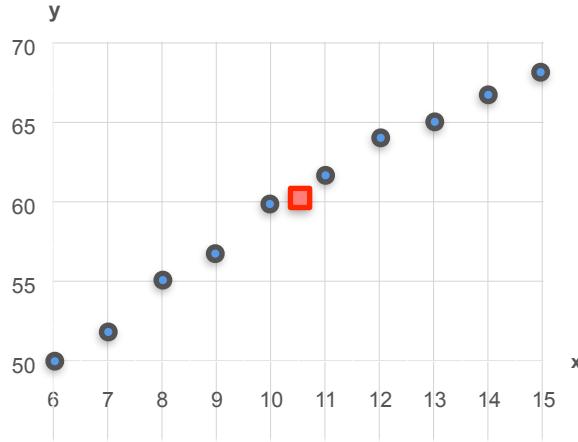


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



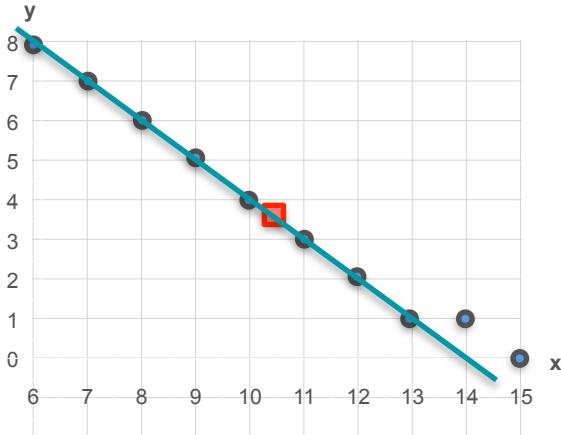
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

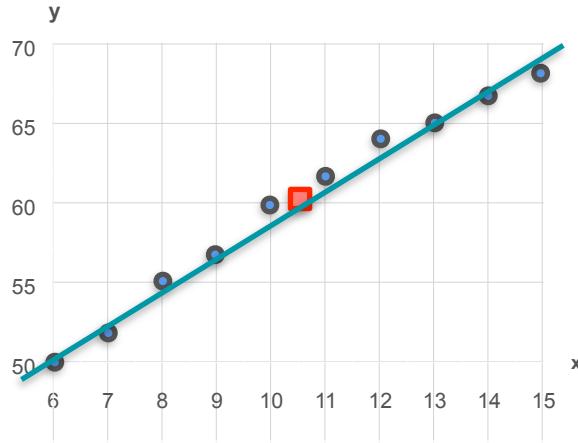


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



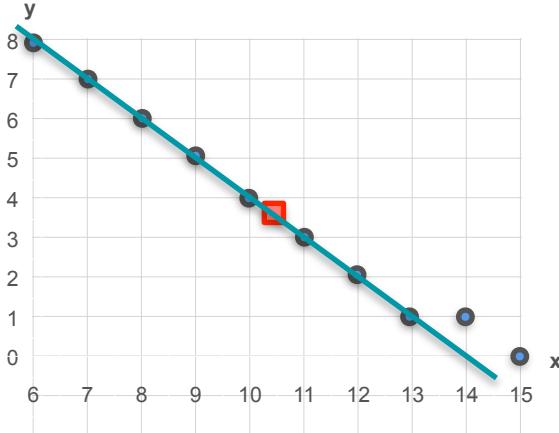
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

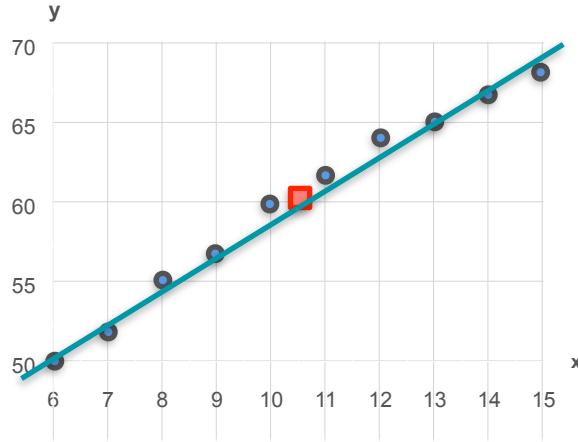


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



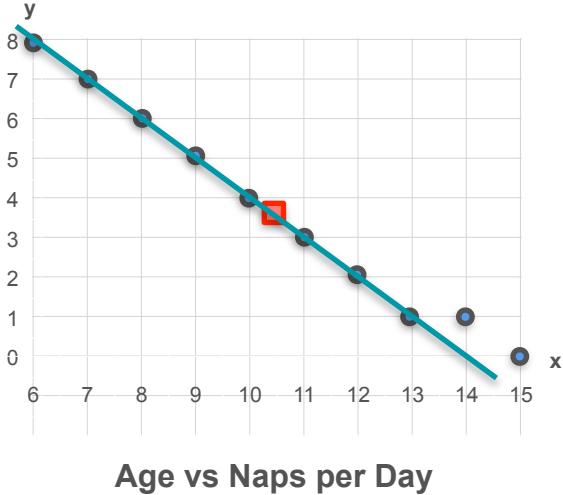
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient



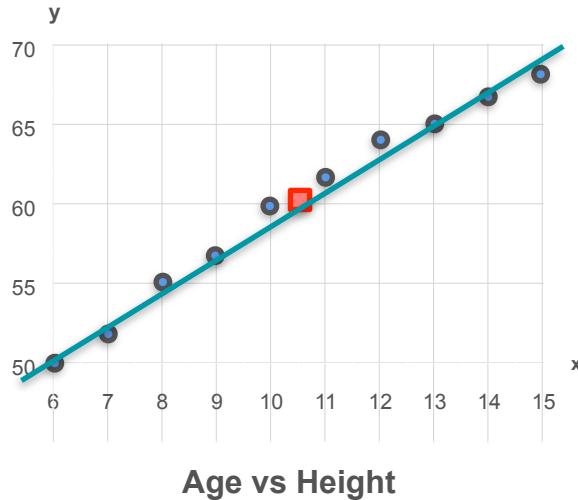
Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

Is the correlation
here stronger?



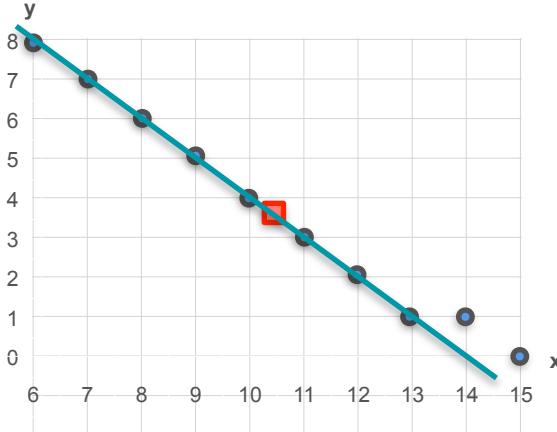
Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

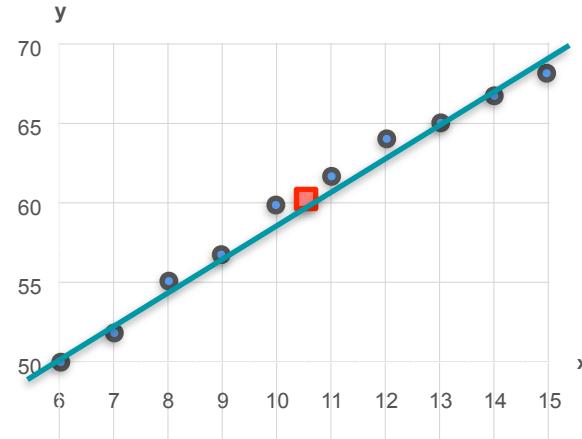
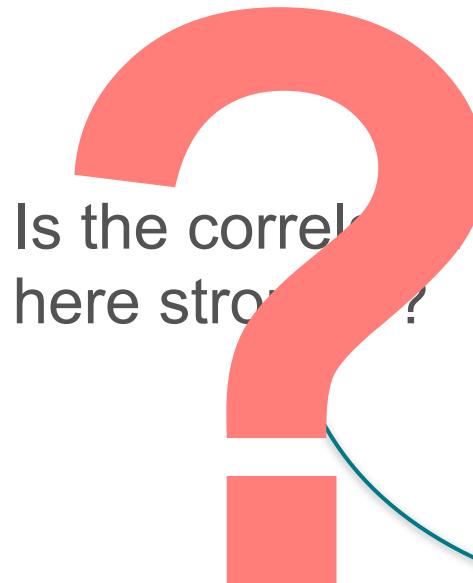


Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

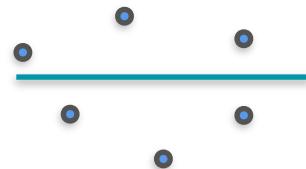
$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

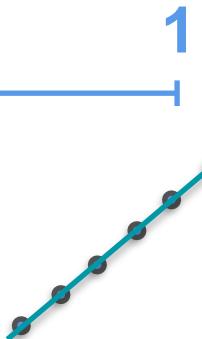


Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



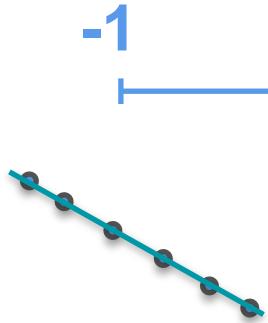
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



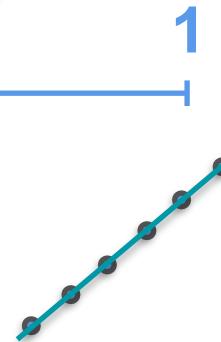
0

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

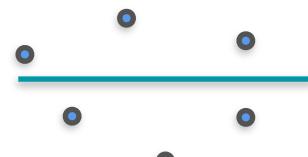
$$Cov(X, Y) = -7.45$$

-1



Correlation Coefficient

0



Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

1



Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation
Coefficient

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation
Coefficient =

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation
Coefficient = $Cov(X, Y)$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient = $\frac{Cov(X, Y)}{\sqrt{Var(X) \cdot Var(Y)}}$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sigma_x}$$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

**Correlation
Coefficient** = $\frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y}$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

**Correlation
Coefficient** = $\frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y}$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient = $\frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y}$ =

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

-1

0

1

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

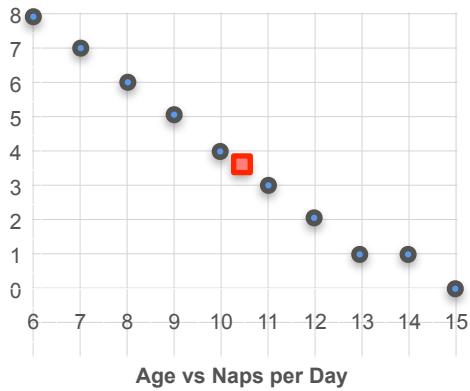
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

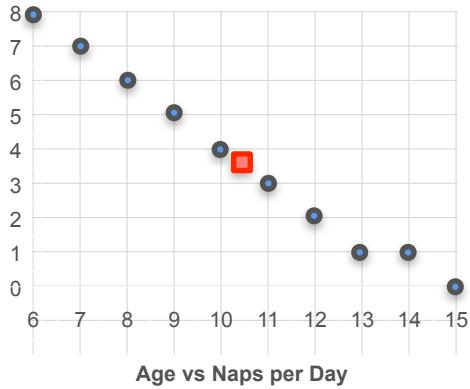
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

=

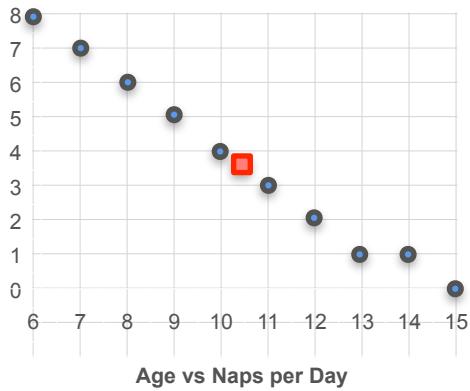
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{-7.45}{\sqrt{9.17} \cdot \sqrt{7.57}}\end{aligned}$$

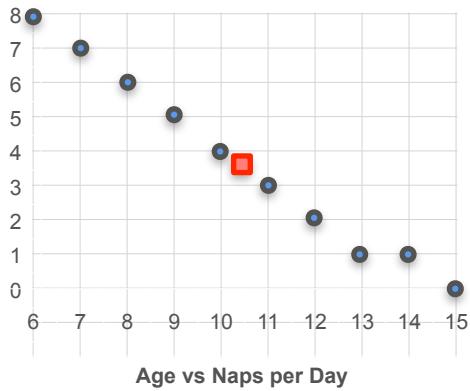
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

$$= \frac{-7.45}{\sqrt{9.17} \cdot \sqrt{7.57}}$$

$$\approx -0.894$$

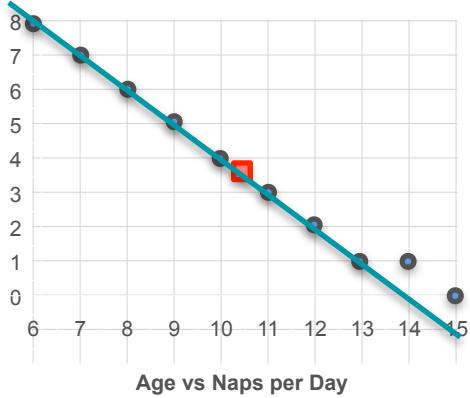
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

$$= \frac{-7.45}{\sqrt{9.17} \cdot \sqrt{7.57}}$$

$$\approx -0.894$$

Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

Correlation Coefficient = $\frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$

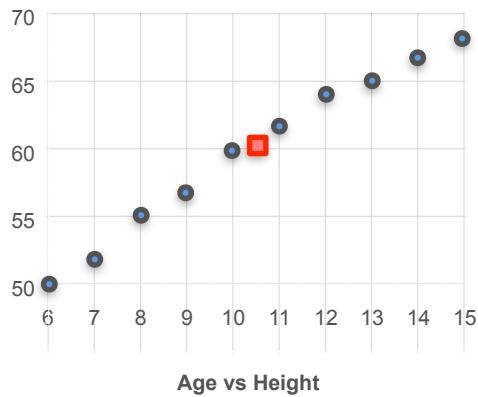
Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



Correlation Coefficient =
$$\frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

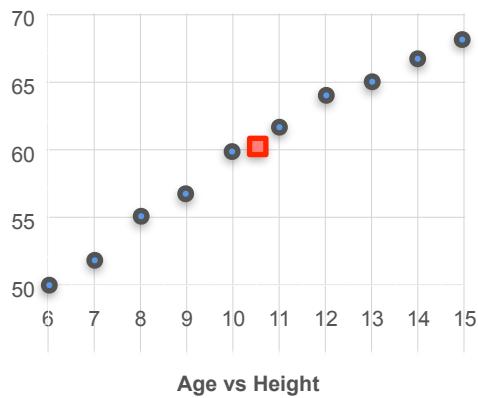
Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

=

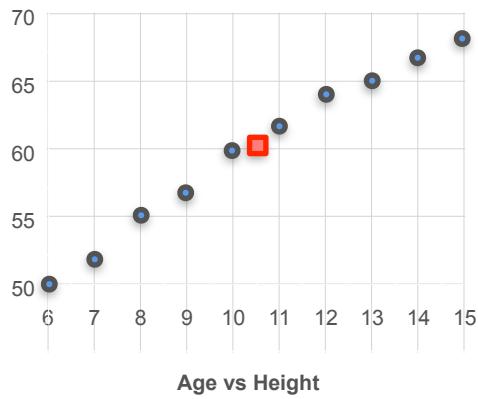
Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{17}{\sqrt{9.17} \cdot \sqrt{39.56}}\end{aligned}$$

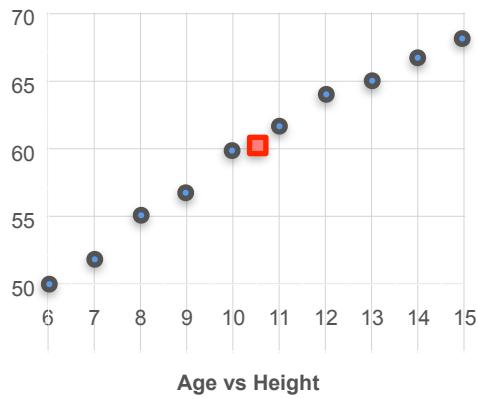
Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

$$= \frac{17}{\sqrt{9.17} \cdot \sqrt{39.56}}$$

$$\approx 0.893$$

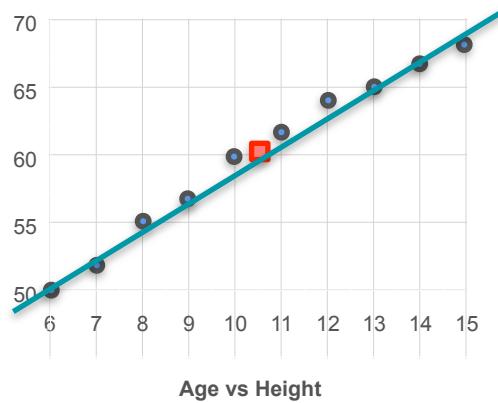
Correlation Coefficient

Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{17}{\sqrt{9.17} \cdot \sqrt{39.56}} \\ &\approx 0.893\end{aligned}$$

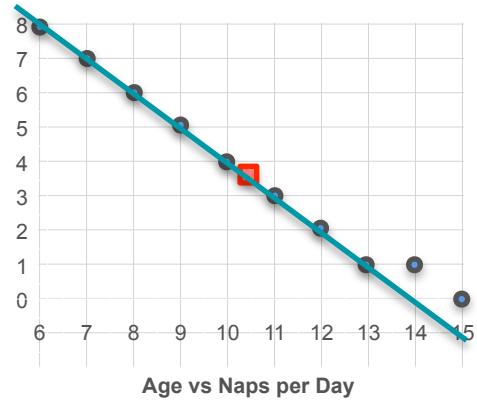
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$



Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$

$$\approx 0.893$$

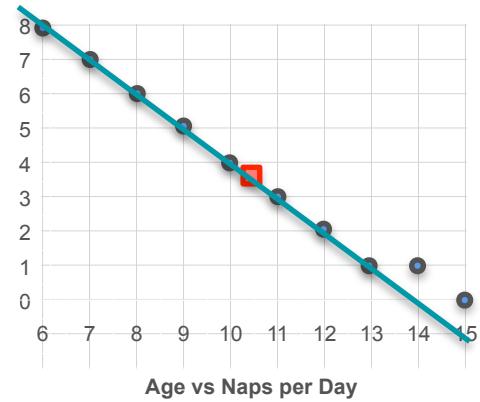
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

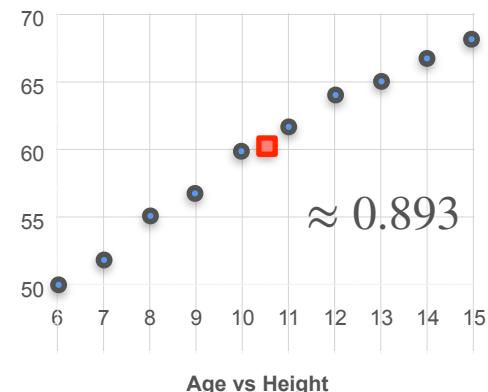


Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



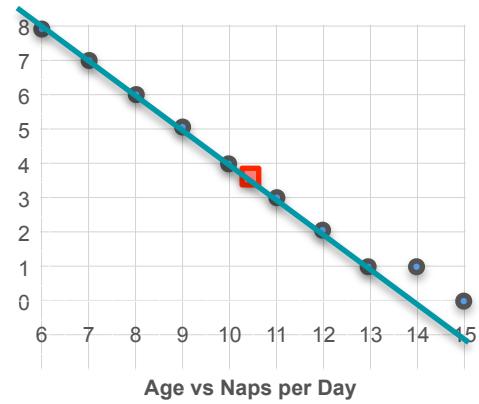
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

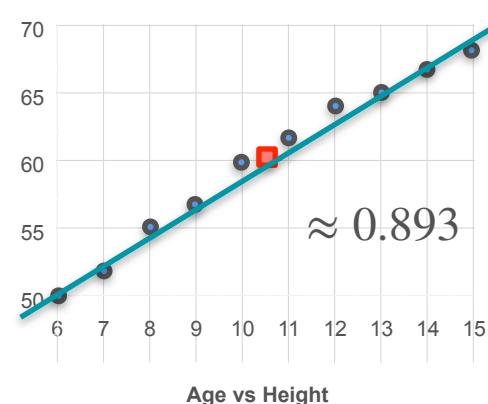


Age vs Height

$$Var(X) = 9.17$$

$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



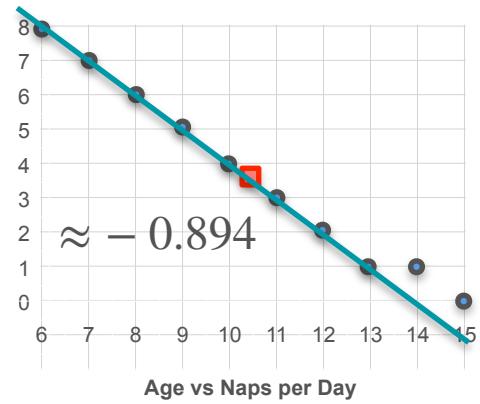
Correlation Coefficient

Age vs Naps per Day

$$Var(X) = 9.17$$

$$Var(Y) = 7.57$$

$$Cov(X, Y) = -7.45$$

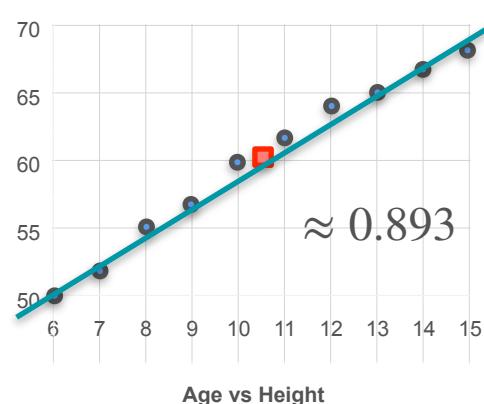


Age vs Height

$$Var(X) = 9.17$$

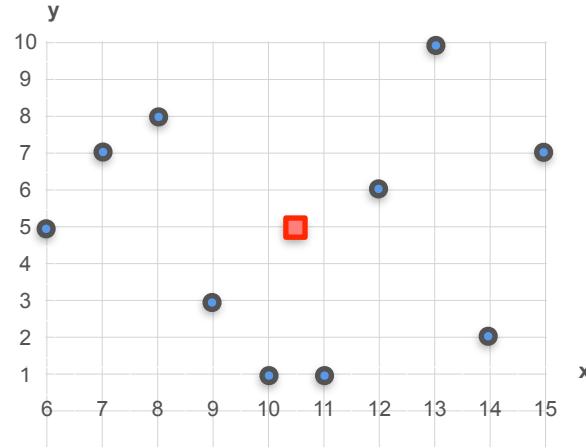
$$Var(Y) = 39.56$$

$$Cov(X, Y) = 17$$



Correlation Coefficient

Correlation Coefficient



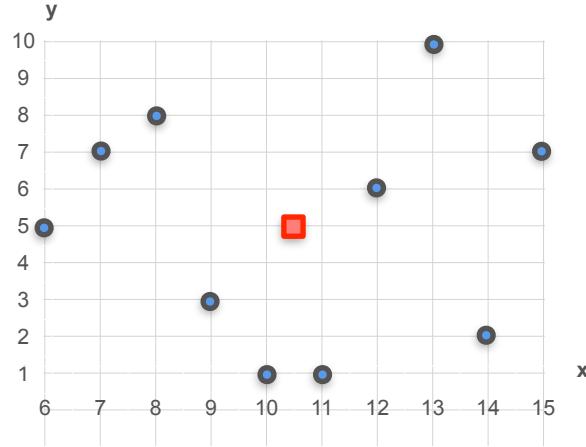
Age vs Grades

$$Var(X) = 9.17$$

$$Var(Y) = 9.78$$

$$Cov(X, Y) = 0.1$$

Correlation Coefficient



Age vs Grades

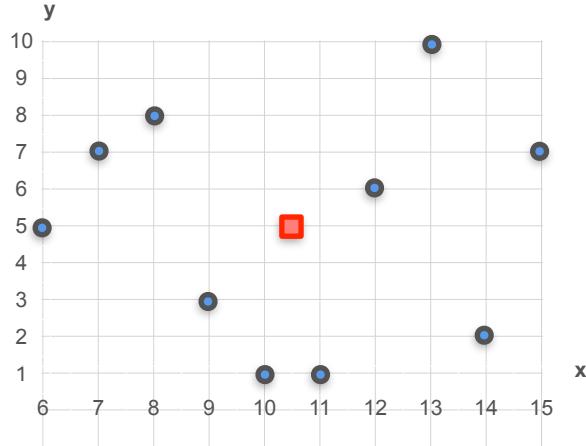
$$Var(X) = 9.17$$

$$Var(Y) = 9.78$$

$$Cov(X, Y) = 0.1$$

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

Correlation Coefficient



Age vs Grades

$$Var(X) = 9.17$$

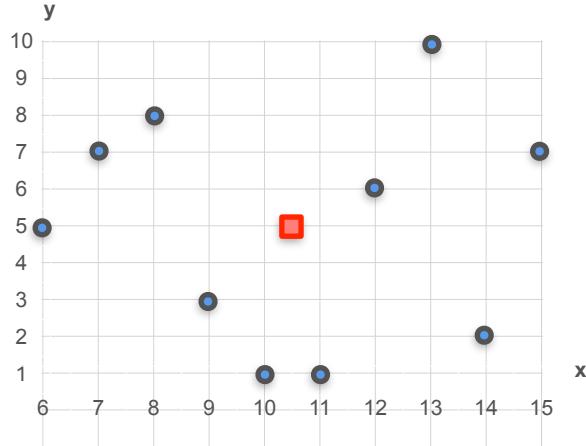
$$Var(Y) = 9.78$$

$$Cov(X, Y) = 0.1$$

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

=

Correlation Coefficient



Age vs Grades

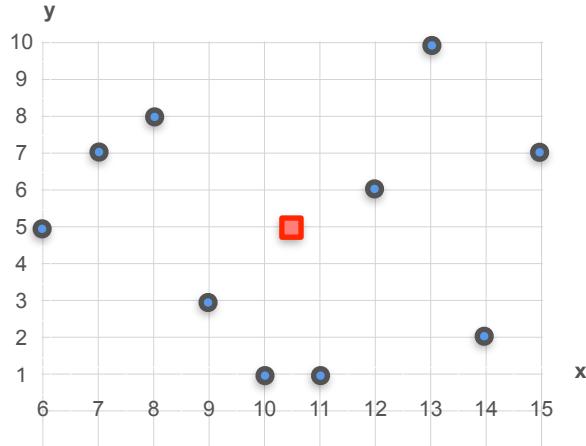
$$Var(X) = 9.17$$

$$Var(Y) = 9.78$$

$$Cov(X, Y) = 0.1$$

$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{0.1}{\sqrt{9.17} \cdot \sqrt{9.78}}\end{aligned}$$

Correlation Coefficient



Age vs Grades

$$Var(X) = 9.17$$

$$Var(Y) = 9.78$$

$$Cov(X, Y) = 0.1$$

$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{0.1}{\sqrt{9.17} \cdot \sqrt{9.78}}\end{aligned}$$

$$\approx 0.01$$

Correlation Coefficient

Correlation Coefficient

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

Correlation Coefficient

$$\text{Correlation Coefficient} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \cdot \sqrt{\text{Var}(Y)}}$$

=

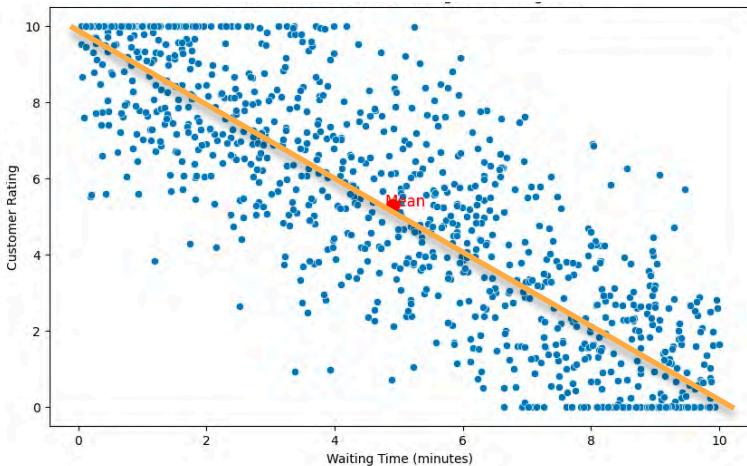
Correlation Coefficient

$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{-7.878}{\sqrt{8.562} \cdot \sqrt{10.163}}\end{aligned}$$

Correlation Coefficient

$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{-7.878}{\sqrt{8.562} \cdot \sqrt{10.163}} \\ &\approx -0.845\end{aligned}$$

Correlation Coefficient



$$Var(X) = 8.526$$

$$Var(Y) = 10.163$$

$$Cov(X, Y) = -7.878$$

$$\begin{aligned}\text{Correlation Coefficient} &= \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} \\ &= \frac{-7.878}{\sqrt{8.562} \cdot \sqrt{10.163}}\end{aligned}$$

≈ -0.845

Correlation Coefficient

Correlation Coefficient

Correlation
Coefficient =

Correlation Coefficient

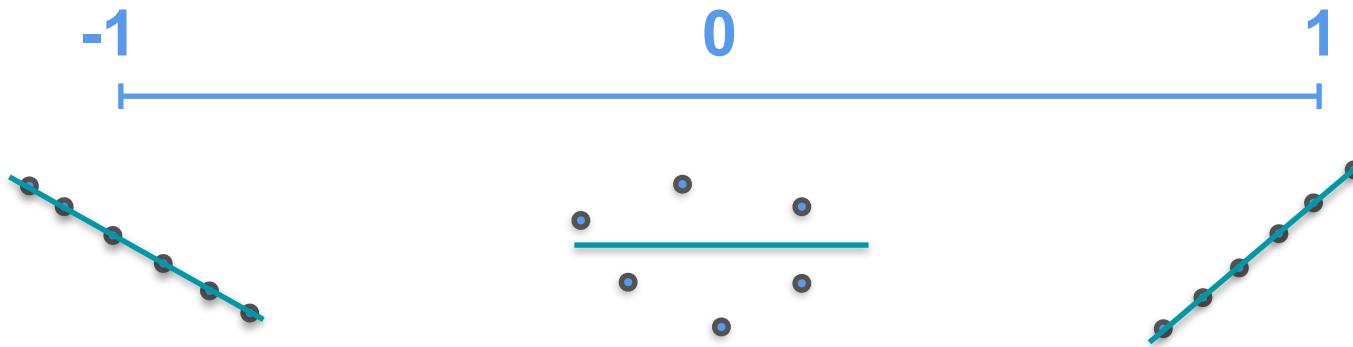
$$\text{Correlation Coefficient} = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y}$$

Correlation Coefficient

$$\text{Correlation Coefficient} = \frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

Correlation Coefficient

Correlation Coefficient = $\frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y}$ = $\frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$





DeepLearning.AI

Probability Distributions with Multiple Variables

Multivariate Gaussian Distribution

Multivariate Gaussian Distribution

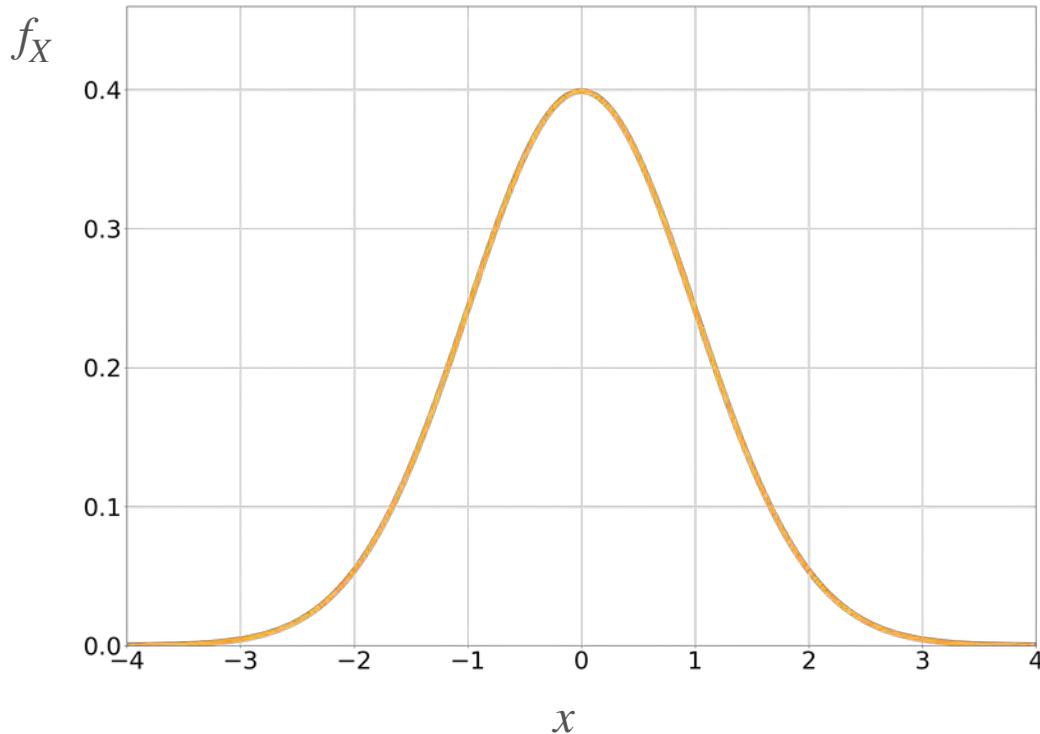
Multivariate Gaussian Distribution

For a single variable, X

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

Parameters:

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



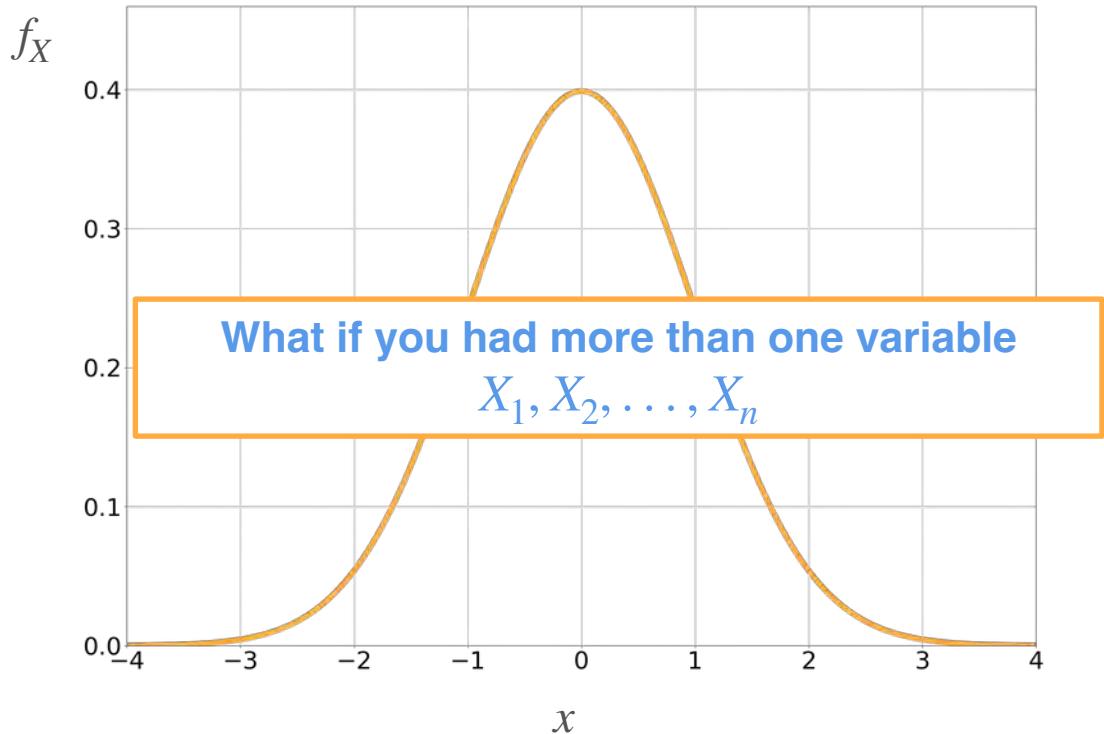
Multivariate Gaussian Distribution

For a single variable, X

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

Parameters:

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



Multivariate Gaussian Distribution: An Example

Multivariate Gaussian Distribution: An Example

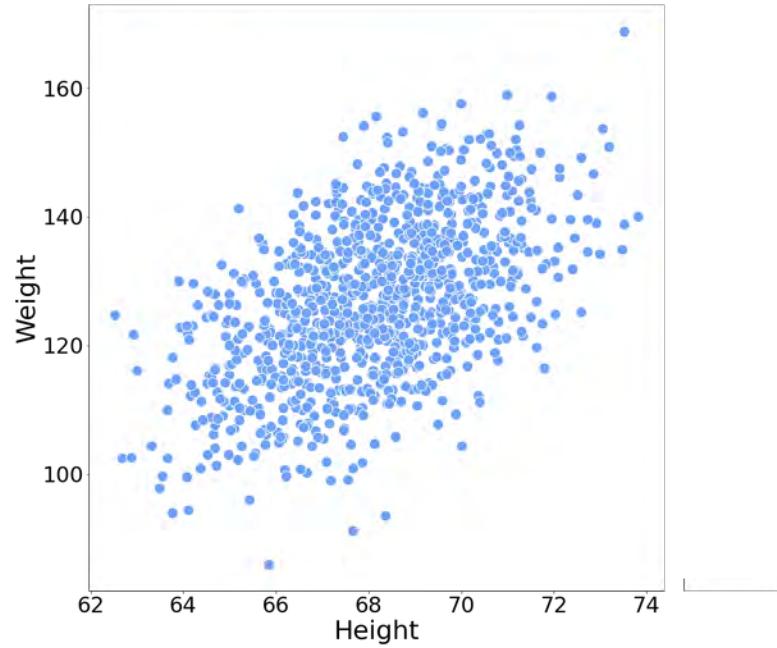
Two variables

Multivariate Gaussian Distribution: An Example

Two variables

H : Height of an adult in inches

W : Weight of an adult in pounds

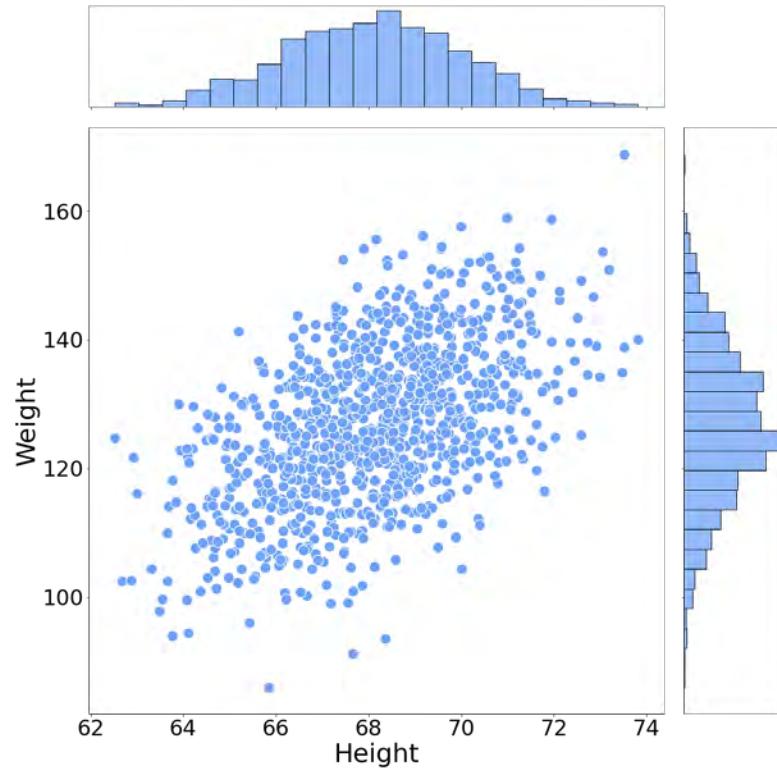


Multivariate Gaussian Distribution: An Example

Two variables

H : Height of an adult in inches

W : Weight of an adult in pounds



Multivariate Gaussian Distribution: An Example

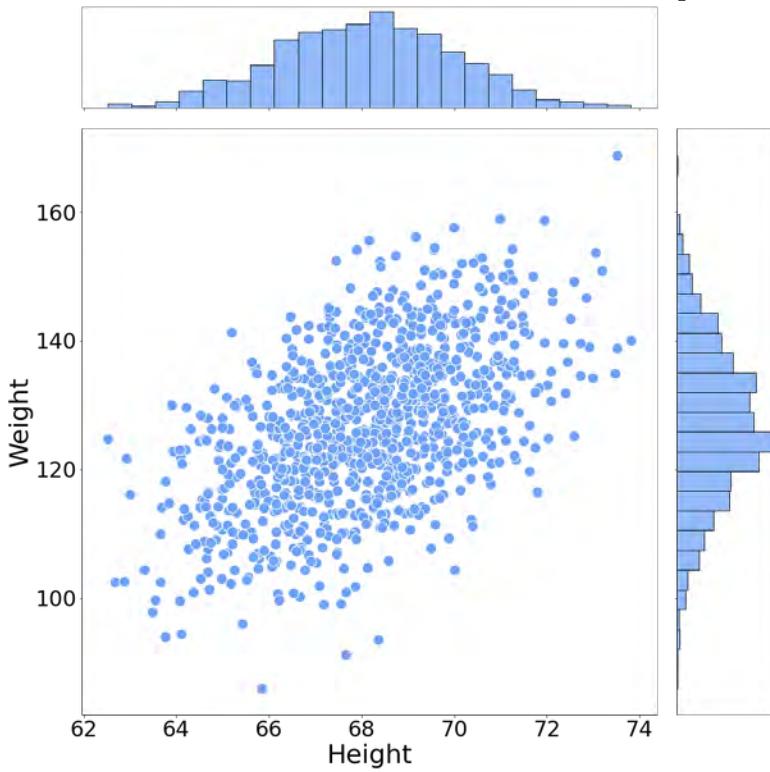
Two variables

H : Height of an adult in inches

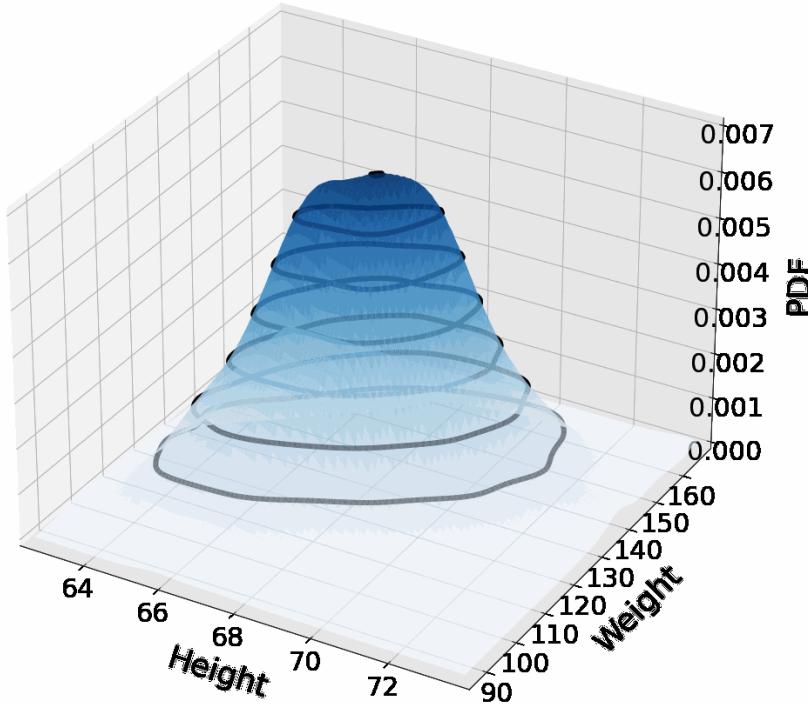
$$H \sim \mathcal{N}(\mu_H, \sigma_H)$$

W : Weight of an adult in pounds

$$W \sim \mathcal{N}(\mu_W, \sigma_W)$$



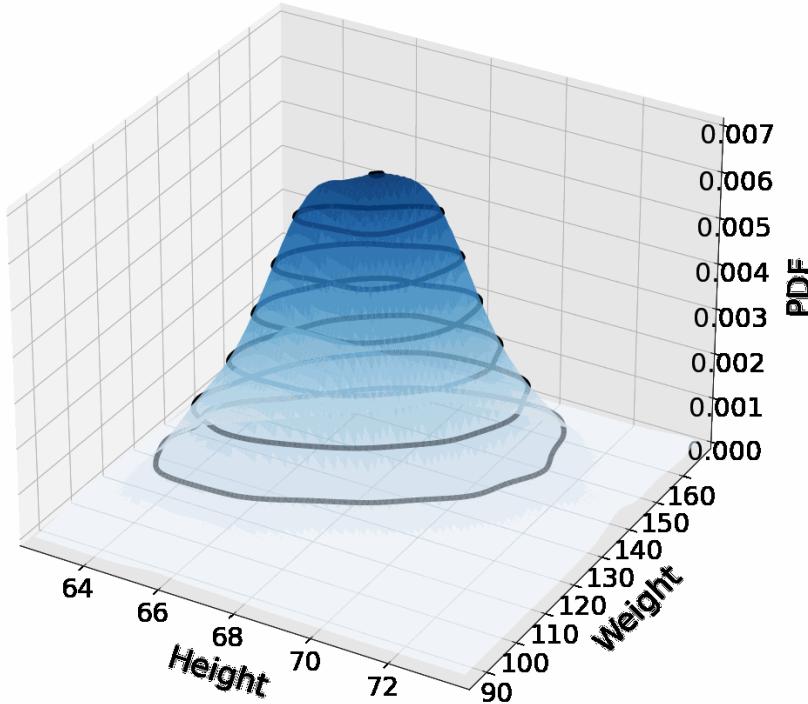
Multivariate Gaussian Distribution: An Example



If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

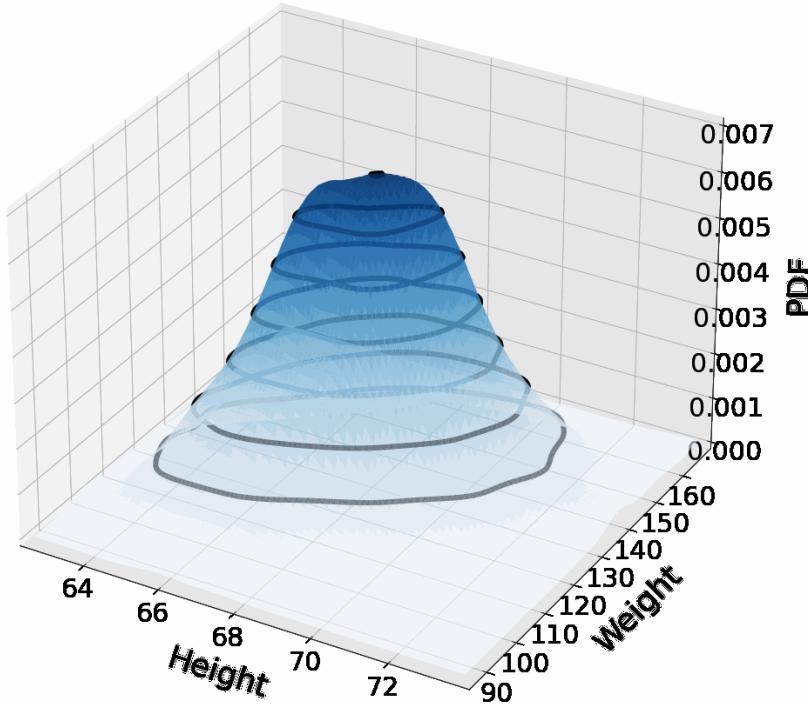
Multivariate Gaussian Distribution: An Example



If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

Multivariate Gaussian Distribution: An Example

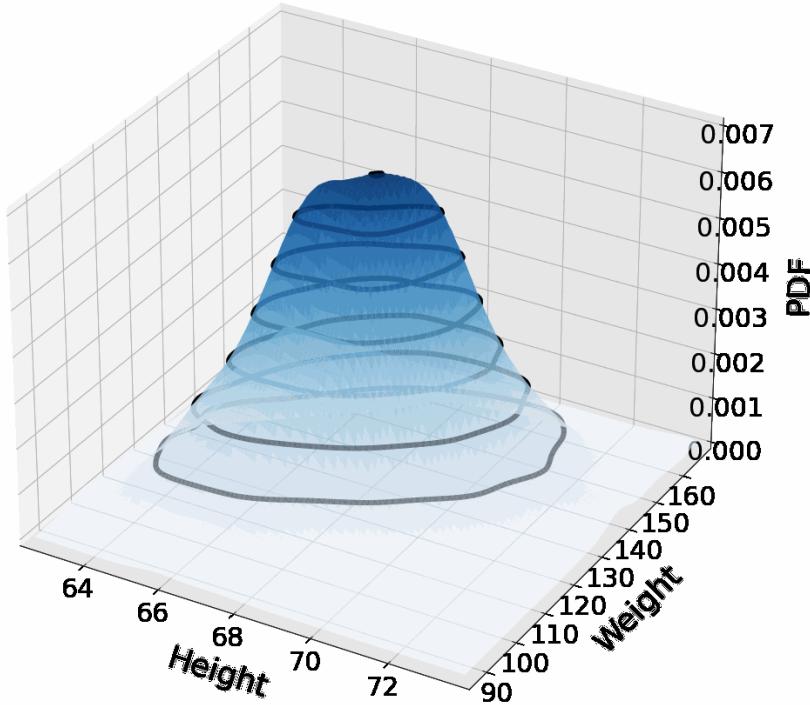


If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_H} e^{-\frac{1}{2}\frac{(h - \mu_H)^2}{\sigma_H^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_W} e^{-\frac{1}{2}\frac{(w - \mu_W)^2}{\sigma_W^2}}$$

Multivariate Gaussian Distribution: An Example



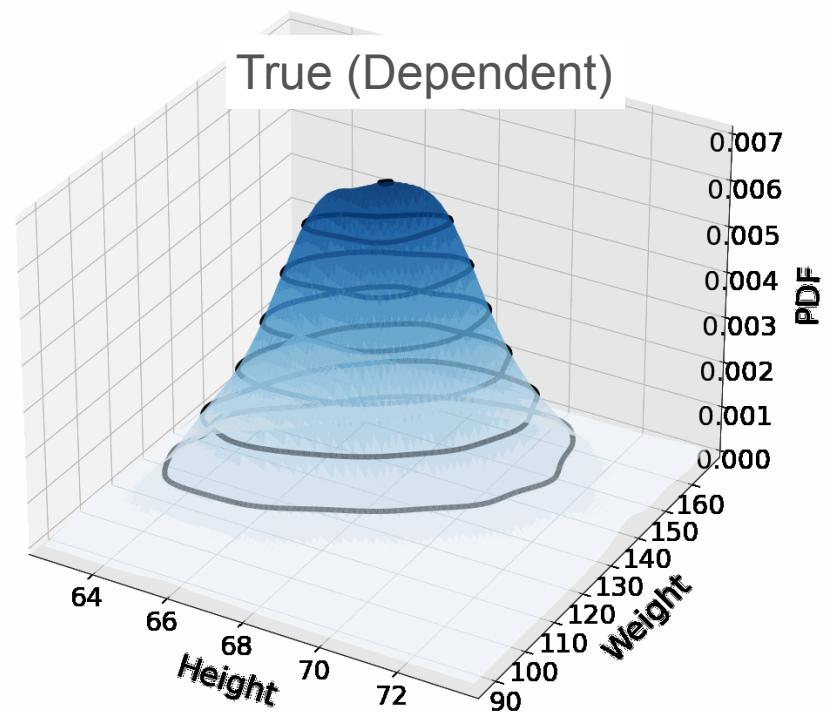
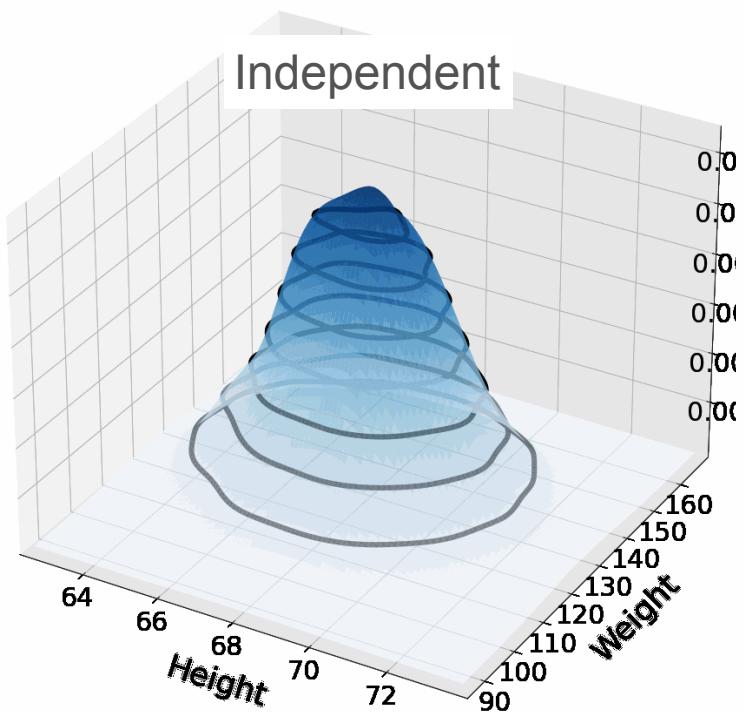
If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

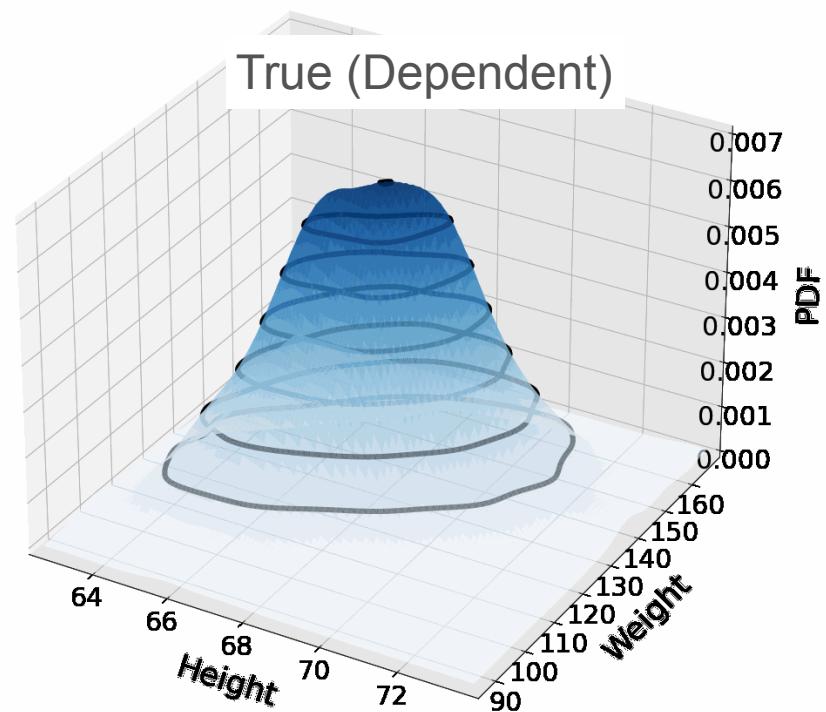
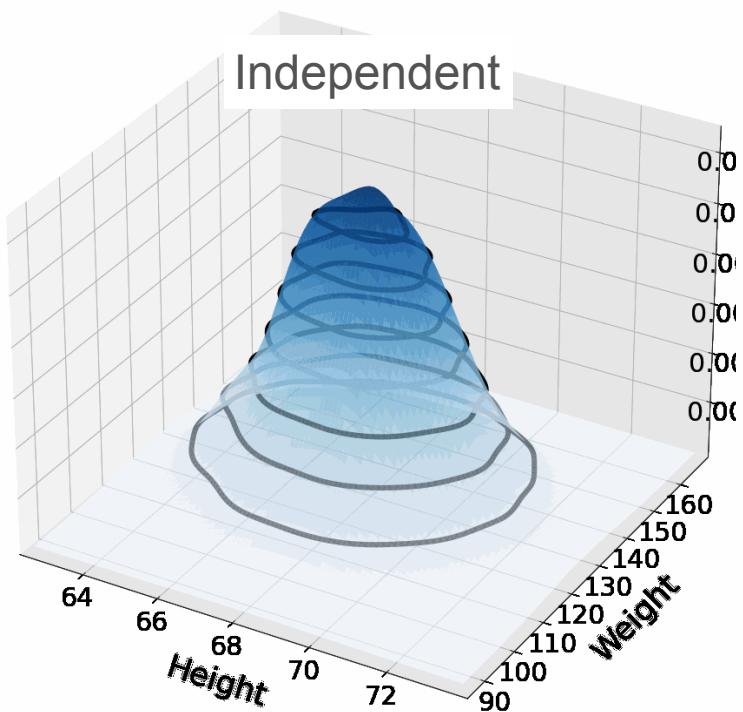
$$= \frac{1}{\sqrt{2\pi}\sigma_H} e^{-\frac{1}{2}\frac{(h-\mu_H)^2}{\sigma_H^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_W} e^{-\frac{1}{2}\frac{(w-\mu_W)^2}{\sigma_W^2}}$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

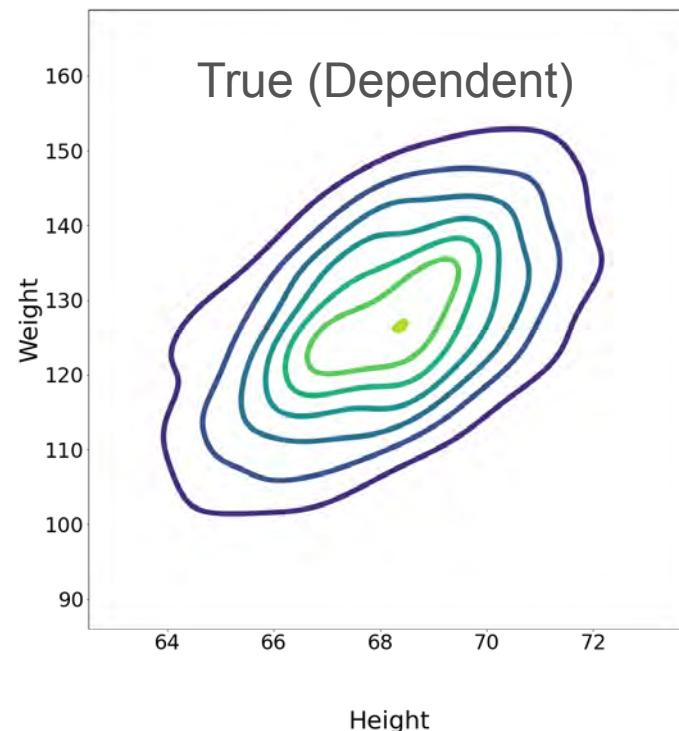
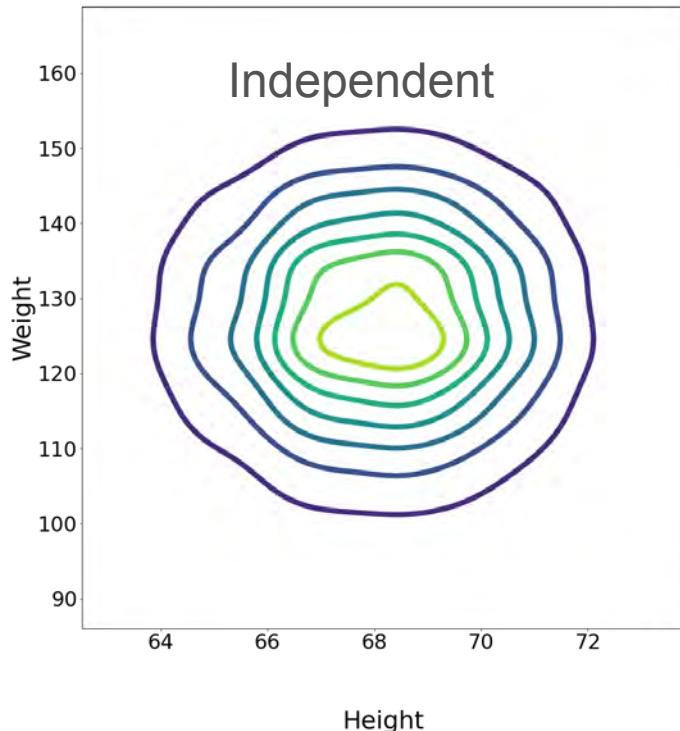
Multivariate Gaussian Distribution: An Example



Multivariate Gaussian Distribution: An Example



Multivariate Gaussian Distribution: An Example



Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 = \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} & \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix}$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|^2 = \begin{bmatrix} h-\mu_H & w-\mu_W \\ \sigma_H & \sigma_W \end{bmatrix} \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix}$$
$$\left[\begin{array}{c} h - \mu_h \\ w - \mu_w \end{array} \right] = \left[\begin{array}{c} h \\ w \end{array} \right] - \left[\begin{array}{c} \mu_h \\ \mu_w \end{array} \right]$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|^2 = \frac{h-\mu_H}{\sigma_H} \frac{w-\mu_W}{\sigma_W}$$
$$\left[\begin{bmatrix} h - \mu_h \\ w - \mu_w \end{bmatrix} \right] = \left[\begin{bmatrix} h \\ w \end{bmatrix} \right] - \left[\begin{bmatrix} \mu_h \\ \mu_w \end{bmatrix} \right]$$

Multiply by diagonal matrix

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\begin{aligned} & \left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 \\ &= ([h \ w] - [\mu_H \ \mu_W]) \begin{bmatrix} \frac{1}{\sigma_H^2} & 0 \\ 0 & \frac{1}{\sigma_W^2} \end{bmatrix} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right) \end{aligned}$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\begin{aligned} & \left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 \\ &= \left[\frac{h-\mu_H}{\sigma_H} \quad \frac{w-\mu_W}{\sigma_W} \right] \begin{bmatrix} \frac{1}{\sigma_H^2} & 0 \\ 0 & \frac{1}{\sigma_W^2} \end{bmatrix} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right) \\ &= \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right) \end{aligned}$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 = \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} & \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_H^2} & 0 \\ 0 & \frac{1}{\sigma_W^2} \end{bmatrix} \begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix}$$

Covariance matrix!
(Σ)

$$= \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 = \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} & \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_H^2} & 0 \\ 0 & \frac{1}{\sigma_W^2} \end{bmatrix} \begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix}$$

Covariance matrix!
(Σ)

$$= \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$\det(\Sigma)^{1/2}$

Covariance matrix!
(Σ)

$$\begin{aligned} & \left\| \begin{bmatrix} \frac{h-\mu_H}{\sigma_H} \\ \frac{w-\mu_W}{\sigma_W} \end{bmatrix} \right\|_2^2 \\ &= \left[\frac{h-\mu_H}{\sigma_H} \quad \frac{w-\mu_W}{\sigma_W} \right] \begin{bmatrix} \frac{1}{\sigma_H^2} & 0 \\ 0 & \frac{1}{\sigma_W^2} \end{bmatrix} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right) \end{aligned}$$

$$= \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} \exp\left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)^T \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \begin{bmatrix} \mu_H \\ \mu_W \end{bmatrix} \right)\right)$$

Multivariate Gaussian Distribution: An Example

If W, H were independent

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$= \frac{1}{2\pi \det \Sigma^{1/2}} \exp \left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right)^T \boldsymbol{\Sigma^{-1}} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right) \right)$$

Multivariate Gaussian Distribution: An Example

Dependent case:

$$f_{HW}(h, w) = f_H(h)f_W(w)$$

$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$= \frac{1}{2\pi \det \Sigma^{1/2}} \exp \left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right)^T \boldsymbol{\Sigma^{-1}} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right) \right)$$

Multivariate Gaussian Distribution: An Example

Dependent case:

$$\begin{aligned} f_{HW}(h, w) &= \cancel{f_H(h)f_W(w)} \\ &= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)} \\ &= \frac{1}{2\pi\det\Sigma^{1/2}} \exp\left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right)^T \Sigma^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right) \right) \end{aligned}$$

Multivariate Gaussian Distribution: An Example

Dependent case:

$$f_{HW}(h, w) = \cancel{f_H(h)f_W(w)}$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$= \frac{1}{2\pi \det \Sigma^{1/2}} \exp \left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right)^T \boldsymbol{\Sigma^{-1}} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu} \right) \right)$$

Multivariate Gaussian Distribution: An Example

Dependent case:

$$f_{HW}(h, w) = \cancel{f_H(h)f_W(w)}$$
$$= \frac{1}{2\pi\sigma_H\sigma_W} e^{-\frac{1}{2}\left(\frac{(h-\mu_H)^2}{\sigma_H^2} + \frac{(w-\mu_W)^2}{\sigma_W^2}\right)}$$

$$\Sigma = \begin{bmatrix} \sigma_H^2 & Cov(H, W) \\ Cov(H, W) & \sigma_W^2 \end{bmatrix}$$

$$= \frac{1}{2\pi\det\Sigma^{1/2}} \exp\left(-\frac{1}{2} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu}\right)^T \Sigma^{-1} \left(\begin{bmatrix} h \\ w \end{bmatrix} - \boldsymbol{\mu}\right)\right)$$

Multivariate Gaussian Distribution: General Definition

Multivariate Gaussian Distribution: General Definition

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

Multivariate Gaussian Distribution: General Definition

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

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) =$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) =$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{\dots}$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{\sqrt{2\pi}\sigma}$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2}}$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2}}$$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}}$$

covariance matrix
spread of the bell



$|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}}$$

covariance matrix
spread of the bell

↑
 $|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}}$$

covariance matrix
spread of the bell

↑
 $|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}}$$

covariance matrix
spread of the bell

↑
 $|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}$$
$$\boldsymbol{x} = [x_1 \ x_2 \ \dots \ x_n]^T$$
$$\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_n]^T$$
$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}$$

covariance matrix
spread of the bell

$|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}$$

covariance matrix
spread of the bell

$|\Sigma|$ determinant of the covariance matrix

$$x = [x_1 \ x_2 \ \dots \ x_n]^T$$

Mean vector
 $\mu = [\mu_1, \mu_2, \dots, \mu_n]^T$

Multivariate Gaussian Distribution: General Definition

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

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}$$

covariance matrix
spread of the bell

Mean vector
 $\mu = [\mu_1, \mu_2, \dots, \mu_n]^T$

Rescaling / standardization

$|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: General Definition

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

$$\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$$

Mean vector
 $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_n]^T$

$f_X(x_1, x_2)$
random
 $X = [X_1 \ X_2 \ \dots \ X_n]$

- For univariate, we work with scalar values and variances
- For multivariate, we work with vectors and the covariance matrix

covariance matrix
spread of the bell

$|\Sigma|$ determinant of the covariance matrix

Multivariate Gaussian Distribution: Conditionals



DeepLearning.AI

Probability Distributions with Multiple Variables

Conclusion