

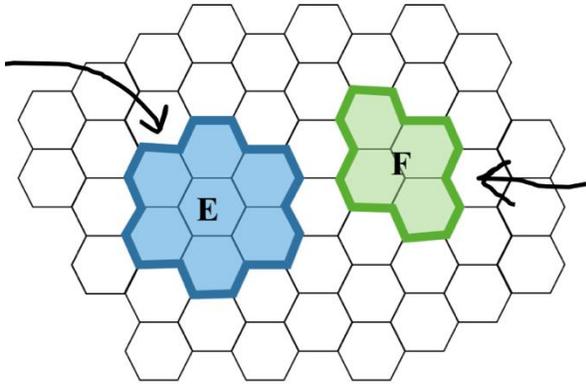
**Binomial**  
CS109

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



Review

# Review: Last Week



Mutually Exclusive Events

make **OR** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$



Independent Events

make **AND** easy:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

# Summary of Combinatorics

Counting tasks on  $n$  objects

Sort objects  
(permutations)

Distinct  
(distinguishable)

$$n!$$

Some  
distinct

$$\frac{n!}{n_1! n_2! \cdots n_r!}$$

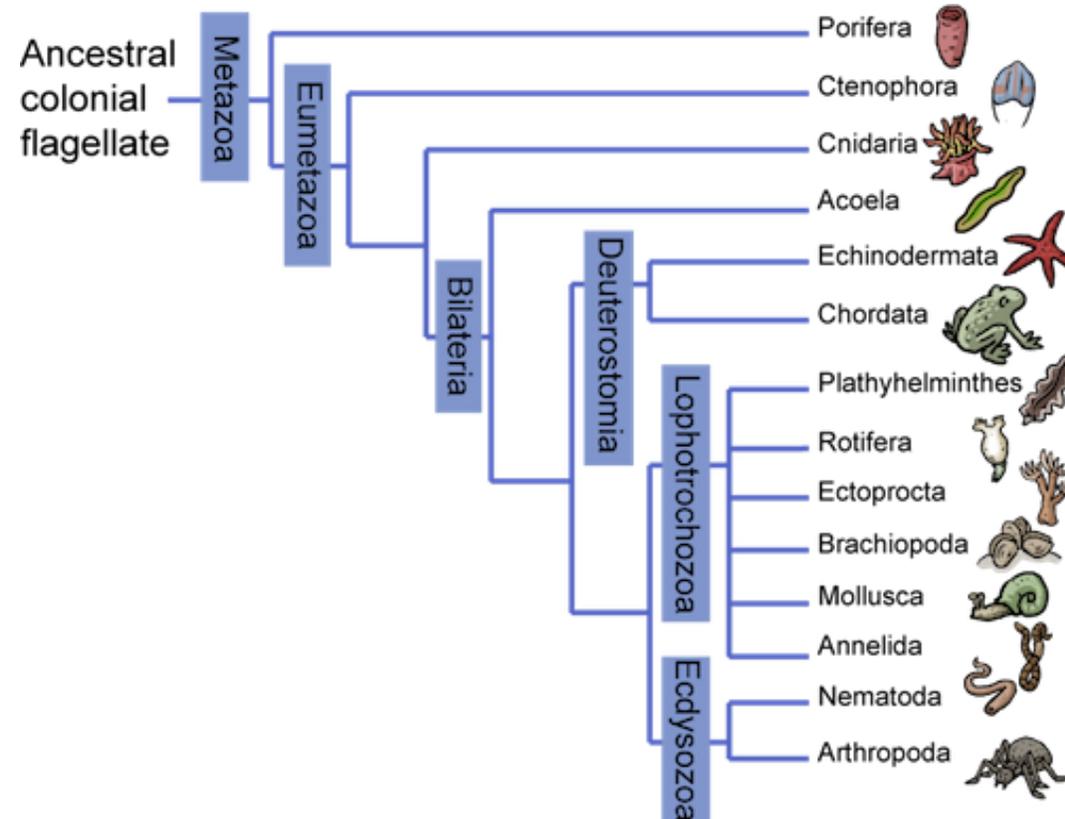
Choose  $k$  objects  
(combinations)

Distinct

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

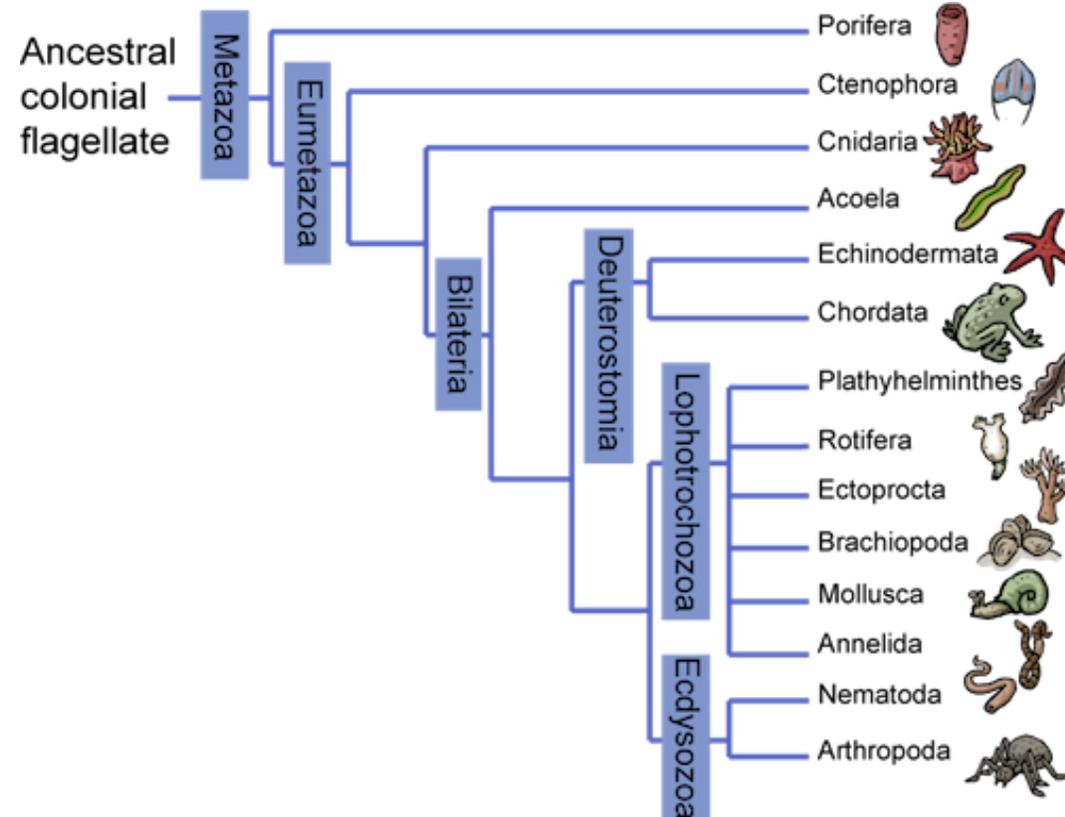
# Counting Review

For a DNA tree we need to calculate the DNA distance between each pair of animals. How many calculations are needed?

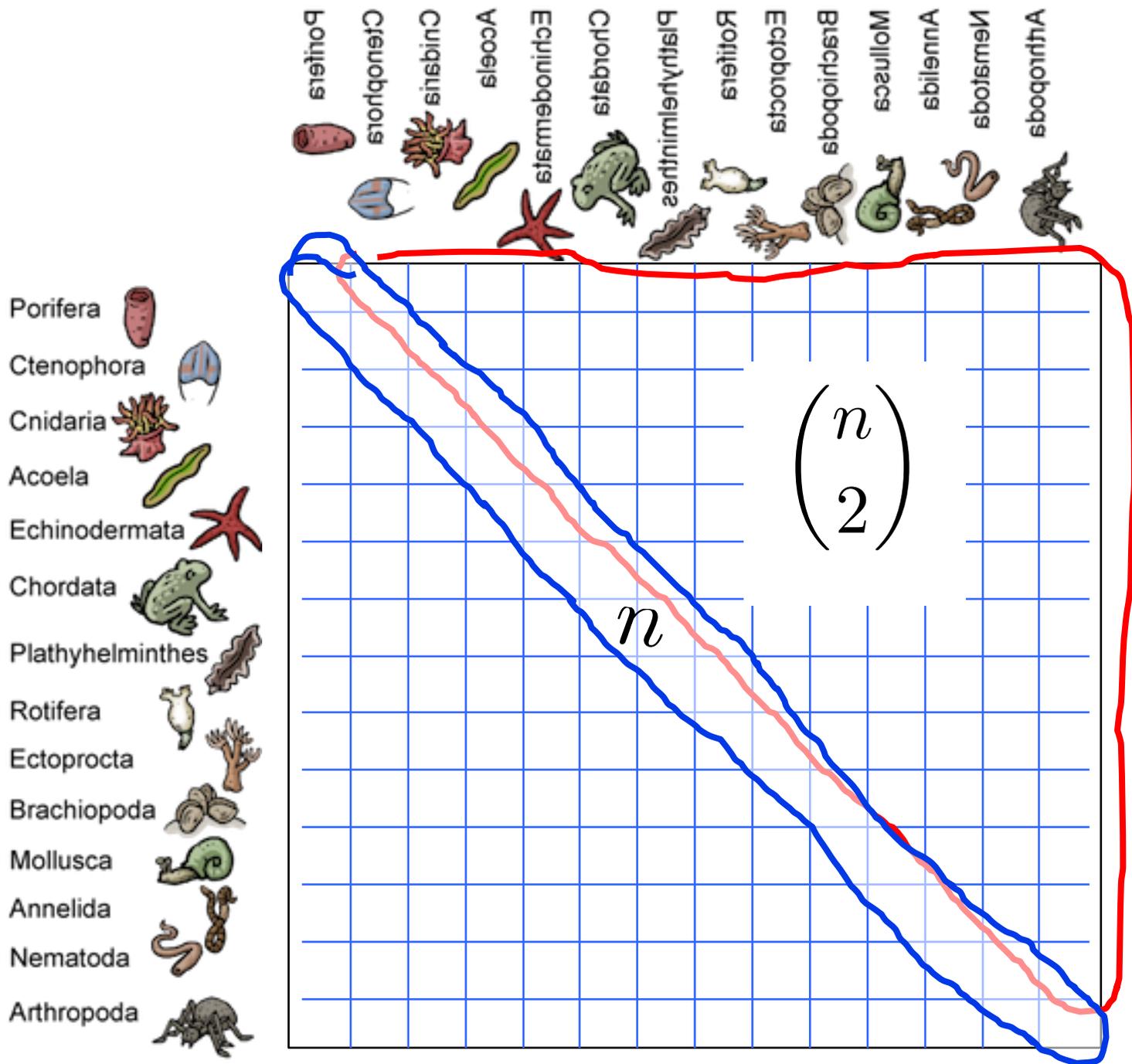


# Counting Review

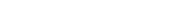
Q: There are  $n$  animals.  
How many **distinct pairs** of animals are there?

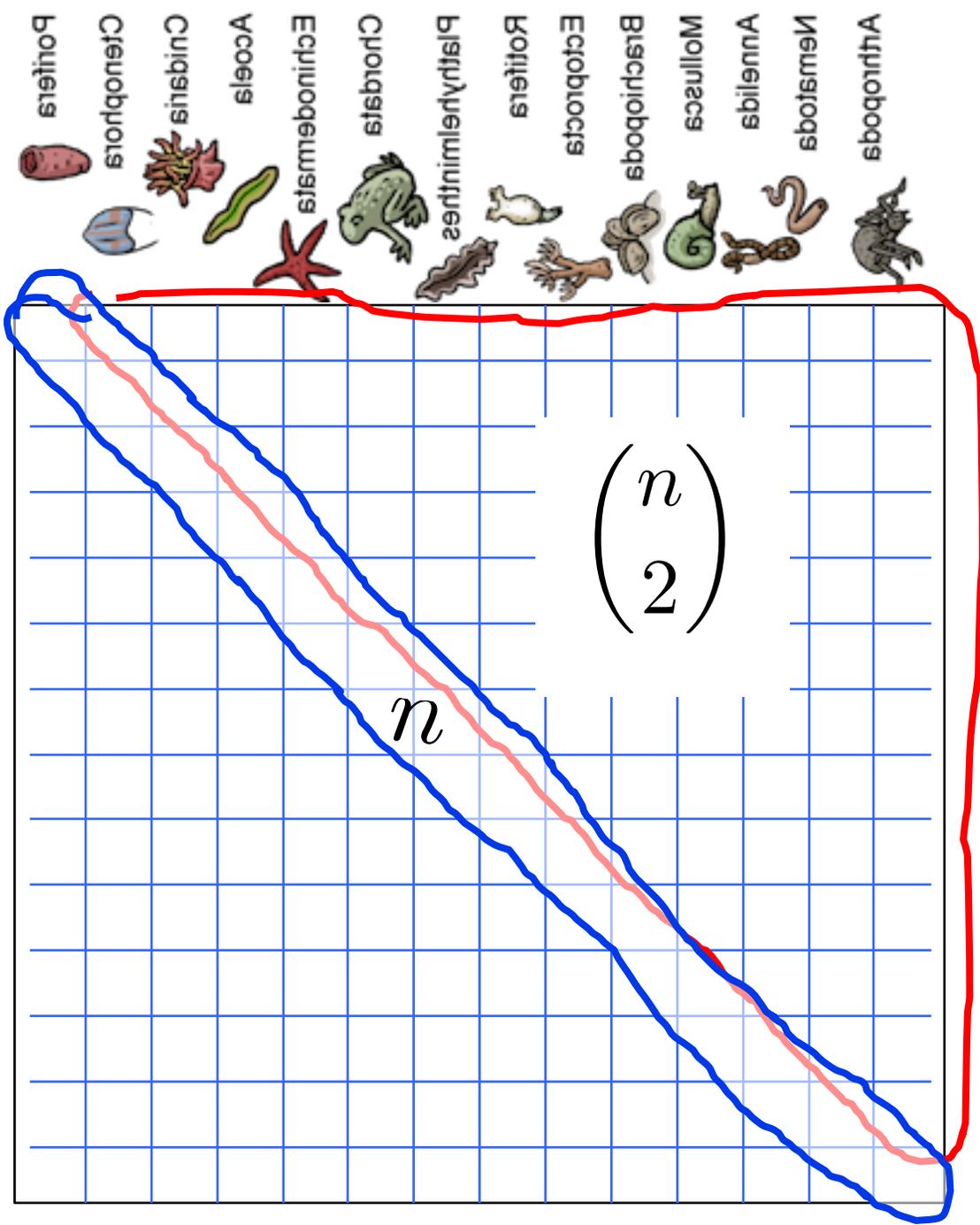


There are  $n$  animals.  
How many **distinct pairs** of  
animals are there?



There are  $n$  animals.  
 How many **distinct pairs** of animals are there?

- Porifera 
- Ctenophora 
- Cnidaria 
- Acoela 
- Echinodermata 
- Chordata 
- Platyhelminthes 
- Rotifera 
- Ectoprocta 
- Brachiopoda 
- Mollusca 
- Annelida 
- Nematoda 
- Arthropoda 



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

$$= \frac{(n)(n-1)}{2}$$

$$= \frac{n^2 - n}{2}$$

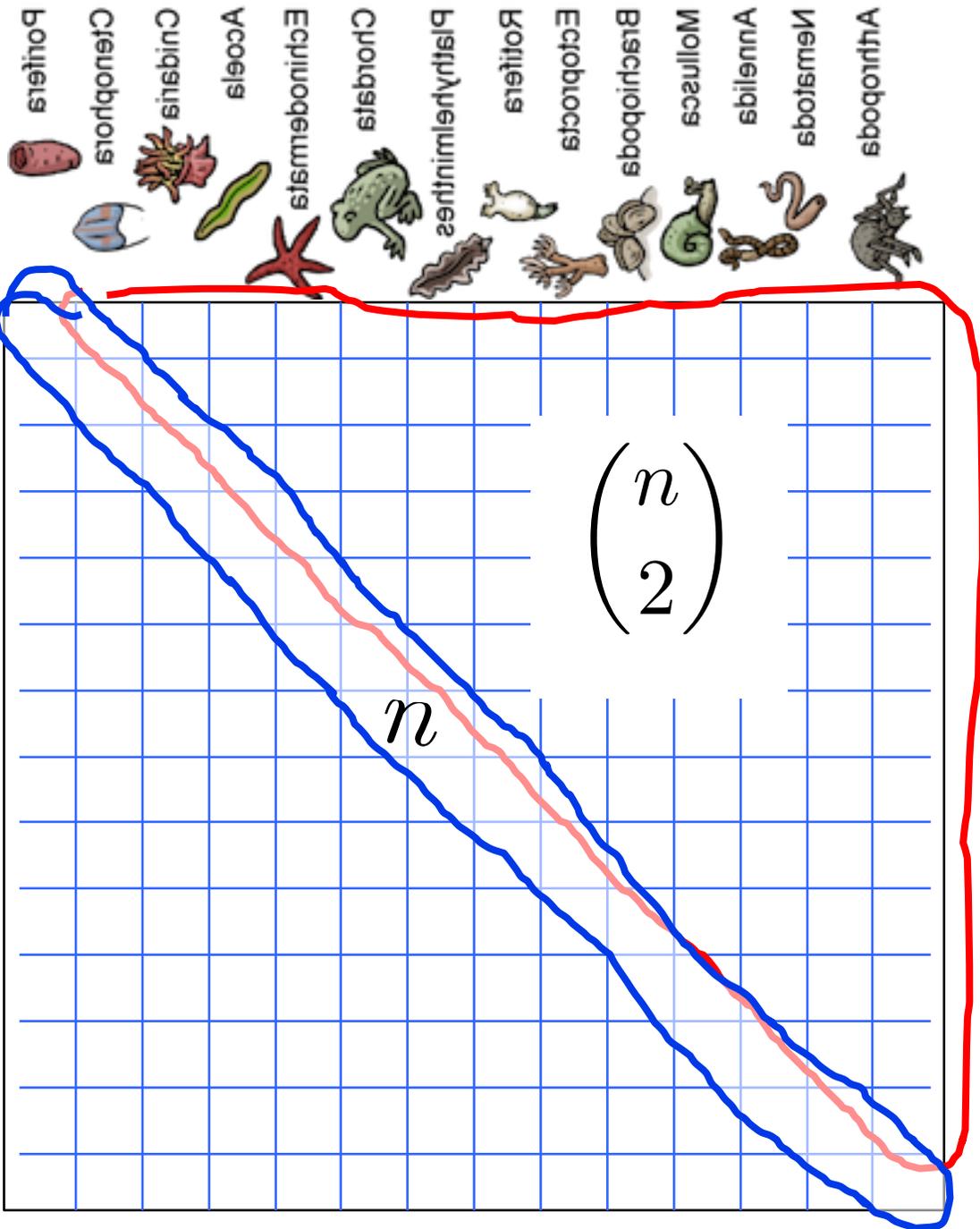
There are  $n$  animals.  
 How many **distinct pairs** of animals are there?

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

$$= \frac{(n)(n-1)}{2}$$

$$= \frac{n^2 - n}{2}$$

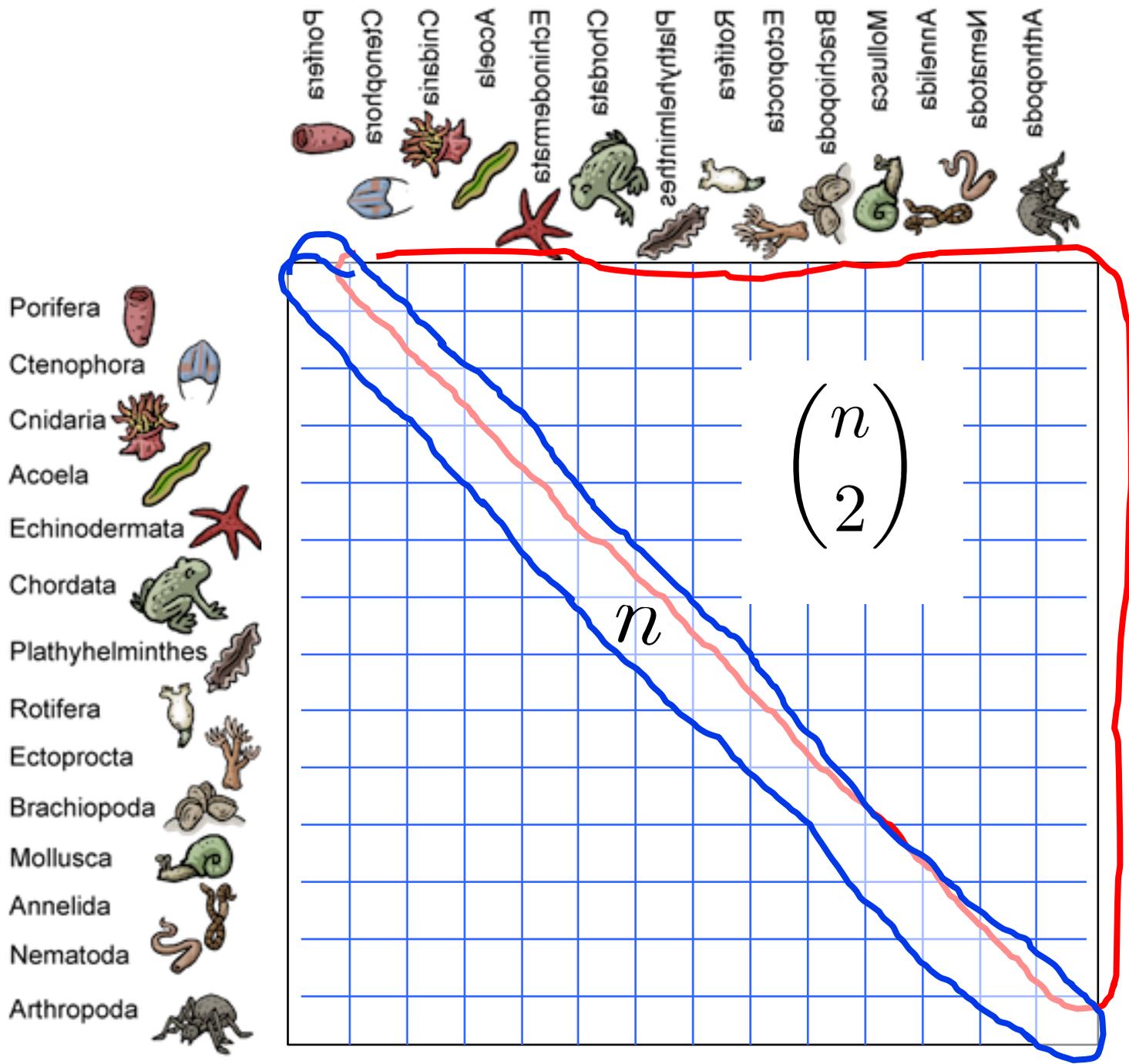
Same as the number of pairs in the upper right triangle here



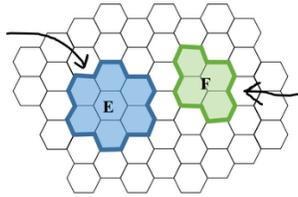
- Porifera
- Ctenophora
- Cnidaria
- Acoela
- Echinodermata
- Chordata
- Platyhelminthes
- Rotifera
- Ectoprocta
- Brachiopoda
- Mollusca
- Annelida
- Nematoda
- Arthropoda

- Porifera
- Ctenophora
- Cnidaria
- Acoela
- Echinodermata
- Chordata
- Platyhelminthes
- Rotifera
- Ectoprocta
- Brachiopoda
- Mollusca
- Annelida
- Nematoda
- Arthropoda

There are  $n$  animals.  
How many **distinct pairs** of  
animals are there?



# Are you ready to put them together??



Mutually Exclusive Events

make **OR** easy:

$$P(A \text{ or } B) = P(A) + P(B)$$



Independent Events

make **AND** easy:

$$P(A \text{ and } B) = P(A) \cdot P(B)$$



Choose  $k$  objects  
(combinations)

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

End Review

# Learning Goals for Today



A **random variable** is a number which takes on values probabilistically.



A discrete random variable is fully described by a **probability mass function**.



A **binomial** is a particular random variable which represents number of heads in  $n$  coin flips.

Super Critical Problem!

# To the Course Reader!!

The screenshot shows a web browser window with the URL `probabilitycoders.stanford.edu/win26/many_flips`. The page title is "Many Coin Flips".

**Probability for Computer Science**  
Winter 2026  
Stanford University

Reference  
Notation Reference  
Core Probability Reference  
Random Variable Reference  
Calculators  
Python Reference  
Calculus Reference  
Language Model Tool

Part 1: Core Probability  
Probability  
Equally Likely Outcomes  
Axioms of Probability  
Probability of or  
Conditional Probability  
Law of Total Probability  
Monty Hall Example  
Bayes' Theorem  
Independence  
Probability of and  
De Morgan's Law  
Router Example  
Log Probabilities  
Many Coin Flips  
Counting  
Combinatorics  
Stories  
Bacteria Evolution  
Google Rain Prediction  
Random Walks  
Binomial with Different Probs  
Netflix Genres  
Poker  
Beam Search  
Serendipity  
Core Probability Practice

## Many Coin Flips

In this section we are going to consider the number of heads on  $n$  coin flips. This thought experiment is going to be a basis for much probability theory! It goes far beyond coin flips.

Say a coin comes up heads with probability  $p$ . Most coins are fair and as such come up heads with probability  $p = 0.5$ . There are many events for which coin flips are a great analogy that have different values of  $p$  so let's leave  $p$  as a variable. You can try simulating flipping coins here. Note that **H** is short for Heads and **T** is short for Tails. We think of each coin as distinct:

**Coin Flip Simulator**

Number of flips  $n$ :       Probability of heads  $p$ :      

Simulator results:  
**T, T, H, T, H, H, H, H, H, H**

Total number of heads: 7  
 $P(\text{exactly 7 heads in 10 flips}) = 0.215$

Using the math in this chapter we will be able to calculate the probability of different numbers of heads. For example, if you flip  $n = 10$  coins which have a  $p = 0.6$  probability of landing heads, the probability of getting exactly 7 heads is 0.215. This chapter is organized into the following sections:

- [Warmups](#): We calculate the probability of a few exact outcomes.
- [Exactly  \$k\$  heads](#): We derive the general formula.
- [More than  \$k\$  heads](#): We explore this interesting related problem

## Warmups

In all of these warmups we are going to consider the probability of different outcomes when you flip a coin  $n$  times and each time the probability of heads is  $p$ . In each solution we will consider the case where  $n = 10$  and  $p = 0.6$ .

**What is the probability that all  $n$  flips are heads?**

This question is asking what is the probability of getting the outcome:  
**H, H, H, H, H, H, H, H, H, H**

Where each flip lands in heads (H). Each coin flip is independent so we can use the rule for probability of **and** with

# Probability of Exactly $k$ Heads in $n$ Coin Flips

We flip the coin 10 times. Probability of heads is  $p$ .

We want to know the probability of getting 4 heads.

What is the probability of the outcome below?

(H, H, H, H, T, T, T, T, T, T)

# Probability of Exactly $k$ Heads in $n$ Coin Flips

We flip the coin 10 times. Probability of heads is  $p$ .

We want to know the probability of getting 4 heads.

What is the probability of the outcome below?

(H, H, H, H, T, T, T, T, T, T)

$$p^4(1 - p)^6$$

# Probability of Exactly $k$ Heads in $n$ Coin Flips

We flip the coin 10 times. Probability of heads is  $p$ .

We want to know the probability of getting 4 heads.

What is the probability of the outcome below?

(H, H, H, H, T, T, T, T, T, T)

(H, H, H, T, H, T, T, T, T, T)

$$p^4(1 - p)^6$$

# Probability of Exactly $k$ Heads in $n$ Coin Flips

We flip the coin 10 times. Probability of heads is  $p$ .

We want to know the probability of getting 4 heads.

What is the probability of the outcome below?

(H, H, H, H, T, T, T, T, T, T)

$$p^4(1 - p)^6$$

(H, H, H, T, H, T, T, T, T, T)

$$p^4(1 - p)^6$$

All of the outcomes with exactly 4 heads have the same probability

# Probability of Exactly $k$ Heads in $n$ Coin Flips

(H, H, H, H, T, T, T, T, T, T)  
(H, H, H, T, H, T, T, T, T, T)  
(H, H, H, T, T, H, T, T, T, T)  
(H, H, H, T, T, T, H, T, T, T)  
(H, H, H, T, T, T, T, H, T, T)  
(H, H, H, T, T, T, T, T, H, T)  
(H, H, H, T, T, T, T, T, T, H)  
(H, H, T, H, H, T, T, T, T, T)  
(H, H, T, H, T, H, T, T, T, T)  
(H, H, T, H, T, T, H, T, T, T)  
(H, H, T, H, T, T, T, H, T, T)  
(H, H, T, H, T, T, T, T, H, T)  
(H, H, T, H, T, T, T, T, T, H)  
(H, H, T, T, H, H, T, T, T, T)  
(H, H, T, T, H, T, H, T, T, T)  
(H, H, T, T, H, T, T, H, T, T)  
(H, H, T, T, H, T, T, T, H, T)  
(H, H, T, T, H, T, T, T, T, H)



The probability of getting **4** heads, *in any ordering*, is the “**or**” of all these **mutually exclusive** cases

Each outcome has probability  $p^4(1-p)^6$

How many cases are there?

# Probability of Exactly $k$ Heads in $n$ Coin Flips

(H, H, H, H, T, T, T, T, T, T)  
(H, H, H, T, H, T, T, T, T, T)  
(H, H, H, T, T, H, T, T, T, T)  
(H, H, H, T, T, T, H, T, T, T)  
(H, H, H, T, T, T, T, H, T, T)  
(H, H, H, T, T, T, T, T, H, T)  
(H, H, H, T, T, T, T, T, T, H)  
(H, H, T, H, H, T, T, T, T, T)  
(H, H, T, H, T, H, T, T, T, T)  
(H, H, T, H, T, T, H, T, T, T)  
(H, H, T, H, T, T, T, H, T, T)  
(H, H, T, H, T, T, T, T, H, T)  
(H, H, T, H, T, T, T, T, T, H)  
(H, H, T, T, H, H, T, T, T, T)  
(H, H, T, T, H, T, H, T, T, T)  
(H, H, T, T, H, T, T, H, T, T)  
(H, H, T, T, H, T, T, T, H, T)  
(H, H, T, T, H, T, T, T, T, H)

← The probability of getting **4** heads, *in any ordering*, is the “**or**” of all these **mutually exclusive** cases

Each outcome has probability  $p^4(1-p)^6$

How many cases are there?  $\binom{10}{4}$

# Probability of Exactly $k$ Heads in $n$ Coin Flips

(H, H, H, H, T, T, T, T, T, T)  
(H, H, H, T, H, T, T, T, T, T)  
(H, H, H, T, T, H, T, T, T, T)  
(H, H, H, T, T, T, H, T, T, T)  
(H, H, H, T, T, T, T, H, T, T)  
(H, H, H, T, T, T, T, T, H, T)  
(H, H, H, T, T, T, T, T, T, H)  
(H, H, T, H, H, T, T, T, T, T)  
(H, H, T, H, T, H, T, T, T, T)  
(H, H, T, H, T, T, H, T, T, T)  
(H, H, T, H, T, T, T, H, T, T)  
(H, H, T, H, T, T, T, T, H, T)  
(H, H, T, H, T, T, T, T, T, H)  
(H, H, T, T, H, H, T, T, T, T)  
(H, H, T, T, H, T, H, T, T, T)  
(H, H, T, T, H, T, T, H, T, T)  
(H, H, T, T, H, T, T, T, H, T)  
(H, H, T, T, H, T, T, T, T, H)

← The probability of getting **4** heads, *in any ordering*, is the “**or**” of all these **mutually exclusive** cases

Each outcome has probability  $p^4(1-p)^6$

How many cases are there?  $\binom{10}{4}$

$$P(4 \text{ heads}) = \binom{10}{4} p^4 (1-p)^6$$

Pedagogical Pause

How do we formalize this  
result?

It's Time...

It's Time...

*X*

...For Random Variables

Random Variables Are Variables...That Are Random

# Random Variables Are Variables...That Are Random

Check out the variable **result** in the code below.

```
import random

def flip_coin():
    # returns 0 or 1 with prob. 0.5
    return random.choice([0,1])

result = flip_coin()
```

# Random Variables Are Variables...That Are Random

Check out the variable **result** in the code below.

```
import random

def flip_coin():
    # returns 0 or 1 with prob. 0.5
    return random.choice([0,1])

result = flip_coin()
```

- Do we know the value of **result** before we run the code?

# Random Variables Are Variables...That Are Random

Check out the variable **result** in the code below.

```
import random

def flip_coin():
    # returns 0 or 1 with prob. 0.5
    return random.choice([0,1])

result = flip_coin()
```

- Do we know the value of **result** before we run the code? **Nope!**
- Is the value of **result** the same every time we run the code?

# Random Variables Are Variables...That Are Random

Check out the variable **result** in the code below.

```
import random

def flip_coin():
    # returns 0 or 1 with prob. 0.5
    return random.choice([0,1])

result = flip_coin()
```

- Do we know the value of **result** before we run the code? **Nope!**
- Is the value of **result** the same every time we run the code? **Nope!**

Like **result**, a random variable is a variable whose value is uncertain.

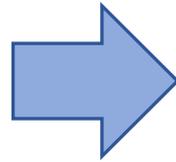
# Random Variables Are Variables...That Are Random

A **random variable** is a variable whose value is uncertain.

```
import random

def flip_coin():
    # returns 0 or 1 with prob. 0.5
    return random.choice([0,1])

result = flip_coin()
```



“Let  $X$  be the result of flipping a coin.”

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

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

Random variables are an abstraction on top of events

Random variables are *not* events

# Random Variables vs. Events

$X$

Let  $X$  be a  
random variable

# Random Variables vs. Events

It is an event when  
 $X$  takes on a value

$$X \quad X = 2$$

Let  $X$  be a  
random variable

# Random Variables vs. Events

It is an event when  
 $X$  takes on a value

 $X$  $X = 2$  $P(X = 2)$ 

Let  $X$  be a  
random variable

So we can still work with  
probabilities of events

# Examples of Random Variables

"Let  $X$  be the result of rolling a dice."

- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

# Examples of Random Variables

"Let  $X$  be the result of rolling a dice."

- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

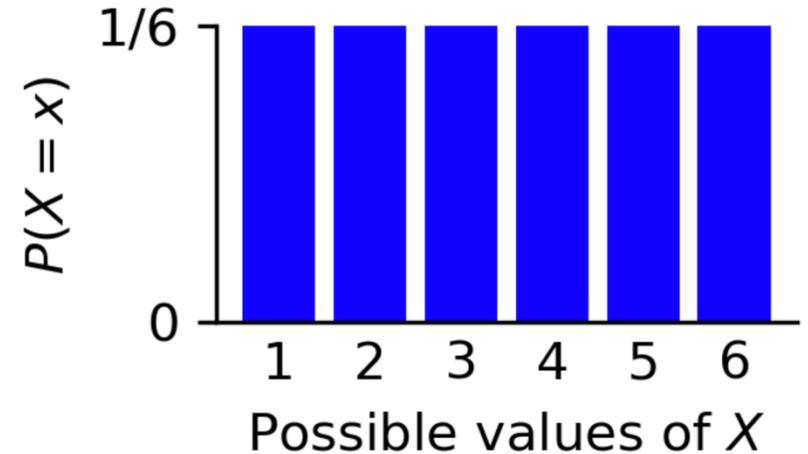
...or,  $P(X = x) = 1/6$  for  $1 \leq x \leq 6$

# Examples of Random Variables

"Let  $X$  be the result of rolling a dice."

- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

...or,  $P(X = x) = 1/6$  for  $1 \leq x \leq 6$

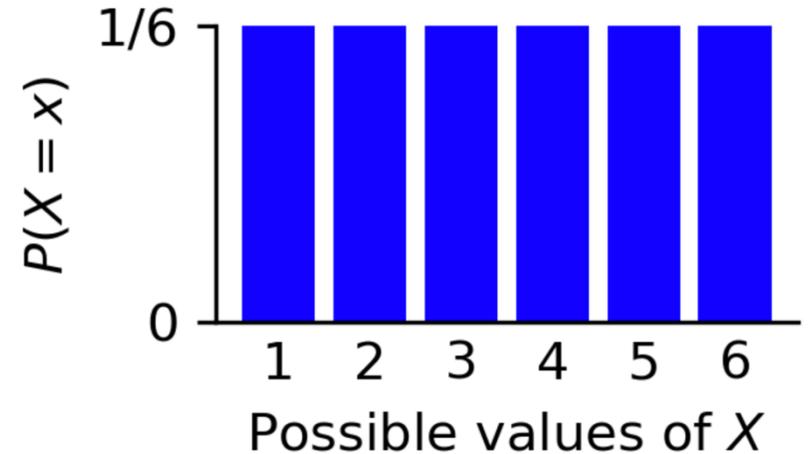


# Examples of Random Variables

"Let  $X$  be the result of rolling a dice."

- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

...or,  $P(X = x) = 1/6$  for  $1 \leq x \leq 6$



"Let  $Y$  be the number of heads seen in 2 coin flips."

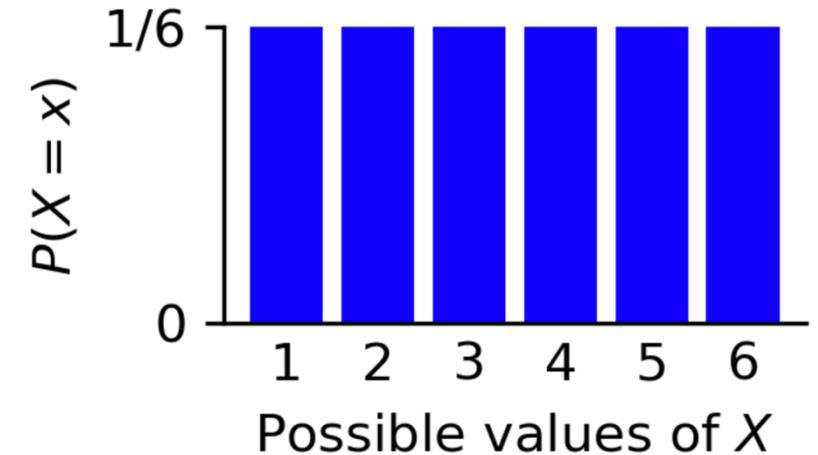
- $P(Y = 0) = 1/4$  (T, T)
- $P(Y = 1) = 1/2$  (H, T), (T, H)
- $P(Y = 2) = 1/4$  (H, H)

# Examples of Random Variables

"Let  $X$  be the result of rolling a dice."

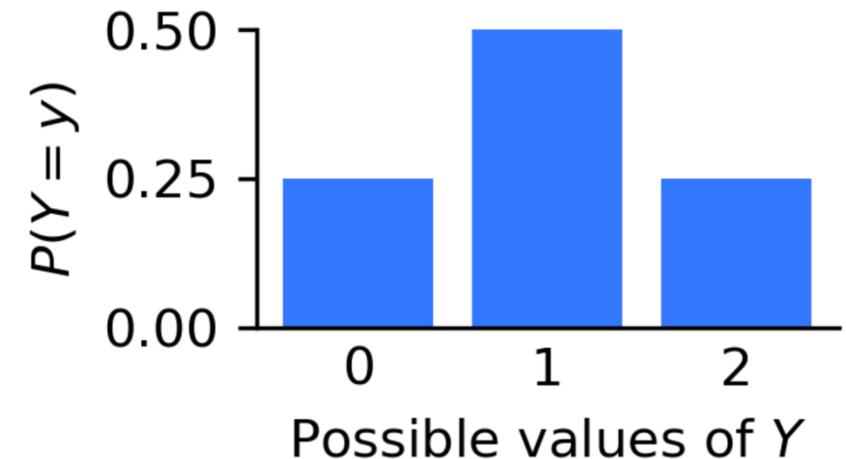
- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

...or,  $P(X = x) = 1/6$  for  $1 \leq x \leq 6$



"Let  $Y$  be the number of heads seen in 2 coin flips."

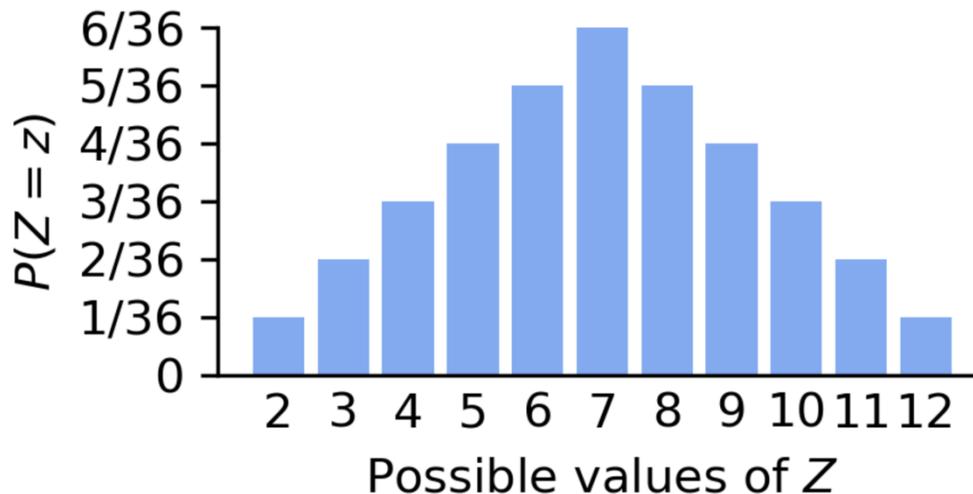
- $P(Y = 0) = 1/4$  (T, T)
- $P(Y = 1) = 1/2$  (H, T), (T, H)
- $P(Y = 2) = 1/4$  (H, H)



# Examples of Random Variables

"Let  $Z$  be the sum of the result of rolling two dice."

- $P(Z = 2) = 1/36$
- $P(Z = 3) = 2/36$
- $P(Z = 4) = 3/36$
- $P(Z = 5) = 4/36$
- $P(Z = 6) = 5/36$
- $P(Z = 7) = 6/36$
- $P(Z = 8) = 5/36$
- $P(Z = 9) = 4/36$
- $P(Z = 10) = 3/36$
- $P(Z = 11) = 2/36$
- $P(Z = 12) = 1/36$



$$P(Z = z) = \begin{cases} \frac{z-1}{36} & z \in \mathbb{Z}, 1 \leq z \leq 6 \\ \frac{13-z}{36} & z \in \mathbb{Z}, 7 \leq z \leq 12 \\ 0 & \text{else} \end{cases}$$

# Examples of Random Variables

"Let  $Z$  be the sum of the result of rolling two dice."

- $P(Z = 2) = 1/36$
- $P(Z = 3) = 2/36$
- $P(Z = 4) = 3/36$
- $P(Z = 5) = 4/36$
- $P(Z = 6) = 5/36$
- $P(Z = 7) = 6/36$
- $P(Z = 8) = 5/36$
- $P(Z = 9) = 4/36$
- $P(Z = 10) = 3/36$
- $P(Z = 11) = 2/36$
- $P(Z = 12) = 1/36$

There's a name for what we're describing, when we list out all possible outcomes + their probabilities:

## Probability Mass Function (PMF)



$$P(Z = z) = \begin{cases} \frac{13-z}{36} & z \in \mathbb{Z}, 1 \leq z \leq 6 \\ \frac{13-z}{36} & z \in \mathbb{Z}, 7 \leq z \leq 12 \\ 0 & \text{else} \end{cases}$$

# Probability Mass Functions

# Random Variables & Functions

"Let  $Y$  be the number of heads seen in 2 coin flips."

If this is a number

$$P(Y = 2)$$

Then this is a number  
(between 0 and 1)

# Random Variables & Functions

"Let  $Y$  be the number of heads seen in 2 coin flips."

If this is a variable

$$P(Y = k)$$

Then this is a function

# Random Variables & Functions

"Let  $Y$  be the number of heads seen in 2 coin flips."

...and get out their probabilities!

$$P(Y = k)$$

0.5

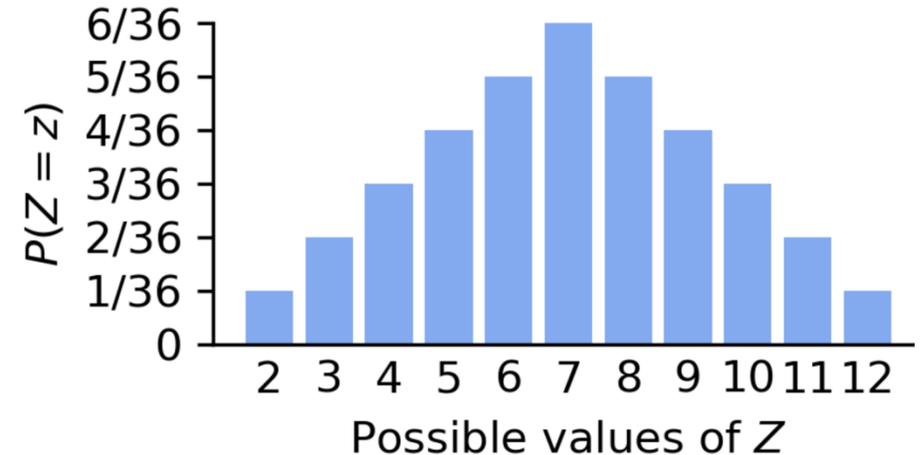
We can put in different inputs...

$$k = 1$$

The relationship between values a random variable can take on, and the corresponding probability, is a *function*!

# Probability Mass Function: Representations

$$P(Z = z) = \begin{cases} \frac{z-1}{36} & z \in \mathbb{Z}, 1 \leq z \leq 6 \\ \frac{13-z}{36} & z \in \mathbb{Z}, 7 \leq z \leq 12 \\ 0 & \text{else} \end{cases}$$



```
def event_probability(z):  
    # probability mass function of Z  
    if not z.is_integer() or z > 12 or z < 1:  
        return 0  
  
    if z < 7:  
        return (z - 1) / 36  
    else:  
        return (13 - z) / 36
```

All of these are different ways we can represent probability mass functions!

# Can We Formalize Many Coin Flips?

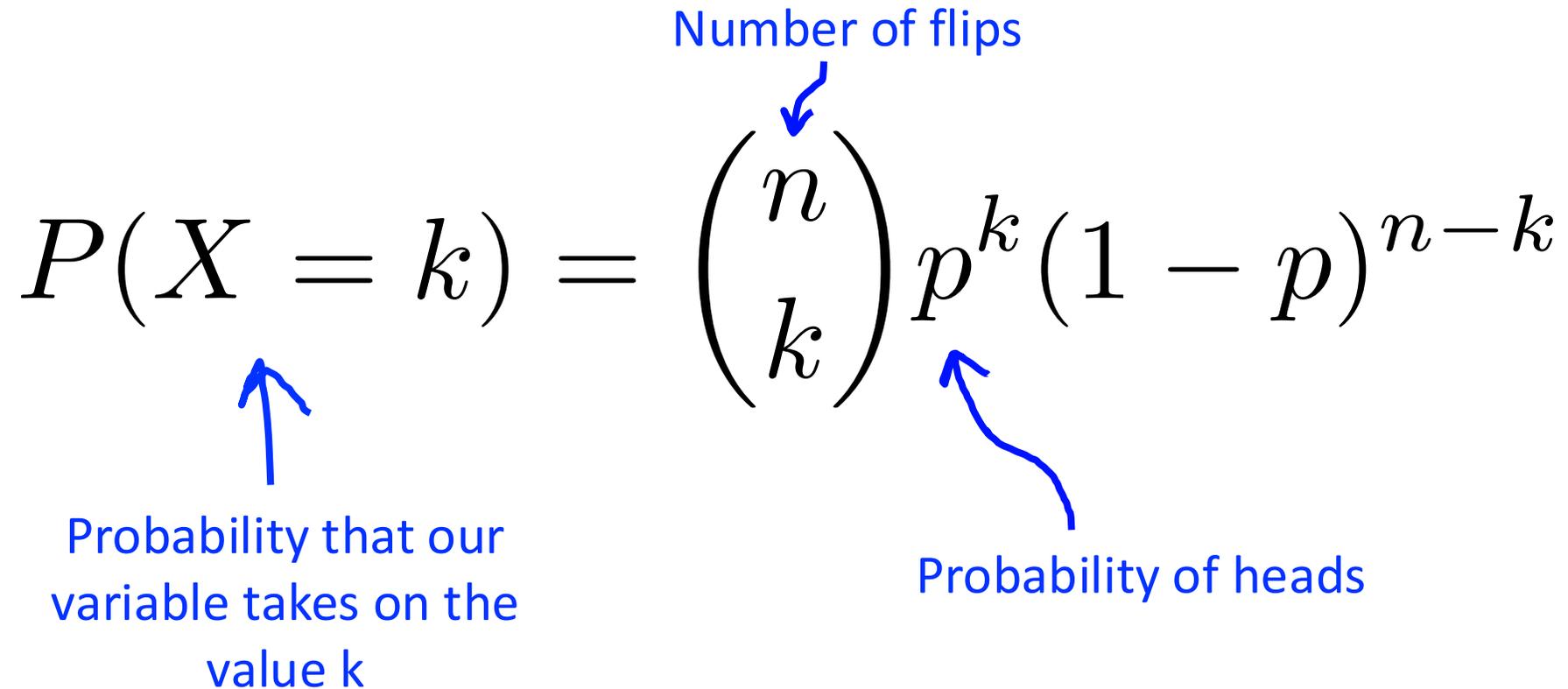
$X$  = number of heads on  $n$  coin flips. Each flip is independent and has prob  $p$  of being heads.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Number of flips

Probability that our variable takes on the value  $k$

Probability of heads



# Quick Understanding Check

$$\sum_{\text{all } k} P(Y = k) \stackrel{?}{=} \underline{\hspace{2cm}}$$

What is the sum of the probabilities of all possible outcomes for  $Y$ ?

# Quick Understanding Check

$$\sum_{\text{all } k} P(Y = k) = 1$$

What is the sum of the probabilities of all possible outcomes for  $Y$ ?

1

# Can You Calculate A PMF From Data? Yes

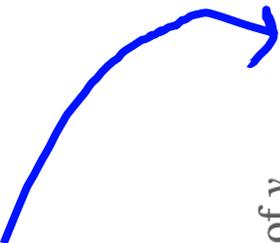
Say this is your dataset, a list of observations of random variable  $Y$ :

[2, 4, 5, 12, 9, 7, 9, 4, 6, 4, 3, 7, 5, 8, 9, 6, 5, 10, 10, 10, 7, 11, 4, 6, 7, 9, 10, 6, 11, 6, 5, 12, 7, 3, 11, 6, 4, 7, 8, 2, 7, 8, 6, 6, 8, 3, 2, 8, 6, 9, 5, 11, 8, 6, 9, 7, 10, 10, 10, 6, 5, 9, 4, 5, 8, 8, 6, 6, 6, 10, 4, 7, 7, 5, 7, 9, 12, 6, 7, 5, 5, 10, 7, 5, 4, 7, 6, 6, 5, 5, 8, 9, 7, 7, 7, 9, 9, 8, 9, 11, 11, 10, 5, 3, 8, 10, 9, 7, 11, 6, 12, 6, 3, 8, 6, 3, 11, 11, 9, 6, 5, 7, 9, 7, 9, 6, 8, 9, 3, 7, 9, 10, 8, 9, 9, 7, 6, 9, 7, 5, 5, 5, 3, 8, 10, 6, 10, 8, 10, 8, 4, 11, 4, 12, 6, 7, 3, 9, 5, 11, 5, 7, 4, 7, 8, 12, 9, 8, 10, 4, 4, 5, 6, 4, 5, 6, 7, 3, 3, 11, 8, 9, 2, 8, 4, 8, 7, 8, 9, 10, 5, 10, 7, 9, 8, 8, 6, 7, 5, 6, 11, 2, 5, 3, 8, 4, 7, 7, 4, 7, 2, 7, 10, 10, 7, 9, 3, 5, 8, 6, 4, 8, 7, 7, 6, 8, 6, 11, 7, 3, 6, 6, 6, 9, 11, 6, 5, 7, 3, 12, 7, 10, 4, 6, 7, 4, 11, 3, 3, 6, 6, 12, 11, 12, 10, 11, 7, 9, 7, 5, 12, 6, 3, 6, 4, 5, 10, 6, 11, 11, 7, 6, 8, 11, 5, 12, 4, 7, 9, 9, 9, 10, 7, 9, 7, 4, 4, 6, 8, 6, 3, 4, 9, 7, 11, 8, 6, 11, 5, 7, 11, 7, 7, 6, 4, 9, 12, 9, 8, 8, 8, 9, 6, 8, 5, 11, 6, 8, 6, 5, 8, 5, 8, 6, 11, 5, 8, 3, 7, 8, 8, 10, 9, 8, 9, 8, 4, 7, 9, 5, 8, 8, 9, 7, 3, 9, 3, 4, 6, 9, 9, 5, 6, 4, 8, 9, 7, 5, 10, 5, 8, 5, 5, 5, 8, 9, 3, 9, 10, 10, 6, 4, 6, 2, 6, 2, 8, 7, 4, 6, 6, 7, 9, 4, 6, 8, 5, 7, 7, 7, 9, 2, 6, 7, 3, 10, 10, 7, 3, 5, 3, 6, 6, 7, 12, 9, 9, 11, 9, 4, 4, 10, 8, 8, 9, 8, 4, 4, 6, 2, 7, 5, 7, 7, 10, 4, 11, 5, 7, 8, 8, 2, 8, 6, 9, 8, 7, 8, 8, 10, 4, 7, 10, 10, 10, 4, 6, 12, 11, 4, 9, 12, 2, 3, 5, 3, 3, 11, 7, 8, 8, 5, 10, 8, 9, 4, 7, 7, 2, 5, 10, 7, 10, 9, 9, 4, 7, 8, 9, 8, 7, 7, 6, 12, 2, 7, 11, 10, 8, 7, 9, 11, 7, 9, 6, 8, 9, 10, 7, 3, 8, 10, 6, 6, 4, 2, 7, 11, 5, 6, 5, 4, 3, 2, 8]

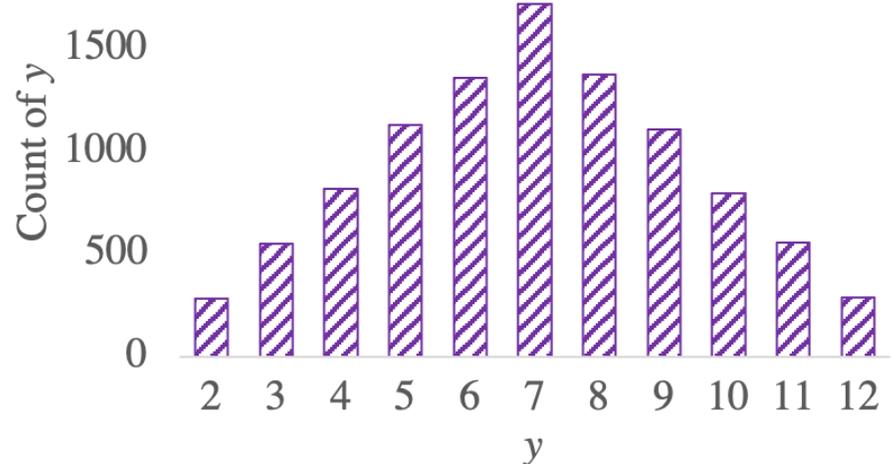
# Can You Calculate A PMF From Data? Yes

Say this is your dataset, a list of observations of random variable  $Y$ :

[2, 4, 5, 12, 9, 7, 9, 4, 6, 4, 3, 7, 5, 8, 9, 6, 5, 10, 10, 10, 7, 11, 4, 6, 7, 9, 10, 6, 11, 6, 5, 12, 7, 3, 11, 6, 4, 7, 8, 2, 7, 8, 6, 6, 8, 3, 2, 8, 6, 9, 5, 11, 8, 6, 9, 7, 10, 10, 10, 6, 5, 9, 4, 5, 8, 8, 6, 6, 6, 10, 4, 7, 7, 5, 7, 9, 12, 6, 7, 5, 5, 10, 7, 5, 4, 7, 6, 6, 5, 5, 8, 9, 7, 7, 7, 9, 9, 8, 9, 11, 11, 10, 5, 3, 8, 10, 9, 7, 11, 6, 12, 6, 3, 8, 6, 3, 11, 11, 9, 6, 5, 7, 9, 7, 9, 6, 8, 9, 3, 7, 9, 10, 8, 9, 9, 7, 6, 9, 7, 5, 5, 5, 3, 8, 10, 6, 10, 8, 10, 8, 4, 11, 4, 12, 6, 7, 3, 9, 5, 11, 5, 7, 4, 7, 8, 12, 9, 8, 10, 4, 4, 5, 6, 4, 5, 6, 7, 3, 3, 11, 8, 9, 2, 8, 4, 8, 7, 8, 9, 10, 5, 10, 7, 9, 8, 8, 6, 7, 5, 6, 11, 2, 5, 3, 8, 4, 7, 7, 4, 7, 2, 7, 10, 10, 7, 9, 3, 5, 8, 6, 4, 8, 7, 7, 6, 8, 6, 11, 7, 3, 6, 6, 6, 9, 11, 6, 5, 7, 3, 12, 7, 10, 4, 6, 7, 4, 11, 3, 3, 6, 6, 12, 11, 12, 10, 11, 7, 9, 7, 5, 12, 6, 3, 6, 4, 5, 10, 6, 11, 11, 7, 6, 8, 11, 5, 12, 4, 7, 9, 9, 9, 10, 7, 9, 7, 4, 4, 6, 8, 6, 3, 4, 9, 7, 11, 8, 6, 11, 5, 7, 11, 7, 7, 6, 4, 9, 12, 9, 8, 8, 8, 9, 6, 8, 5, 11, 6, 8, 6, 5, 8, 5, 8, 6, 11, 5, 8, 3, 7, 8, 8, 10, 9, 8, 9, 8, 4, 7, 9, 5, 8, 8, 9, 7, 3, 9, 3, 4, 6, 9, 9, 5, 6, 4, 8, 9, 7, 5, 10, 5, 8, 5, 5, 5, 8, 9, 3, 9, 10, 10, 6, 4, 6, 2, 6, 2, 8, 7, 4, 6, 6, 7, 9, 4, 6, 8, 5, 7, 7, 7, 9, 2, 6, 7, 3, 10, 10, 7, 3, 5, 3, 6, 6, 7, 12, 9, 9, 11, 9, 4, 4, 10, 8, 8, 9, 8, 4, 4, 6, 2, 7, 5, 7, 7, 10, 4, 11, 5, 7, 8, 8, 2, 8, 6, 9, 8, 7, 8, 8, 10, 4, 7, 10, 10, 10, 4, 6, 12, 11, 4, 9, 12, 2, 3, 5, 3, 3, 11, 7, 8, 8, 5, 10, 8, 9, 4, 7, 7, 2, 5, 10, 7, 10, 9, 9, 4, 7, 8, 9, 8, 7, 7, 6, 12, 2, 7, 11, 10, 8, 7, 9, 11, 7, 9, 6, 8, 9, 10, 7, 3, 8, 10, 6, 6, 4, 2, 7, 11, 5, 6, 5, 4, 3, 2, 8]



Just convert your data into counts:

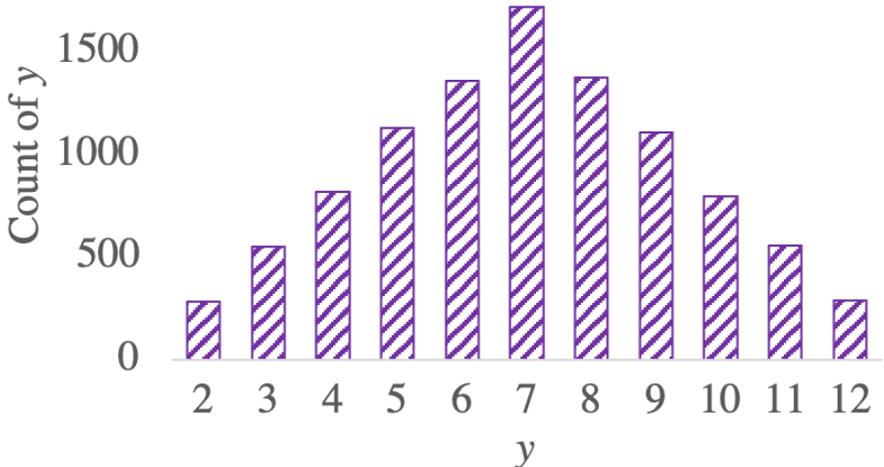


# Can You Calculate A PMF From Data? Yes

Say this is your dataset, a list of observations of random variable  $Y$ :

[2, 4, 5, 12, 9, 7, 9, 4, 6, 4, 3, 7, 5, 8, 9, 6, 5, 10, 10, 10, 7, 11, 4, 6, 7, 9, 10, 6, 11, 6, 5, 12, 7, 3, 11, 6, 4, 7, 8, 2, 7, 8, 6, 6, 8, 3, 2, 8, 6, 9, 5, 11, 8, 6, 9, 7, 10, 10, 10, 6, 5, 9, 4, 5, 8, 8, 6, 6, 6, 10, 4, 7, 7, 5, 7, 9, 12, 6, 7, 5, 5, 10, 7, 5, 4, 7, 6, 6, 5, 5, 8, 9, 7, 7, 7, 9, 9, 8, 9, 11, 11, 10, 5, 3, 8, 10, 9, 7, 11, 6, 12, 6, 3, 8, 6, 3, 11, 11, 9, 6, 5, 7, 9, 7, 9, 6, 8, 9, 3, 7, 9, 10, 8, 9, 9, 7, 6, 9, 7, 5, 5, 5, 3, 8, 10, 6, 10, 8, 10, 8, 4, 11, 4, 12, 6, 7, 3, 9, 5, 11, 5, 7, 4, 7, 8, 12, 9, 8, 10, 4, 4, 5, 6, 4, 5, 6, 7, 3, 3, 11, 8, 9, 2, 8, 4, 8, 7, 8, 9, 10, 5, 10, 7, 9, 8, 8, 6, 7, 5, 6, 11, 2, 5, 3, 8, 4, 7, 7, 4, 7, 2, 7, 10, 10, 7, 9, 3, 5, 8, 6, 4, 8, 7, 7, 6, 8, 6, 11, 7, 3, 6, 6, 6, 9, 11, 6, 5, 7, 3, 12, 7, 10, 4, 6, 7, 4, 11, 3, 3, 6, 6, 12, 11, 12, 10, 11, 7, 9, 7, 5, 12, 6, 3, 6, 4, 5, 10, 6, 11, 11, 7, 6, 8, 11, 5, 12, 4, 7, 9, 9, 9, 10, 7, 9, 7, 4, 4, 6, 8, 6, 3, 4, 9, 7, 11, 8, 6, 11, 5, 7, 11, 7, 7, 6, 4, 9, 12, 9, 8, 8, 8, 9, 6, 8, 5, 11, 6, 8, 6, 5, 8, 5, 8, 6, 11, 5, 8, 3, 7, 8, 8, 10, 9, 8, 9, 8, 4, 7, 9, 5, 8, 8, 9, 7, 3, 9, 3, 4, 6, 9, 9, 5, 6, 4, 8, 9, 7, 5, 10, 5, 8, 5, 5, 5, 8, 9, 3, 9, 10, 10, 6, 4, 6, 2, 6, 2, 8, 7, 4, 6, 6, 7, 9, 4, 6, 8, 5, 7, 7, 7, 9, 2, 6, 7, 3, 10, 10, 7, 3, 5, 3, 6, 6, 7, 12, 9, 9, 11, 9, 4, 4, 10, 8, 8, 9, 8, 4, 4, 6, 2, 7, 5, 7, 7, 10, 4, 11, 5, 7, 8, 8, 2, 8, 6, 9, 8, 7, 8, 8, 10, 4, 7, 10, 10, 10, 4, 6, 12, 11, 4, 9, 12, 2, 3, 5, 3, 3, 11, 7, 8, 8, 5, 10, 8, 9, 4, 7, 7, 2, 5, 10, 7, 10, 9, 9, 4, 7, 8, 9, 8, 7, 7, 6, 12, 2, 7, 11, 10, 8, 7, 9, 11, 7, 9, 6, 8, 9, 10, 7, 3, 8, 10, 6, 6, 4, 2, 7, 11, 5, 6, 5, 4, 3, 2, 8]

Just convert your data into counts:



And then use those counts to calculate probabilities for each outcome:

$$P(Y = 3) \approx \frac{\text{count}(Y = 3)}{n} = \frac{552}{10000} = 0.0552$$

# You Can Use PMFs Other People Give You

Let  $X$  be the number of earthquakes that happen in California next year.

Here's the PMF for  $X$ :

$$P(X = x) = \frac{69^x e^{-69}}{x!}$$

What is the probability that there are 60 earthquakes in California next year?

# You Can Use PMFs Other People Give You

Let  $X$  be the number of earthquakes that happen in California next year.

Here's the PMF for  $X$ :

$$P(X = x) = \frac{69^x e^{-69}}{x!}$$

What is the probability that there are 60 earthquakes in California next year?

$$P(X = 60) = \frac{69^{60} e^{-69}}{60!} \approx 0.028$$

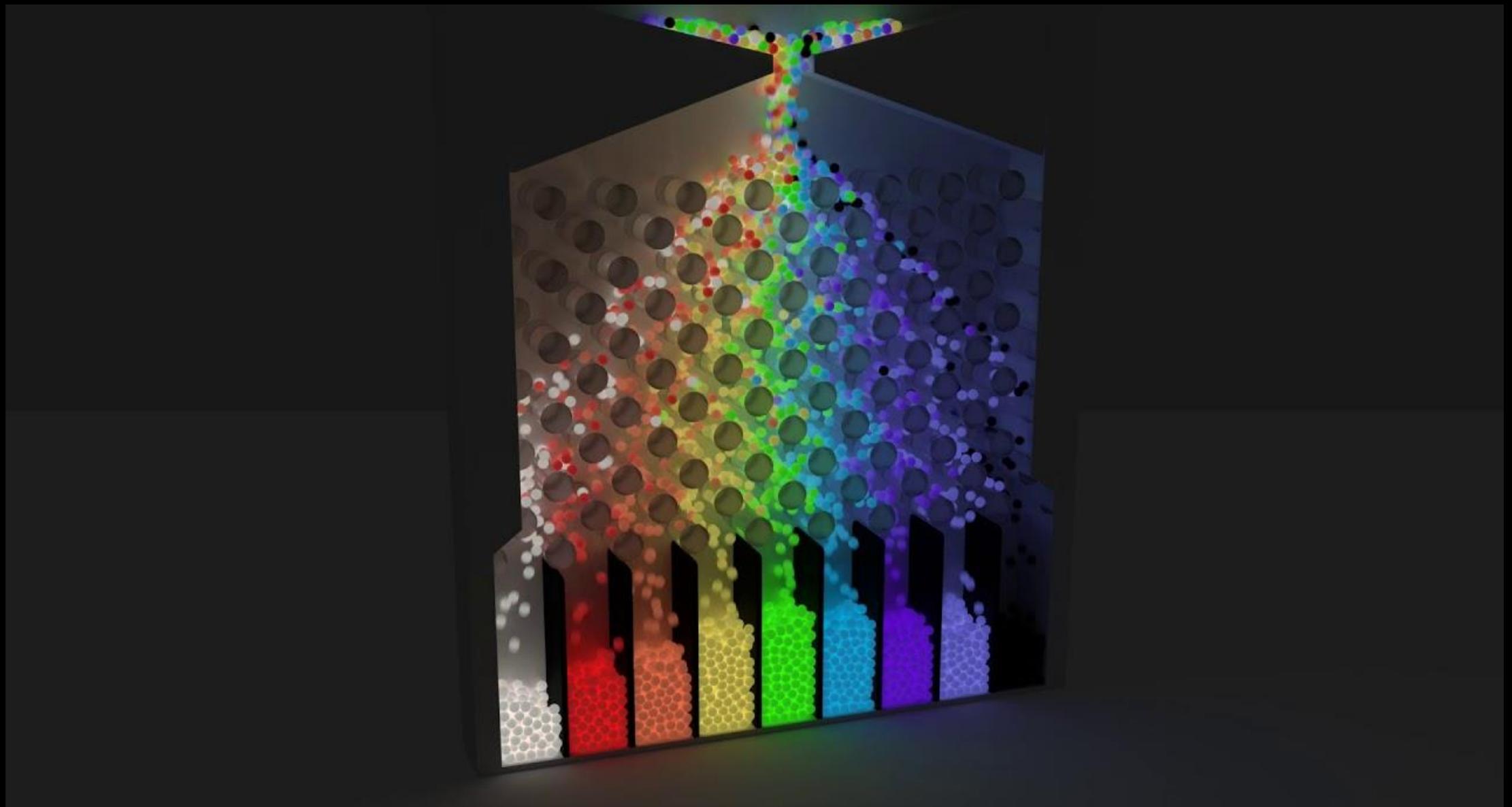
Just plug numbers in!

What is the most important thing to know about a discrete random variable?

# Probability Mass Function

Random Variables Are Awesome





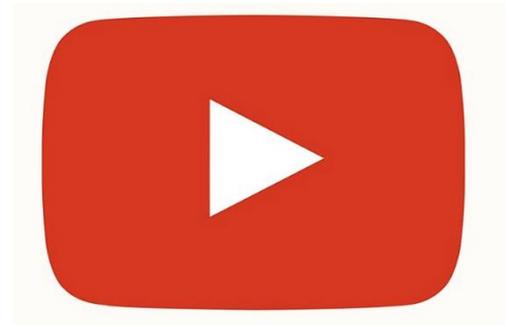
Some Random Variables Are “Classics”

# Practice: Ad Clicks

Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

What is the probability of this ad getting 10 clicks?



# Practice: Ad Clicks

Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

What is the probability of this ad getting 10 clicks?

Same as asking: “What is the probability of getting 10 heads out of 1000 coin flips, when the probability of each head is 0.01”



# We Have Discovered The Binomial

The binomial describes scenarios with:

1.  $n$  independent trials (coin flips)
2. A consistent probability  $p$  of success on each trial (heads)
3. What we want: what is the probability of exactly  $k$  successes?

Many, many scenarios fit this description:

- # of 1's in randomly generated in length  $n$  bit string
- # of servers working in a large computer cluster
- # of votes for one of two candidates in an election
- # of jury members selected from a particular demographic
- # of CS109 students who decided to come to lecture today

# We Have Discovered The Binomial



*Here yee. This type of random variable is so common it needs a name so that I can talk about it generally.*

*I shall call it: the **Binomial** Random Variable. Huzzah.*

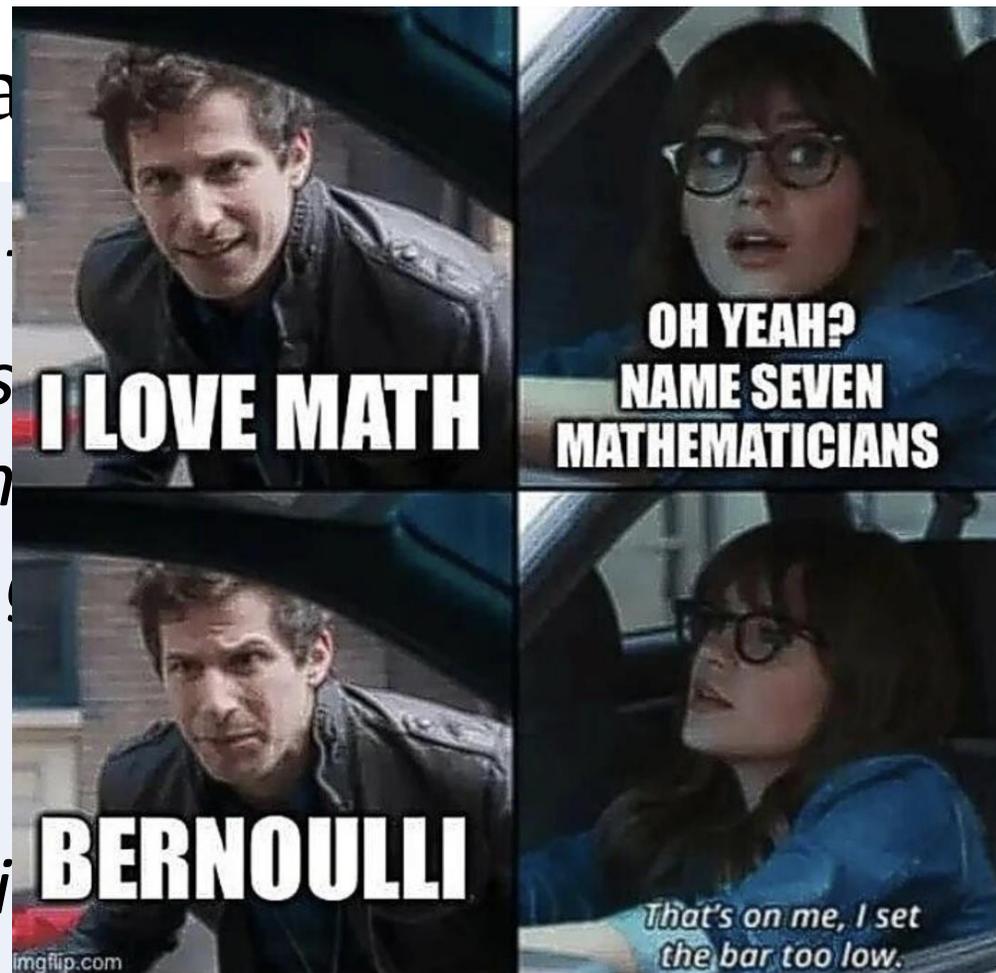
Jacob “James” Bernoulli (1654-1705): Swiss mathematician  
One of many mathematicians in the Bernoulli family

# We Have Discovered The Binomial



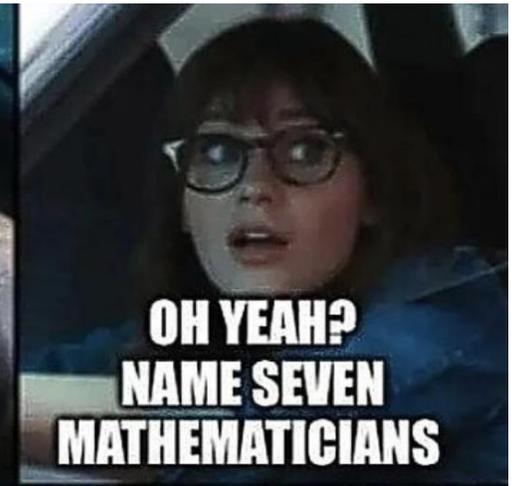
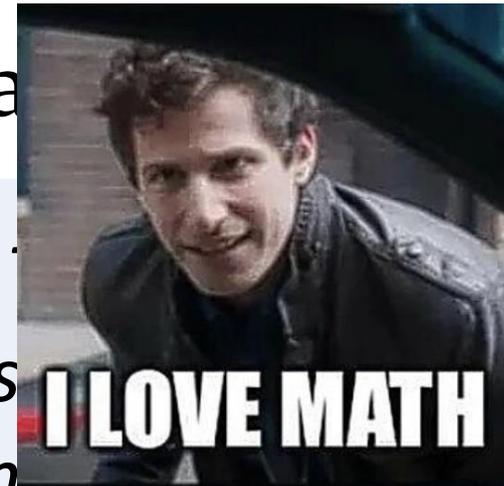
*Here yee.  
variable is s  
name so th*

*I shall call it:  
Vari*



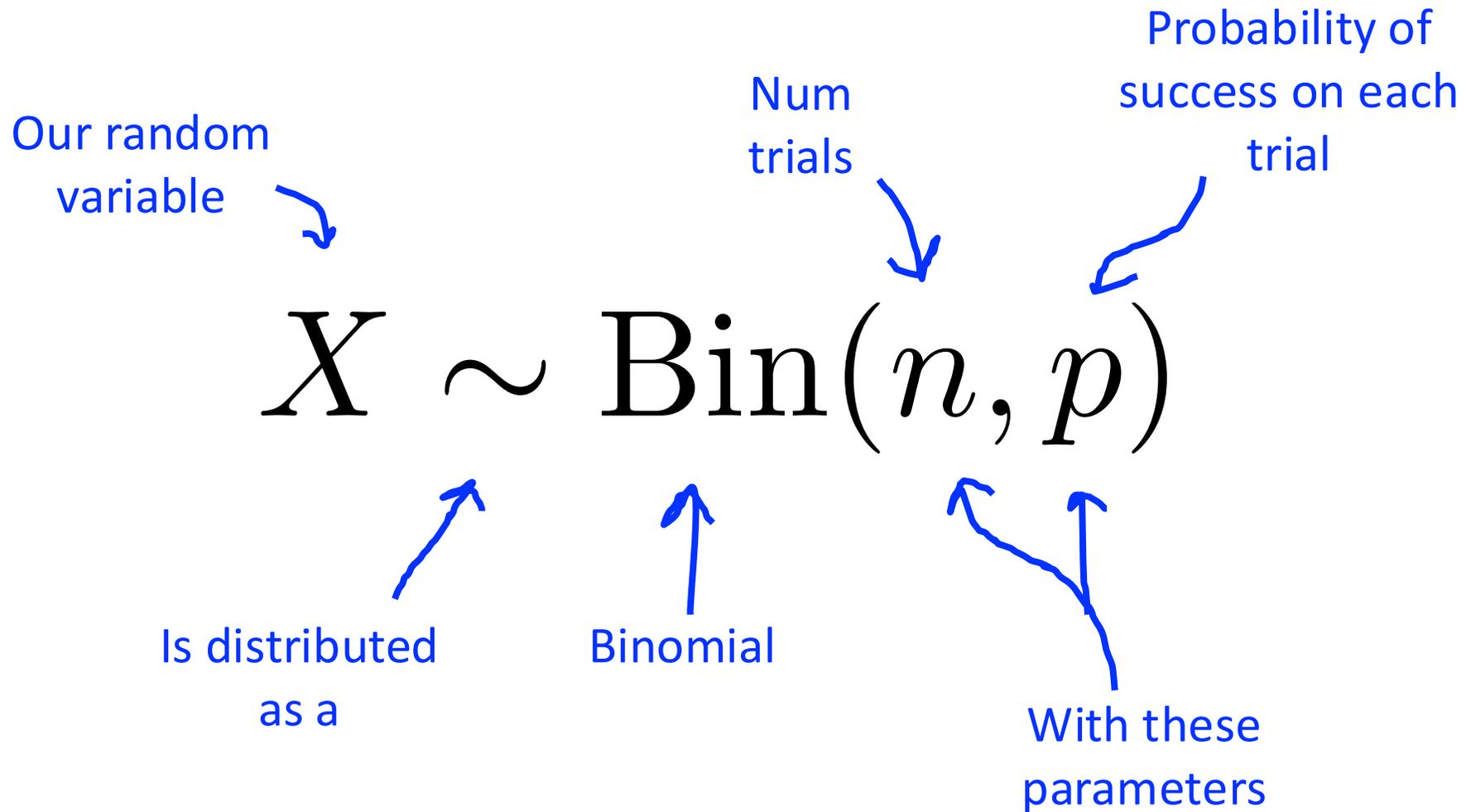
Jacob “James” Bernoulli (1654-1705): Swiss mathematician  
One of many mathematicians in the Bernoulli family

# We Have Discovered The Binomial



Jacob “James” Bernoulli (1654-1705): Swiss mathematician  
One of many mathematicians in the Bernoulli family

# Declaring a Random Variable to be Binomial



# Then We Automatically Know the PMF!

Probability Mass Function for a  
Binomial

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

↑  
Probability that our  
variable takes on the  
value  $k$

# New Recipe For Solving Problems!

1. Recognize a classic random variable type



# New Recipe For Solving Problems!

1. Recognize a classic random variable type

2. Define a random variable to be that type,  
with parameters



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

# New Recipe For Solving Problems!

1. Recognize a classic random variable type

2. Define a random variable to be that type, with parameters

3. Profit off the PMF



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



# You Get So Much For Free!

## Binomial Random Variable

**Notation:**  $X \sim \text{Bin}(n, p)$

**Description:** Number of "successes" in  $n$  identical, independent experiments each with probability of success  $p$ .

**Parameters:**  $n \in \{0, 1, \dots\}$ , the number of experiments.  
 $p \in [0, 1]$ , the probability that a single experiment gives a "success".

**Support:**  $x \in \{0, 1, \dots, n\}$

**PMF equation:**  $\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$

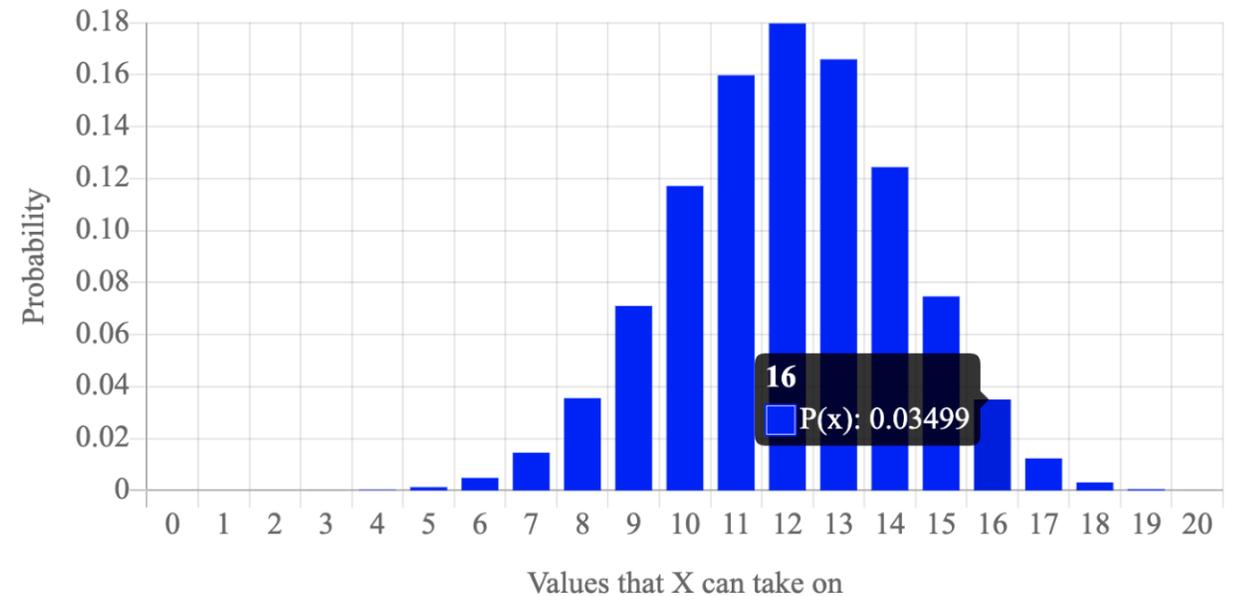
**Expectation:**  $E[X] = n \cdot p$

**Variance:**  $\text{Var}(X) = n \cdot p \cdot (1 - p)$

**PMF graph:**

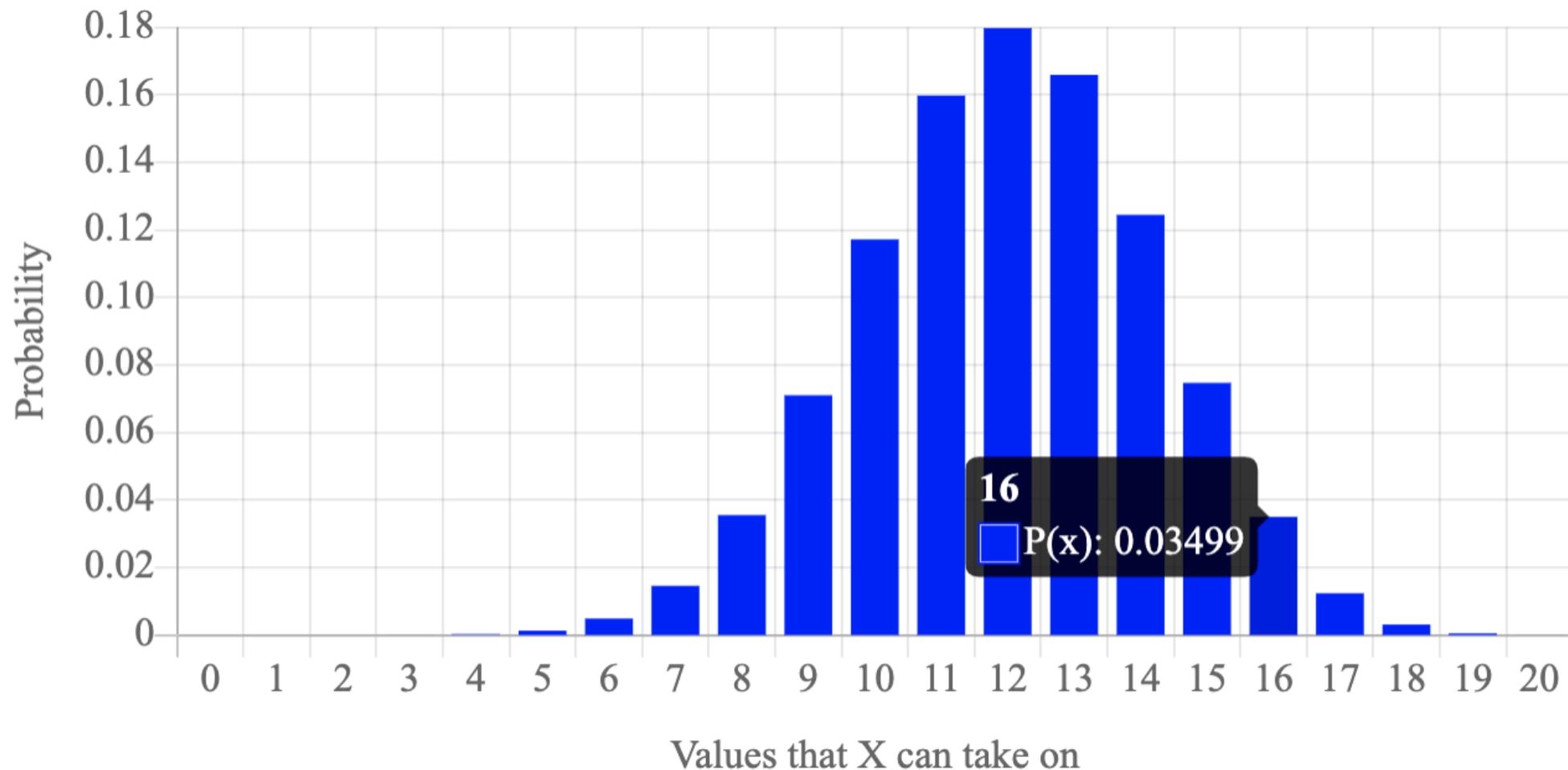
Parameter  $n$ :

Parameter  $p$ :



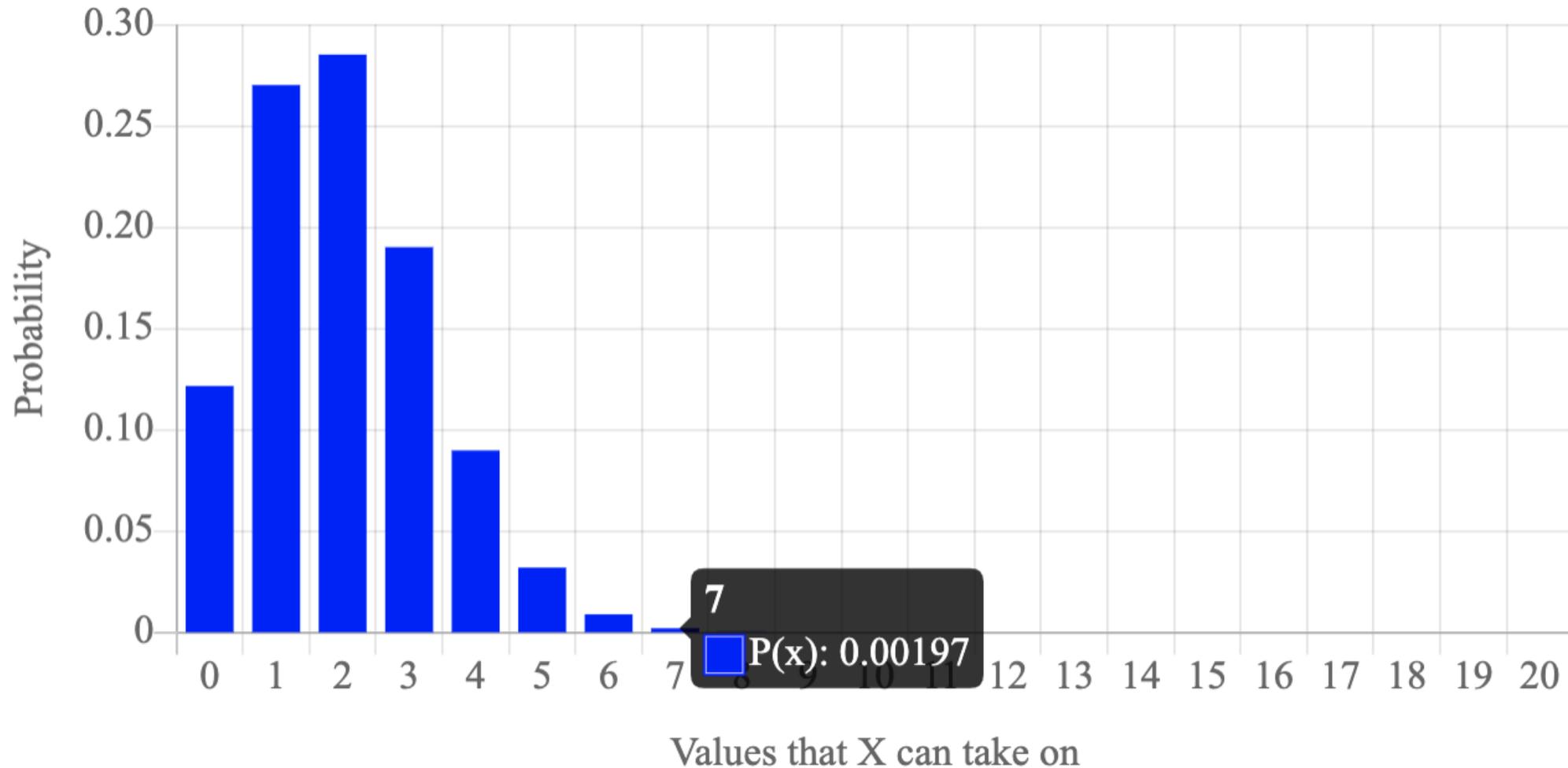
# The PMF as a Graph: $X \sim \text{Bin}(n = 20, p = 0.6)$

Parameter  $n$ :  Parameter  $p$ :



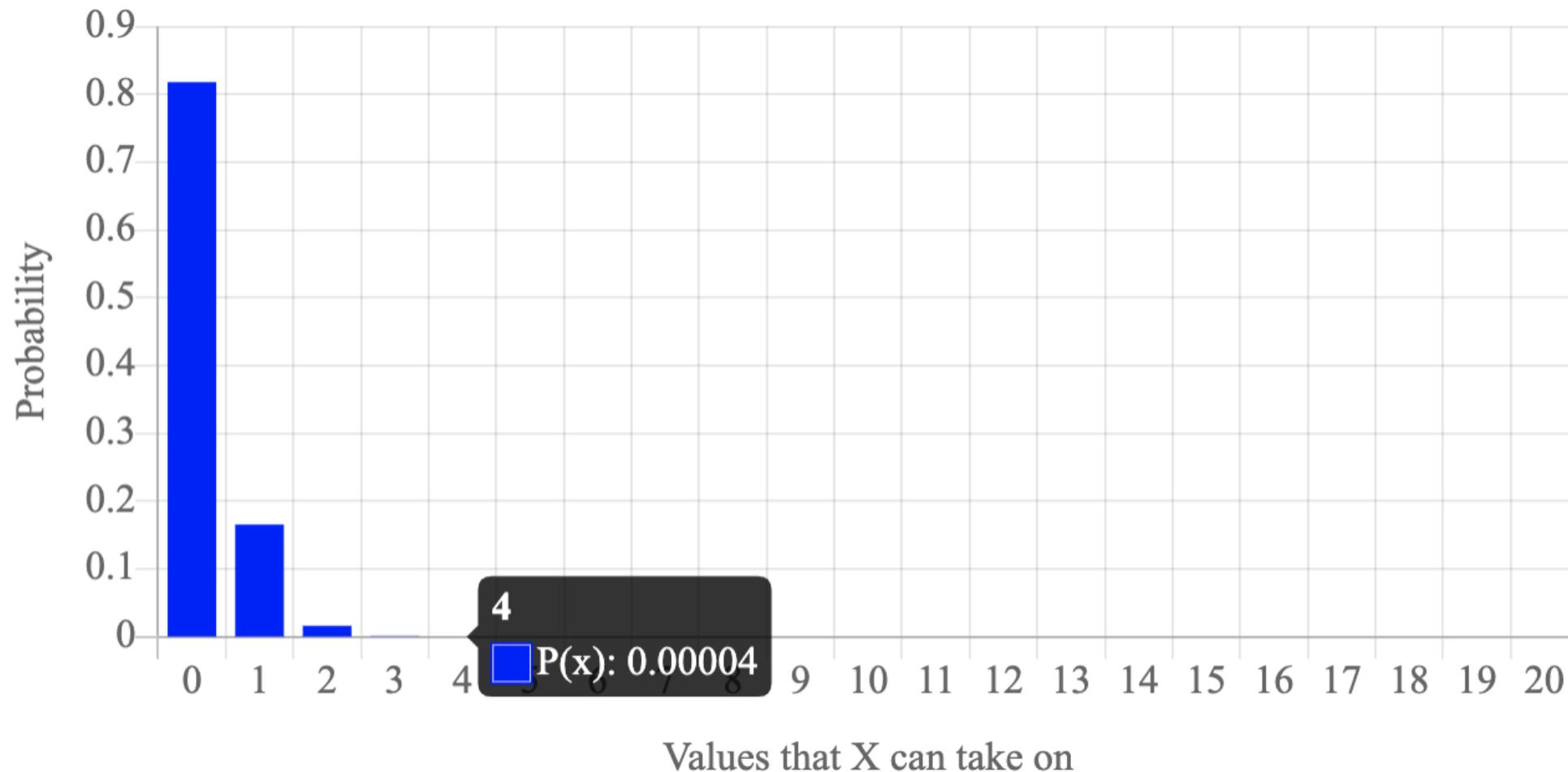
# The PMF as a Graph: $X \sim \text{Bin}(n = 20, p = 0.1)$

Parameter  $n$ :  Parameter  $p$ :



# The PMF as a Graph: $X \sim \text{Bin}(n = 20, p = 0.01)$

Parameter  $n$ :  Parameter  $p$ :



# Practice: Ad Clicks



Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

What is the probability of this ad getting 10 clicks?

Let  $X$  be the number of ad clicks.  $X \sim \text{Bin}(n = 1000, p = 0.01)$ .

$$P(X = k) = \binom{1000}{k} (0.01)^k (0.99)^{1000-k}$$

$$P(X = 10) = \binom{1000}{10} (0.01)^{10} (0.99)^{990} \approx 0.125$$

# Practice: Ad Clicks



Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

What is the probability of this ad getting **20** clicks?

Let  $X$  be the number of ad clicks.  $X \sim \text{Bin}(n = 1000, p = 0.01)$ .

$$P(X = k) = \binom{1000}{k} (0.01)^k (0.99)^{1000-k}$$

$$P(X = \mathbf{20}) = \binom{1000}{\mathbf{20}} (0.01)^{\mathbf{20}} (0.99)^{\mathbf{980}} \approx 0.0018$$

# Practice: Ad Clicks



Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

What is the probability of this ad getting **20** clicks?

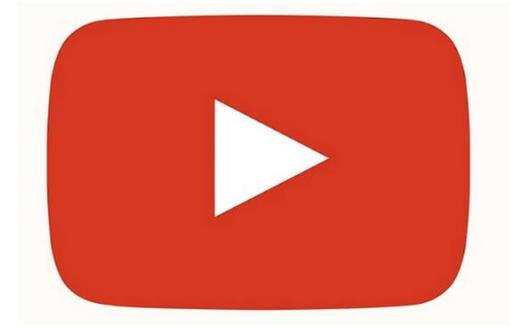
Let  $X$  be the number of ad clicks.  $X \sim \text{Bin}(n = 1000, p = 0.01)$ .

```
>>> from scipy import stats
>>> X = stats.binom(1000, 0.01)
>>> X.pmf(20)
0.0017918782400182223
```

$k$   $n$   $p$

Hand-drawn blue arrows: one points from  $k$  to the value 20 in the pmf call; another points from  $n$  to the value 1000; a third points from  $p$  to the value 0.01.

# Practice: Ad Clicks



Every day, Youtube shows a particular ad 1000 times.

Each ad served is clicked with  $p = 0.01$  (otherwise it's ignored).

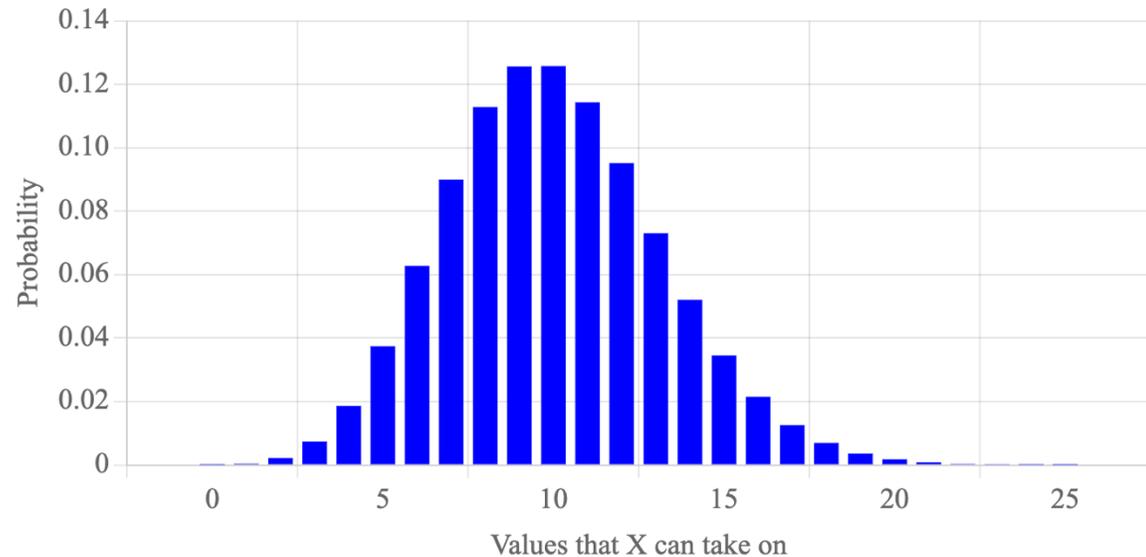
What is the probability of this ad getting **20** clicks?

Let  $X$  be the number of ad clicks.

$$X \sim \text{Bin}(n = 1000, p = 0.01).$$

**PMF graph:**

Parameter  $n$ :  Parameter  $p$ :



# Server Redundancy



A network can remain functional as long as at least 2 out of 7 servers are alive.

The probability of any server working is 0.8.

What is the probability that less than 2 servers are alive?

# Server Redundancy



A network can remain functional as long as at least 2 out of 7 servers are alive.

The probability of any server working is 0.8.

What is the probability that less than 2 servers are alive?

Let  $X$  be the number of servers alive.

$$X \sim \text{Bin}(n = 7, p = 0.8).$$

# Server Redundancy



A network can remain functional as long as at least 2 out of 7 servers are alive.

The probability of any server working is 0.8.

What is the probability that less than 2 servers are alive?

Let  $X$  be the number of servers alive.  $X \sim \text{Bin}(n = 7, p = 0.8)$ .

$$P(X = k) = \binom{7}{k} (0.8)^k (0.2)^{7-k}$$

# Server Redundancy



A network can remain functional as long as at least 2 out of 7 servers are alive.

The probability of any server working is 0.8.

What is the probability that less than 2 servers are alive?

Let  $X$  be the number of servers alive.  $X \sim \text{Bin}(n = 7, p = 0.8)$ .

$$P(X = k) = \binom{7}{k} (0.8)^k (0.2)^{7-k}$$

$$P(X < 2) = P(X = 0) + P(X = 1)$$

# Server Redundancy



A network can remain functional as long as at least 2 out of 7 servers are alive.

The probability of any server working is 0.8.

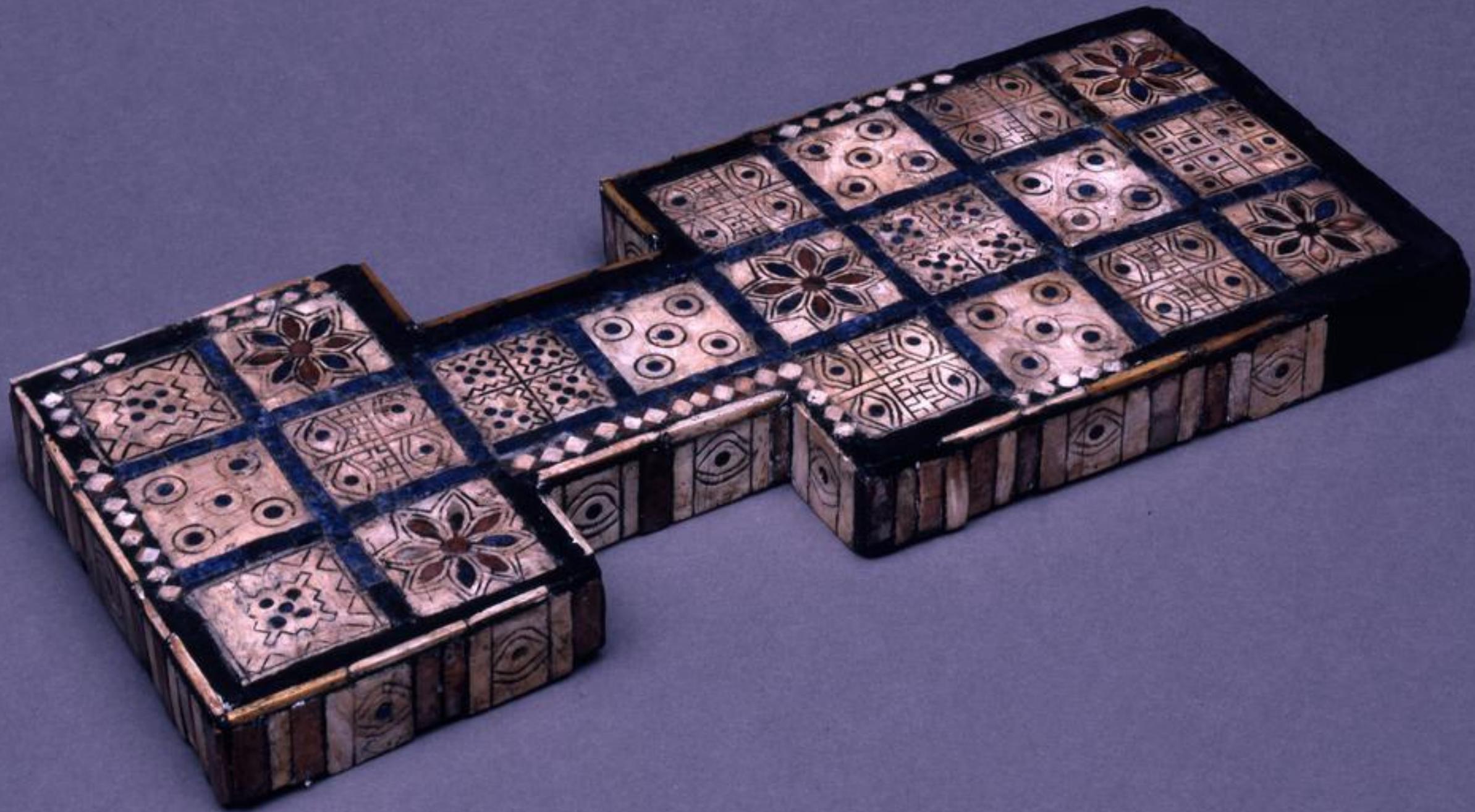
What is the probability that less than 2 servers are alive?

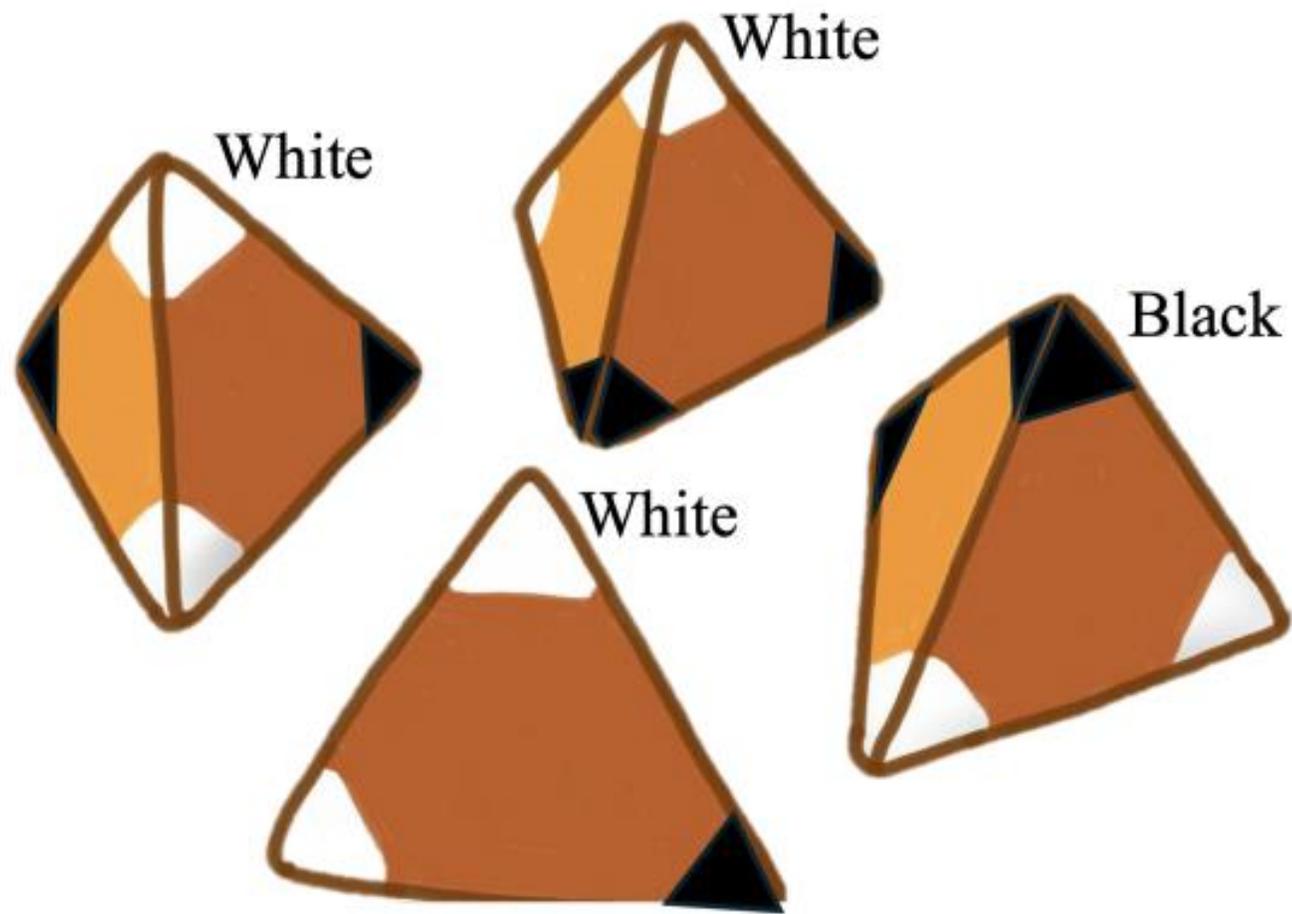
Let  $X$  be the number of servers alive.  $X \sim \text{Bin}(n = 7, p = 0.8)$ .

$$P(X = k) = \binom{7}{k} (0.8)^k (0.2)^{7-k}$$

$$P(X < 2) = P(X = 0) + P(X = 1) = \binom{7}{0} (0.8)^0 (0.2)^{7-0} + \binom{7}{1} (0.8)^1 (0.2)^{7-1} \approx 0.0004$$









How many of these questions do you answer “yes”?

1. Would you rather have unlimited time than unlimited money?
2. Do you enjoy surprises?
3. Would you rather go to the beach than the mountains?
4. Do you think physical books are better than e-books?
5. Would you rather take a vacation in your home country than abroad?
6. Do you support the dictator?

Let  $p$  be the probability that a person supports the dictator

What is the probability that someone would answer “yes” to 4 questions?



How many of these questions do you answer “yes”?

1. Would you rather have unlimited time than unlimited money?
2. Do you enjoy surprises?
3. Would you rather go to the beach than the mountains?
4. Do you think physical books are better than e-books?
5. Would you rather take a vacation in your home country than abroad?
6. Do you support the dictator?

50/50  
questions

Let  $p$  be the probability that a person supports the dictator

What is the probability that someone would answer “yes” to 4 questions?

What is the most important thing to know about a discrete random variable?

# Probability Mass Function



# What is the probability of winning a 7 game series?

Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games. *Assume: All 7 games are played, even if a team has already won 4.*

---

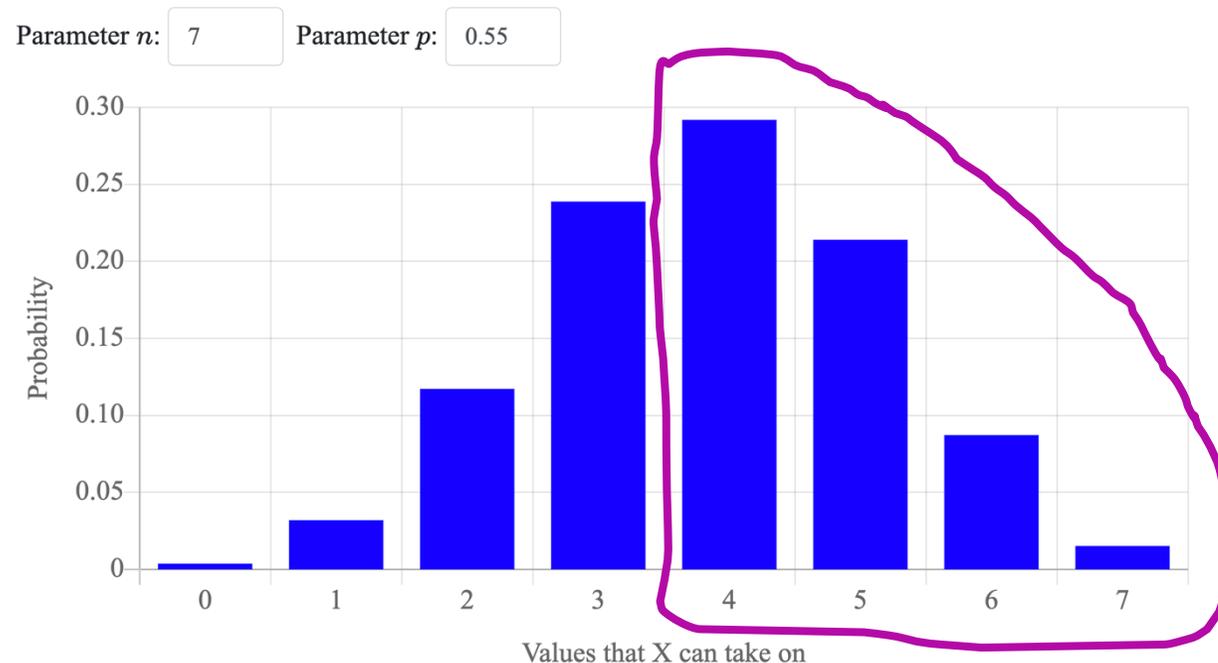
Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 7, p = 0.55)$ .  $P(X > 3)$ ?

# What is the probability of winning a 7 game series?

Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games. *Assume: All 7 games are played, even if a team has already won 4.*

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 7, p = 0.55)$ .  $P(X > 3)$ ?



# What is the probability of winning a 7 game series?

Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games. *Assume: All 7 games are played, even if a team has already won 4.*

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 7, p = 0.55)$ .  $P(X > 3)$ ?

$$\begin{aligned} P(X \geq 4) &= \sum_{i=4}^7 P(X = i) \\ &= \sum_{i=4}^7 \binom{7}{i} p^i (1-p)^{7-i} \\ &= \sum_{i=4}^7 \binom{7}{i} 0.55^i (0.45)^{7-i} \approx 0.61 \end{aligned}$$

# What is the probability of winning a 7 game series?

Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games. *Assume: All 7 games are played, even if a team has already won 4.*

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 7, p = 0.55)$ .  $P(X > 3)$ ?

$$\begin{aligned} P(X \geq 4) &= \sum_{i=4}^7 P(X = i) \\ &= \sum_{i=4}^7 \binom{7}{i} p^i (1-p)^{7-i} \\ &= \sum_{i=4}^7 \binom{7}{i} 0.55^i (0.45)^{7-i} \approx 0.61 \end{aligned}$$

# What is the probability of winning a 7 game series?

Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games. ***Assume: All 7 games are played, even if a team has already won 4.***

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 7, p = 0.55)$ .  $P(X > 3)$ ?

$$\begin{aligned} P(X \geq 4) &= \sum_{i=4}^7 P(X = i) \\ &= \sum_{i=4}^7 \binom{7}{i} 0.55^i (0.45)^{7-i} \\ &= \sum_{i=4}^7 \binom{7}{i} 0.55^i (0.45)^{7-i} \approx 0.61 \end{aligned}$$

**My Claim: Even if we drop this assumption, we will still get to the same answer.**

To the code !!!

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games.

---

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: All 3 games are played, even if a team has already won 2.*

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 3, p = 0.55)$ .  $P(X > 1)$ ?

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) \\ &= \binom{3}{2} p^2 (1 - p) + \binom{3}{3} p^3 (1 - p)^0 \\ &= \binom{3}{2} (0.55)^2 (0.45) + \binom{3}{3} (0.55)^3 \approx 0.575 \end{aligned}$$

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. ~~*Assume: All 3 games are played, even if a team has already won 2.*~~

---

# Debugging Probability



# Debugging Probability

How to calculate the probability of at least  $k$  successes in  $n$  independent trials?

- $X$  is number of successes in  $n$  trials each with probability  $p$

- $$P(X \geq k) = \binom{n}{k} p^k$$

Chose slots for success, don't care about rest

# ways to choose slots for success

Probability that each is success

# Debugging Probability

How to calculate the probability of at least  $k$  successes in  $n$  independent trials?

- $X$  is number of successes in  $n$  trials each with probability  $p$

- $$P(X \geq k) = \binom{n}{k} p^k$$

Chose slots for success, don't care about rest

# ways to choose slots for success

Probability that each is success

$$P(X \geq 2) = \binom{3}{2} p^2$$

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

Plugging in the numbers from our warrior mini example.

# Debugging Probability

How to calculate the probability of at least  $k$  successes in  $n$  independent trials?

- $X$  is number of successes in  $n$  trials each with probability  $p$

- $$P(X \geq k) = \binom{n}{k} p^k$$

Chose slots for success, don't care about rest

# ways to choose slots for success

Probability that each is success

$$P(X \geq 2) = \binom{3}{2} p^2$$

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$$= 1.65$$

Plugging in the numbers from our warrior mini example.

= 1.65 ??????????



= 1.65 ??????????

But our very first rule of probability:  
P(Event) must be a number between 0  
and 1.

# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$H = \text{Win}$

$T = \text{Lose}$

? = Could be either

H      H      ?

H      ?      H

?      H      H

# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$H = \text{Win}$

$T = \text{Lose}$

? = Could be either

H      H      ?

H      ?      H

?      H      H

This equation says, sum over the probabilities of each of these cases.

# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$



$H$  = Win

$T$  = Lose

? = Could be either

H      H      ?

H      ?      H

?      H      H

This equation says, sum over the probabilities of each of these cases.



# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$



This equation says, sum over the probabilities of each of these cases.

H      H      ?

H      ?      H

?      H      H

These cases are not mutually exclusive !! Cannot just sum their probabilities.

$H = \text{Win}$

$T = \text{Lose}$

? = Could be either



# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$H = \text{Win}$

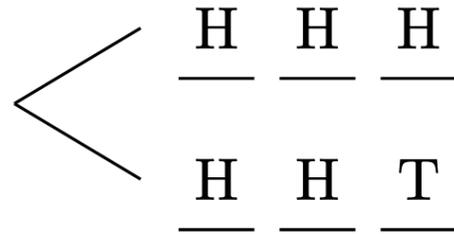
$T = \text{Lose}$

? = Could be either

H      H      ?

H      ?      H

?      H      H



# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$H = \text{Win}$

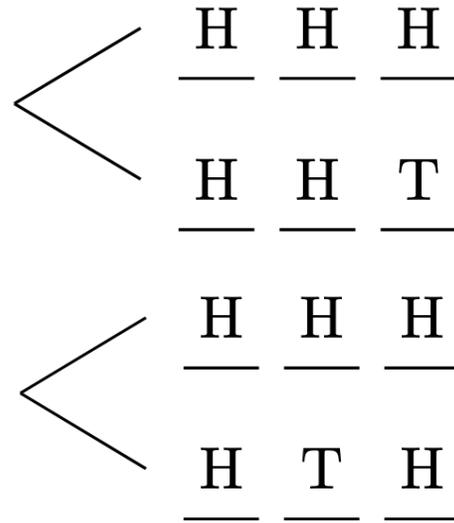
$T = \text{Lose}$

? = Could be either

H      H      ?

H      ?      H

?      H      H





# Debugging Probability

Buggy Strategy: Chose slots for success, don't care about rest

$$P(X \geq 2) = \binom{3}{2} (0.55)^2$$

$H = \text{Win}$

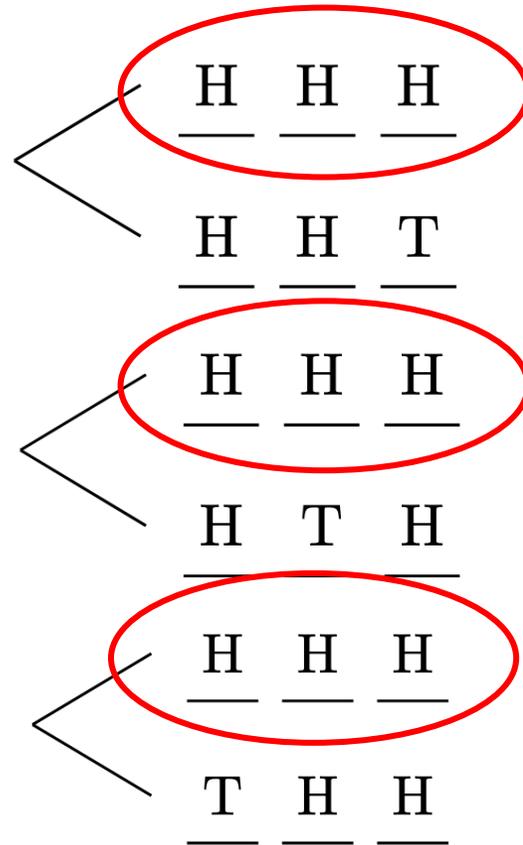
$T = \text{Lose}$

? = Could be either

H      H      ?

H      ?      H

?      H      H



We are counting this case 3 times!!

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

$$P(\text{Warriors win series}) = P(\text{win and ends in 2}) + P(\text{win and ends in 3})$$

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

$$P(\text{Warriors win series}) = P(\text{win and ends in 2}) + P(\text{win and ends in 3})$$

$P(\overset{\text{win and}}{\text{ends in 2}})$

H    H

1 way to win in 2 games.

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

$$P(\text{Warriors win series}) = P(\text{win and ends in 2}) + P(\text{win and ends in 3})$$

$P(\overset{\text{win and}}{\text{ends in 2}})$

$\frac{\text{H}}{\text{—}} \quad \frac{\text{H}}{\text{—}}$

1 way to win in 2 games.

$P(\overset{\text{win and}}{\text{ends in 3}})$

$\frac{\text{H}}{\text{—}} \quad \frac{\text{T}}{\text{—}} \quad \frac{\text{H}}{\text{—}}$   
 $\frac{\text{T}}{\text{—}} \quad \frac{\text{H}}{\text{—}} \quad \frac{\text{H}}{\text{—}}$

2 ways to win in 3 games (where you didn't already win in 2)

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

$$P(\text{Warriors win series}) = P(\text{win and ends in 2}) + P(\text{win and ends in 3})$$

$$\begin{aligned} P(\text{win and ends in 2}) &= (0.55)^2 \\ &= 0.3025 \end{aligned}$$

$\frac{H}{\quad} \quad \frac{H}{\quad}$

1 way to win in 2 games.

$$\begin{aligned} P(\text{win and ends in 3}) &= 2 \cdot (0.55)^2(1 - 0.55) \\ &= 0.27225 \end{aligned}$$

$\frac{H}{\quad} \quad \frac{T}{\quad} \quad \frac{H}{\quad}$   
 $\frac{T}{\quad} \quad \frac{H}{\quad} \quad \frac{H}{\quad}$

2 ways to win in 3 games (where you didn't already win in 2)

# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: They stop once they win!!*

---

$$P(\text{Warriors win series}) = P(\text{win and ends in 2}) + P(\text{win and ends in 3})$$

$$\begin{aligned} P(\text{win and ends in 2}) &= (0.55)^2 \\ &= 0.3025 \end{aligned}$$

H      H

1 way to win in 2 games.

$$\begin{aligned} P(\text{win and ends in 3}) &= 2 \cdot (0.55)^2(1 - 0.55) \\ &= 0.27225 \end{aligned}$$

H T H  
T H H

2 ways to win in 3 games (where you didn't already win in 2)

$$0.3025 + 0.27225 = 0.575$$



# What is the probability of winning a 3 game series?

Same setting as before – but we are going to make the problem a bit simpler to think about a few hard cases and common bugs. Warriors are playing a series, and they win if they win at least 2 out of the 3 games. *Assume: All 3 games are played, even if a team has already won 2.*

---

Let  $X$  be the number of games won.  $X \sim \text{Bin}(n = 3, p = 0.55)$ .  $P(X > 1)$ ?

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) \\ &= \binom{3}{2} p^2 (1 - p) + \binom{3}{3} p^3 (1 - p)^0 \\ &= \binom{3}{2} (0.55)^2 (0.45) + \binom{3}{3} (0.55)^3 = 0.575 \end{aligned}$$





$P(X \geq K)$  for a binomial random variable is always a sum of each value between  $k$  and  $n$ !! No closed form way.

# Planting Tomatoes



You are planting tomatoes in your garden. Each tomato seed has a 70% chance of germinating. You plant 5 seeds. What is the probability that more than 3 tomatoes sprout?

# Planting Tomatoes



You are planting tomatoes in your garden. Each tomato seed has a 70% chance of germinating. You plant 5 seeds. What is the probability that more than 3 tomatoes sprout?

Let  $X$  be the number of tomato seeds that sprout.  $X \sim \text{Bin}(n = 5, p = 0.7)$ .

# Planting Tomatoes



You are planting tomatoes in your garden. Each tomato seed has a 70% chance of germinating. You plant 5 seeds. What is the probability that more than 3 tomatoes sprout?

Let  $X$  be the number of tomato seeds that sprout.  $X \sim \text{Bin}(n = 5, p = 0.7)$ .

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

# Planting Tomatoes

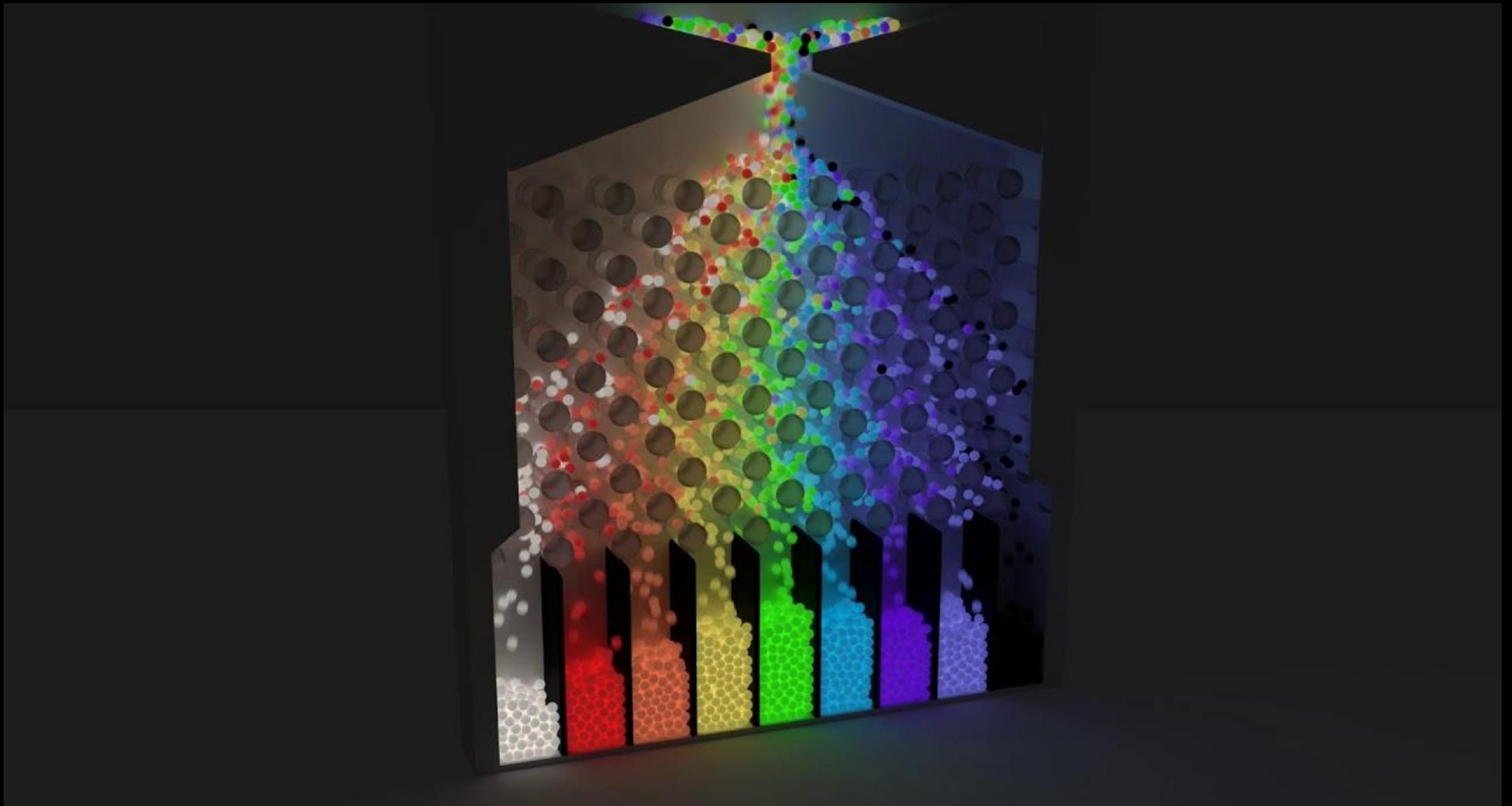


You are planting tomatoes in your garden. Each tomato seed has a 70% chance of germinating. You plant 5 seeds. What is the probability that more than 3 tomatoes sprout?

Let  $X$  be the number of tomato seeds that sprout.  $X \sim \text{Bin}(n = 5, p = 0.7)$ .

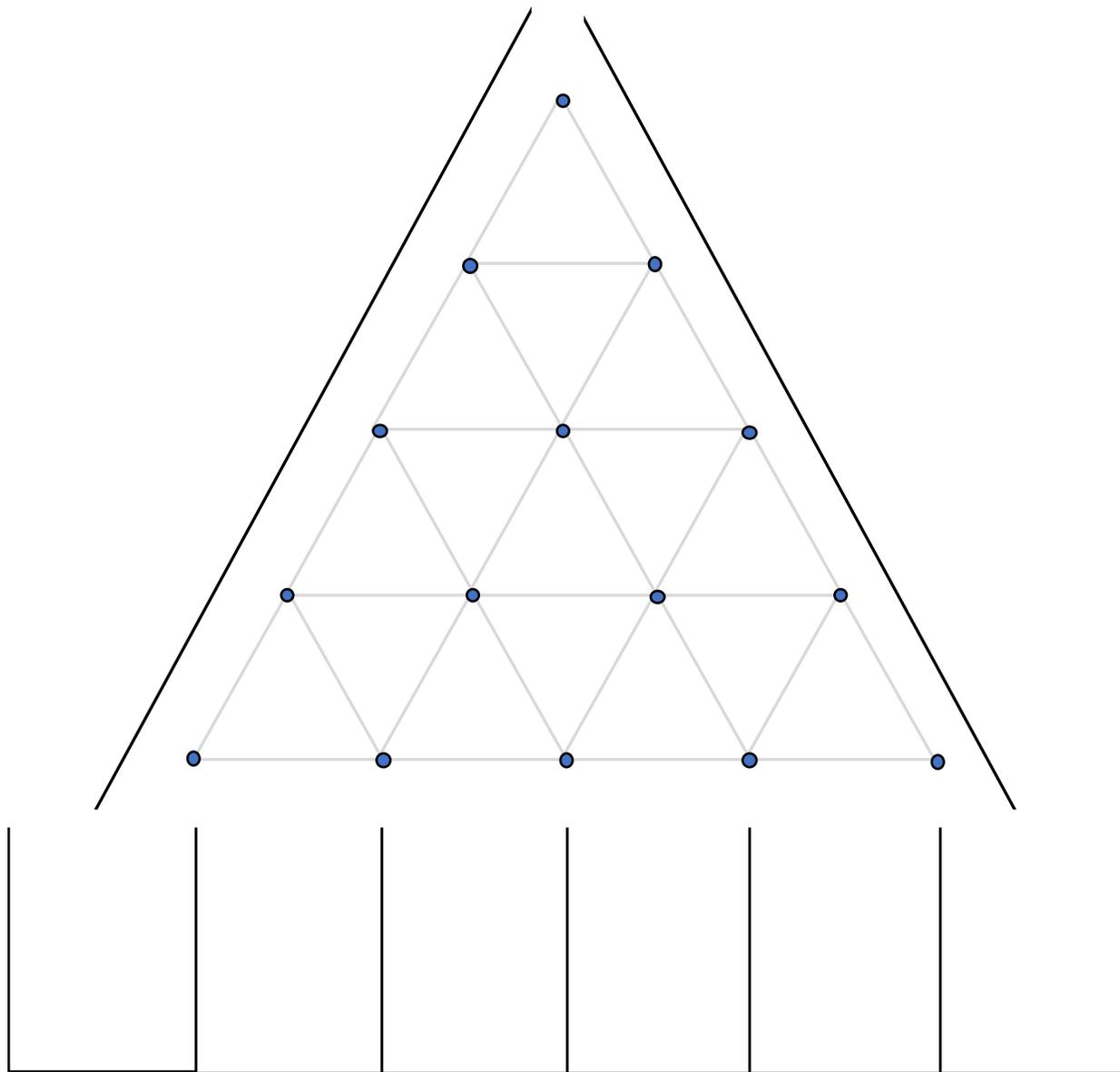
$$P(X = k) = \binom{5}{k} (0.7)^k (0.3)^{5-k}$$

$$P(X > 3) = \binom{5}{4} (0.7)^4 (0.3)^1 + \binom{5}{5} (0.7)^5 (0.3)^0 \approx 0.53$$

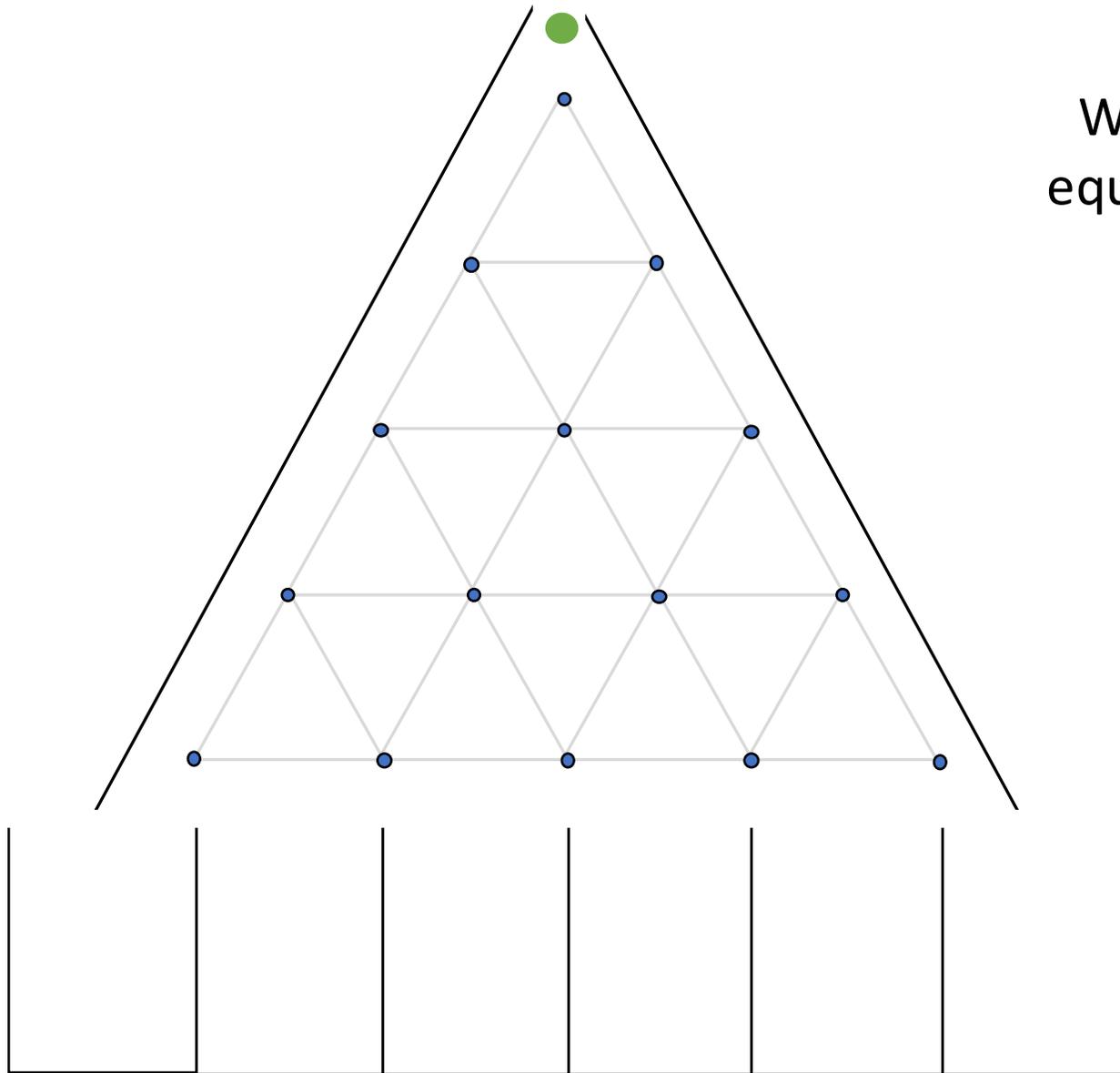


Galton Board Time!

# Galton Board Fun

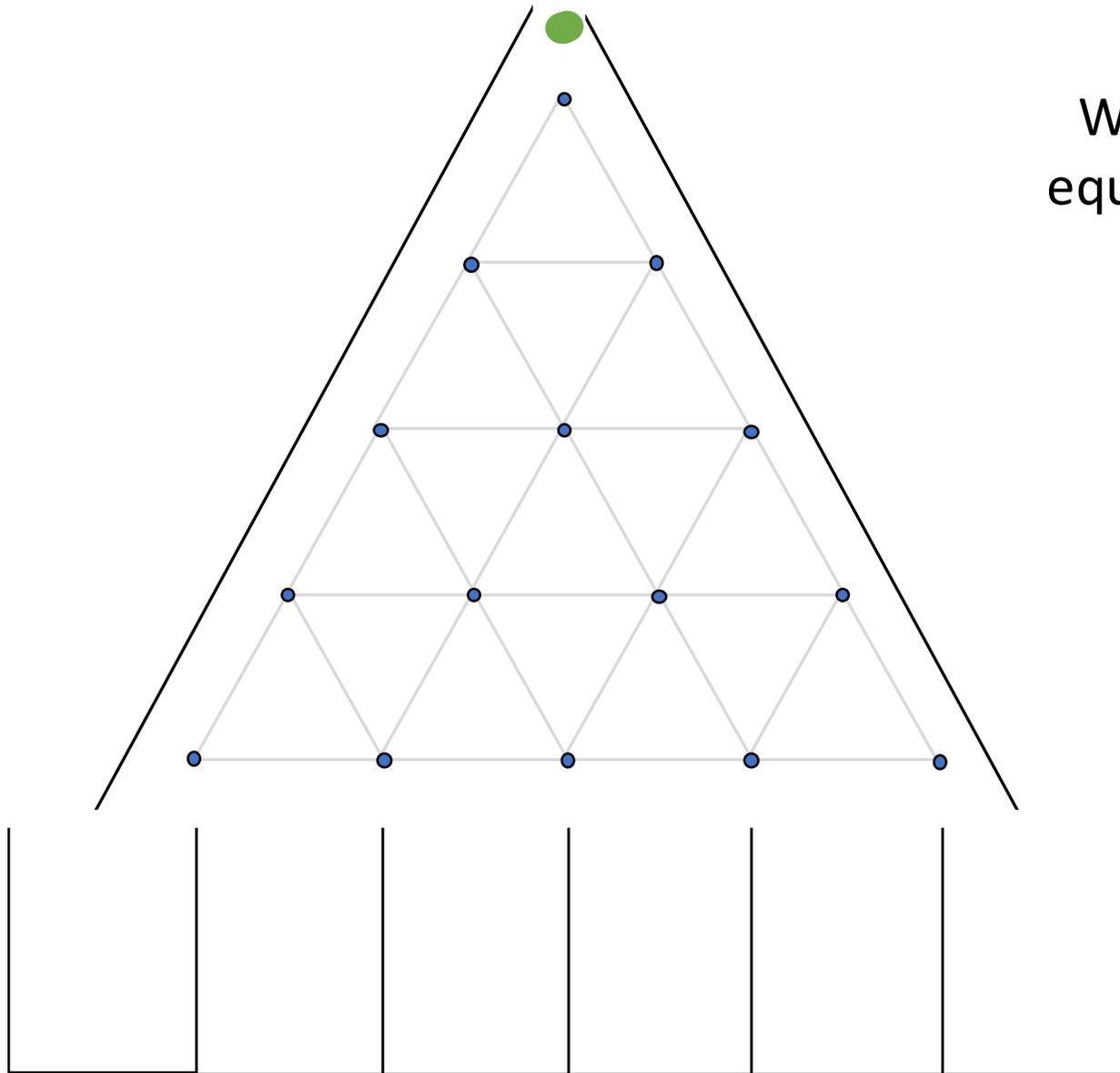


# Galton Board Fun



When a marble hits a pin, it has equal chance of going left or right.

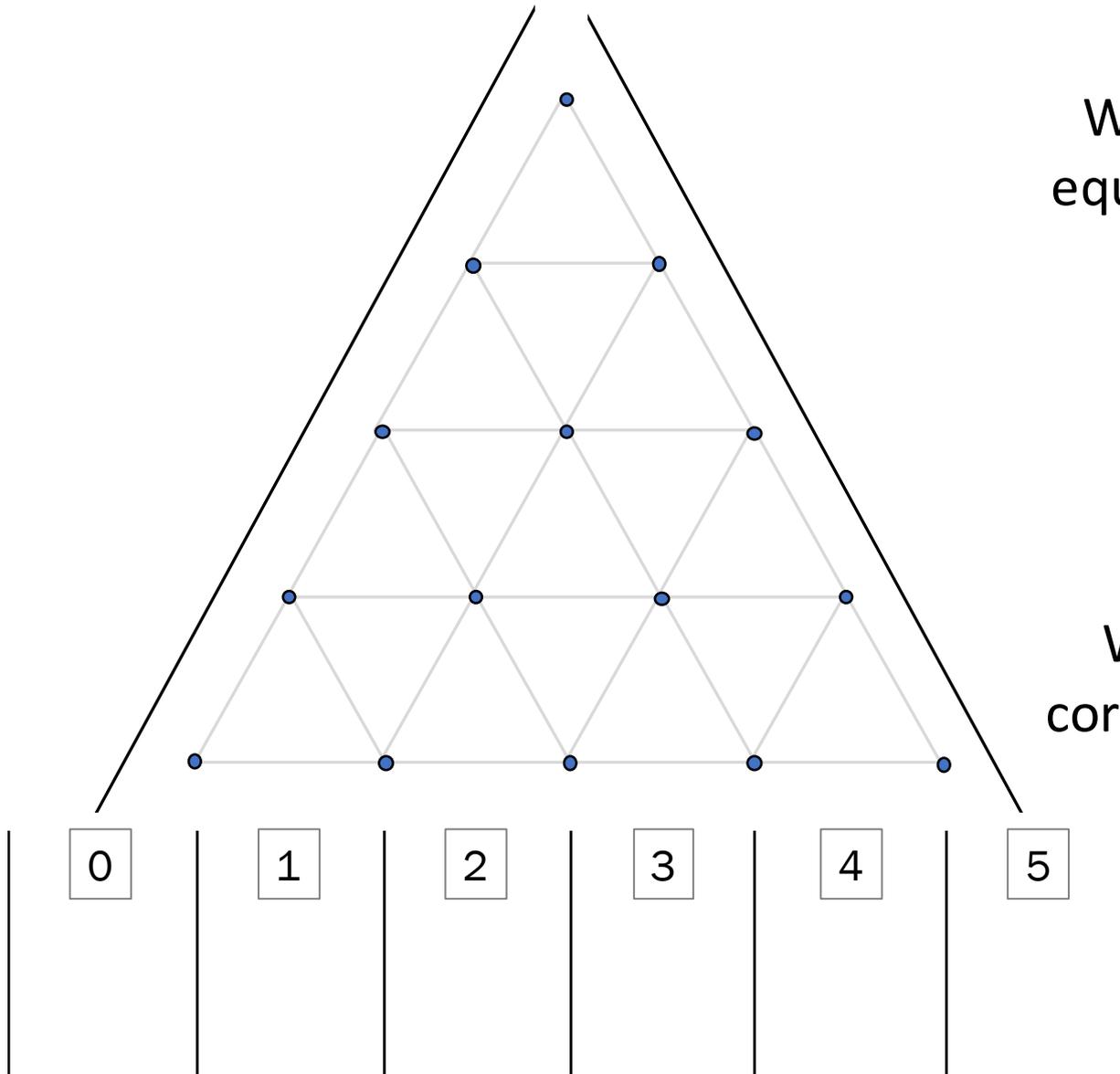
# Galton Board Fun



When a marble hits a pin, it has equal chance of going left or right.

Each pin represents an independent event.

# Galton Board Fun

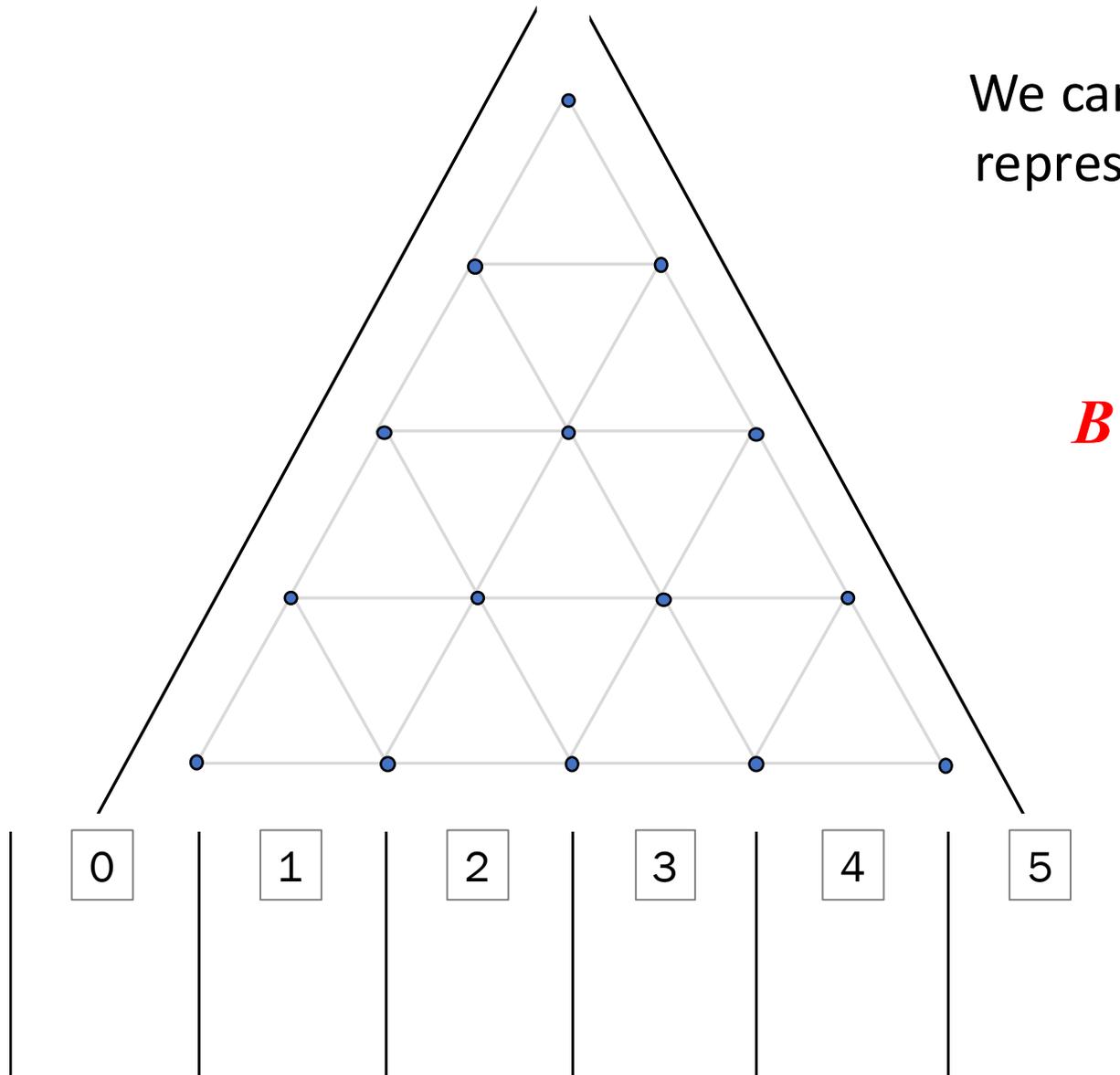


When a marble hits a pin, it has equal chance of going left or right.

Each pin represents an independent event.

Which bucket a marble lands in corresponds to the number of times the marble went right.

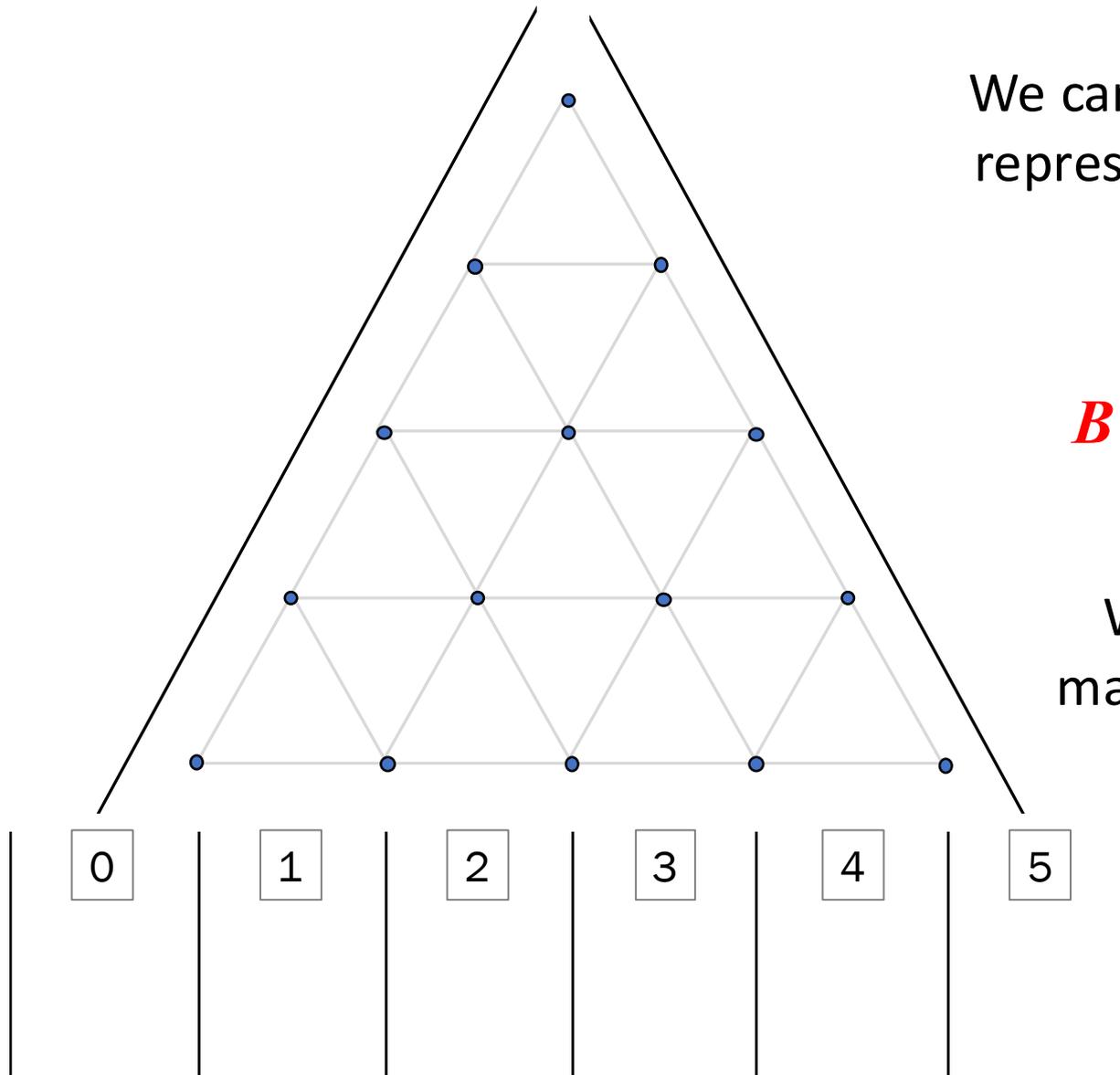
# Galton Board Fun



We can define a random variable ( $B$ ) representing which bucket a marble lands in.

$$B \sim \text{Bin}(n = \text{levels}, p = 0.5)$$

# Galton Board Fun

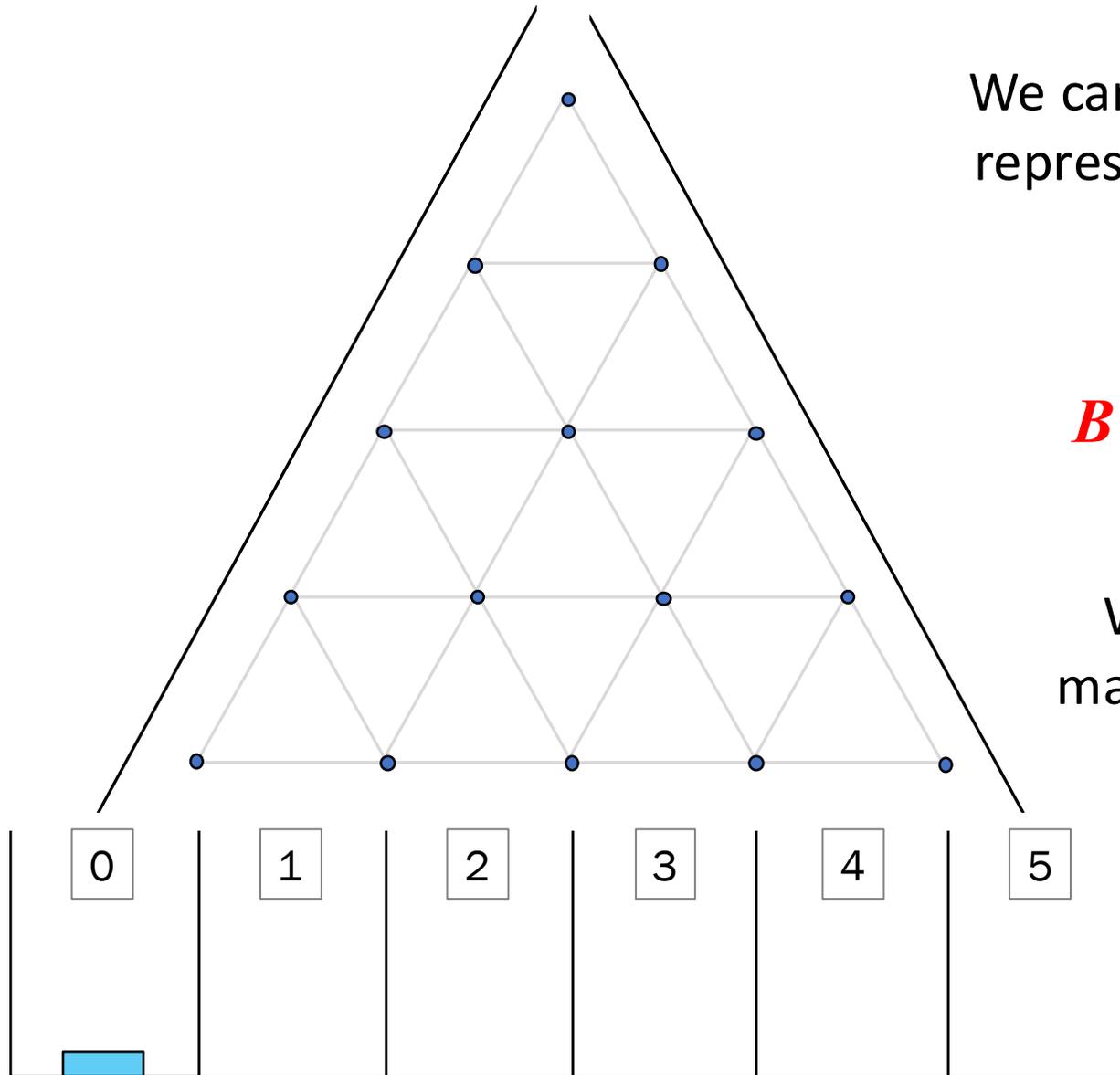


We can define a random variable ( $B$ ) representing which bucket a marble lands in.

$$B \sim \text{Bin}(n = \text{levels}, p = 0.5)$$

What is the probability of a marble landing in each bucket?

# Galton Board Fun



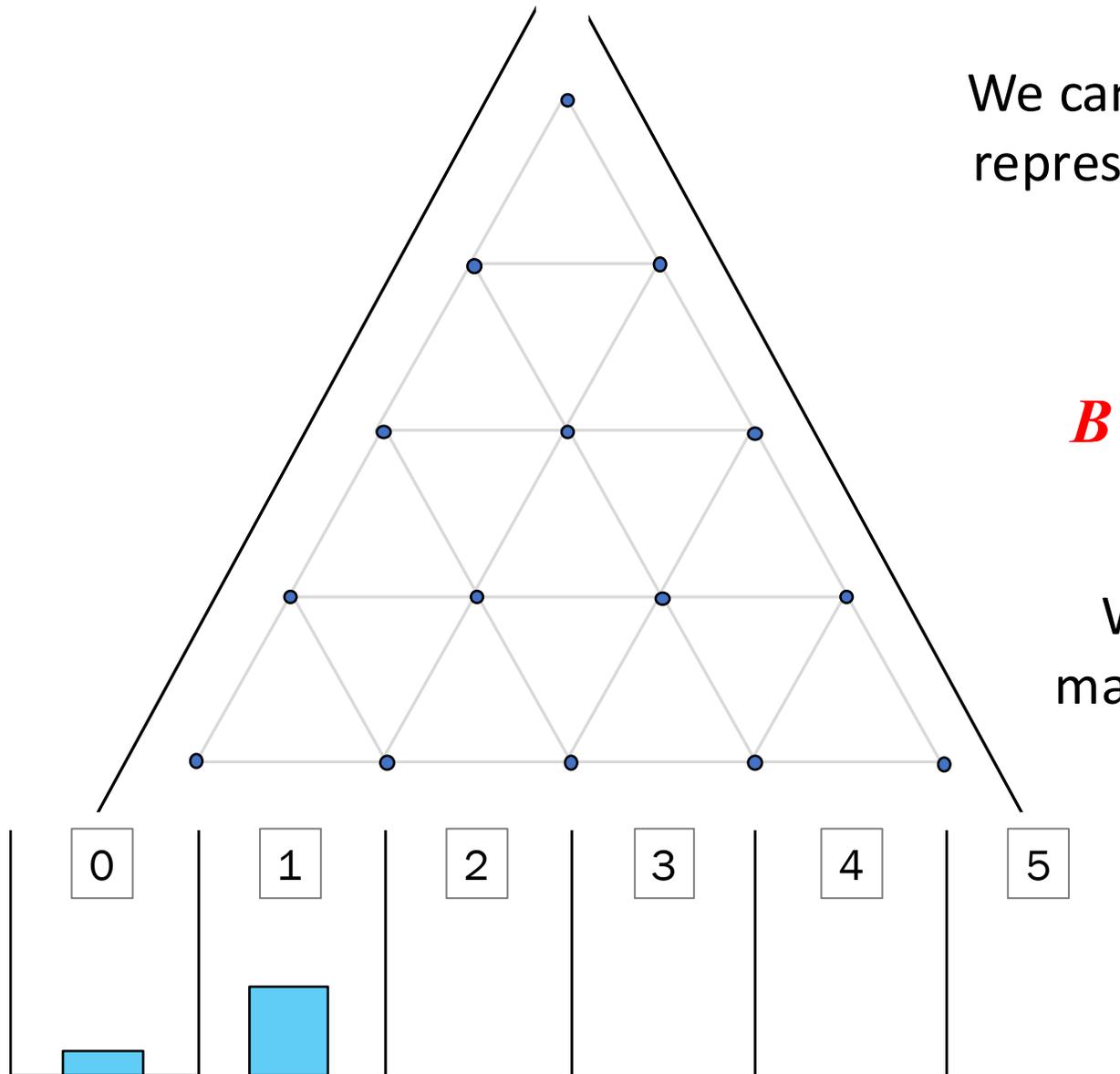
We can define a random variable ( $B$ ) representing which bucket a marble lands in.

$$B \sim \text{Bin}(n = \text{levels}, p = 0.5)$$

What is the probability of a marble landing in each bucket?

$$P(B = 0) = \binom{5}{0} \frac{1}{2}^5 \approx 0.03$$

# Galton Board Fun



We can define a random variable ( $B$ ) representing which bucket a marble lands in.

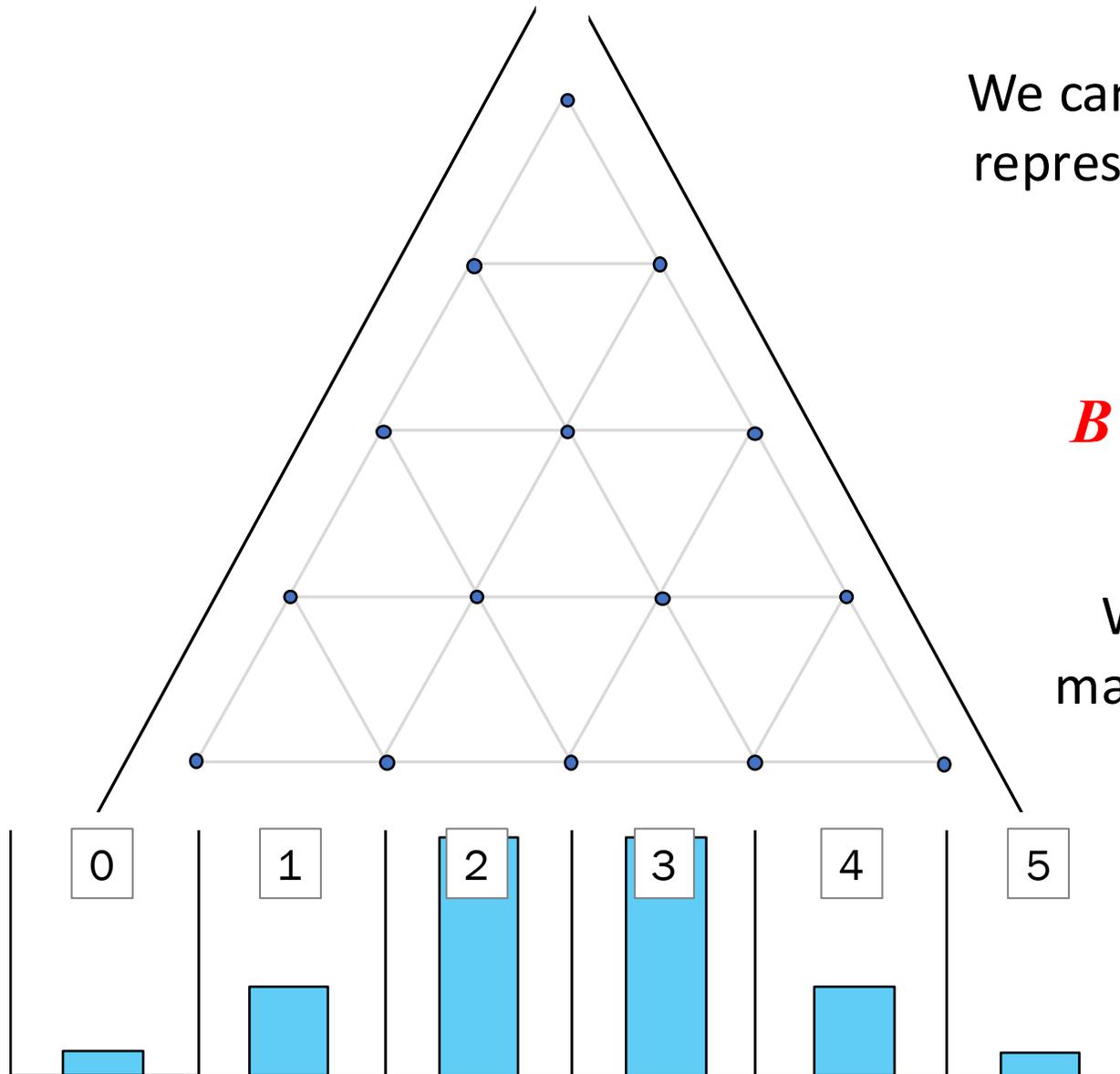
$$B \sim \text{Bin}(n = \text{levels}, p = 0.5)$$

What is the probability of a marble landing in each bucket?

$$P(B = 0) = \binom{5}{0} \frac{1}{2}^5 \approx 0.03$$

$$P(B = 1) = \binom{5}{1} \frac{1}{2}^5 \approx 0.16$$

# Galton Board Fun



We can define a random variable ( $B$ ) representing which bucket a marble lands in.

$$B \sim \text{Bin}(n = \text{levels}, p = 0.5)$$

What is the probability of a marble landing in each bucket?

This is the PMF of the binomial!

*FROM CHAOS TO ORDER*

Probability is *Everywhere*

# Learning Goals for Today



A **random variable** is a number which takes on values probabilistically.



A discrete random variable is fully described by a **probability mass function**.



A **binomial** is a particular random variable which represents number of heads in  $n$  coin flips.

See you on Friday!