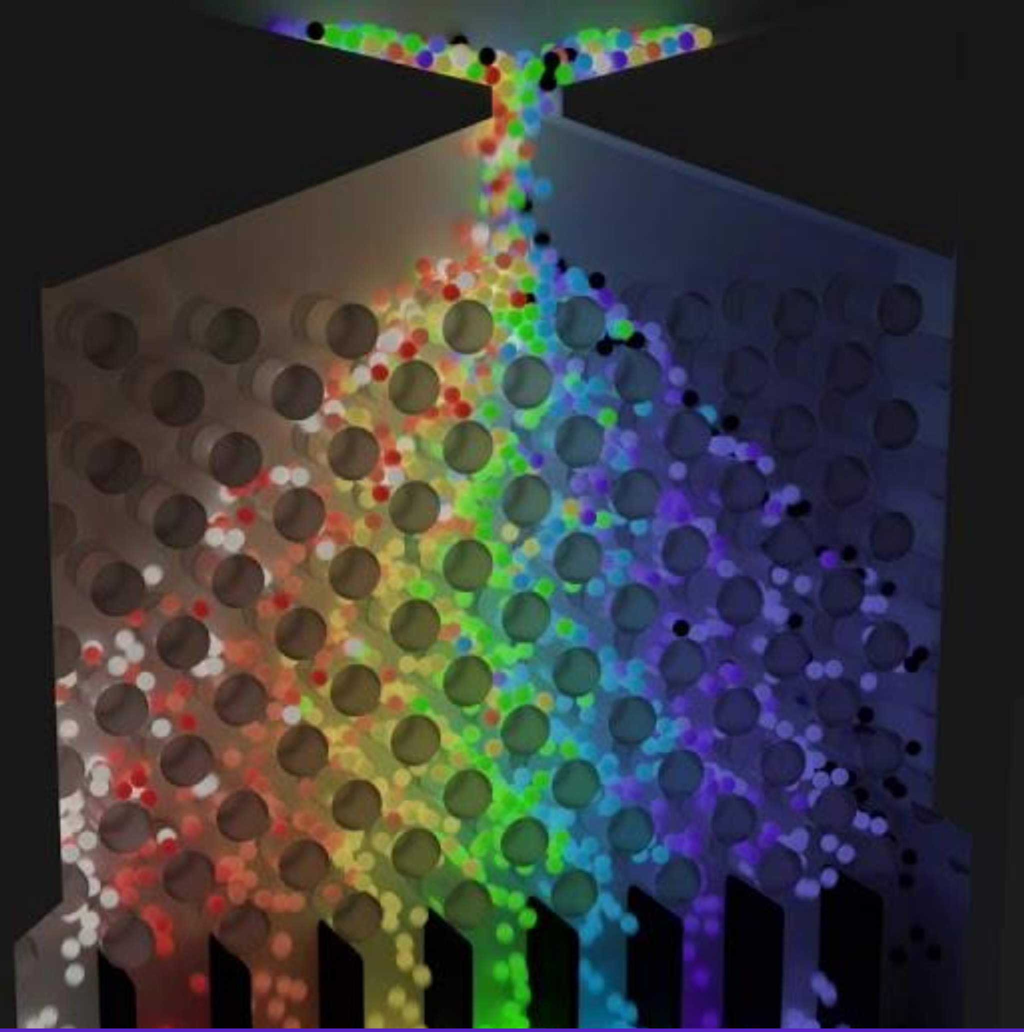


Random Variables & Binomial

CS109 – Chris Piech and Kelly!

Review

The Core Probability Toolkit ✓

The Law of Total Probability

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^C) \qquad P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i)$$

$$P(E) = P(E|F)\,P(F) + P(E|F^C)\,P(F^C) \qquad = \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^C) \cdot P(B^C)}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: $P(S) = 1$

$$P(E^C) = 1 - P(E)$$

De Morgan's Laws

$(A \text{ or } B)^C = A^C \text{ and } B^C$

$(A \text{ and } B)^C = A^C \text{ or } B^C$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:
$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$

Chain Rule

$$P(E \text{ and } F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Review: Conditional Probability (Monday)

$P(E \text{ and } F)$

Definition of Conditional Probability

Chain Rule

$P(E|F)$

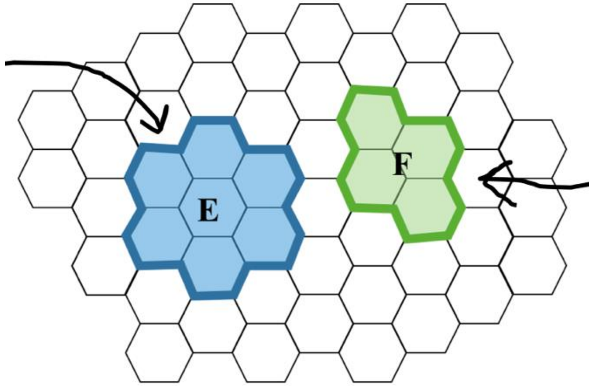
Law of Total Probability

Bayes' Theorem

$P(E)$

$P(F|E)$

Review: Last Lecture



Mutually Exclusive Events

make ***OR*** easy:

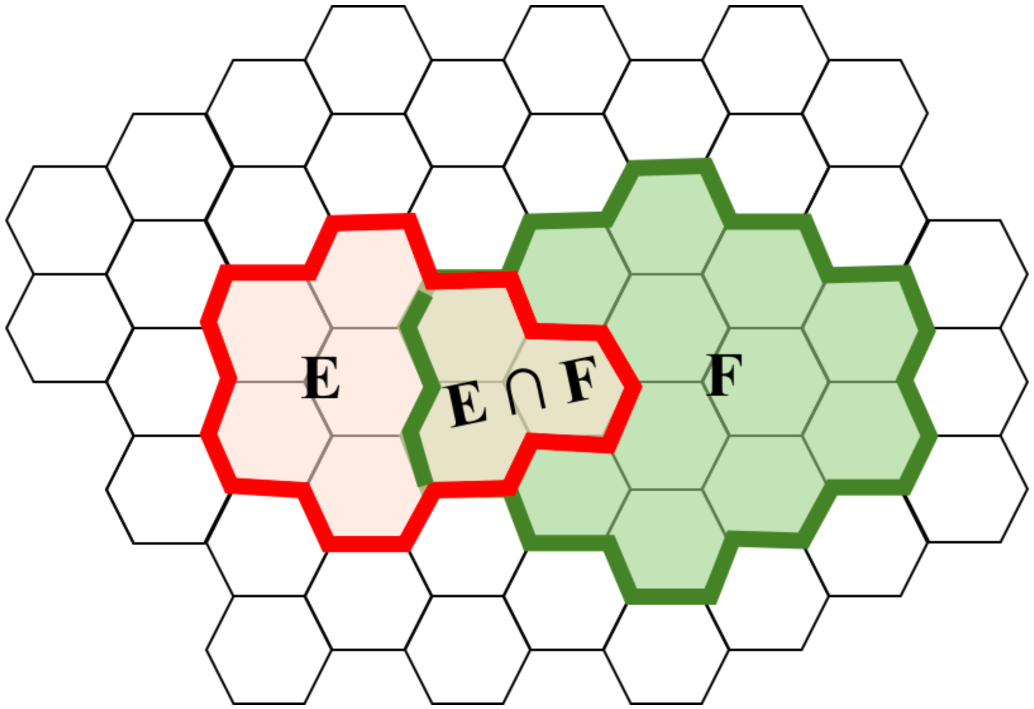
$$P(A \text{ or } B) = P(A) + P(B)$$

Independent Events

make ***AND*** easy:

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

Probability of Or *Without* Mutually Exclusive Events



If events have outcomes in common, we correct for double-counting them:

$$P(E \text{ or } F) = P(E) + P(F) - P(EF)$$

The "Inclusion Exclusion" Rule

Independence

Two events A and B are **independent** if:

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

Intuitive Definition:

Knowing that B happened doesn't change our belief that A happens.

With independence, we can simplify the chain rule:

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

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)

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

(H, H, H, T, H, 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) $\quad p^4(1-p)^6$

(H, H, H, T, H, T, T, T, T, T) $\quad p^4(1-p)^6$

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? $\binom{10}{4}$

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

End Review

What’s Missing?

What's Missing?

Events are nice, but…

- They're binary: "did E happen, yes or no"

What’s Missing?

Events are nice, but…

- They’re binary: “did E happen, yes or no”
- They’re *clunky*

What's Missing?

Events are nice, but…

- They're binary: "did E happen, yes or no"
- They're *clunky*

…this feels silly

Example: let's write out all the outcomes + probabilities of rolling a dice.

Let E_1 be the event we get a 1. $P(E_1) = 1/6$

Let E_2 be the event we get a 2. $P(E_2) = 1/6$

…

What's Missing?

Events are nice, but…

- They're binary: "did E happen, yes or no"
- They're *clunky*

…this feels silly

Example: let's write out all the outcomes + probabilities of rolling a dice.

Let E_1 be the event we get a 1. $P(E_1) = 1/6$

Let E_2 be the event we get a 2. $P(E_2) = 1/6$

…

Doesn't say that E_1, E_2,…
are mutually exclusive outcomes
from the same experiment

What's Missing?

Events are nice, but…

- They're binary: "did E happen, yes or no"
- They're *clunky*

…this feels silly

Example: let's write out all the outcomes + probabilities of rolling a dice.

Let E_1 be the event we get a 1. $P(E_1) = 1/6$

Let E_2 be the event we get a 2. $P(E_2) = 1/6$

…

Doesn't say that E_1, E_2,…
are mutually exclusive outcomes
from the same experiment

"Clunky" keeps us from recognizing
patterns across problems

What's Missing?

Events are nice, but…

- They're binary: "did E happen, yes or no"
- They're *clunky*

…this feels silly

Example: let's write out all the outcomes + probabilities of rolling a dice.

Let E_1 be the event we get a 1. $P(E_1) = 1/6$

Let E_2 be the event we get a 2. $P(E_2) = 1/6$

…

Doesn't say that $E_1, E_2,\ldots$
are mutually exclusive outcomes
from the same experiment

"Clunky" keeps us from recognizing
patterns across problems

With just events, it would be hard to scale all the way up to machine learning.

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.

```python
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.

```python
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.

```python
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.

```python
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 Variables…That Are Random

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

```python
import random

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

result = flip_coin()
```

"Let $\boldsymbol{X}$ be the result of flipping a coin."

$$P(X = 0) = 0.5$$
$$P(X = 1) = 0.5$$

- Random variables store the outcome of an experiment
- Random variables can be described by their possible outcomes + probabilities
  - Note: random variables can only be numbers (not "heads" or "tails")

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

$$X \qquad\qquad X = 2$$

Let X be a random variable

Random Variables vs. Events

It is an event when
X takes on a value

$$X \qquad X = 2 \qquad P(X = 2)$$

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."

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

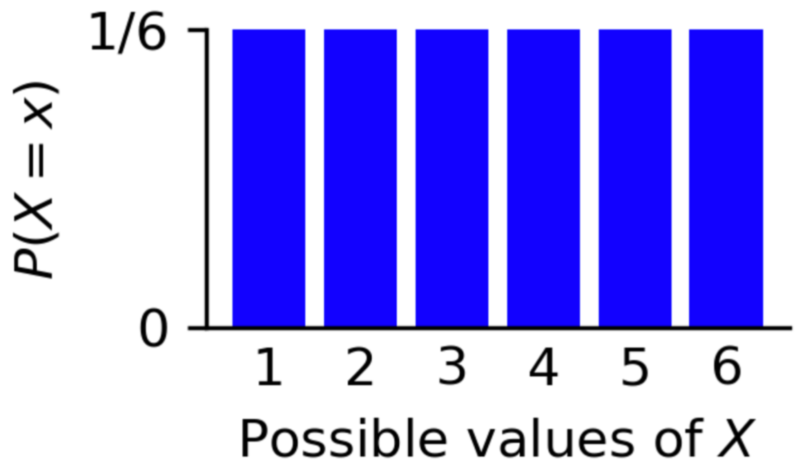
...or, $P(X = x) = 1/6$ for $1 \leq x \leq 6$

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 $1 \leq x \leq 6$

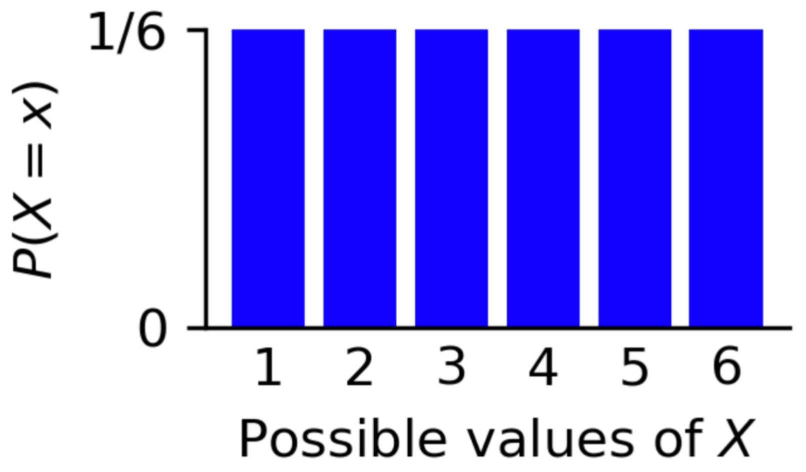


Examples of Random Variables

"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."

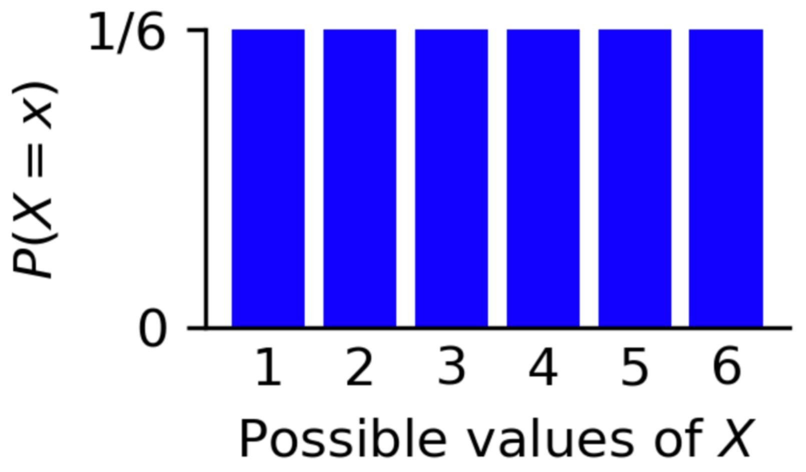
- $P(Y = 0) = 1/4$ (T, T)
- $P(Y = 1) = 1/2$ (H, T), (T, H)
- $P(Y = 2) = 1/4$ (H, H)

Examples of Random Variables

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

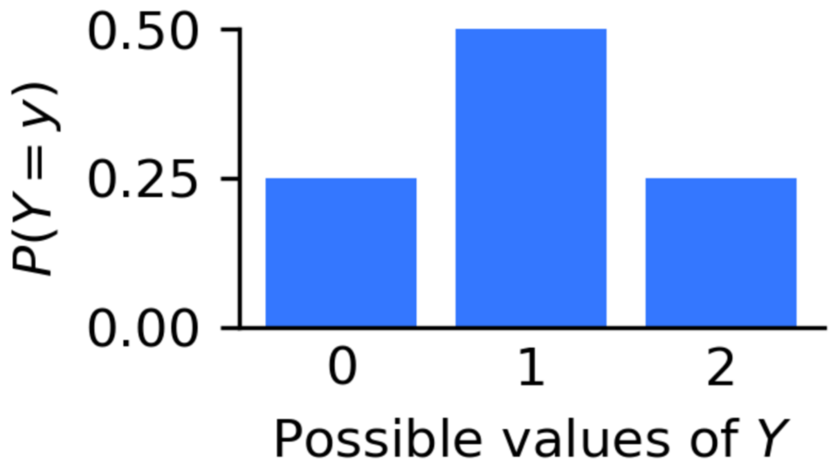
- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

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



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

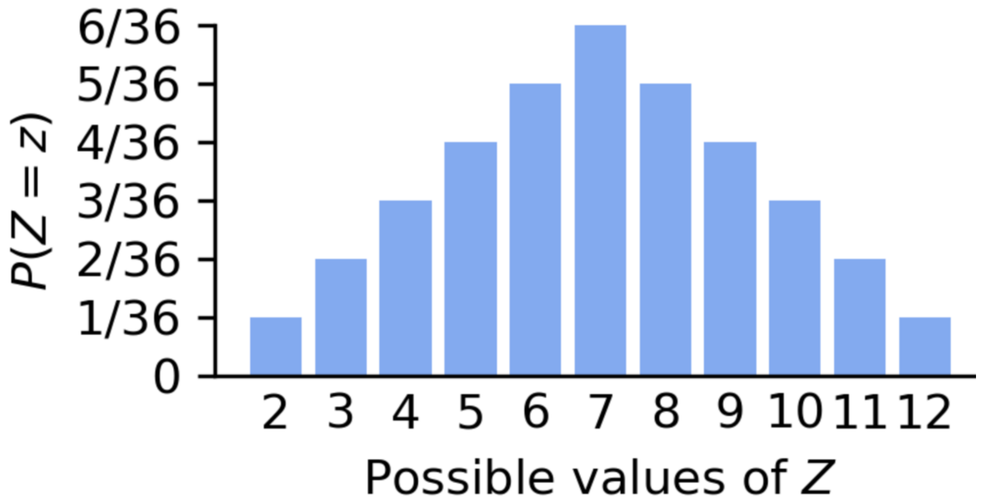
- $P(Y = 0) = 1/4$ (T, T)
- $P(Y = 1) = 1/2$ (H, T), (T, H)
- $P(Y = 2) = 1/4$ (H, H)



Examples of Random Variables

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

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



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

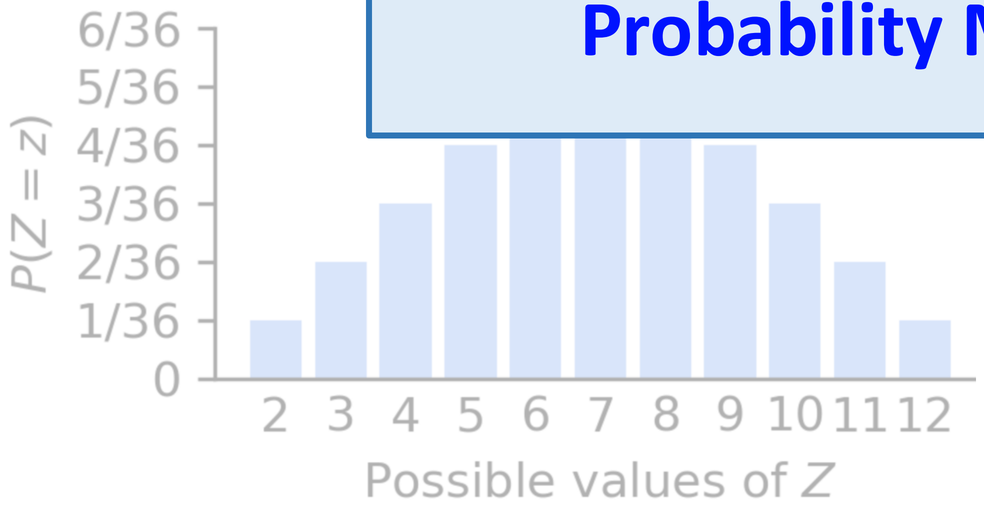
Examples of Random Variables

"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
- $P(Z =$ 5
- $P(Z = 6) = 5/36$
- $P(Z = 7) = 6/36$
- $P(Z = 10) = 3/36$
- $P(Z = 11) = 2/36$
- /36

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

Probability Mass Function (PMF)



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

Probability Mass Functions

Random Variables & Functions

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

If this is a number

$$\underbrace{P(Y = 2)}$$

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!

$$0.5$$

$$P(Y = k)$$

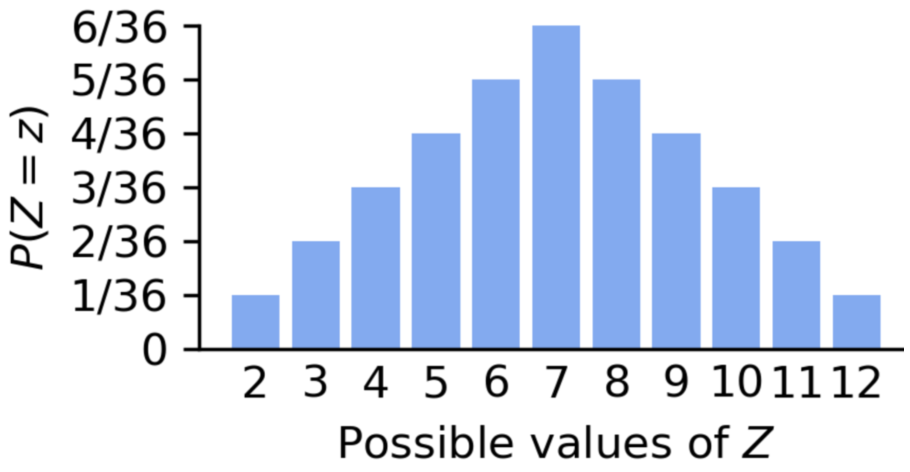
$$k = 1$$

We can put in different inputs...

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!

Quick Understanding Check

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

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:

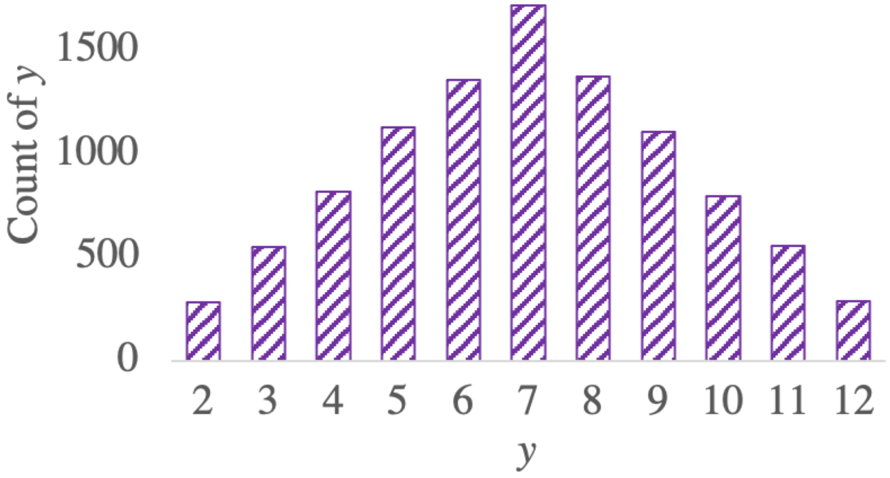
[2, 4, 5, 12, 9, 7, 9, 4, 6, 4, 3, 7, 5, 8, 9, 6, 5, 10, 10, 10, 7, 11, 4, 6, 7, 9, 10, 6, 11, 6, 5, 12, 7, 3, 11, 6, 4, 7, 8, 2, 7, 8, 6, 6, 8, 3, 2, 8, 6, 9, 5, 11, 8, 6, 9, 7, 10, 10, 10, 6, 5, 9, 4, 5, 8, 8, 6, 6, 6, 10, 4, 7, 7, 5, 7, 9, 12, 6, 7, 5, 5, 10, 7, 5, 4, 7, 6, 6, 5, 5, 8, 9, 7, 7, 7,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

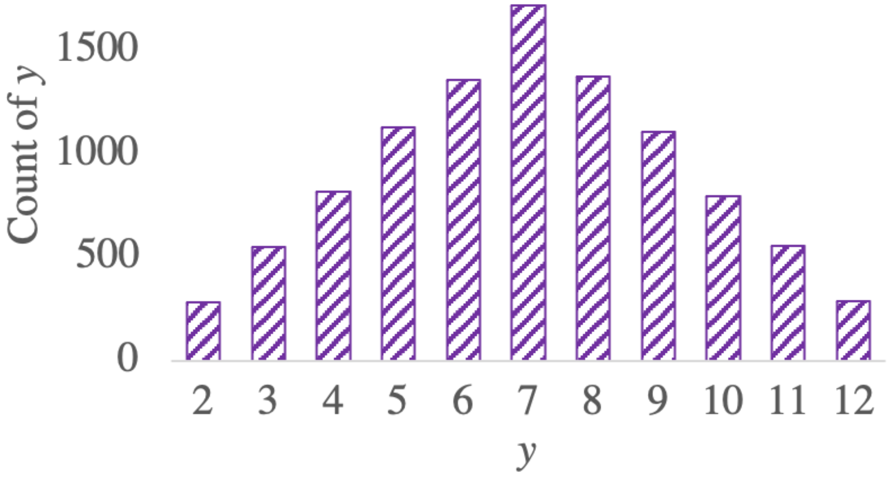


Can You Calculate A PMF From Data? Yes

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:

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

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

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

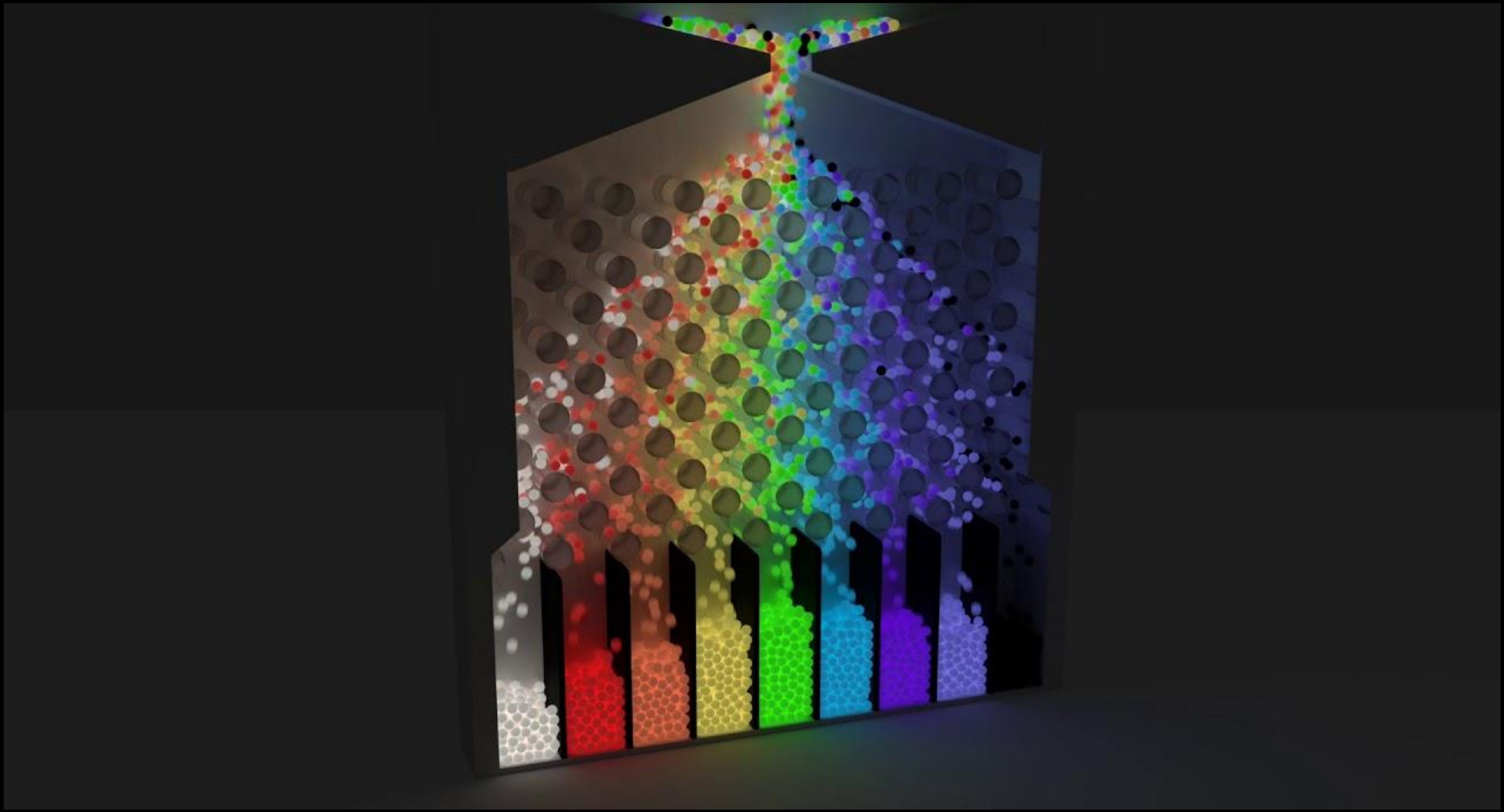
$$P(X = 60) = \frac{69^{60} e^{-69}}{60!} \approx 0.028$$

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”

Let’s Derive The Most Famous Random Variable

Let's Derive The Most Famous Random Variable

Let's Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability p of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of k cats being adopted?

Gilbert

Butters

Hashbrown

Patty

Let’s Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability *p* of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of *k* cats being adopted?

Let X be the number of cats adopted.

$$P(X = 0) = \ ?$$

no adoptions ☹

Let's Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability *p* of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of *k* cats being adopted?

Let X be the number of cats adopted.

$$P(X = 0) = (1 - p)^4$$

$$P(X = 1) = \ ?$$

(1 cat adopted)

Gilbert

or

Butters

or

Hashbrown

or

Patty

Let's Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability *p* of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of *k* cats being adopted?

Let X be the number of cats adopted.

$$P(X = 0) = (1-p)^4$$

$$P(X = 1) = 4 \cdot p(1-p)^3$$

$$P(X = 2) = \ ?$$

(2 cats adopted)


Gilbert

and


Butters

or


Butters

and


Hashbrown

or

...

Let’s Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability *p* of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of *k* cats being adopted?

Let X be the number of cats adopted.

$$P(X = 0) = (1 - p)^4$$

$$P(X = 1) = 4 \cdot p(1 - p)^3$$

$$P(X = 2) = \binom{4}{2} \cdot p^2(1 - p)^2$$

(2 cats adopted)


Gilbert

and


Butters

or


Butters

and


Hashbrown

or

...

Let's Derive The Most Famous Random Variable

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability p of being adopted this week.
- Each cat gets adopted independent of other cats.

What is the probability of k cats being adopted?

Gilbert

This feels like ***the same problem*** *as the coin flips example?!*

Recall: We flip a coin n times and count the number of heads.

1. n coin flips → n cats to adopt: ***n* independent trials**
2. Each coin flip (cat) has a **probability *p*** of being heads (adopted)
3. What we want to know: what is the probability of **exactly *k* heads (adoptions)**?

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**?

For all of these scenarios, the same math applies:

$$P(k \text{ successes}) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

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 Binomia

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 Binomia

Jacob "James" Bernoulli (1654-1705): Swiss mathematician

One of many mathematicians in the Bernoulli family

Declaring a Random Variable to be Binomial

Our random variable

Num trials

Probability of success on each trial

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

Is distributed as a

Binomial

With these parameters

Then We Automatically Know the PMF!

$$P(X = k) = \overbrace{\binom{n}{k} p^k (1-p)^{n-k}}^{\text{Probability Mass Function for a Binomial}}$$

Probability that our variable takes on the value k

Cat Adoptions: Now With Binomial

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability $p = 1/4$ of being adopted this week.
- Each cat gets adopted independent of other cats.

Let X be the number of cats adopted.

Gilbert

Butters

Hashbrown

Patty

Cat Adoptions: Now With Binomial

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability p = 1/4 of being adopted this week.
- Each cat gets adopted independent of other cats.

Let X be the number of cats adopted.

$$X \sim \mathrm{Bin}(n = 4, p = 1/4)$$

What is the probability of…

… 0 adoptions?

… 1 adoptions?

… 2 adoptions?

Gilbert

Butters

Hashbrown

Patty

Cat Adoptions: Now With Binomial

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability p = 1/4 of being adopted this week.
- Each cat gets adopted independent of other cats.

Let X be the number of cats adopted. $X \sim \text{Bin}(n = 4, p = 1/4)$

What is the probability of…

… 0 adoptions? $$P(X=0) = \binom{4}{0} p^0 (1-p)^4$$

… 1 adoptions?

… 2 adoptions?

Gilbert

Butters

Hashbrown

Patty

Cat Adoptions: Now With Binomial

Here are 4 cats at the Palo Alto animal shelter. The shelter wants to model how many of these 4 cats will get adopted in the next week.

- Each cat has probability p = 1/4 of being adopted this week.
- Each cat gets adopted independent of other cats.

Let X be the number of cats adopted. $X \sim \mathrm{Bin}(n = 4, p = 1/4)$

What is the probability of…

… 0 adoptions? $P(X=0) = \binom{4}{0} p^0 (1-p)^4$

… 1 adoptions? $P(X=1) = \binom{4}{1} p^1 (1-p)^3$

… 2 adoptions? $P(X=2) = \binom{4}{2} p^2 (1-p)^2$

Gilbert

Butters

Hashbrown

Patty

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

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

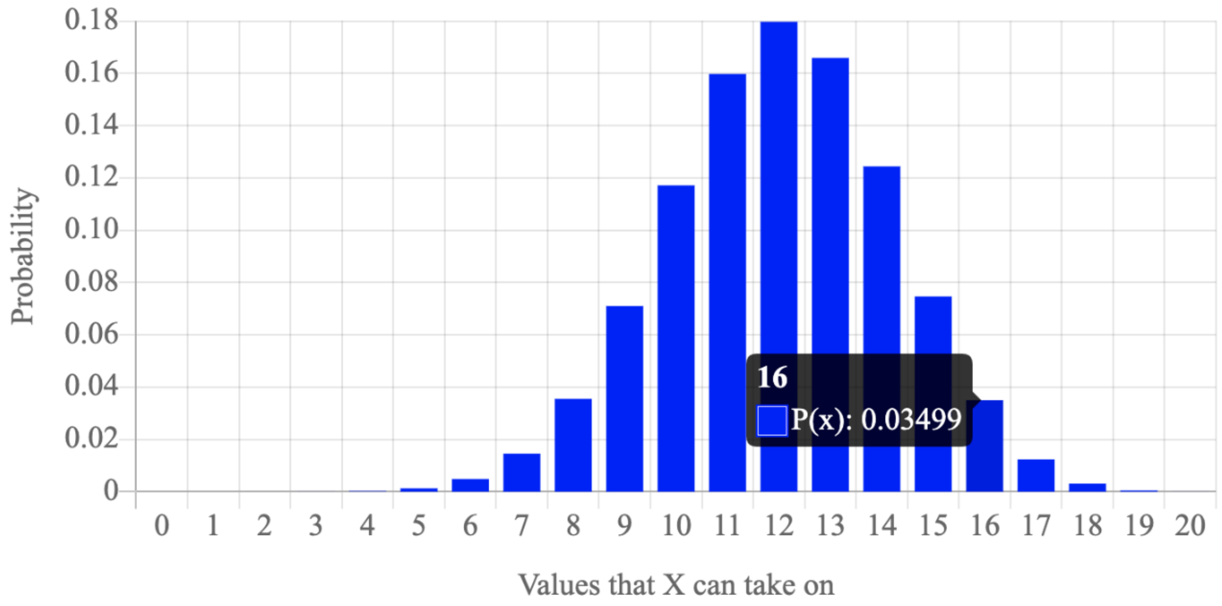
3. Profit off the PMF

You Get So Much For Free!

Binomial Random Variable

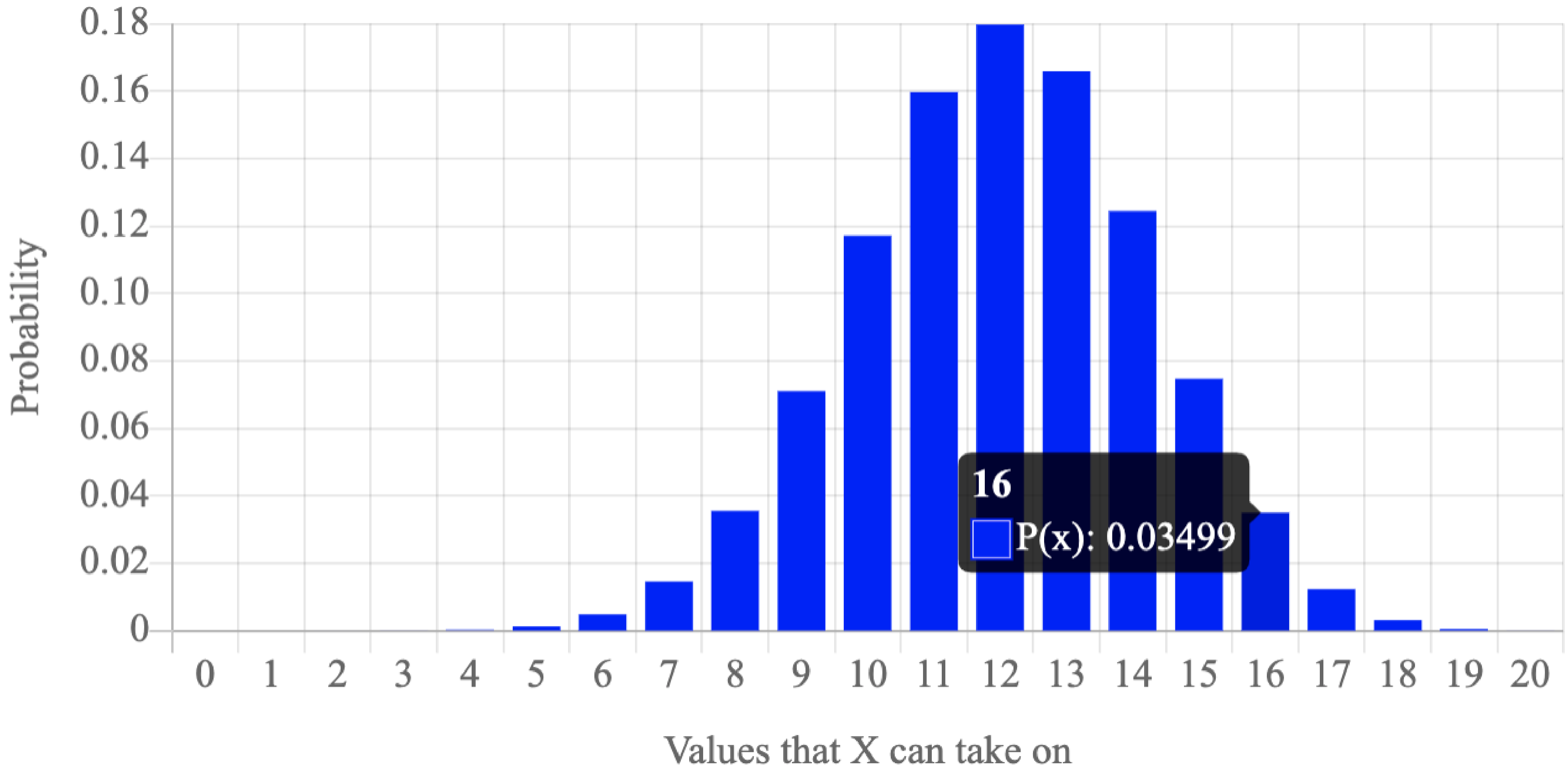
Notation:	$X \sim \text{Bin}(n, p)$
Description:	Number of "successes" in n identical, independent experiments each with probability of success p.
Parameters:	$n \in \{0, 1, \ldots\}$, the number of experiments. $p \in [0, 1]$, the probability that a single experiment gives a "success".
Support:	$x \in \{0, 1, \ldots, n\}$
PMF equation:	$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$
Expectation:	$\text{E}[X] = n \cdot p$
Variance:	$\text{Var}(X) = n \cdot p \cdot (1-p)$
PMF graph:	

Parameter n: 20 Parameter p: 0.60

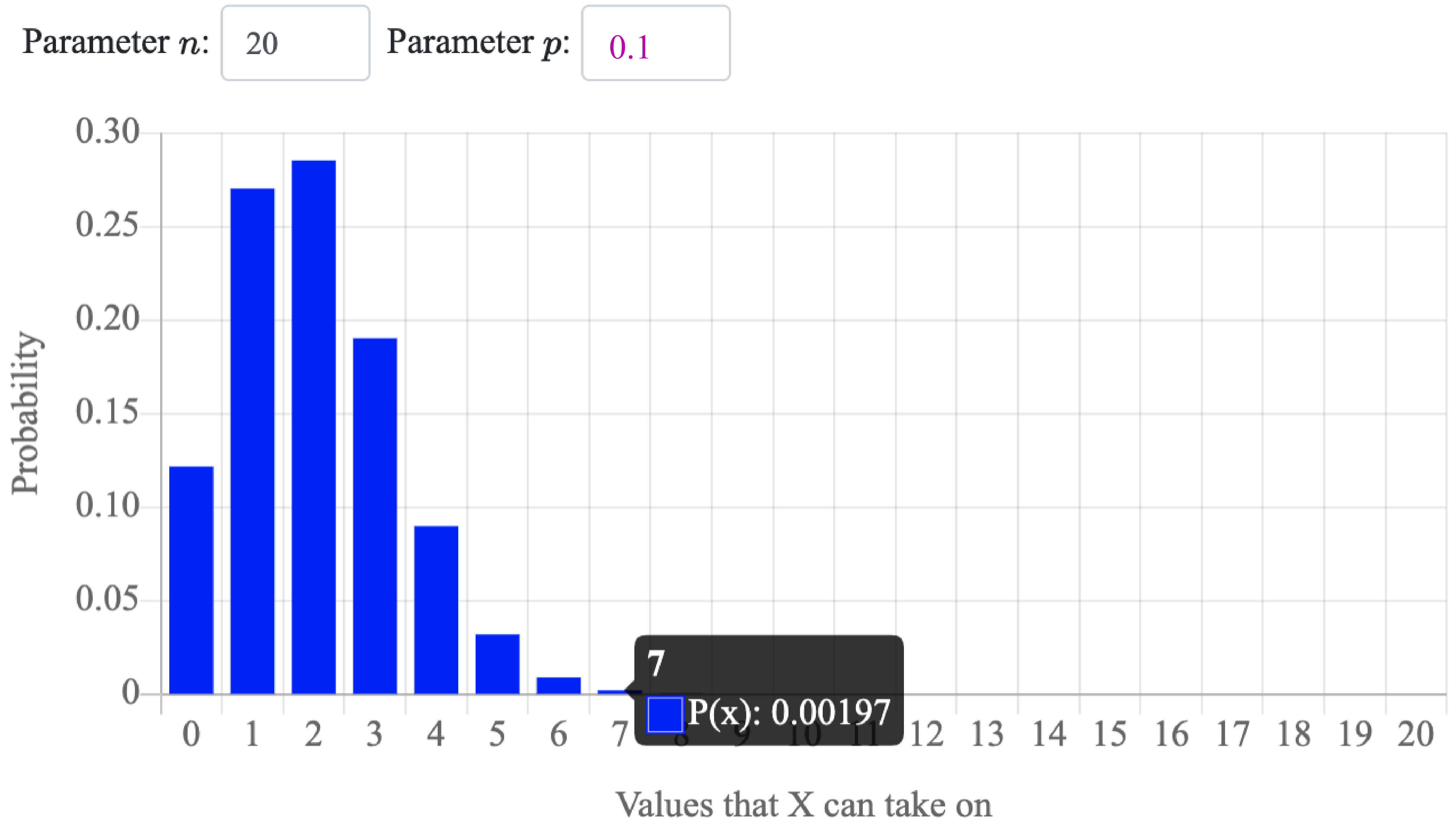


The PMF as a Graph: $X \sim \text{Bin}(n = 20, p = 0.6)$

Parameter n: 20 Parameter p: 0.60

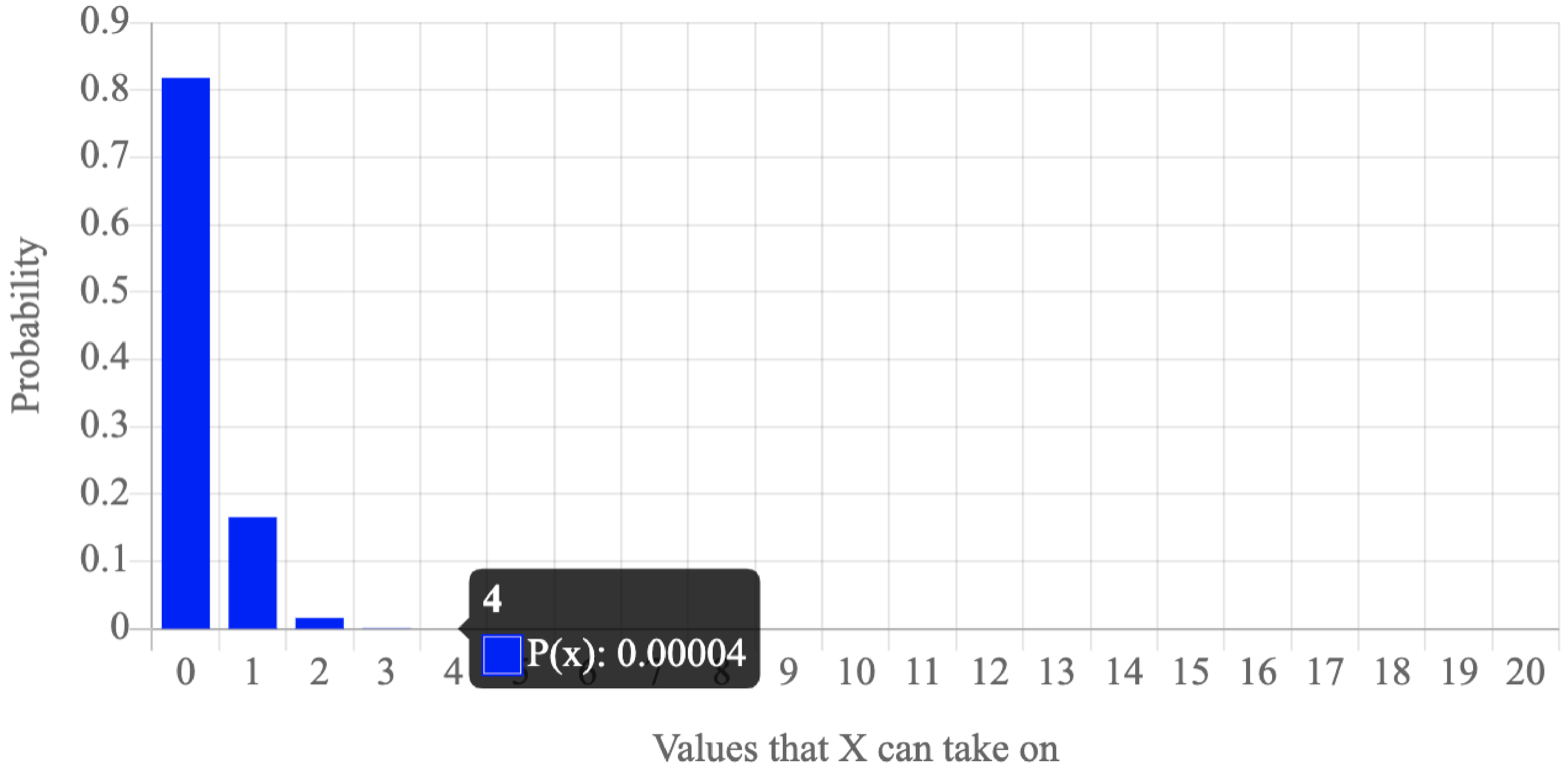


The PMF as a Graph: $X \sim \text{Bin}(n = 20, p = 0.1)$



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

Parameter n: 20 Parameter p: 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 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?

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

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

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

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

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
>>> stats.binom.pmf(10, 1000, 0.01)
0.12574021112620 75
>>> stats.binom.pmf(20, 1000, 0.01)
0.0017918782400182195
```

k n p

Practice: Ad Clicks

Every day, Youtube shows a particular ad 1000 times.

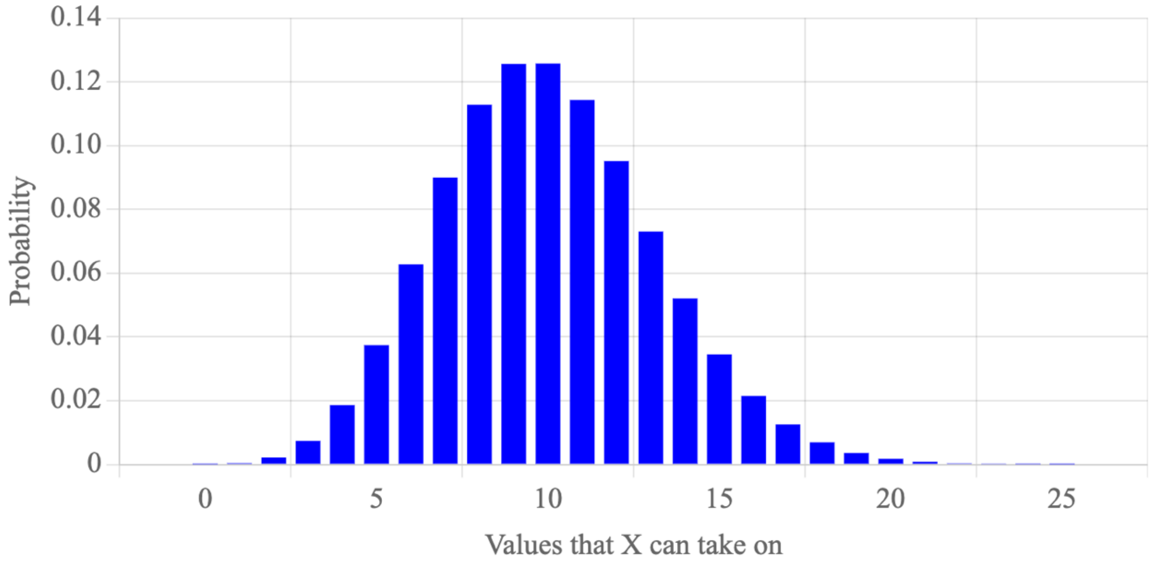
Each ad served is clicked with p = 0.01 (otherwise it's ignored).

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

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

PMF graph:

Parameter n: 1000 Parameter p: 0.01



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

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*** ~ 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*** ~ 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.

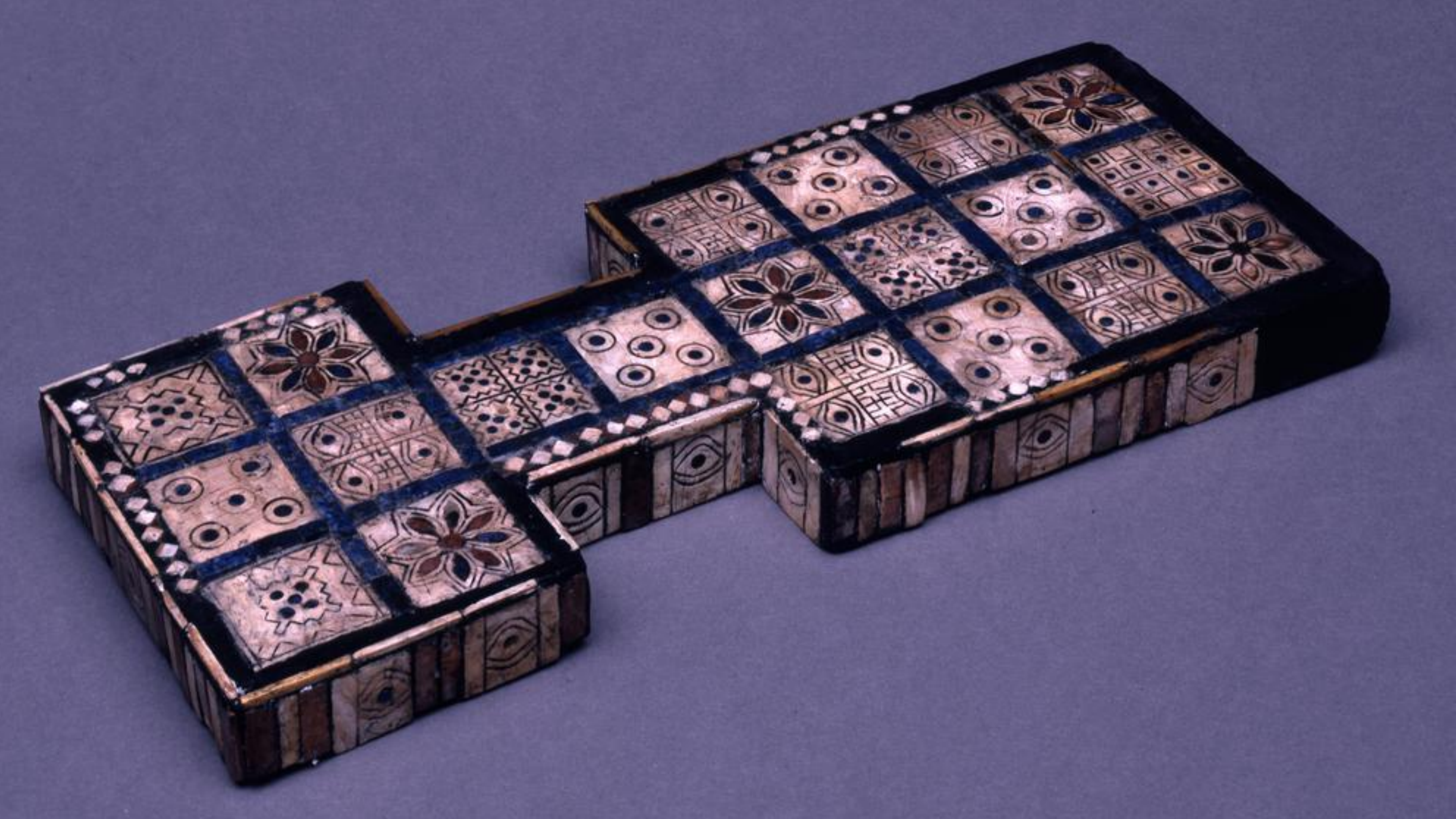
The probability of any server working is 0.8.

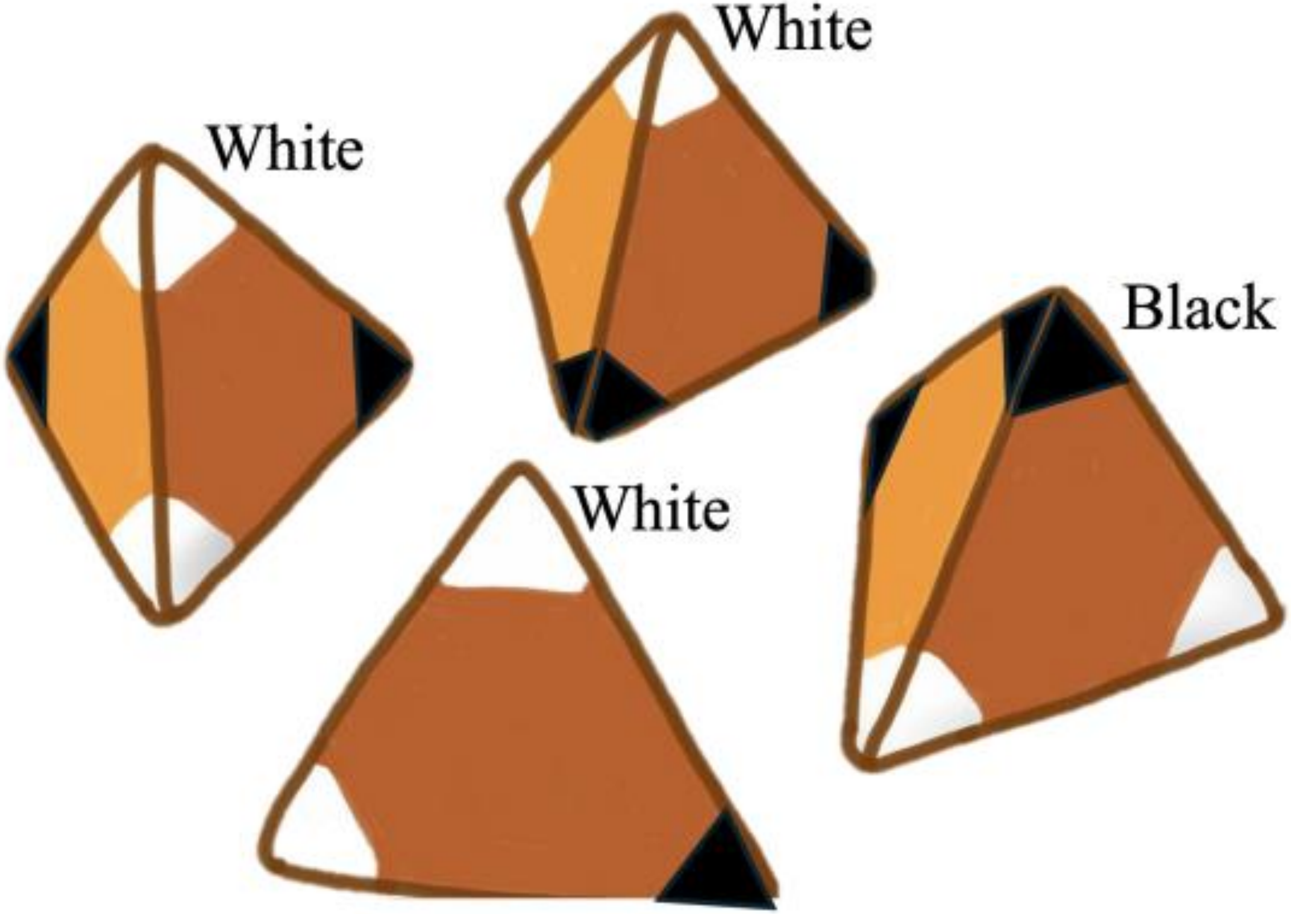
What is the probability that less than 2 servers are alive?

Let ***X*** be the number of servers alive. ***X*** ~ Bin(n = 7, p = 0.8).

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

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




White
White
White
Black

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

1. Would you rather have unlimited time or 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


6
Rakuten
30

What is the probability of winning a 7 game series?

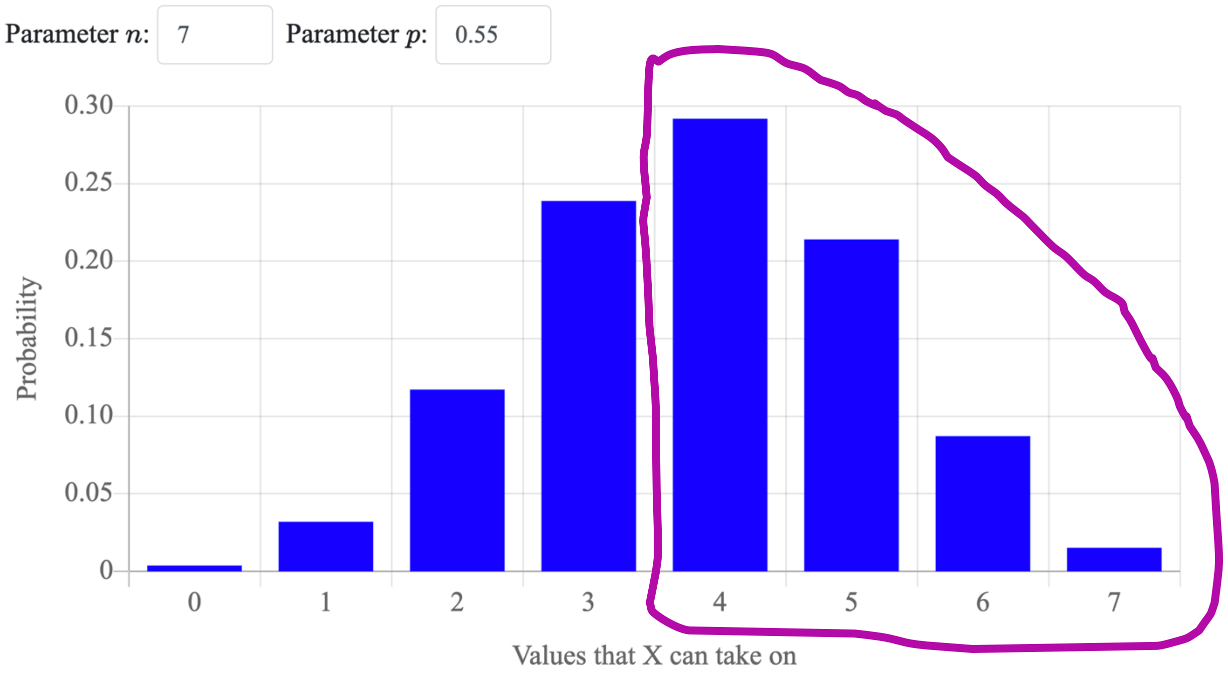
Warriors are going to play the Celtics in a best of 7 series during the 2050 NBA finals. What is the probability that the warriors win the series? Each game is **independent**. Each game, the warriors have a 0.55 probability of winning. Win series if you win at least 4 games.

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.

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.

Let X be the number of games won. X ~ Bin(n= 7, p = 0.55). P(X > 3)?

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

Why are we playing imaginary games?
What if we just considered cases where the tournament stops when one team has 4 wins?

What is the probability of winning a 7 game series?

Warriors are going to play the Bucks in a best of 7 series during the 2022 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.

Warning: these are **not** mutually exclusive

4 wins in a row: Win, Win, Win, Win

4 wins, 1 Loss: Win, Win, Win, Win, Loss

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*
- $P(X \geq k) =$

Chose slots for success, don't care about rest

$$\binom{n}{k} p^k$$

Probability that each is success

ways to choose slots for 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*
- $P(X \geq k) =$

Chose slots for success, don't care about rest

$$\binom{n}{k} p^k$$

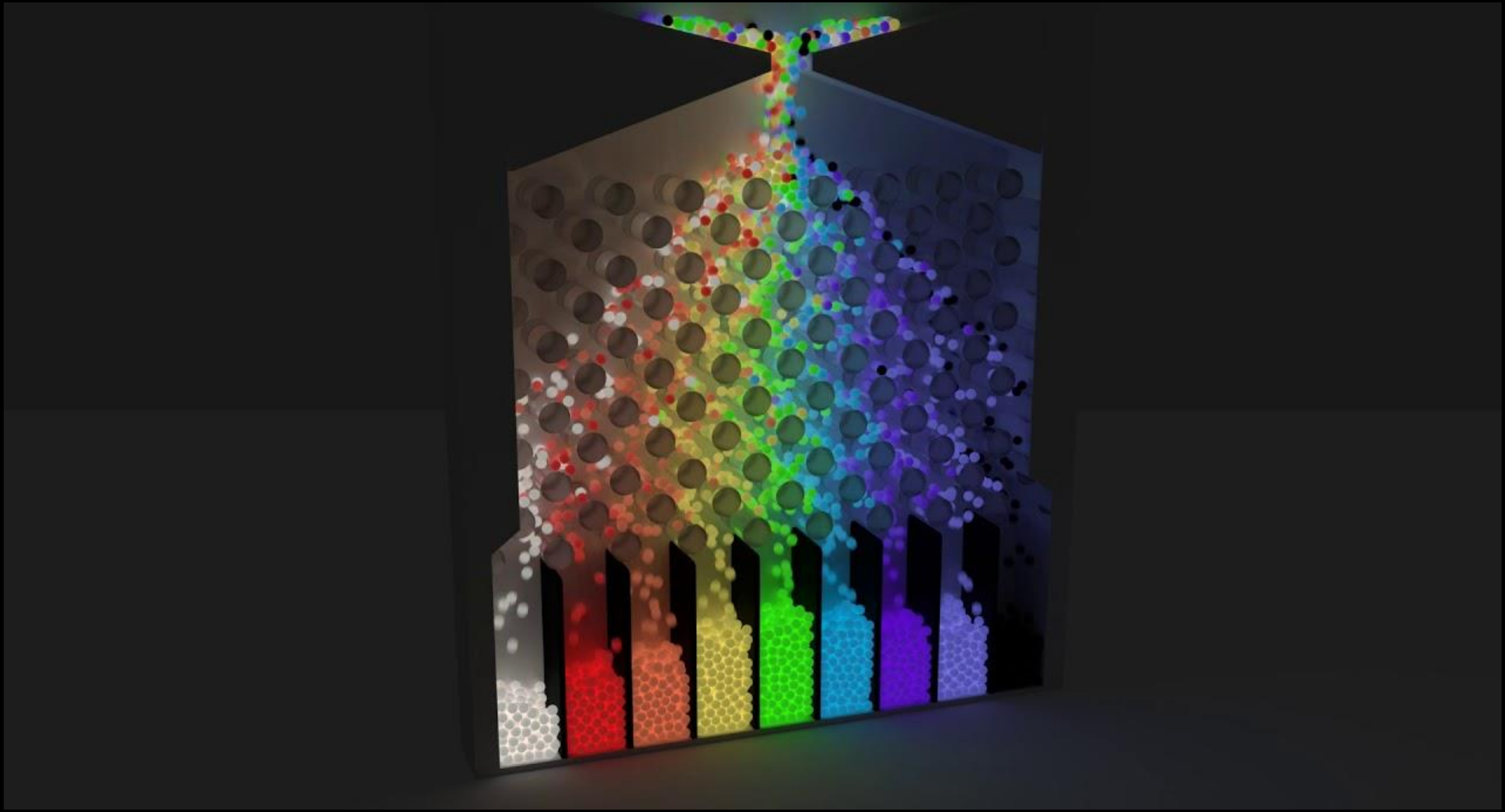
Probability that each is success

ways to choose slots for success

First clue that something is wrong. Think about *p* = 1

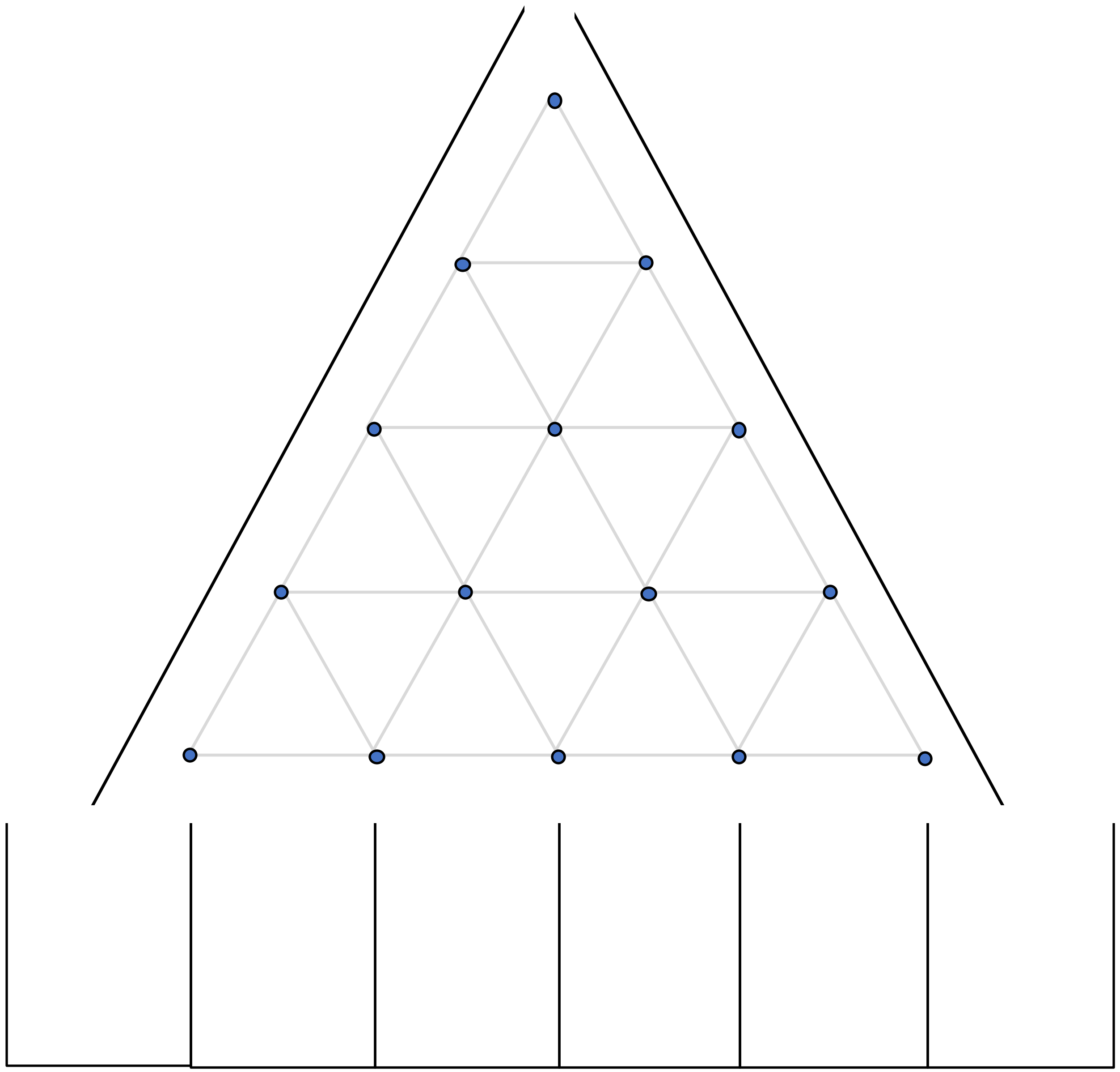
Not mutually exclusive…

Correct: $P(X \geq k) = \sum_{i=k}^{n} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-1}$

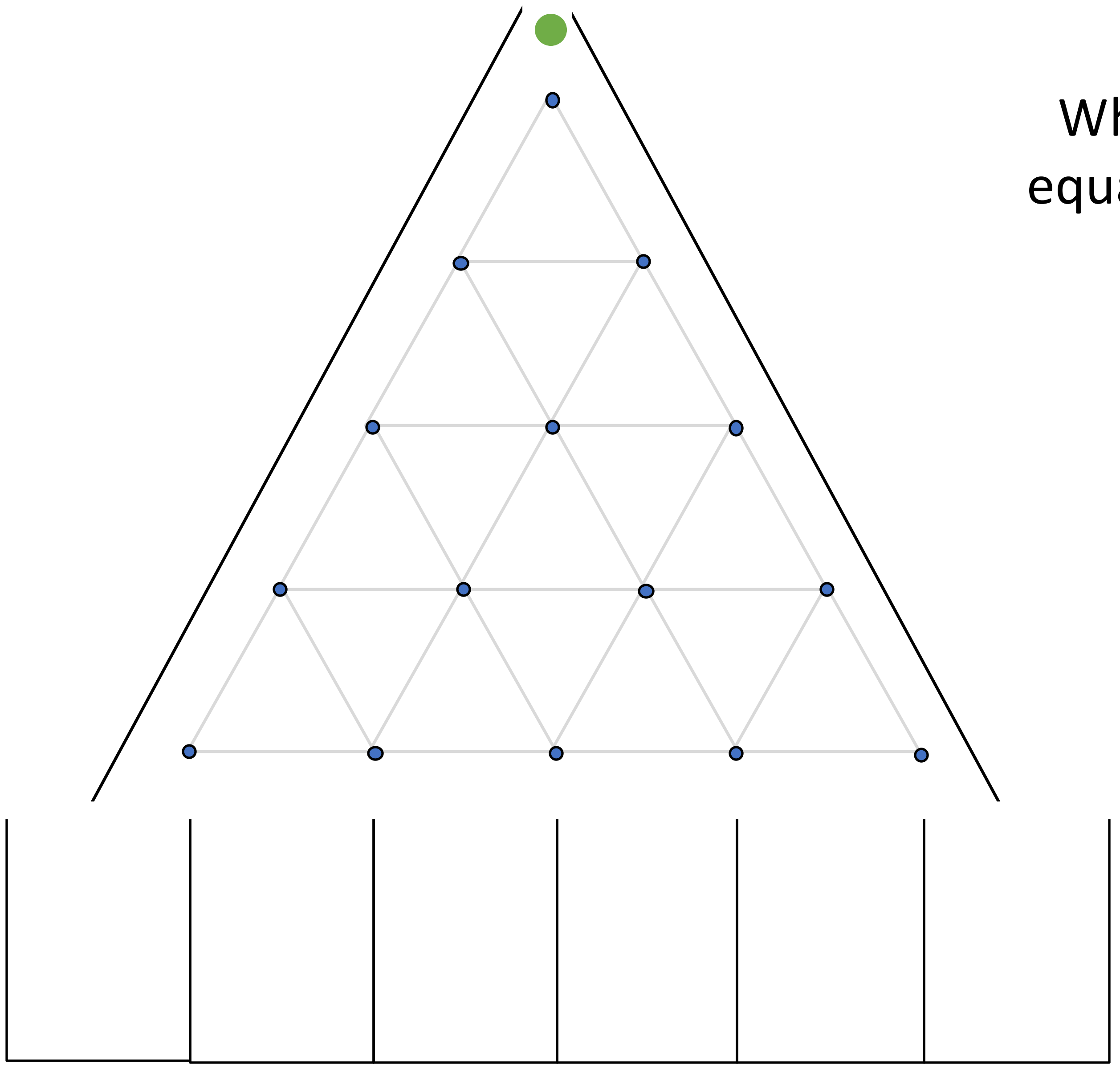


Galton Board Time!

Galton Board Fun

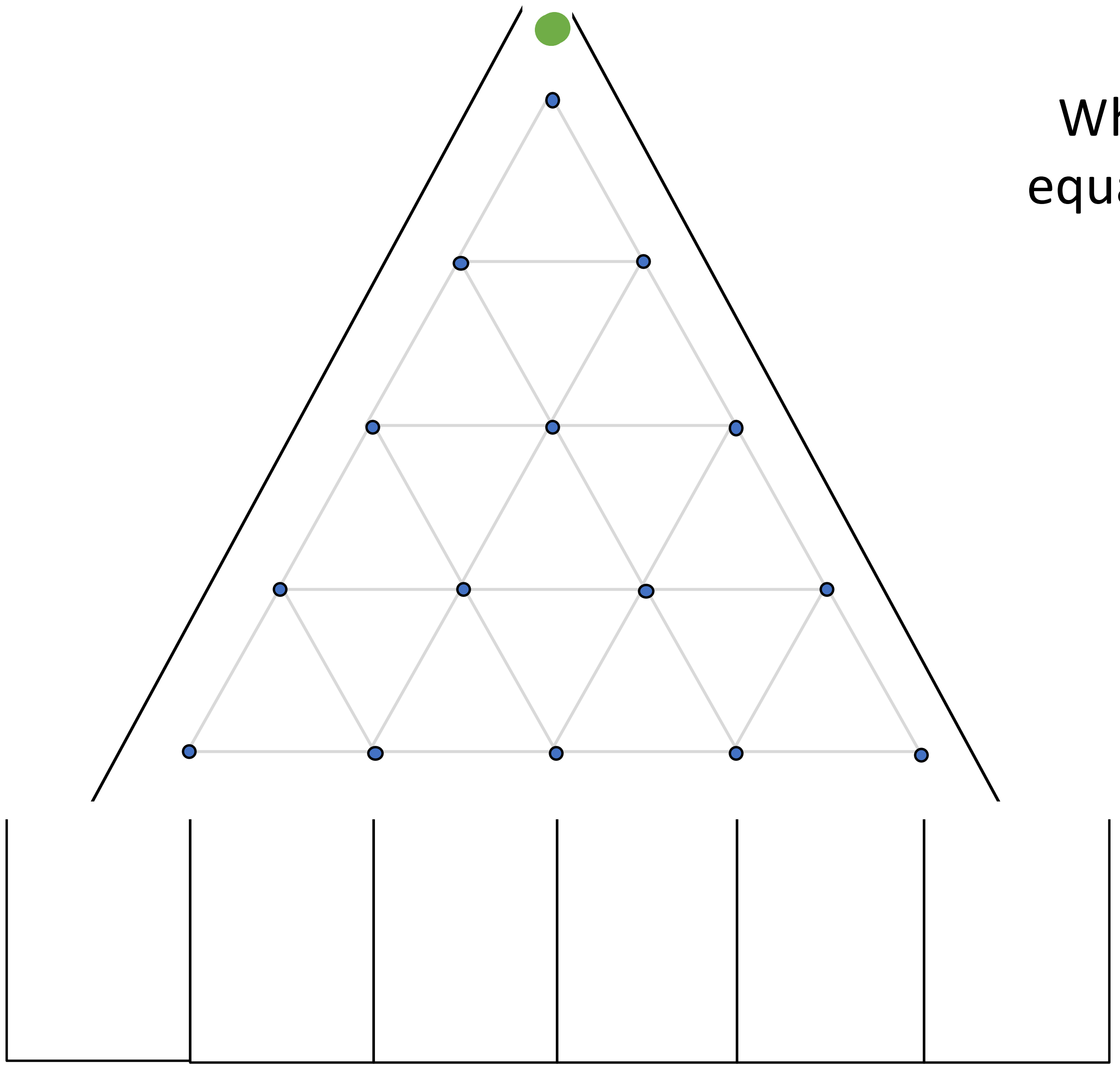


Galton Board Fun



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

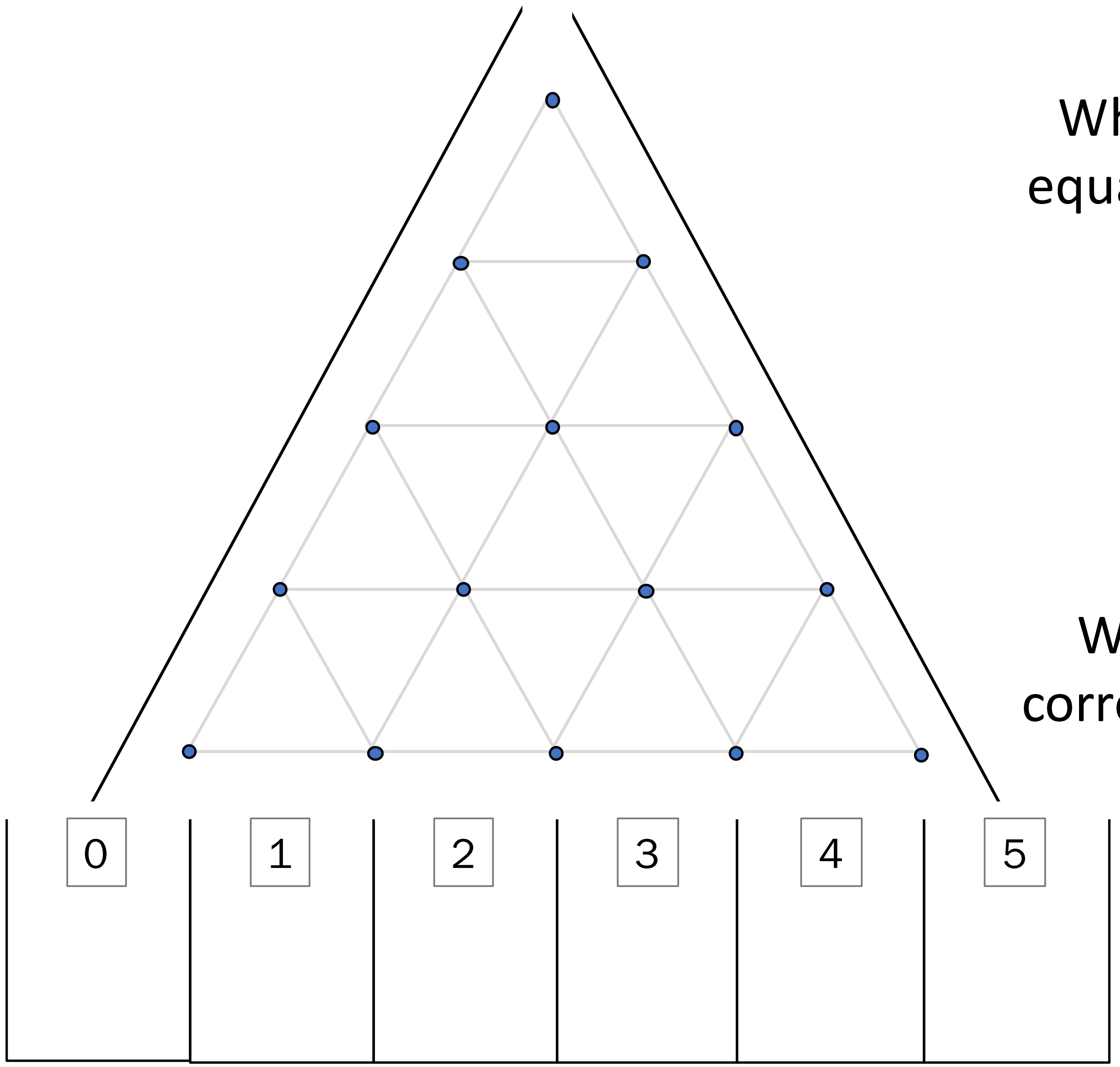
Galton Board Fun



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

Each pin represents an independent event.

Galton Board Fun

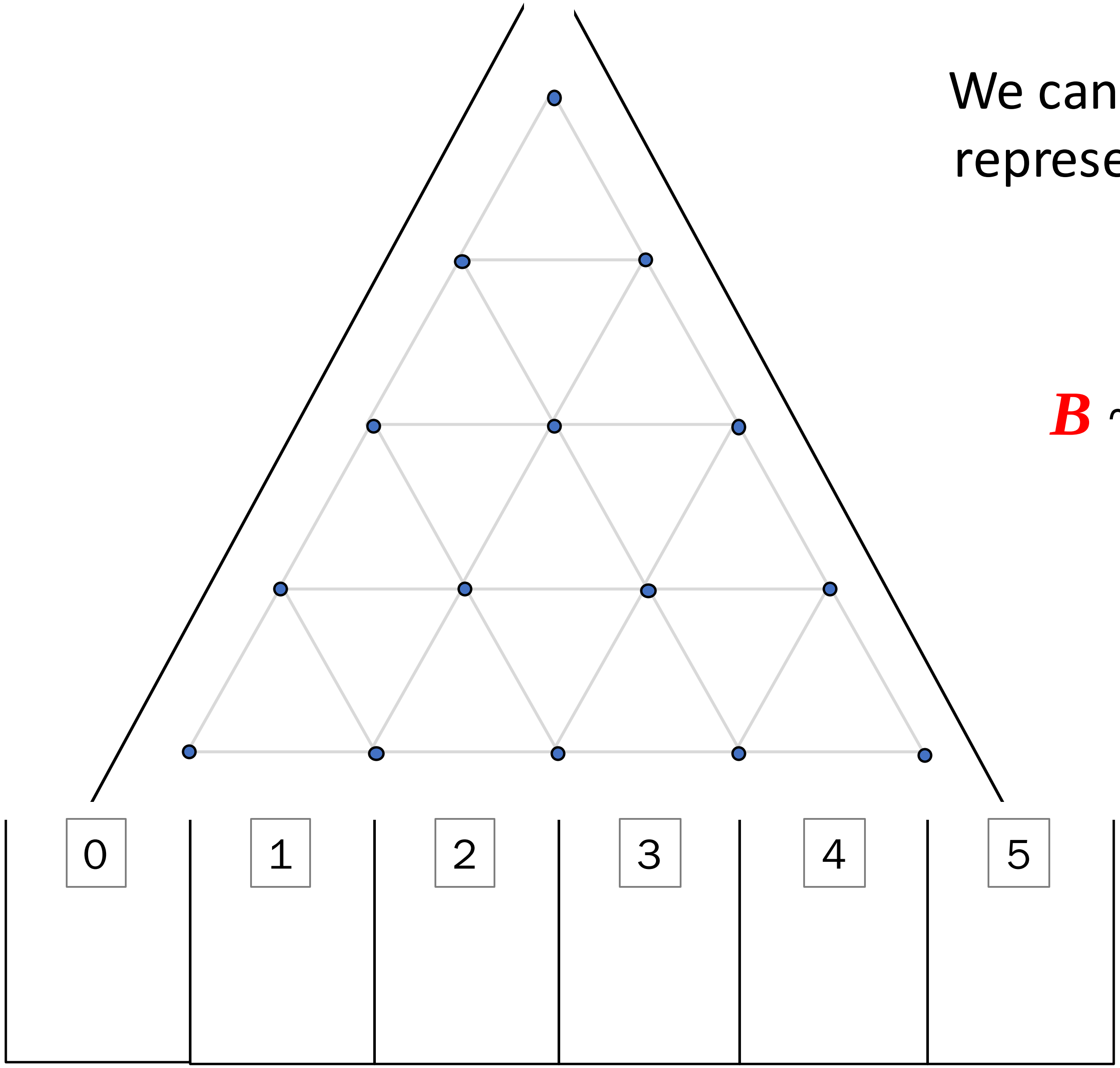


When a marble hits a pin, it has equal chance of going left or right.

Each pin represents an independent event.

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

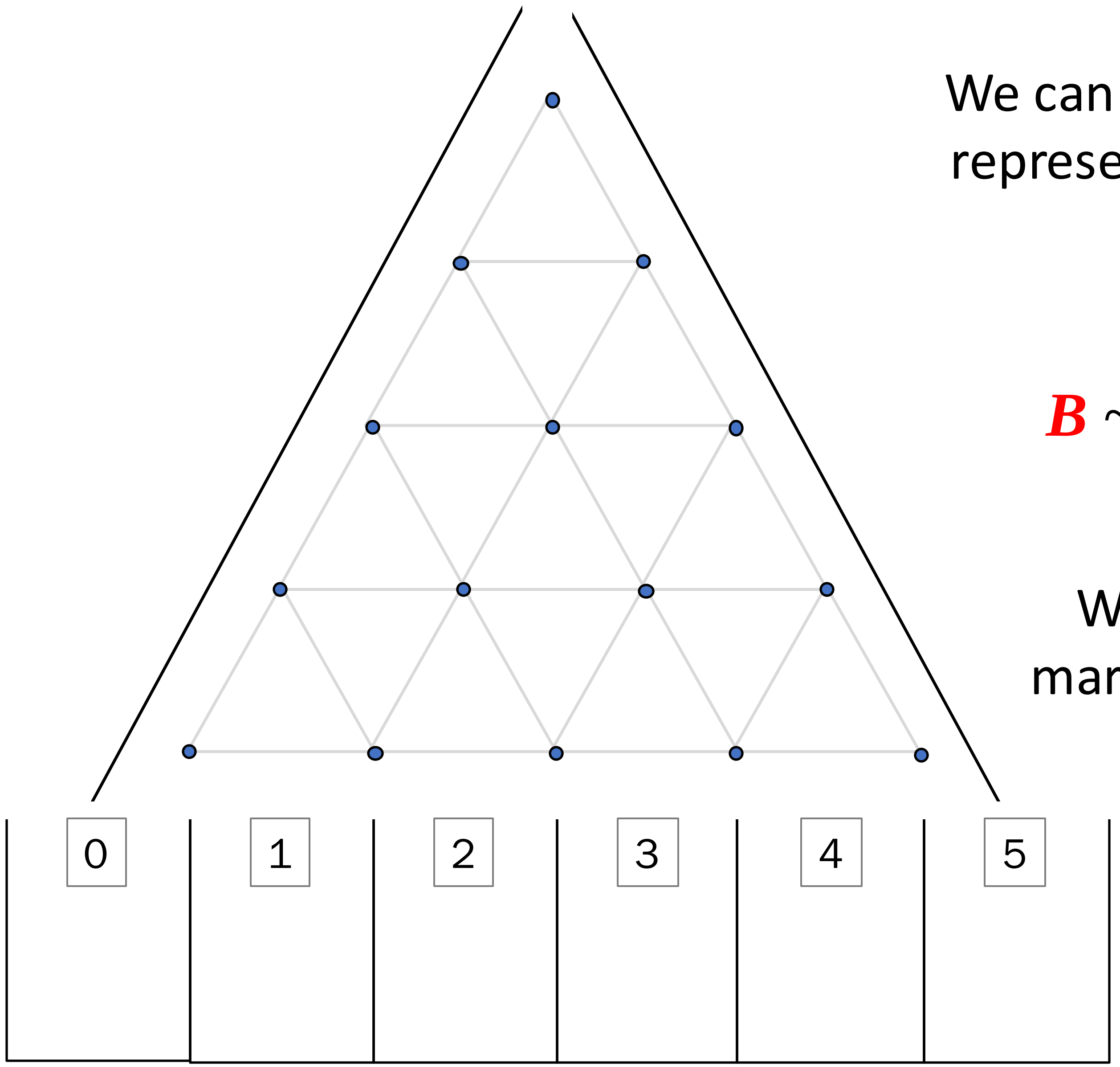
Galton Board Fun



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

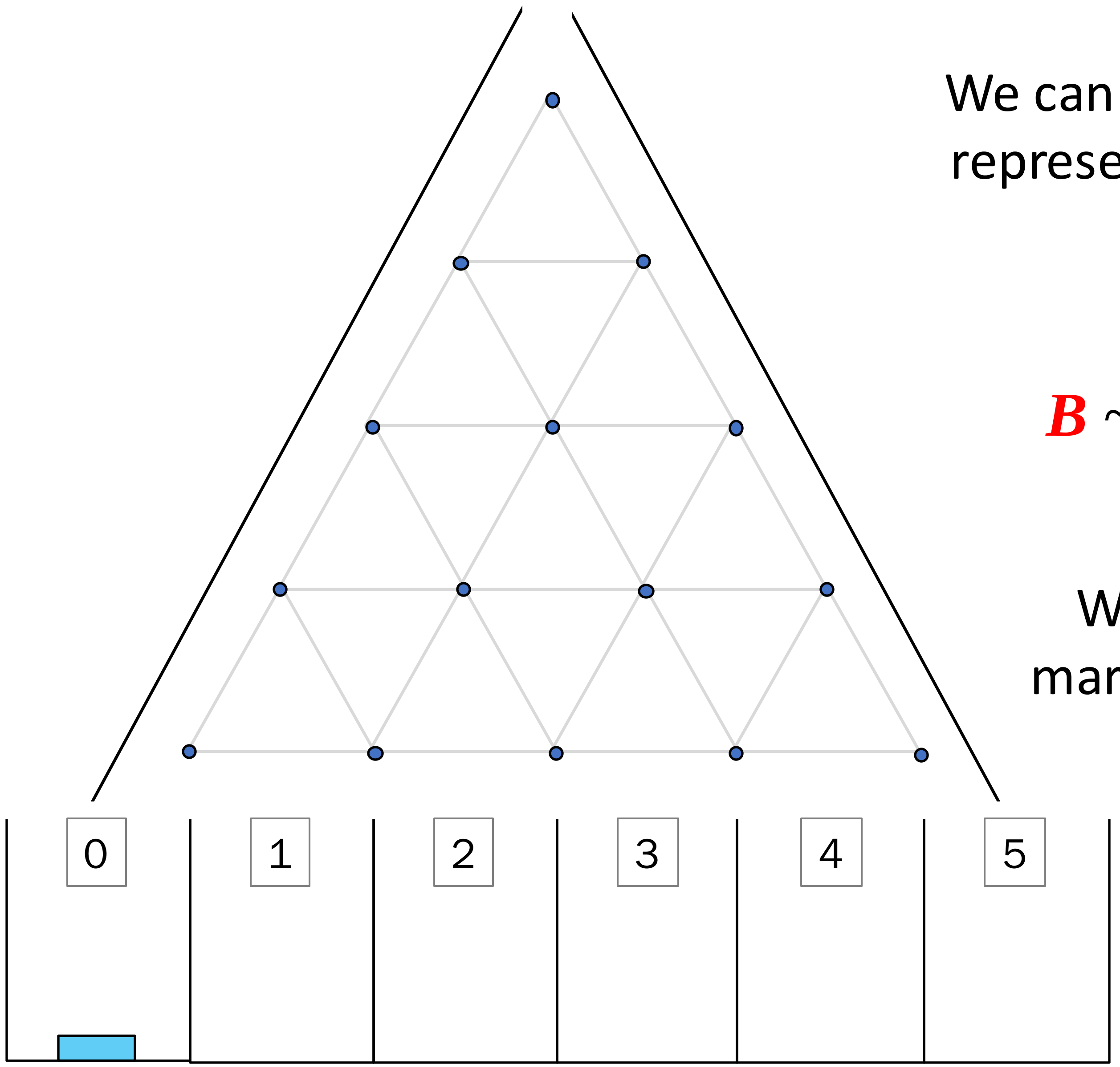


We can define a random variable (***B***) representing which bucket a marble lands in.

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

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

Galton Board Fun



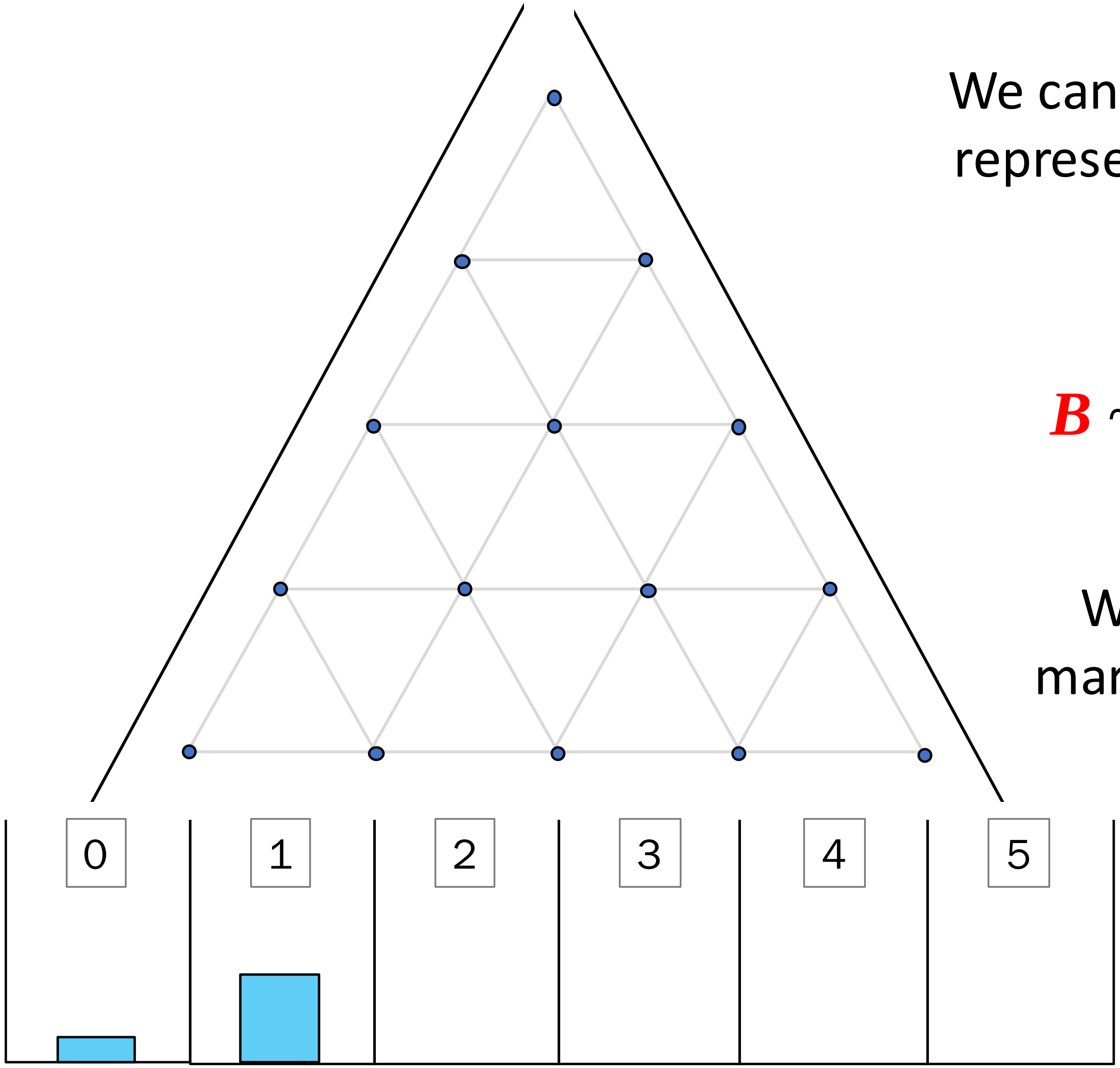
We can define a random variable (***B***) representing which bucket a marble lands in.

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

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.

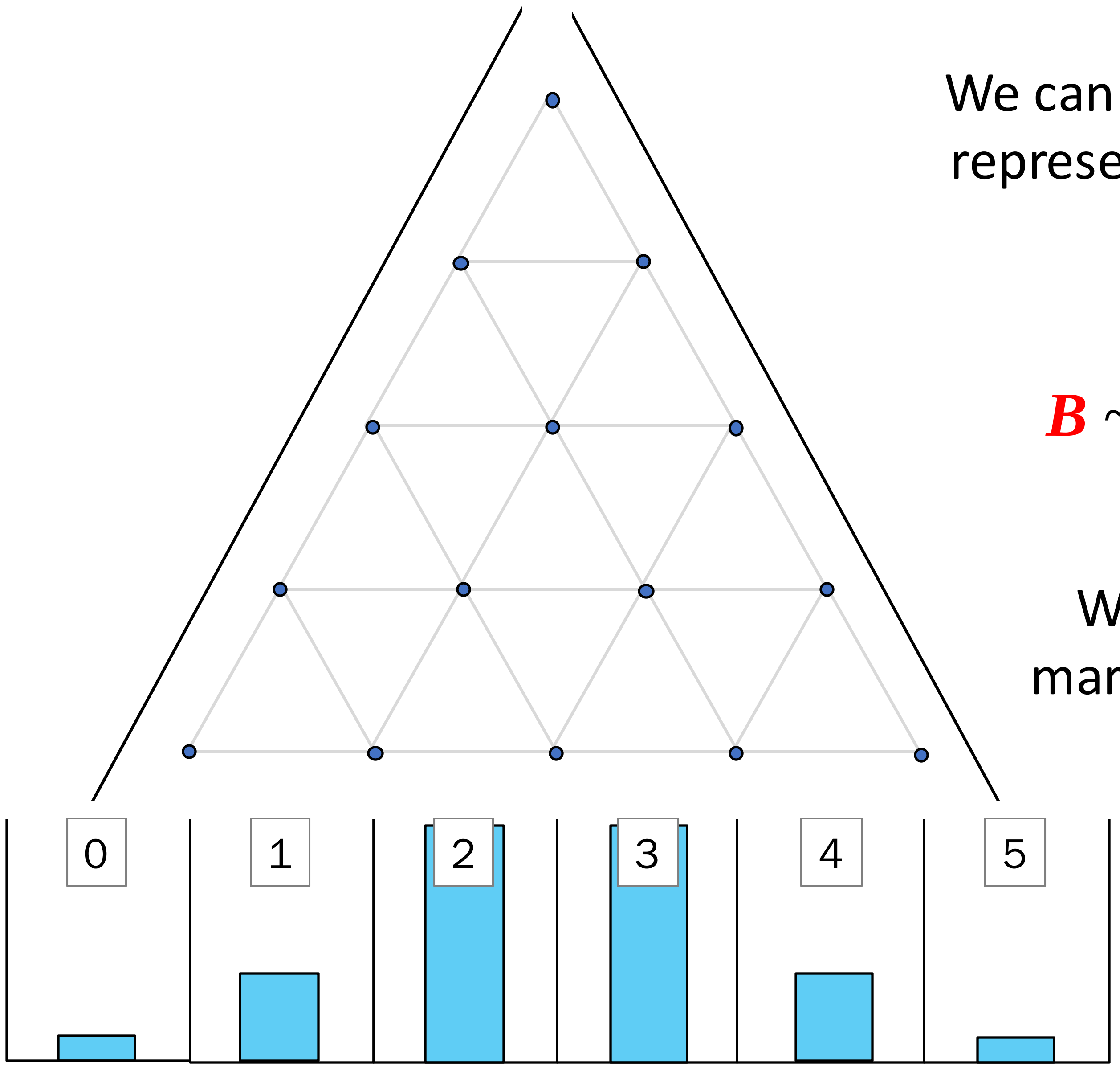
$\boldsymbol{B} \sim \text{Bin}(n = \text{levels}, p = 0.5)$

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*

Have a wonderful weekend!