

REVOLUTION

Adding Random Variables

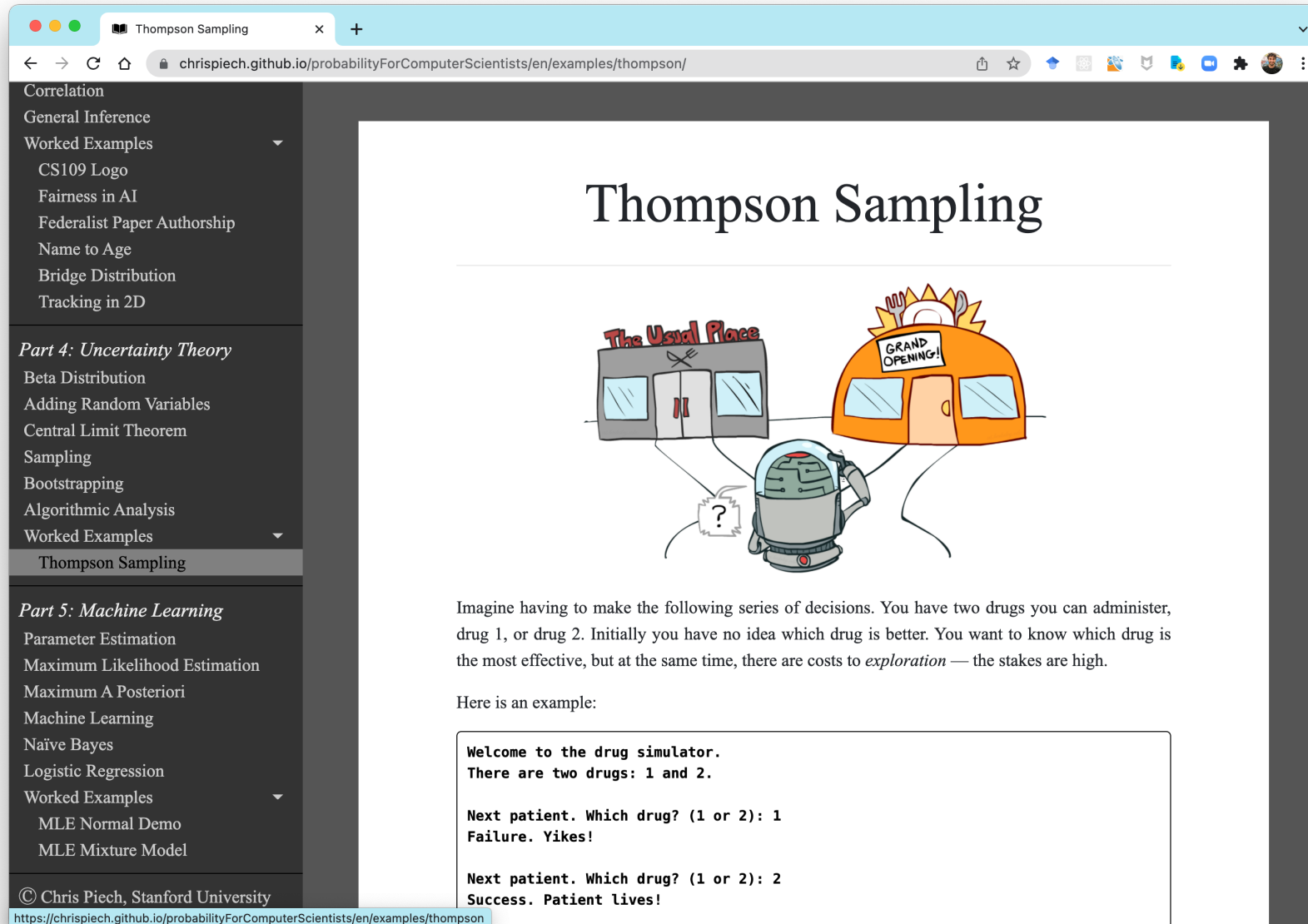
Chris Piech

CS109, Stanford University

Announcements

- New sections in the reader
- Pset 4, the pros of finishing early
- The CS109 Challenge
- Digitally signing pset work???
- Midterm results (end of class)

Thompson Sampling



Correlation

General Inference

Worked Examples

- CS109 Logo
- Fairness in AI
- Federalist Paper Authorship
- Name to Age
- Bridge Distribution
- Tracking in 2D

Part 4: Uncertainty Theory

- Beta Distribution
- Adding Random Variables
- Central Limit Theorem
- Sampling
- Bootstrapping
- Algorithmic Analysis
- Worked Examples
- Thompson Sampling**

Part 5: Machine Learning

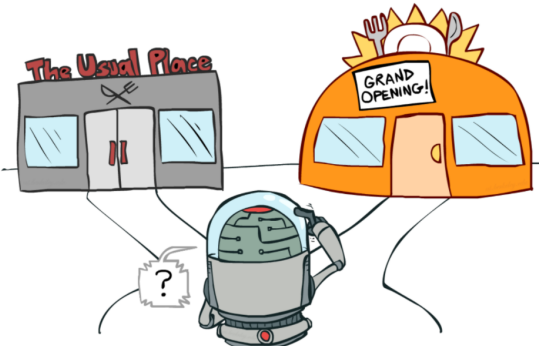
- Parameter Estimation
- Maximum Likelihood Estimation
- Maximum A Posteriori
- Machine Learning
- Naïve Bayes
- Logistic Regression
- Worked Examples

 - MLE Normal Demo
 - MLE Mixture Model

© Chris Piech, Stanford University

<https://chrispiech.github.io/probabilityForComputerScientists/en/examples/thompson>

Thompson Sampling



Imagine having to make the following series of decisions. You have two drugs you can administer, drug 1, or drug 2. Initially you have no idea which drug is better. You want to know which drug is the most effective, but at the same time, there are costs to *exploration* — the stakes are high.

Here is an example:

```
Welcome to the drug simulator.
There are two drugs: 1 and 2.


Next patient. Which drug? (1 or 2): 1
Failure. Yikes!

Next patient. Which drug? (1 or 2): 2
Success. Patient lives!
```

PSet 4: Deadline Coming Up

PS4 Chess.com Puzzles

Chess.com is a website for playing chess. They are trying to estimate how well a player can solve chess puzzles (puzzle ability) as a random variable, A , which can take on integer values in the range 0 to 100 inclusive. Higher abilities mean the player is better at chess puzzles. Note that ability is **discrete**.



Write a function `update_belief` which takes in a prior belief in a player's puzzle ability and an observation of them solving a puzzle. Your function should infer the posterior belief in the player's ability, based on the observation using Bayes' Theorem (with random variables).

Representation of Belief

Both the **prior** and the **posterior** you return should be probability mass functions for ability. These probability mass functions are represented as a dictionary where the keys are all the values that ability can take on $[0, 100]$. The value corresponding to key i represents $P(A = i)$. The posterior that you return should be a dictionary with the exact same keys, where the value corresponding to key i represents $P(A = i | Y = y)$. Y is a Bernoulli random variable which is

Previous Question Next Question

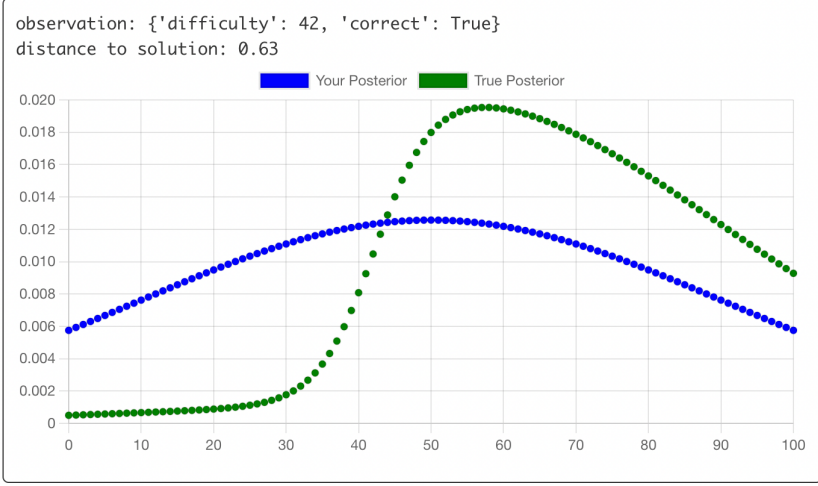
Answer Editor Solution

Program:

```
1 import math
2
3 def update_belief(prior, observation):
4     # TODO: your code here
5     return prior
6
7 #####
8 # Helper Functions!
9 #####
10
11 def p_correct_given_ability(ability, difficulty):
12     """
13     This uses item response theory to model the chance that a
14     patient with a given ability will correctly identify a letter
15     of a given size
16     """
17     p_guess = 0.05
18     p_slip = 0.08
```

Run

observation: {'difficulty': 42, 'correct': True}
distance to solution: 0.63





New Definition

IID Random Variables

- Consider n random variables X_1, X_2, \dots, X_n
 - X_i are all independently and identically distributed (I.I.D.)
 - All have the same PMF (if discrete) or PDF (if continuous)
 - All have the same expectation
 - All have the same variance

IID

iid

Quick check





Are X_1, X_2, \dots, X_n i.i.d. with the following distributions?

1. $X_i \sim \text{Exp}(\lambda)$, X_i independent
2. $X_i \sim \text{Exp}(\lambda_i)$, X_i independent
3. $X_i \sim \text{Exp}(\lambda)$, $X_1 = X_2 = \dots = X_n$
4. $X_i \sim \text{Bin}(n_i, p)$, X_i independent



Quick check

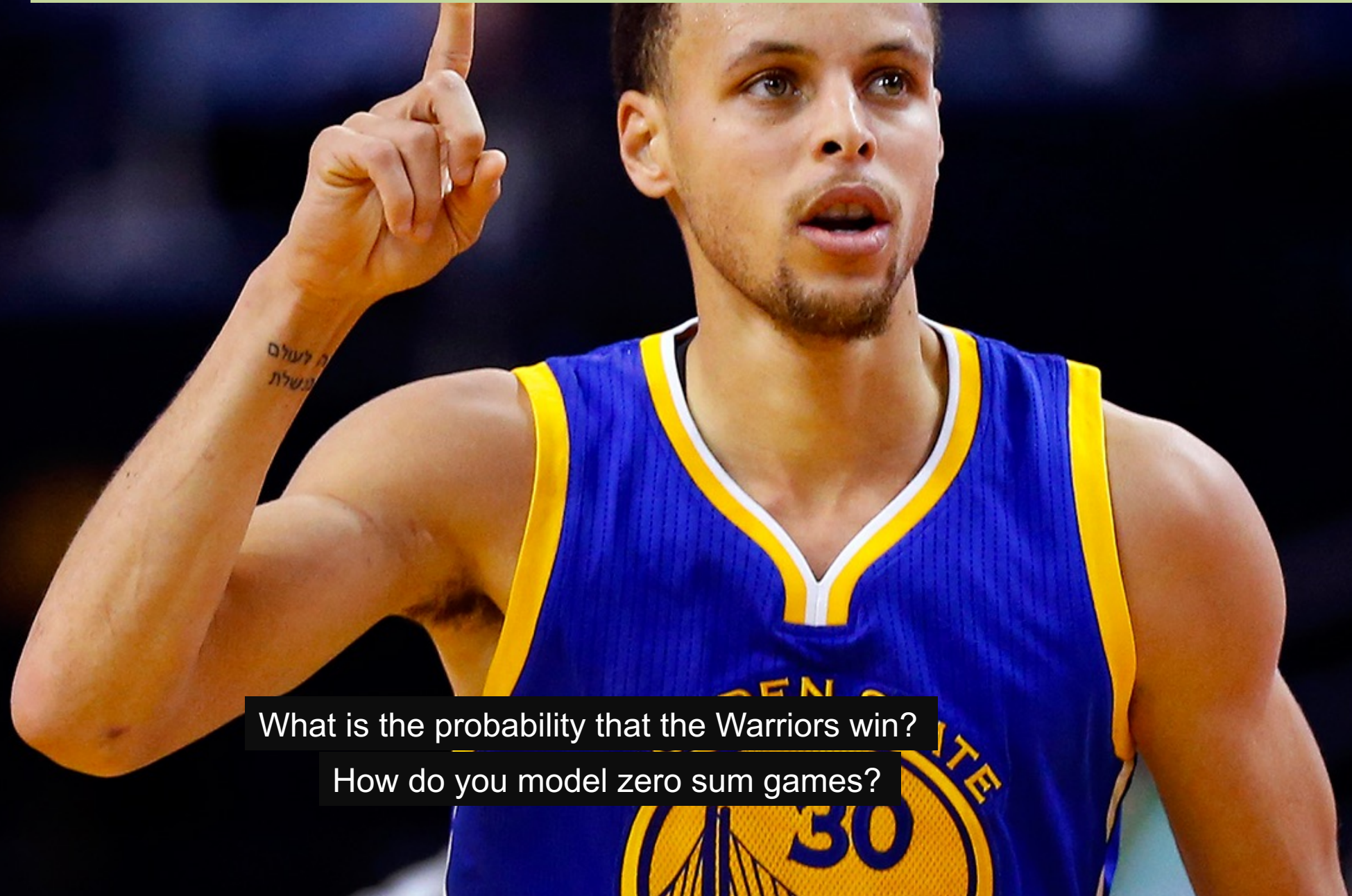
Are X_1, X_2, \dots, X_n i.i.d. with the following distributions?

1. $X_i \sim \text{Exp}(\lambda)$, X_i independent 
2. $X_i \sim \text{Exp}(\lambda_i)$, X_i independent  (unless λ_i equal)
3. $X_i \sim \text{Exp}(\lambda)$, $X_1 = X_2 = \dots = X_n$  dependent: $X_1 = X_2 = \dots = X_n$
4. $X_i \sim \text{Bin}(n_i, p)$, X_i independent  (unless n_i equal)
Note underlying Bernoulli RVs are i.i.d.!

What happens when you add random variables?

Why should you care?

Zero Sum Games



What is the probability that the Warriors win?

How do you model zero sum games?

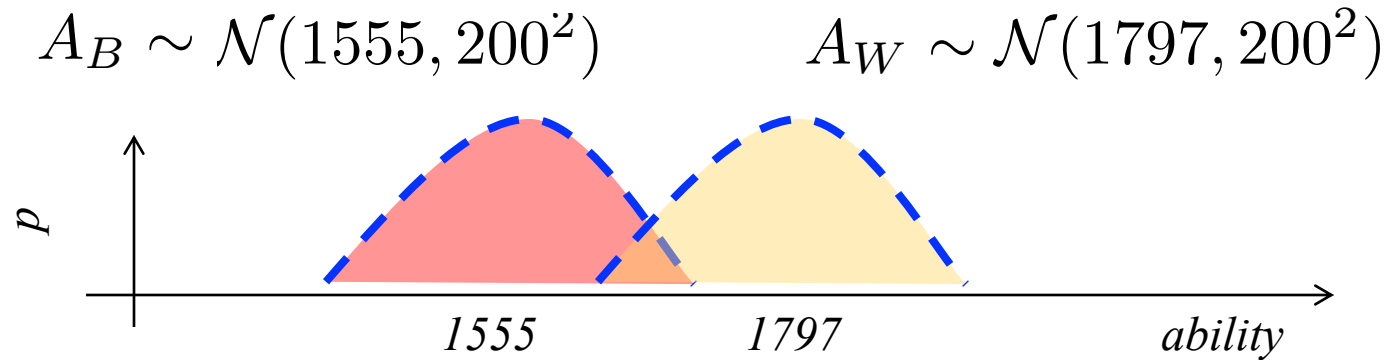
Motivating Idea: Zero Sum Games

How it works:

- Each team has an “ELO” score S , calculated based on their past performance.
- Each game, the team has ability $A \sim \mathcal{N}(S, 200^2)$
- The team with the higher sampled ability wins.



Arpad Elo



$$P(\text{Warriors win}) = P(A_W > A_B)$$

Motivating Idea: Zero Sum Games

$$A_W \sim \mathcal{N}(1797, 200^2)$$

$$A_B \sim \mathcal{N}(1555, 200^2)$$

$$P(\text{Warriors win}) = P(A_W > A_B)$$

How do we do this???

Review

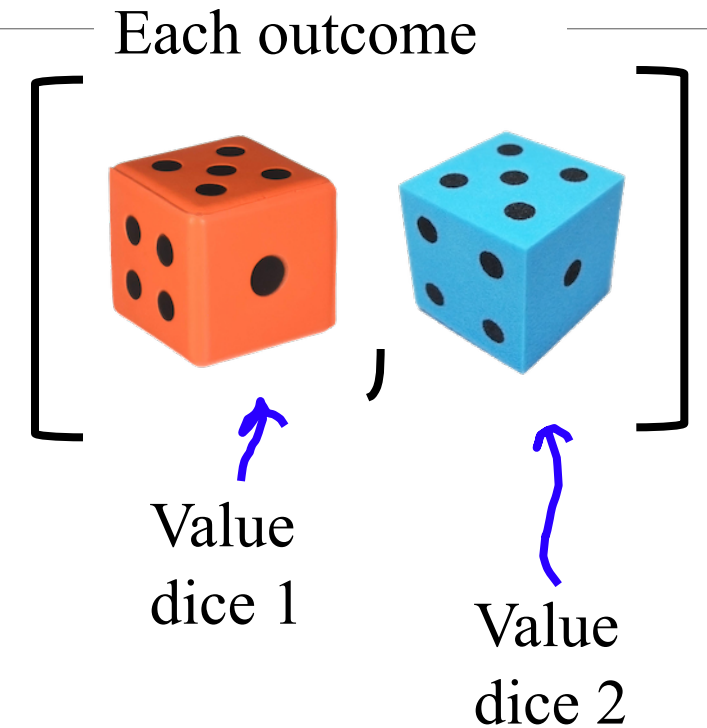
Sum of Two Die?

Roll two 6-sided dice. What is $P(\text{sum} = 7)$?

$S = \{$

[1,1]	[1,2]	[1,3]	[1,4]	[1,5]	[1,6]
[2,1]	[2,2]	[2,3]	[2,4]	[2,5]	[2,6]
[3,1]	[3,2]	[3,3]	[3,4]	[3,5]	[3,6]
[4,1]	[4,2]	[4,3]	[4,4]	[4,5]	[4,6]
[5,1]	[5,2]	[5,3]	[5,4]	[5,5]	[5,6]
[6,1]	[6,2]	[6,3]	[6,4]	[6,5]	[6,6]

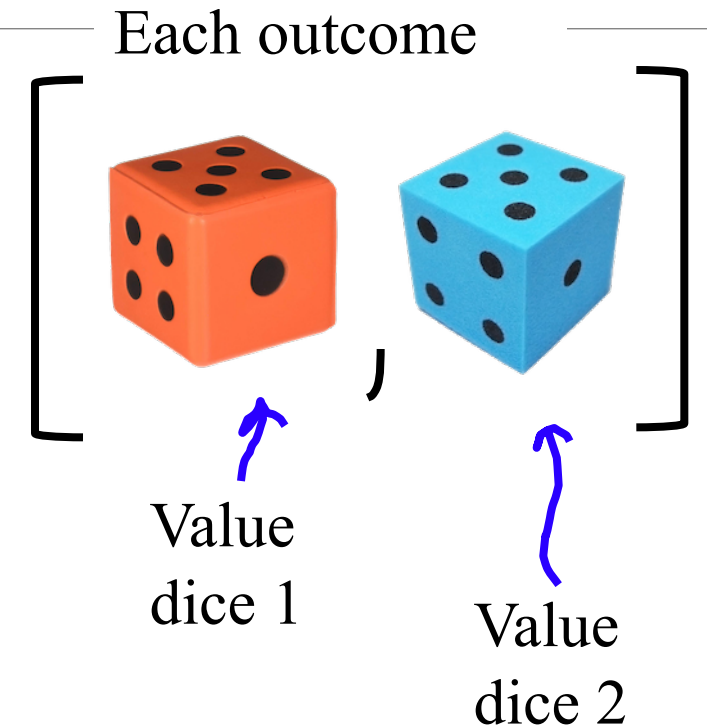
$\}$



Sum of Two Die = 7?

Roll two 6-sided dice. What is $P(\text{sum} = 7)$?

S = {	[1,1]	[1,2]	[1,3]	[1,4]	[1,5]	[1,6]
	[2,1]	[2,2]	[2,3]	[2,4]	[2,5]	[2,6]
	[3,1]	[3,2]	[3,3]	[3,4]	[3,5]	[3,6]
	[4,1]	[4,2]	[4,3]	[4,4]	[4,5]	[4,6]
	[5,1]	[5,2]	[5,3]	[5,4]	[5,5]	[5,6]
	[6,1]	[6,2]	[6,3]	[6,4]	[6,5]	[6,6] }



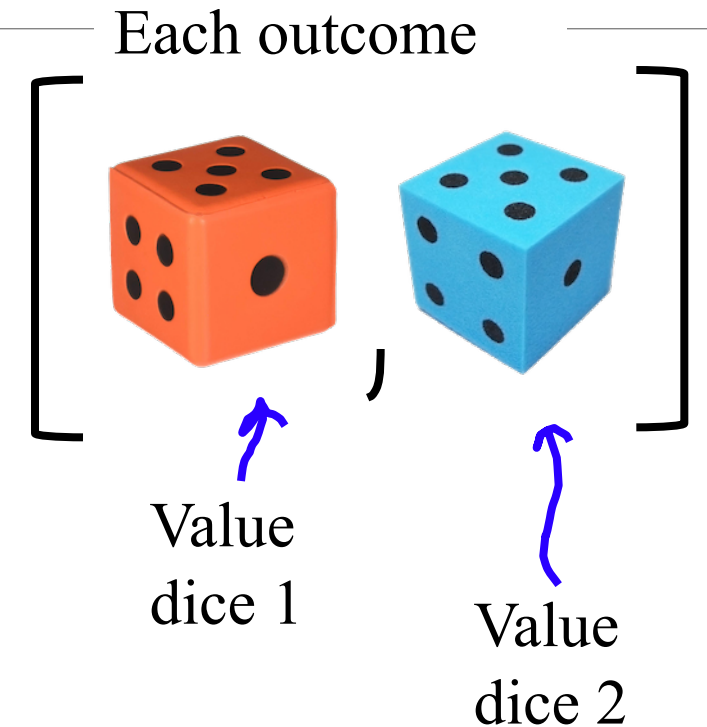
$E = \textit{in blue}$

$$P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.1\overline{6}$$

Sum of Two Die = 10?

Roll two 6-sided dice. What is $P(\text{sum} = 10)$?

S = {	[1,1]	[1,2]	[1,3]	[1,4]	[1,5]	[1,6]
	[2,1]	[2,2]	[2,3]	[2,4]	[2,5]	[2,6]
	[3,1]	[3,2]	[3,3]	[3,4]	[3,5]	[3,6]
	[4,1]	[4,2]	[4,3]	[4,4]	[4,5]	[4,6]
	[5,1]	[5,2]	[5,3]	[5,4]	[5,5]	[5,6]
	[6,1]	[6,2]	[6,3]	[6,4]	[6,5]	[6,6] }



$E = \textit{in blue}$

$$P(E) = \frac{|E|}{|S|} = \frac{3}{36} = 0.08\bar{3}$$

End Review

Sum of Two Dice



$$Y = \sum_{i=1}^2 X_i$$



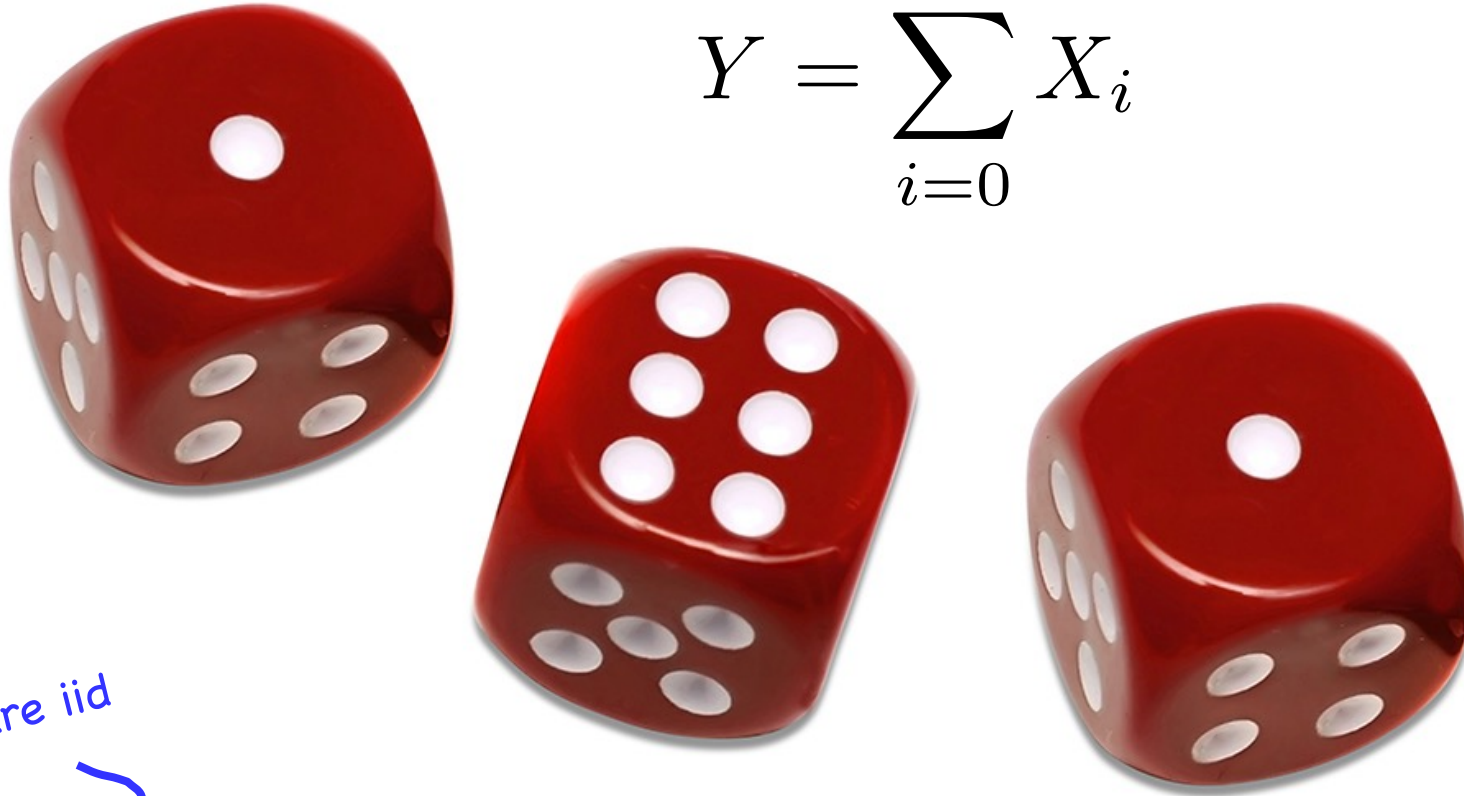
X_i s are iid



*X_i is the outcome of dice roll *i**

Sum of Three Dice

$$Y = \sum_{i=0}^3 X_i$$



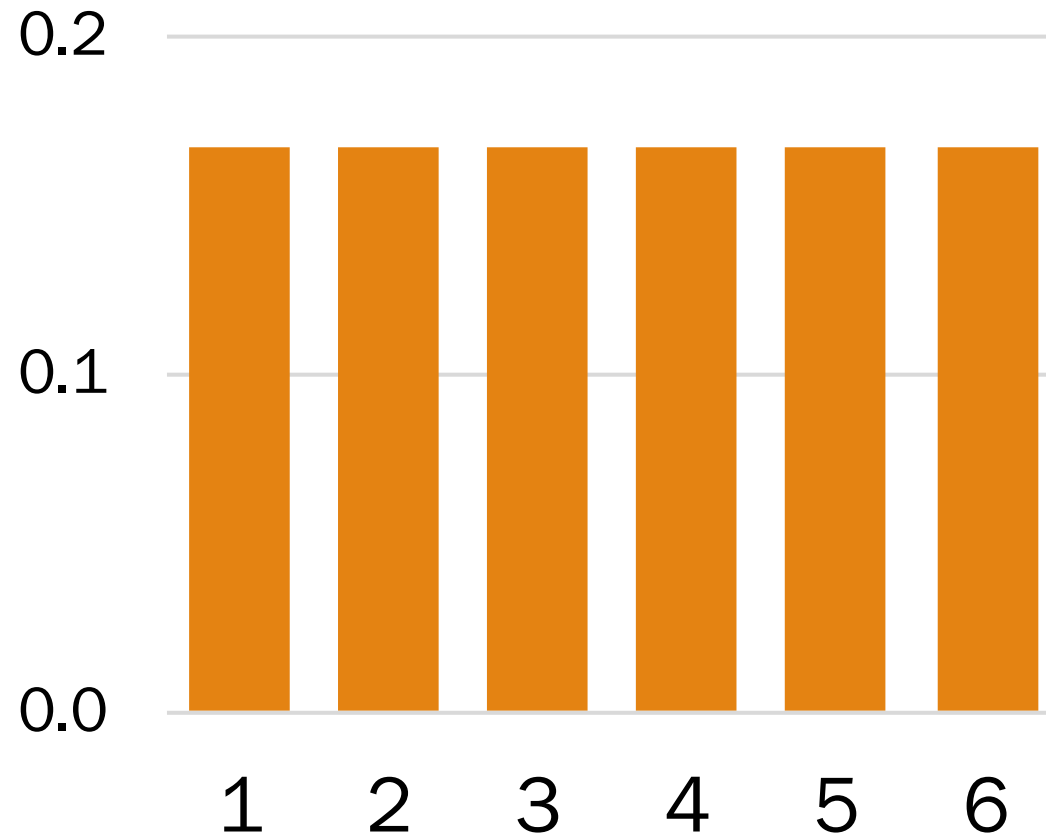
X_i s are iid



X_i is the outcome of dice roll i

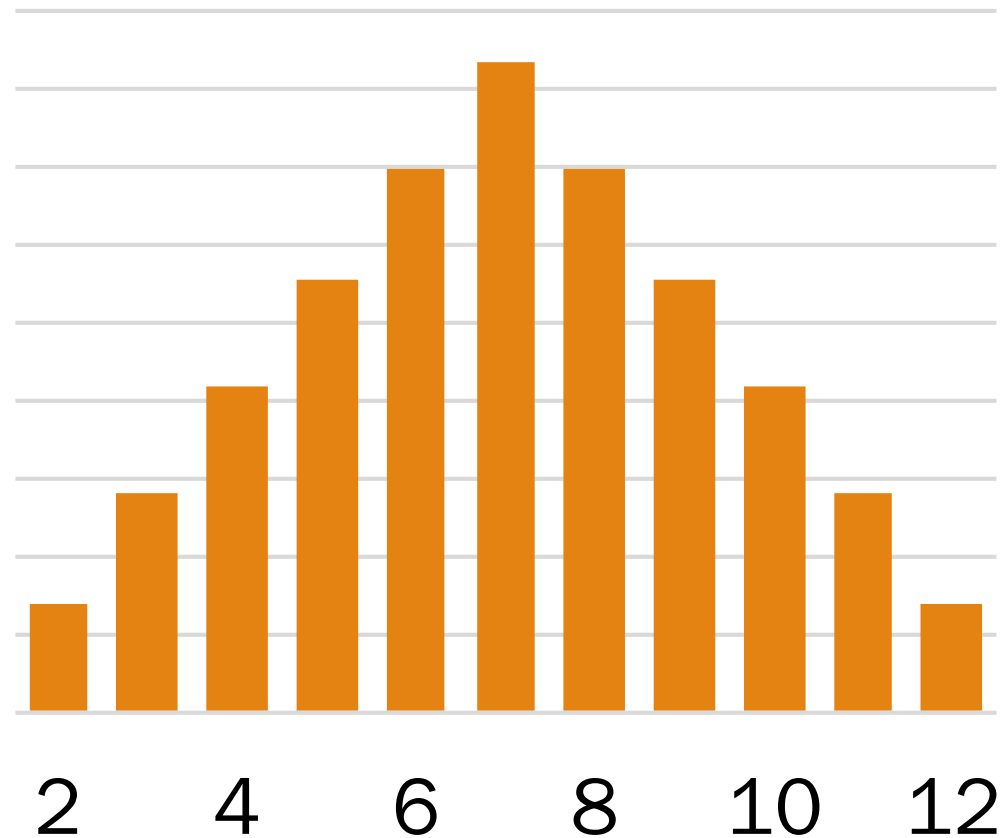
Sum of One Dice

This is the PMF of the sum of one dice



Sum of Two Dice

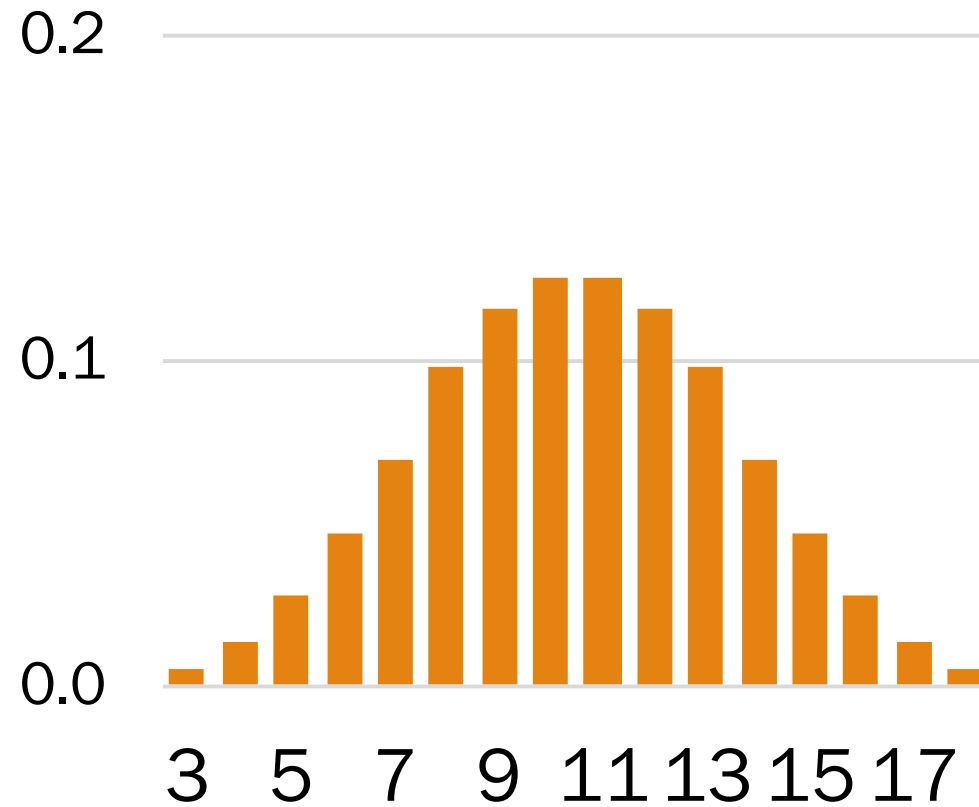
This is the PMF of the sum of two dice



Why is there more mass in the middle?

Sum of Three Dice

This is the PMF of the sum of three dice



Why is there more mass in the middle?

Sum of 50 dice?

The Insight to Convolution

Imagine a game
where each player *independently* scores between 0 and 100 points:

Let X be the amount of points you score.
Let Y be the amount of points your opponent scores.
Let's say you know $P(X = x)$ and $P(Y = y)$.

What is the probability of a tie?

$$\begin{aligned} P(\text{tie}) &= \sum_{i=0}^{100} P(X = i, Y = i) \\ &= \sum_{i=0}^{100} P(X = i)P(Y = i) \end{aligned}$$

The Insight to Convolution Proofs

What is the
probability that $X +$
 $Y = n$?

$$P(X + Y = n)?$$

$$P(X + Y = n) = \sum_{i=0}^n P(X = i, Y = n - i)$$

X	Y	i	
0	n	0	$P(X = 0, Y = n)$
1	$n - 1$	1	$P(X = 1, Y = n - 1)$
2	$n - 2$	2	$P(X = 2, Y = n - 2)$
	...		
n	0	n	$P(X = n, Y = 0)$

The Insight to Convolution Proofs

What is the probability that $X + Y = n$?

$$P(X + Y = n)?$$

$$P(X + Y = n) = \sum_{k=0}^n P(X = k, Y = n - k)$$

Since this is the OR or mutually exclusive events

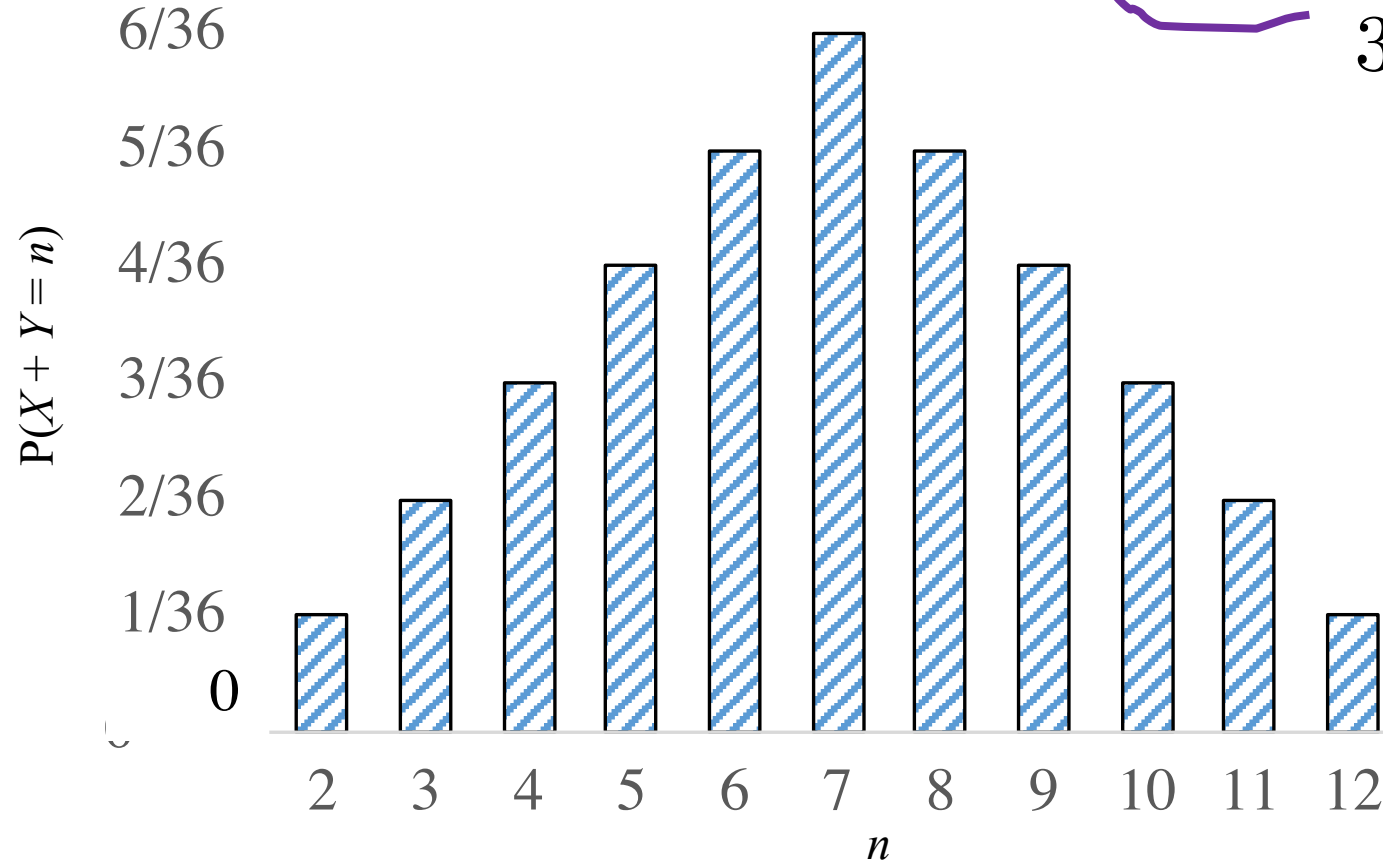
$$= \sum_{k=0}^n P(X = k)P(Y = n - k)$$

If the random variables are independent

Sum of Two Dice

Let $X+Y$ be the value of the sum of two dice
(aka two independent random variables)

$$P(X + Y = n) = \sum_{i=1}^{n-1} P(X = i, Y = n - i) \quad \frac{1}{36}$$



Convolution: The fanciest way to say
“adding random variables”

Sometimes Adding is Easy:

Sum of Independent Binomials

- Let X and Y be independent binomials with the same value for p :
 - $X \sim \text{Bin}(n_1, p)$ and $Y \sim \text{Bin}(n_2, p)$
 - $X + Y \sim \text{Bin}(n_1 + n_2, p)$
- Intuition:
 - X has n_1 trials and Y has n_2 trials
 - Each trial has same “success” probability p
 - Define Z to be $n_1 + n_2$ trials, each with success prob. p
 - $Z \sim \text{Bin}(n_1 + n_2, p)$, and also $Z = X + Y$

Sum of Independent Normals

- Let X and Y be independent random variables
 - $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$
 - $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- Generally, have n independent random variables $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, \dots, n$:

$$\left(\sum_{i=1}^n X_i \right) \sim N \left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2 \right)$$

Sum of Independent Poissons

- Let X and Y be independent random variables
 - $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$
 - $X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$

Virus Infections

- Say you are working with the WHO to plan a response to a the initial conditions of a virus:
 - Two exposed groups
 - P1: 50 people, each independently infected with $p = 0.1$
 - P2: 100 people, each independently infected with $p = 0.4$
 - Question: Probability of more than 40 infections?

Sanity check: Should we use the Binomial Sum-of-RVs shortcut?

- A. YES!
- B. NO!
- C. Other/none/more

Virus Infections

- Say you are working with the WHO to plan a response to a the initial conditions of a virus:
 - Two exposed groups
 - P1: 50 people, each independently infected with $p = 0.1$
 - P2: 100 people, each independently infected with $p = 0.4$
 - $A = \#$ infected in P1 $A \sim \text{Bin}(50, 0.1) \approx X \sim N(5, 4.5)$
 - $B = \#$ infected in P2 $B \sim \text{Bin}(100, 0.4) \approx Y \sim N(40, 24)$
 - What is $P(\geq 40 \text{ people infected})$?
 - $P(A + B \geq 40) \approx P(X + Y \geq 39.5)$
 - $X + Y = W \sim N(5 + 40 = 45, 4.5 + 24 = 28.5)$

$$P(W > 39.5) = 1 - P(X < 39.5)$$

$$= 1 - F_X(39.5) = 1 - \Phi\left(\frac{39.5 - 45}{\sqrt{28.5}}\right) \approx 0.8485$$

Linear Transform

Thinking of Y as a linear transform

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

$$Y = X + X = 2 \cdot X$$

$$Y \sim N(2\mu, 4\sigma^2)$$

$$Y = X + X = 2 \cdot X$$

Thinking of Y as the sum
of independent normals

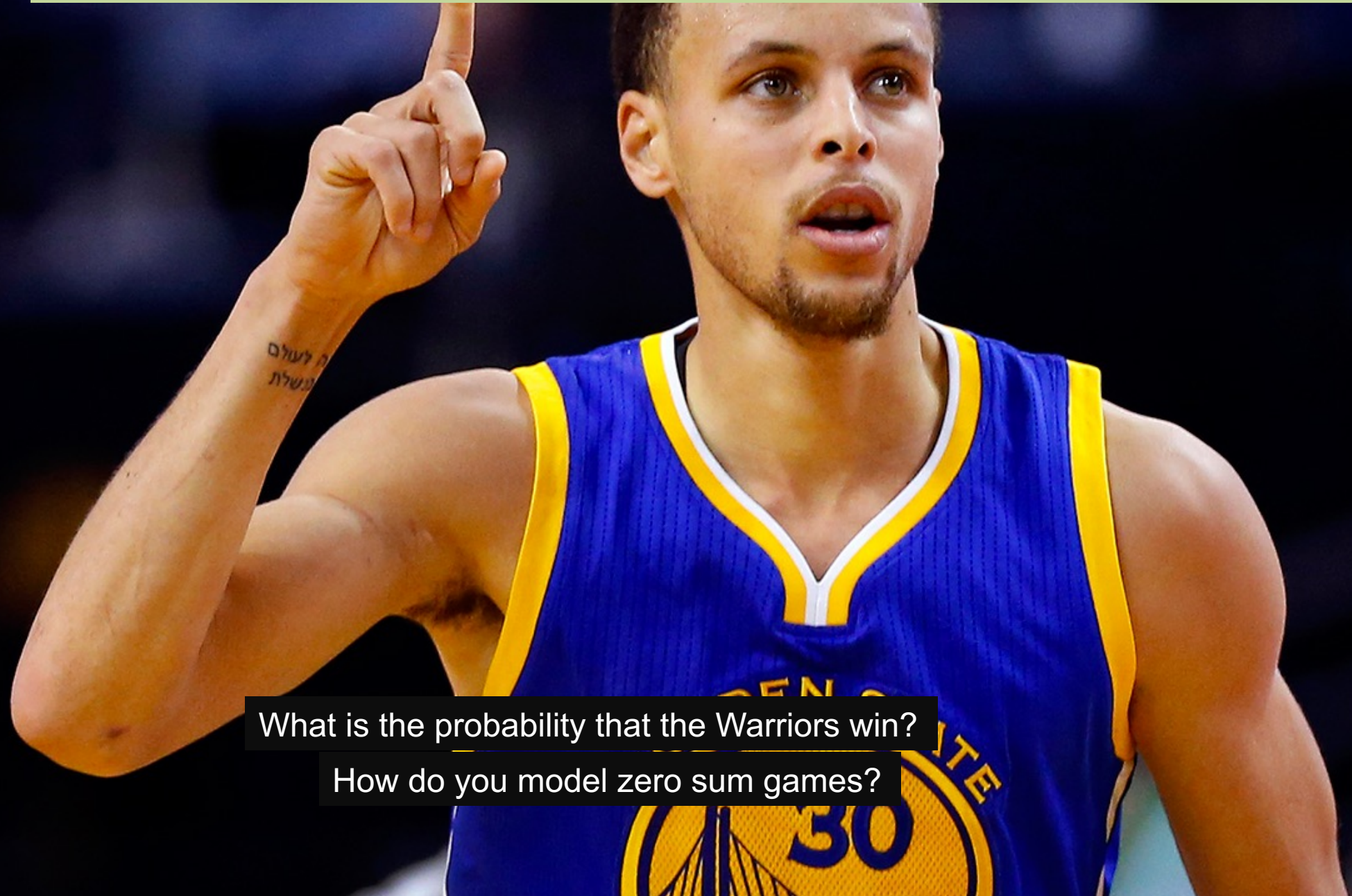
$$X + X \sim N(\mu + \mu, \sigma^2 + \sigma^2)$$

$$Y \sim N(2\mu, 2\sigma^2)$$



X is not independent of
X

Zero Sum Games



What is the probability that the Warriors win?

How do you model zero sum games?

Gaussian Sampling and ELO ratings

Basketball == Stats



What is the probability that the Warriors win?
How do you model zero-sum games?

Gaussian Sampling and ELO ratings

Each team has an ELO score S , calculated based on its past performance.

- Each game, a team has ability $A \sim \mathcal{N}(S, 200^2)$.
- The team with the higher sampled ability wins.

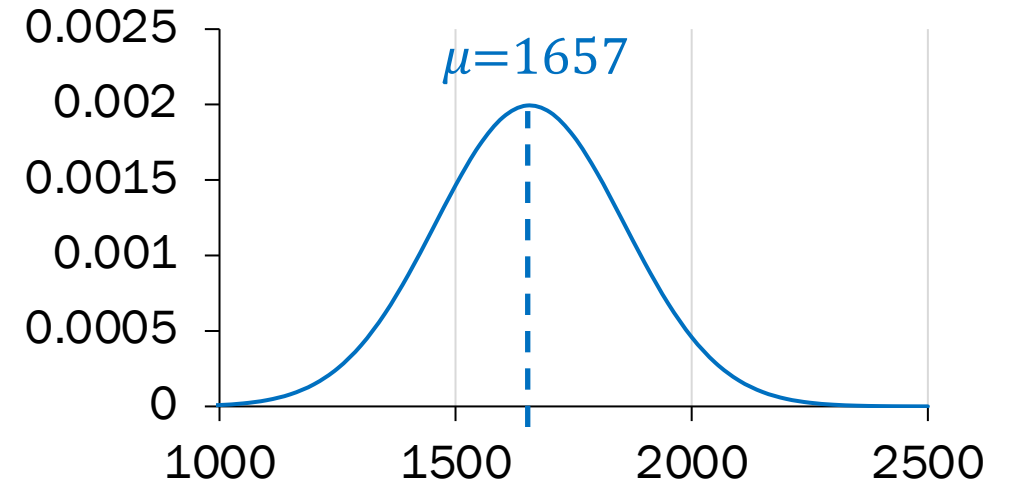


Arpad Elo

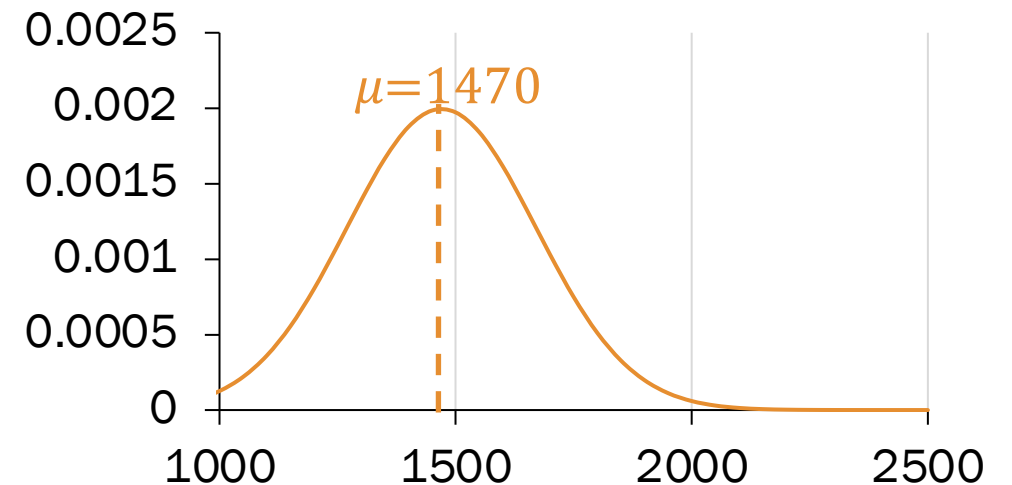
What is the probability that Warriors win this game?

Want: $P(\text{Warriors win}) = P(A_W > A_O)$

Warriors' $A_W \sim \mathcal{N}(S = 1657, 200^2)$



Opponent's $A_O \sim \mathcal{N}(S = 1470, 200^2)$



Probability of Winning a Game



$$A_W \sim N(1797, 200^2)$$

$$A_O \sim N(1555, 200^2)$$

$$P(\text{Warriors win}) = P(A_W > A_O)$$

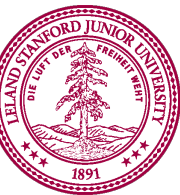
$$P(\text{Warriors win}) = P(A_W - A_O > 0)$$

$$-A_O \sim N(-1555, 200^2)$$

$$D = A_W + (-A_O)$$

$$D \sim N(242, 2 \cdot 200^2)$$

$$P(D > 0) = 1 - F_D(0) \approx 0.804$$





We talked about sum of Binomial, Normal and Poisson...who's missing from this party?

Uniform.

Discrete Vs Continuous

Discrete

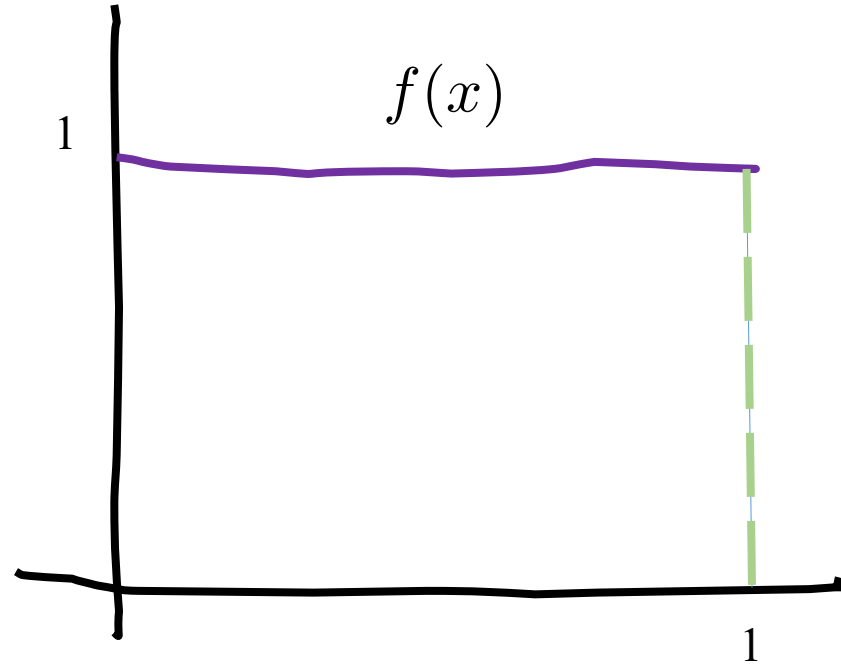
$$P(X + Y = a) = \sum_{y=-\infty}^{\infty} P(X = a - y)P(Y = y) dy$$

Continuous

$$f(X + Y = a) = \int_{y=-\infty}^{\infty} f(X = a - y)f(Y = y) dy$$

Sum of Independent Uniforms

- Let X and Y be independent random variables
 - $X \sim \text{Uni}(0, 1)$ and $Y \sim \text{Uni}(0, 1) \rightarrow f(x) = 1$ for $0 \leq x \leq 1$



For both X and Y

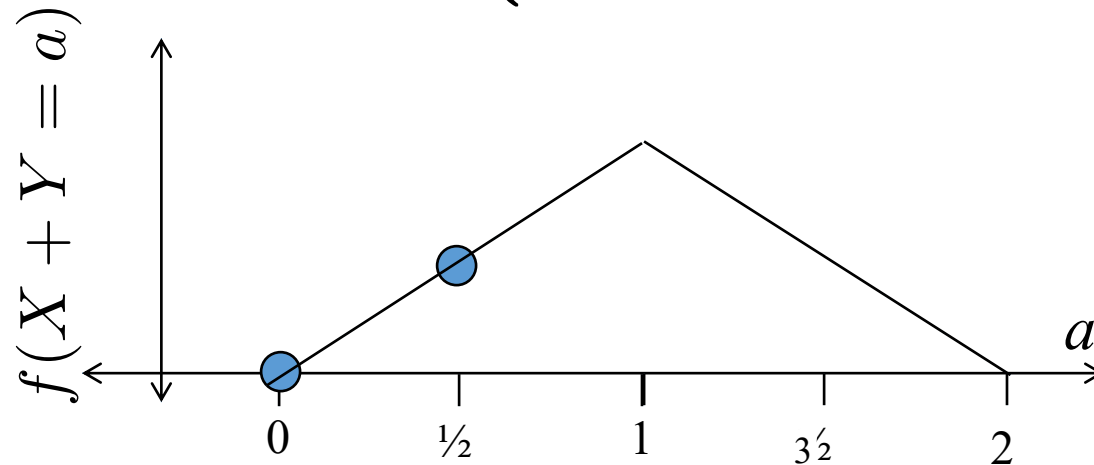
$$1 < a < 2$$

$X \sim \text{Uni}(0, 1)$ $Y \sim \text{Uni}(0, 1)$
 X and Y are independent

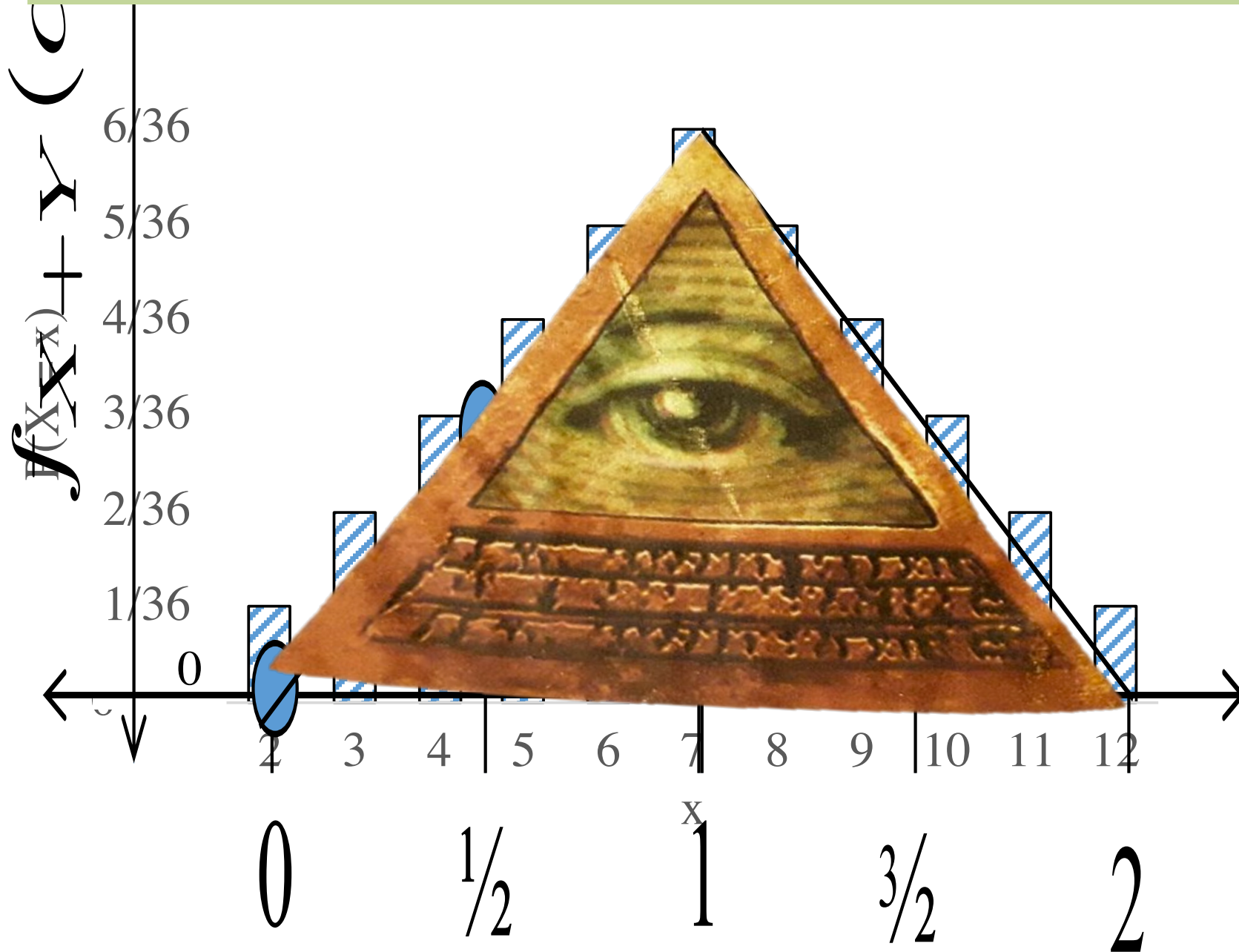
$$f(X + Y = a)?$$

$$f(X + Y = a) = \int_{y=-\infty}^{\infty} f(X = a - y) f(Y = y) dy$$

$$f(X + Y = a) = \begin{cases} a & 0 < a < 1 \\ 2 - a & 1 < a < 2 \\ 0 & \text{otherwise} \end{cases}$$



Sum of Uniforms and Sum of Dice



Sum of 100 uniforms???

Were talking about the sum of uniforms

```
sum.py x
1 import random
2
3 def main():
4     x = random.random()
5     y = random.random()
6     z = x + y
7     print(z)
8
9 if __name__ == '__main__':
10    main()
```

Sum of 100 poissons???

Silence!!



And now a moment of silence...

...before we present...

...a beautiful result of probability theory!



(silent drumroll)

Central Limit Theorem

Consider n **independent and identically distributed (i.i.d)** variables X_1, X_2, \dots, X_n with $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$.

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

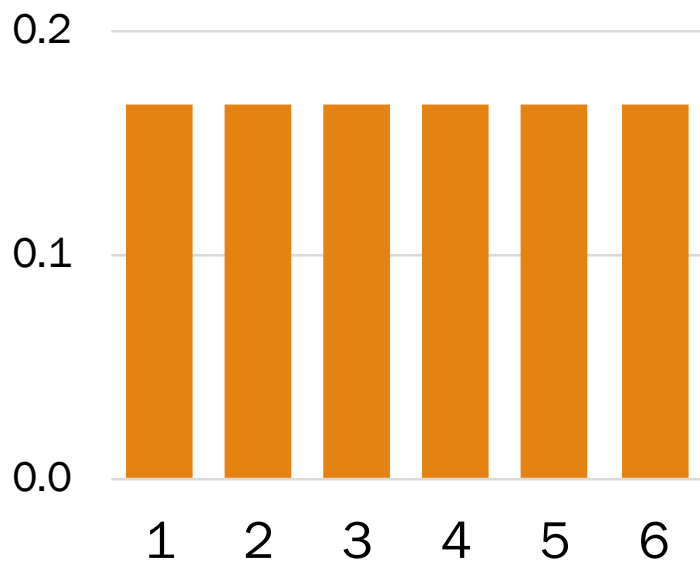
The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.

True happiness



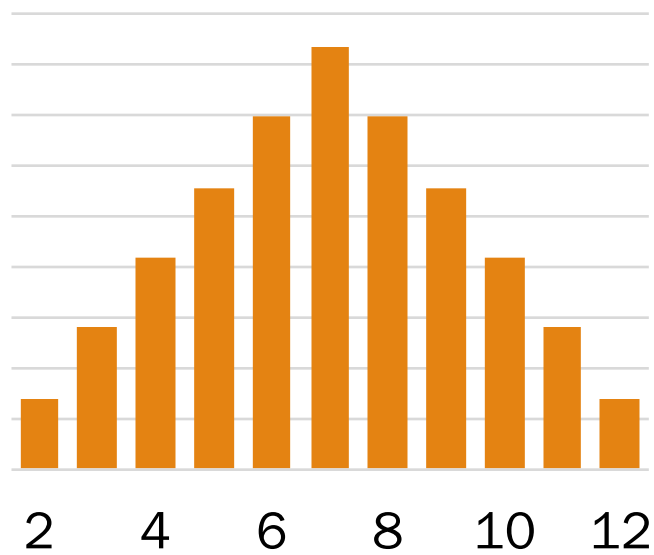
Sum of dice rolls

Roll n independent dice. Let X_i be the outcome of roll i . X_i are i.i.d.



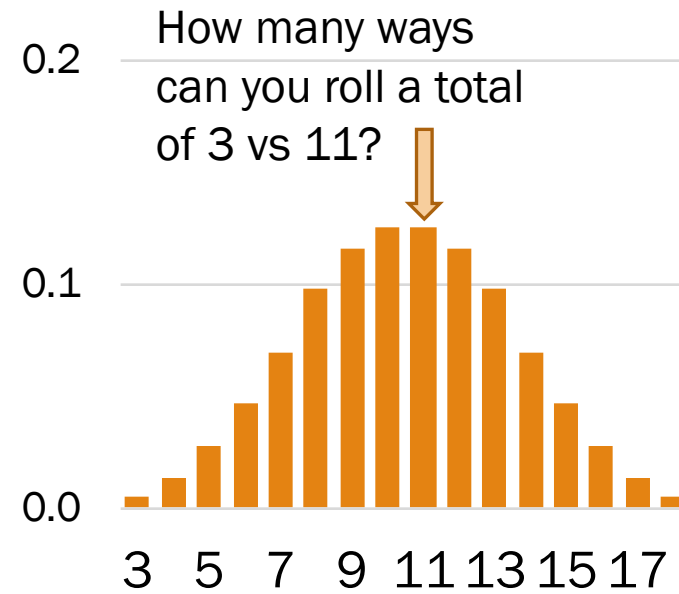
$$\sum_{i=1}^1 X_i$$

Sum of 1 die roll



$$\sum_{i=1}^2 X_i$$

Sum of 2 dice rolls



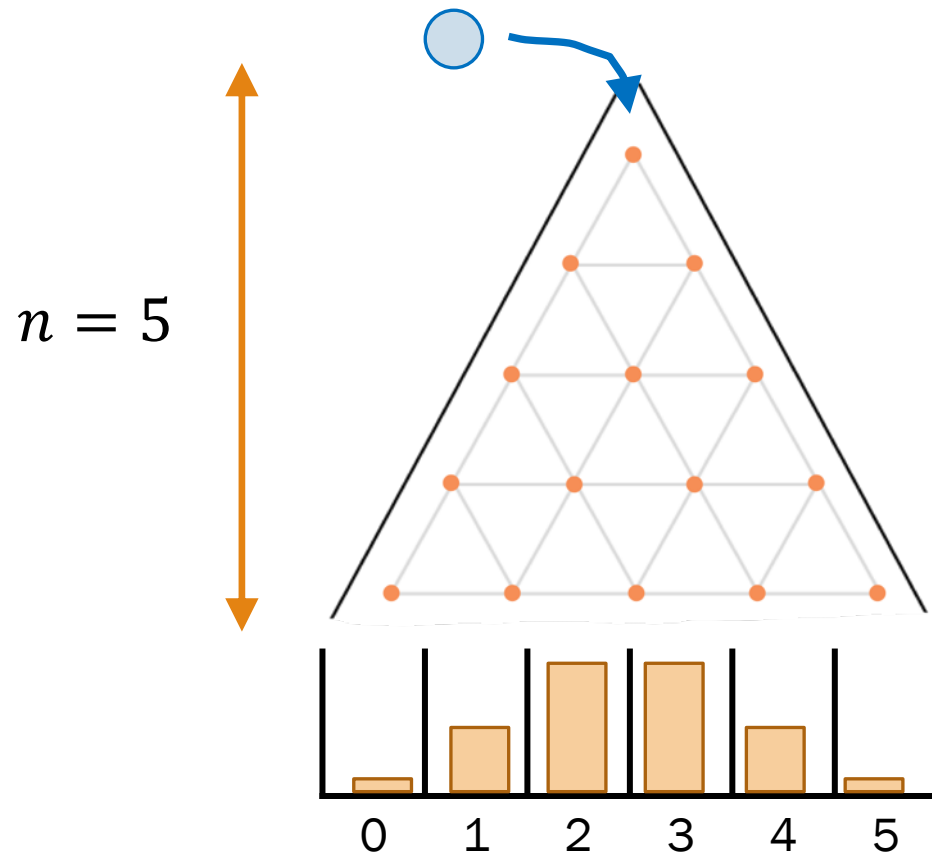
$$\sum_{i=1}^3 X_i$$

Sum of 3 dice rolls

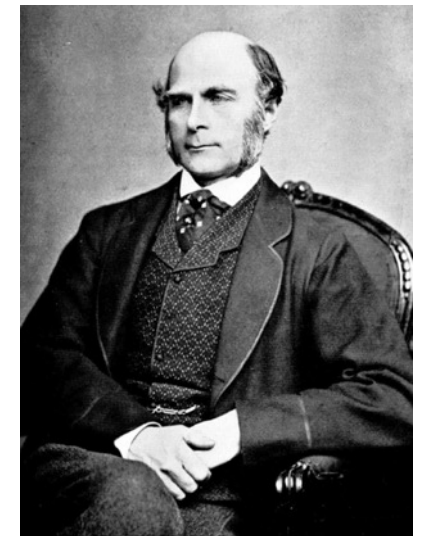
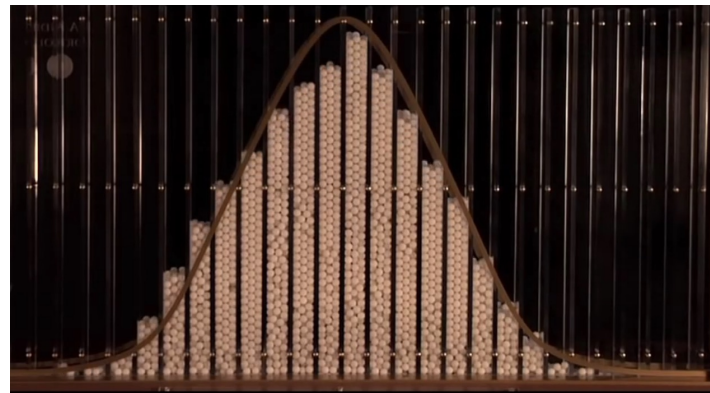
CLT explains a lot

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



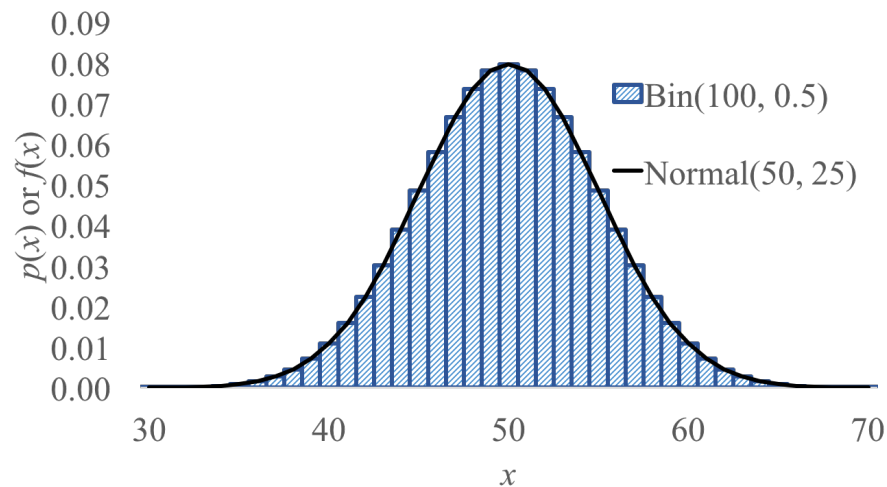
Galton Board, by Sir Francis Galton
(1822-1911)



CLT explains a lot

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



Proof:

Let $X_i \sim \text{Ber}(p)$ for $i = 1, \dots, n$, where X_i are i.i.d.
 $E[X_i] = p, \text{Var}(X_i) = p(1 - p)$

$$X = \sum_{i=1}^n X_i \quad (X \sim \text{Bin}(n, p))$$

$$X \sim \mathcal{N}(n\mu, n\sigma^2) \quad (\text{CLT, as } n \rightarrow \infty)$$

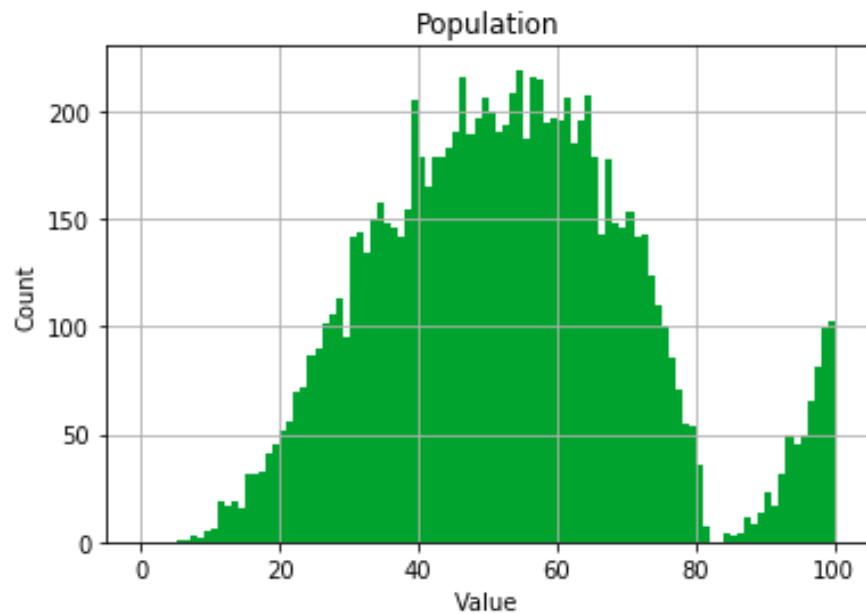
$$X \sim \mathcal{N}(np, np(1 - p)) \quad (\text{substitute mean, variance of Bernoulli})$$

Normal approximation of Binomial
Sum of i.i.d. Bernoulli RVs \approx Normal

CLT explains a lot

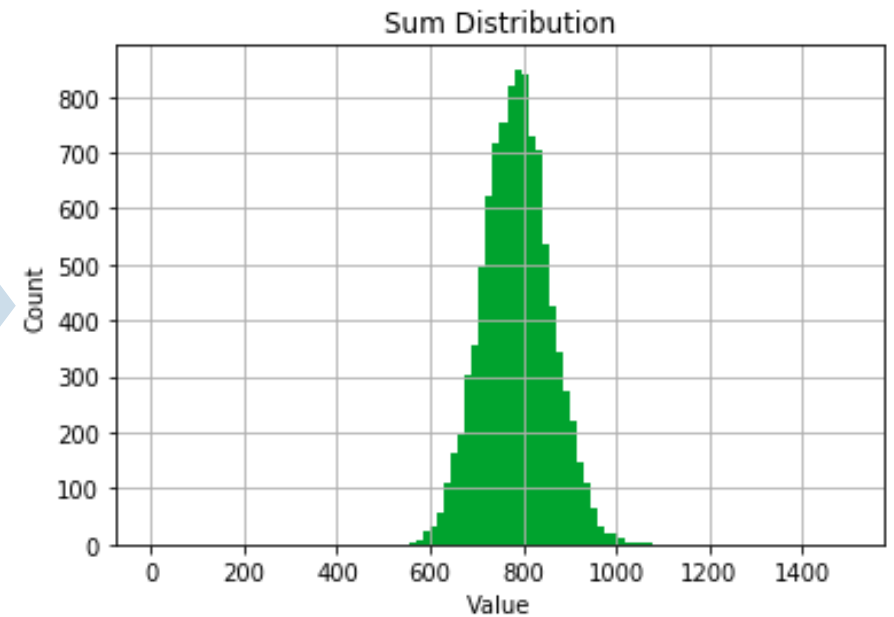
$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



Distribution of X_i

Sample of
size 15,
sum values

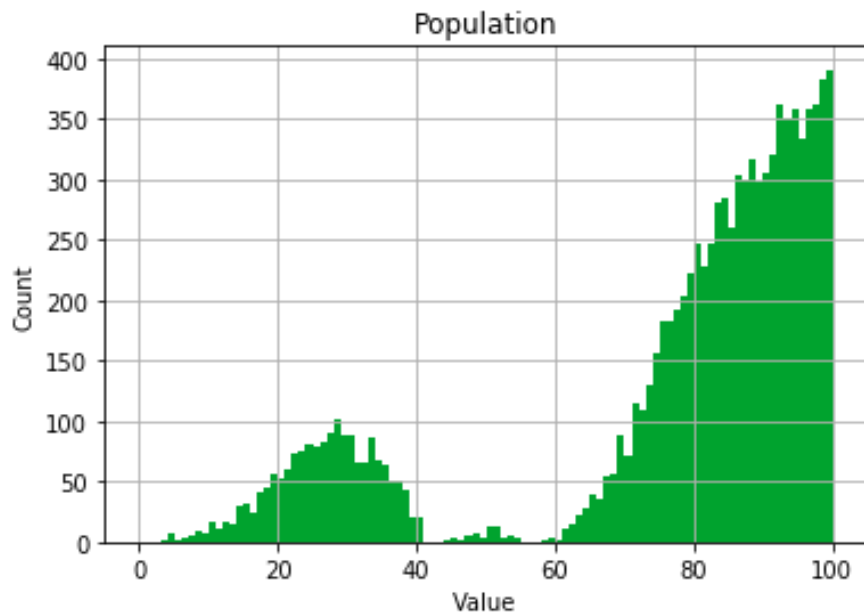


Distribution of $\sum_{i=1}^{15} X_i$

CLT explains a lot

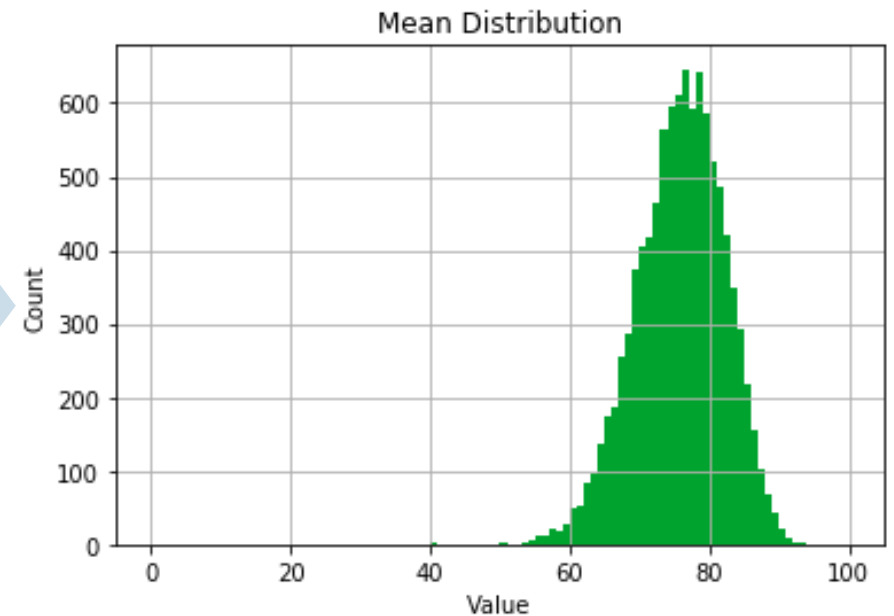
$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



Distribution of X_i

Sample of
size 15,
average values



Distribution of $\frac{1}{15} \sum_{i=1}^{15} X_i$

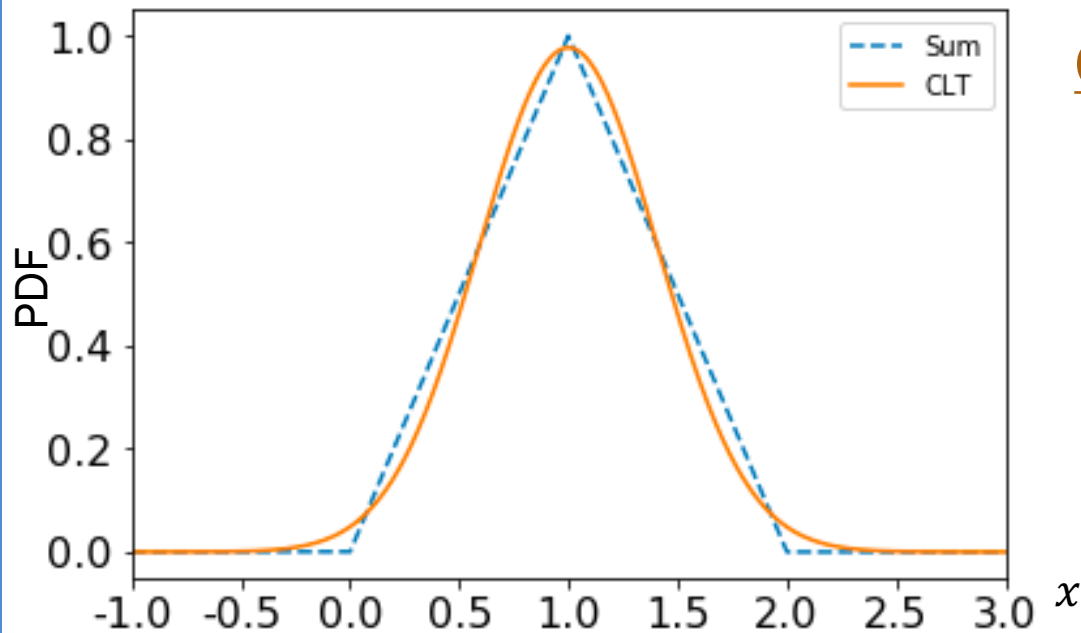
CLT example

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 2$:



Exact

$$P(X \leq 2/3) \approx 0.2222$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(1, 1/6)$$

$$P(X \leq 2/3) \approx P(Y \leq 2/3)$$

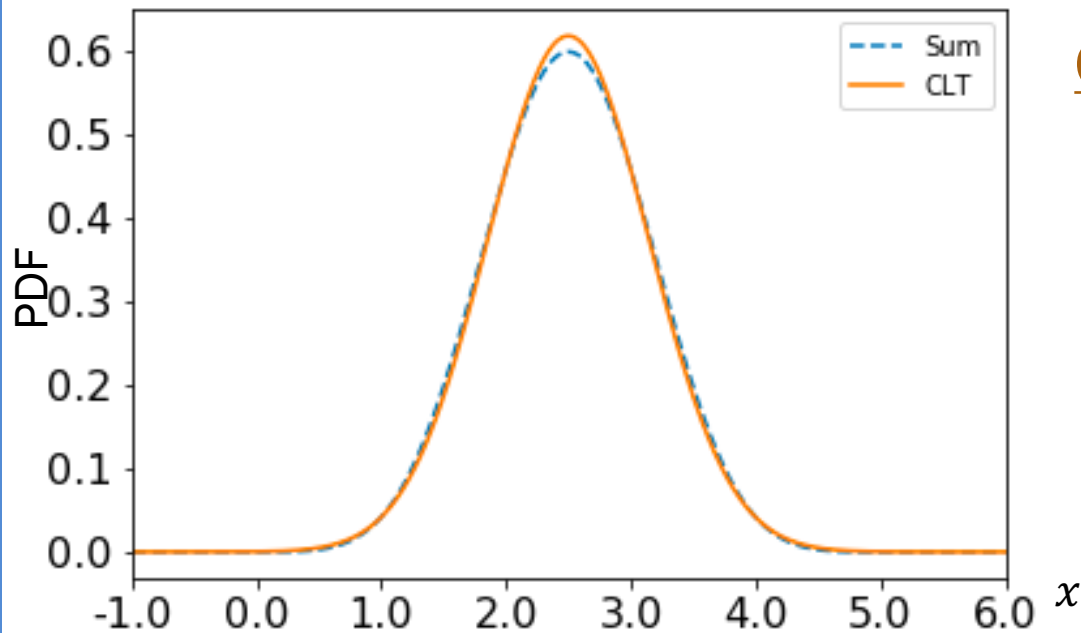
$$= \Phi\left(\frac{2/3 - 1}{\sqrt{1/6}}\right) \approx 0.2071$$

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 5$:



Exact

$$P(X \leq 5/3) \approx 0.1017$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(5/2, 5/12)$$

$$P(X \leq 5/3) \approx P(Y \leq 5/3)$$

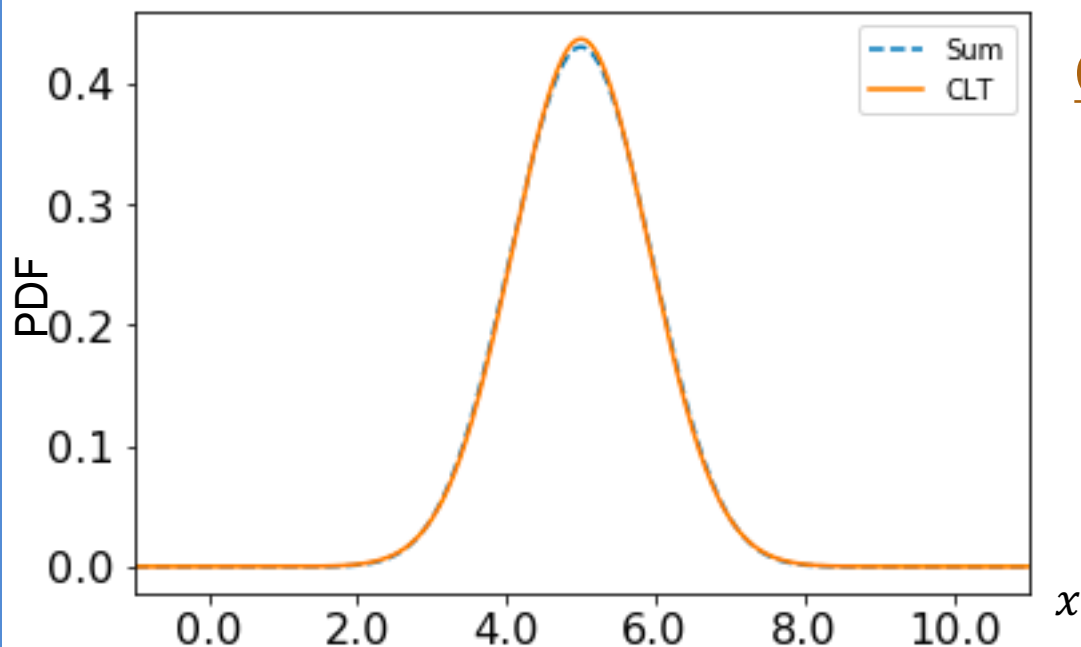
$$= \Phi\left(\frac{5/3 - 5/2}{\sqrt{5/12}}\right) \approx 0.0984$$

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 10$:



Exact

$$P(X \leq 10/3) \approx 0.0337$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(5, 5/6)$$

$$P(X \leq 10/3) \approx P(Y \leq 10/3)$$

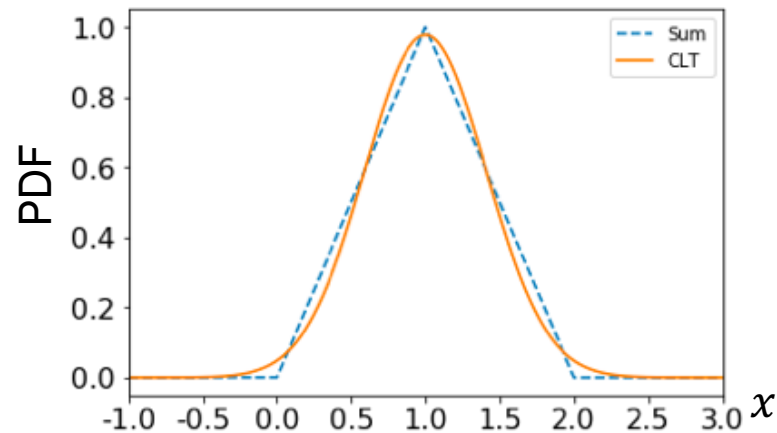
$$= \Phi\left(\frac{10/3 - 5}{\sqrt{5/6}}\right) \approx 0.0339$$

Sum of n independent Uniform RVs

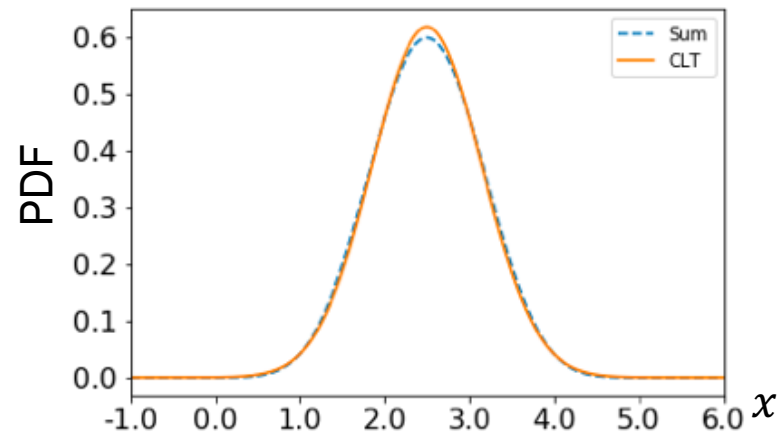
Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

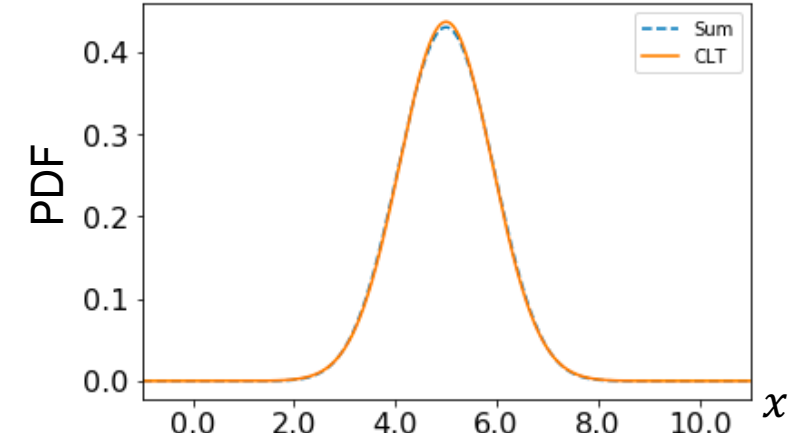
$n = 2$:



$n = 5$:



$n = 10$:



Most books will tell you that CLT holds if $n \geq 30$, but it can hold for smaller n depending on the distribution of your i.i.d. X_i 's.

The sum of independent, identically distributed variables:

$$Y = \sum_{i=0}^n X_i$$

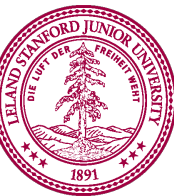


Is normally distributed:

$$Y \sim N(n\mu, n\sigma^2)$$

where $\mu = E[X_i]$

$$\sigma^2 = \text{Var}(X_i)$$

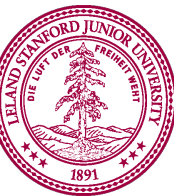


What about other functions?

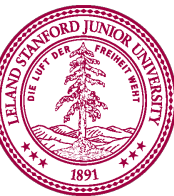
Sum of iid? Normal

Average of iid?

Max of iid?



It's play time!



Sum of Dice

- You will roll 10 6-sided dice (X_1, X_2, \dots, X_{10})
 - X = total value of all 10 dice = $X_1 + X_2 + \dots + X_{10}$
 - Win if: $X \leq 25$ or $X \geq 45$
 - Roll!
- And now the truth (according to the CLT)...



Sum of Dice

- You will roll 10 6-sided dice (X_1, X_2, \dots, X_{10})
 - X = total value of all 10 dice = $X_1 + X_2 + \dots + X_{10}$
 - Win if: $X \leq 25$ or $X \geq 45$
-

- Recall CLT: $X = \sum_i^n X_i \rightarrow N(n\mu, n\sigma^2)$ As $n \rightarrow \infty$
 - Determine $P(X \leq 25 \text{ or } X \geq 45)$ using CLT:

$$\mu = E[X_i] = 3.5 \qquad \sigma^2 = \text{Var}(X_i) = \frac{35}{12} \qquad X \approx N(35, 29.2)$$

$$1 - P(25.5 < X < 44.5) = 1 - P\left(\frac{25.5 - 35}{\sqrt{29.2}} < Z < \frac{44.5 - 35}{\sqrt{29.2}}\right)$$

$$\approx 1 - (2\Phi(1.76) - 1) \approx 2(1 - 0.9608) = 0.0784$$

Wonderful Form of Cosmic Order

I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "[Central limit theorem]". The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement, amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along.

- Sir Francis Galton

