Definition of a Random Variable

Lecture 19, CS70 Fall 2026

Definitions

- **Definition of a Random Variable**
- Distribution of a Random Variable
- Expectation of a Random Variable

Example Distributions

- Bernoulli
- Binomial
- Uniform
- Geometric

CC BY NC SA

Random Variables: Informally

Suppose we flip a fair coin 4 times. Let X be the number of heads we see.

What is X?

- Some indeterminate value between 0 and 4.

What is $X + 3$?

- Some indeterminate value between 3 and 7.

What is $X - X$?

- 0

What is X^2?

- Some indeterminate value from {0, 1, 4, 9, 16}.

CC BY NC SA

Random Variables: Informally

Suppose we flip a fair coin 4 times. Let X be the number of heads we see, and Y be the number of tails we see.

What is $X + Y$?

- 4

What is $X - Y$?

- Some indeterminate number from {-4, -2, 0, 2, 4}.

CC BY NC SA

Random Variables: Informally

Suppose we take homework from 100 students, and randomly give each student one homework. Let X be the number of students who get their own homework.

What is X?

- Some indeterminate value between 0 and 100.

Is X more likely to be 1 or 3?

- Much more likely to be 1.

What is X on average?

- This turns out to be 1.

CC BY NC SA

Random Variables: Informally

Suppose we flip a fair coin 4 times. Let X be the number of heads we see.

- Since the value of the variable X depends on the results of an experiment, we can't give a definitive value for X.
- However, we can enumerate all of the samples that yield each X, e.g.:

ω	X
TTTT	0
HTTT, THTT, TTHT, TTTH	1
HHTT, HTHT, HTTH, THHT, THTH, TTHH	2
HHHT, HHTH, HTHH, THHH	3
HHHH	4

CC BY NC SA

Random Variable: Formally

A **Random Variable** X on a sample space is a function that maps $\Omega \to \mathbb{R}$, i.e. $X(\omega)$ is a real number for every $\omega \in \Omega$.

$\omega \in \Omega$ $\quad$ $0 \in \mathbb{R}$

Example, number of heads in four coin flips: $X(TTTT) = 0$, $X(HHTH) = 3$

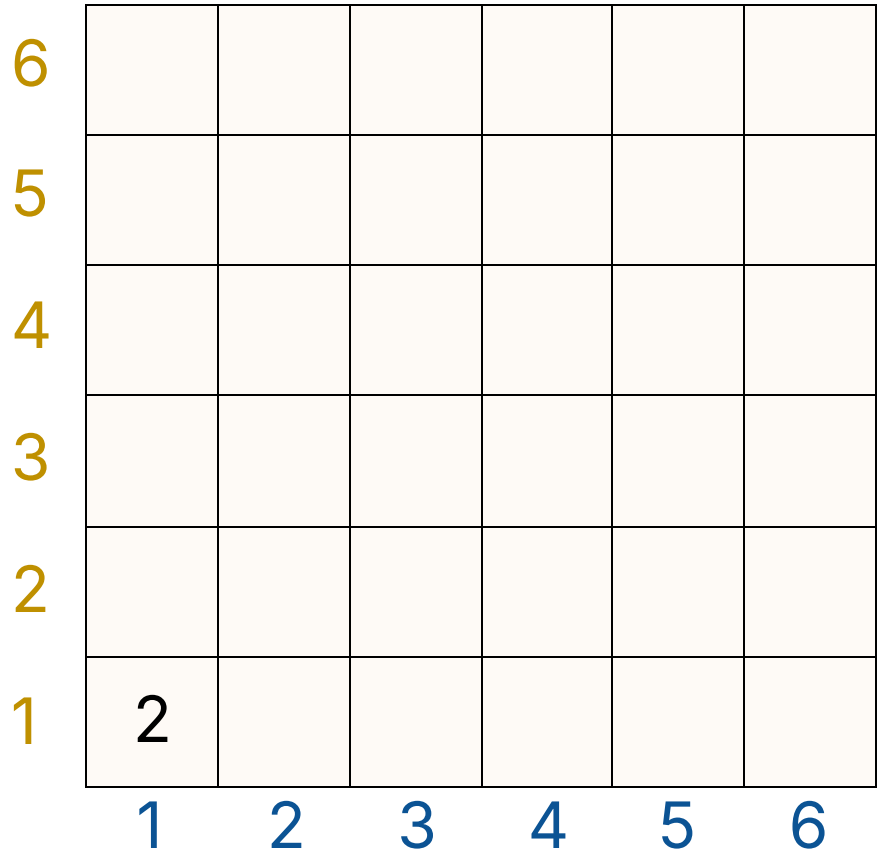
ω	X
TTTT	0
HTTT, THTT, TTHT, TTTH	1
HHTT, HTHT, HTTH, THHT, THTH, TTHH	2
HHHT, HHTH, HTHH, THHH	3
HHHH	4

CC BY NC SA

Formal Random Variable Example 2

Imagine rolling two six sided dice. Let X be their sum.

- $X((1, 1)) = 2$

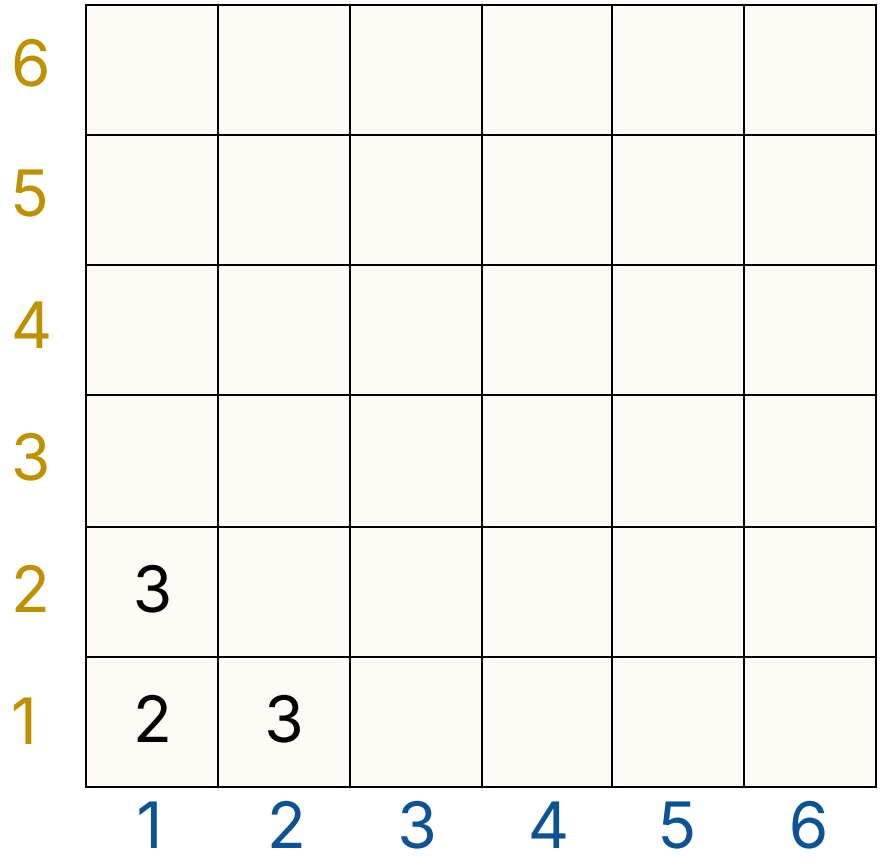


CC BY NC SA

Formal Random Variable Example 2

Imagine rolling two six sided dice. Let X be their sum.

- $X((1, 2)) = 3$ and $X((2, 1)) = 3$



CC BY NC SA

Formal Random Variable Example 2

Imagine rolling two six sided dice. Let X be their sum.

- The domain of X is the set of all tuples of integers $(i,\ j)$ where $1 \leq i, j \leq 6$.
- The range of X is $\{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ \mathbf{10},\ 11,\ 12\}$.

6	7	8	9	10	11	12
5	6	7	8	9	10	11
4	5	6	7	8	9	10
3	4	5	6	7	8	9
2	3	4	5	6	7	8
1	2	3	4	5	6	7
	1	2	3	4	5	6

$$X((4, 6)) = 10$$

cc BY NC SA

Formal Random Variable Example 3

Consider the sample space of all sequences of coin flips consisting of all tails followed by heads.

Let X be the number of flips until we get our first heads.

- $X(H) = 1$
- $X(TH) = 2$
- $X(TTTTTH) = 6$

CC BY NC SA

Formal Random Variable Example 3

Consider the sample space of all sequences of coin flips consisting of all tails followed by heads.

Let X be the number of flips until we get our first heads.

- $X(H) = 1$
- $X(TH) = 2$
- $X(TTTTTH) = 6$

What are the domain and range of X?

What is the cardinality of each?

https://joshh.ug/A



CC BY NC SA

Formal Random Variable Example 3

Consider the sample space of all sequences of coin flips consisting of all tails followed by heads.

Let X be the number of flips until we get our first heads.

- $X(H) = 1$
- $X(TH) = 2$
- $X(TTTTTH) = 6$

The domain of X is infinitely large, it's any number of tails followed by heads.

The range of X is $\mathbb{N}^+$.

CC BY NC SA

Distribution of a Random Variable

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- **Distribution of a Random Variable**
- Expectation of a Random Variable

Example Distributions

- Bernoulli
- Binomial
- Uniform
- Geometric

CC BY NC SA

Events and Probability Spaces

Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.
- Example: $\mathrm{C}(HTTT) = 3$ because the longest run is the three tails at the end.
- Example: $\mathrm{C}(HHTH) = 2$ because the longest run is the two heads.

$$\Omega = \begin{array}{cccc} \mathtt{HHHH} & \mathtt{HHHT} & \mathtt{HHTH} & \mathtt{HHTT} \\ \mathtt{HTHH} & \mathtt{HTHT} & \mathtt{HTTH} & \mathtt{HTTT} \\ \mathtt{THHH} & \mathtt{THHT} & \mathtt{THTH} & \mathtt{THTT} \\ \mathtt{TTHH} & \mathtt{TTHT} & \mathtt{TTTH} & \mathtt{TTTT} \end{array}$$

CC BY NC SA

Events and Probability Spaces

Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

Consider $\{\omega \in \Omega : C(\omega) = 4\}$

- Read this as "the set of all ω such that $C(\omega) = 4$".
- Is this set an event on Ω?

$$\Omega = \begin{array}{cccc} \texttt{HHHH} & \texttt{HHHT} & \texttt{HHTH} & \texttt{HHTT} \\ \texttt{HTHH} & \texttt{HTHT} & \texttt{HTTH} & \texttt{HTTT} \\ \texttt{THHH} & \texttt{THHT} & \texttt{THTH} & \texttt{THTT} \\ \texttt{TTHH} & \texttt{TTHT} & \texttt{TTTH} & \texttt{TTTT} \end{array}$$

Events and Probability Spaces

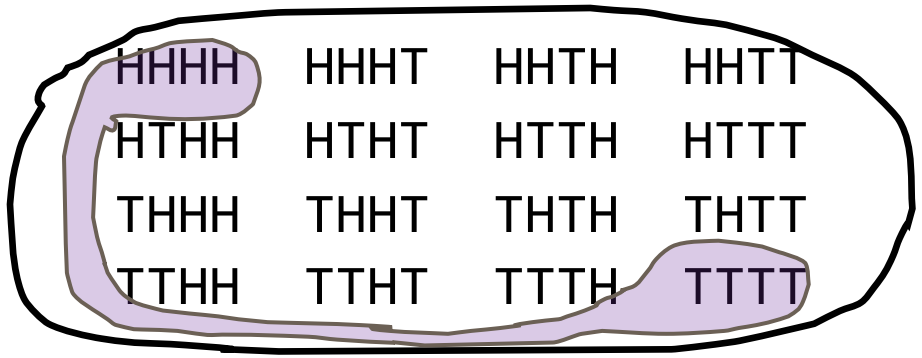
Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

Consider $\{\omega \in \Omega : C(\omega) = 4\}$

- Read this as "the set of all ω such that $C(\omega) = 4$".
- Is this set an event on Ω?
  - Yes, an event is just a subset of the sample space.
  - This event comprises the two outcomes below.

$\Omega =$



CC BY NC SA

Events and Probability Spaces

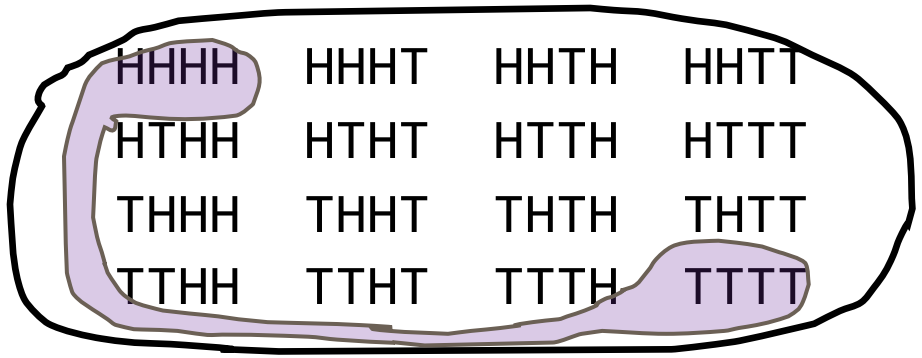
Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

What is $\boldsymbol{P}(\{\omega \in \Omega : C(\omega) = 4\})$?

- Read this as "the probability of the event where $C(\omega) = 4$".

$\Omega =$



CC BY NC SA

Events and Probability Spaces

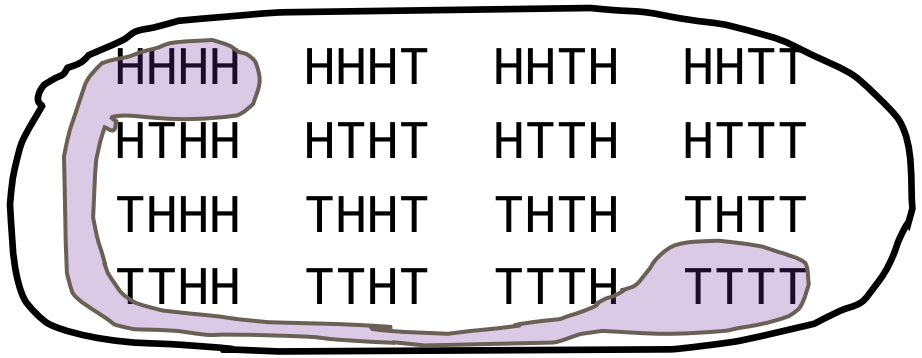
Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

What is $\boldsymbol{P}(\{\omega \in \Omega : C(\omega) = 4\})$?

- Read this as "the probability of the event where $C(\omega) = 4$".
- $\boldsymbol{P}(\{\omega \in \Omega : C(\omega) = 4\}) = 2/16$

$\Omega =$



Events and Probability Spaces

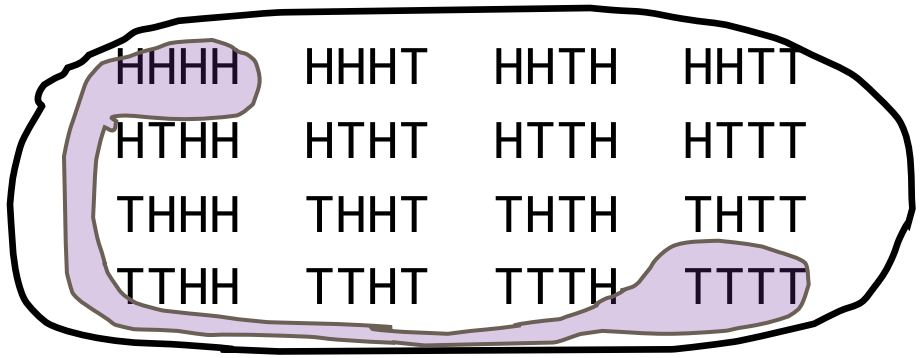
Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

$\{\omega \in \Omega : C(\omega) = 4\}$ is a bit unwieldy from a notation perspective.

- More common to simply represent this set as $C = 4$.
- Example: $\boldsymbol{P}(C = 4)$ means exactly $\boldsymbol{P}(\{\omega \in \Omega : C(\omega) = 4\})$
- Example: $\mathrm{TTTT} \in (C = 4)$

$\Omega =$



CC BY NC SA

Events and Probability Spaces

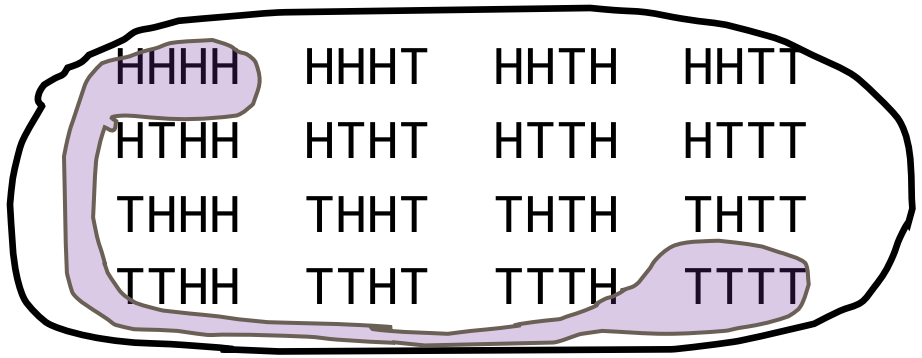
Consider the sample space Ω of four coin flips, shown below.

- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.

$\{\omega \in \Omega : C(\omega) = 4\}$ is a bit unwieldy from a notation perspective.

- More common to simply represent this set as $C = 4$.
- Other notations we will not use:
  - $C^{-1}(4)$
  - C_4

$\Omega =$



CC BY NC SA

Events and Random Variables

Claim: A random variable **partitions** a sample space.

From lecture 16 (paraphrasing): Ω is partitioned into n events $A_1, \ldots, A_n$ if:

- $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$
- $A_i \cap A_j = \emptyset$ for all $i \neq j$. That is, the events are mutually exclusive.

Is this claim true or false?

https://joshh.ug/B

CC BY NC SA

Events and Random Variables

Claim: A random variable **partitions** a sample space.

From lecture 16 (paraphrasing): Ω is partitioned into n events $A_1, \ldots, A_n$ if:

- $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$
- $A_i \cap A_j = \emptyset$ for all $i \neq j$. That is, the events are mutually exclusive.

Is this claim **true** or false?

- For every sample point ω in the sample space, the random variable maps ω to exactly one event.

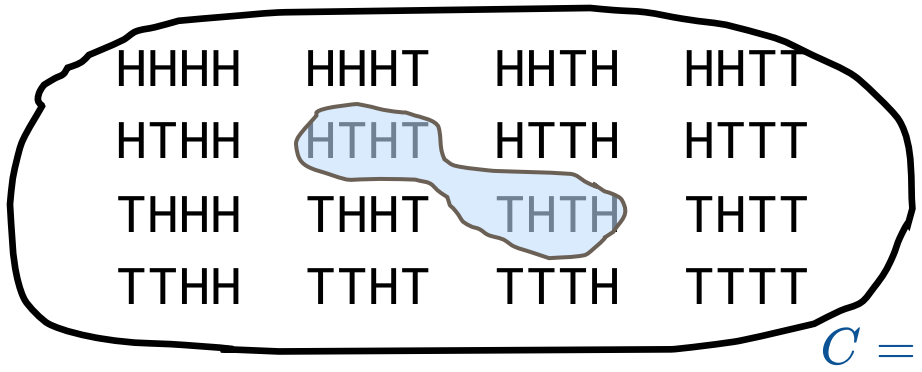
CC BY NC SA

Example: C Partitions Coin Flips

Consider the sample space Ω of four coin flips, shown below.

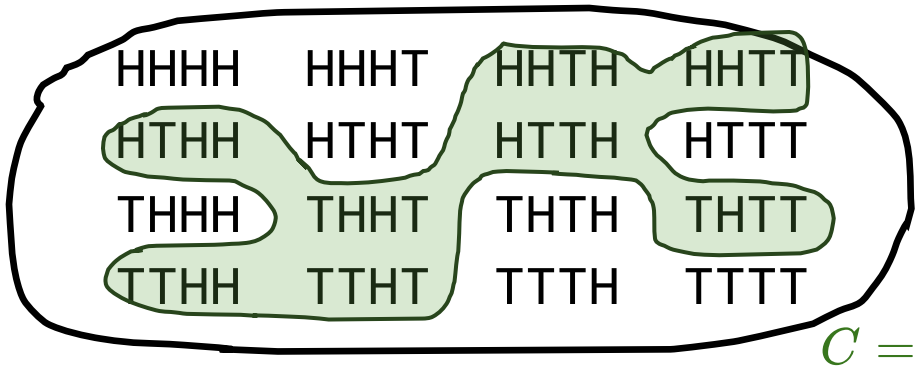
- Define C as a random variable on Ω which counts the longest "run", where a run is a sequence of flips that are all the same.
- C **partitions** the sample space as shown below. All outcomes assigned to exactly one value by C.

$\Omega =$



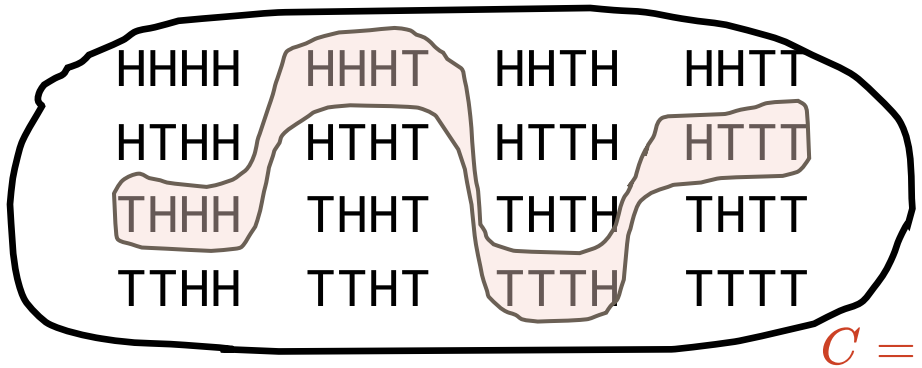
$C = 1$

$\Omega =$



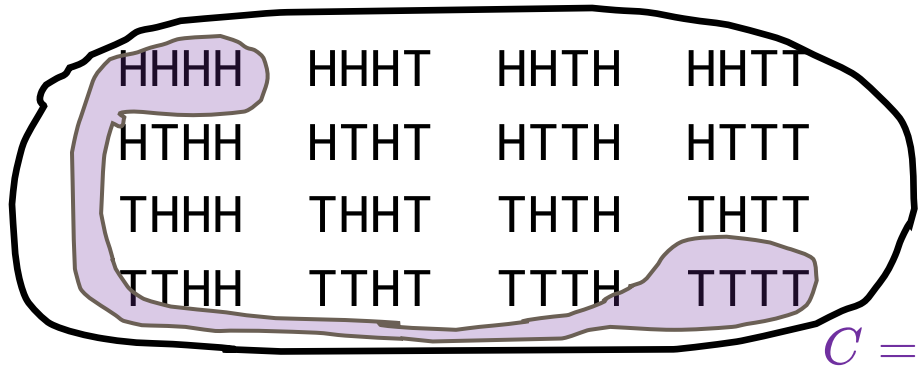
$C = 2$

$\Omega =$



$C = 3$

$\Omega =$



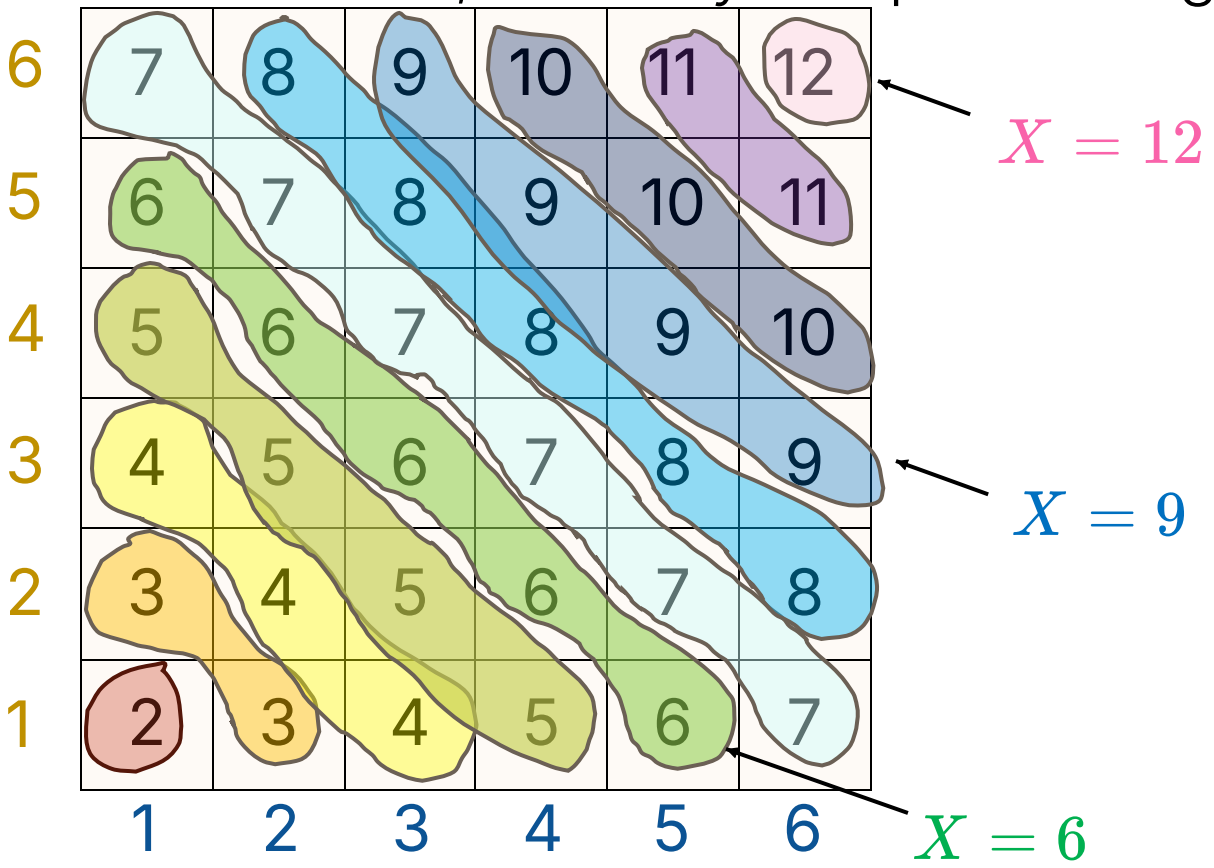
$C = 4$

CC BY NC SA

Another Example: Pairs of Dice

Let X be a random variable that sums the result of two dice.

- X **partitions** the sample space of pairs of dice rolls into events. Why?
- The domain of X is the entire sample space.
- X is a function, so every sample is assigned to exactly one set.



Reminder: A partition of a sample space is defined as a set of sets whose union is the entire sample space, and whose intersections are all $\emptyset$.

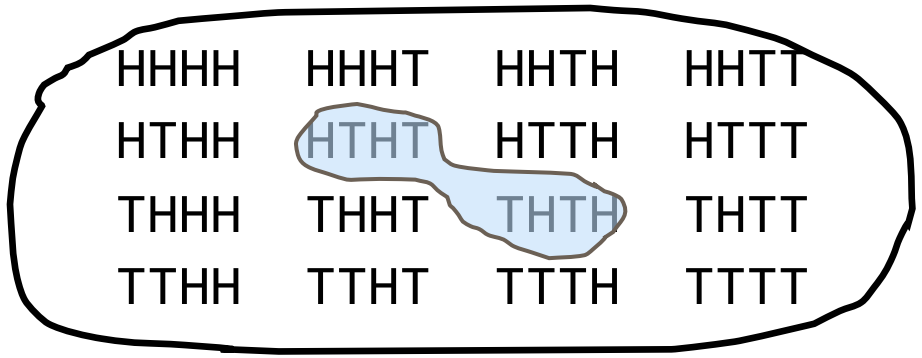
CC BY NC SA

The Distribution of a Random Variable

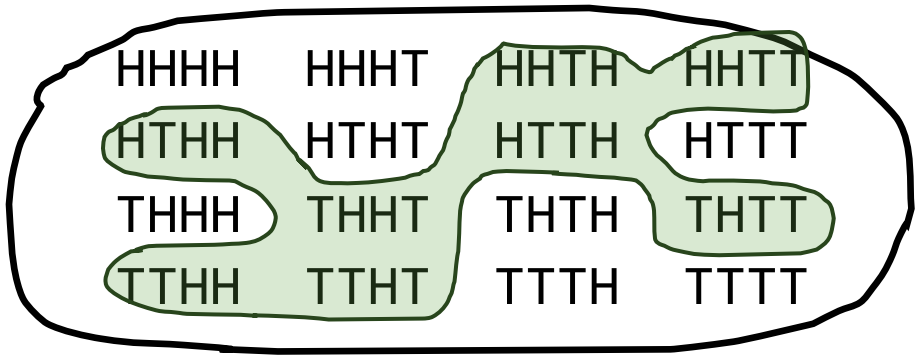
The **distribution** of a random variable X is the set of all values in $\mathbf{range}(X)$ and their probabilities, that is $\{(a, P(X = a)) : a \in \mathbf{range}(X)\}$. For runs of coins we have:

- $\{(1, \frac{2}{16}), \quad (2, \frac{8}{16}), (3, \frac{4}{16}), (4, \frac{2}{16})\}$

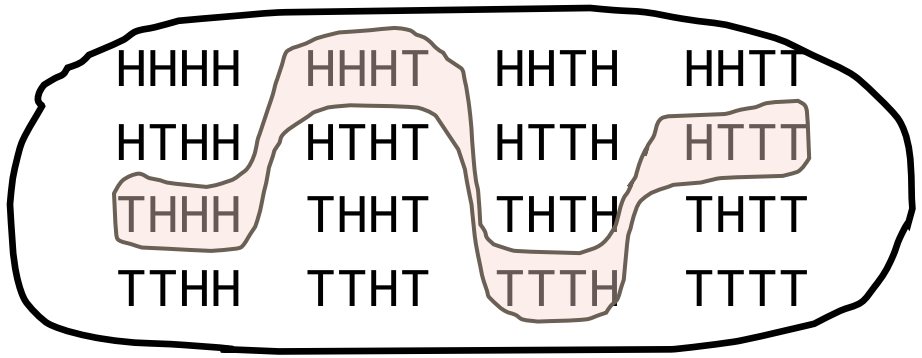
$\Omega =$



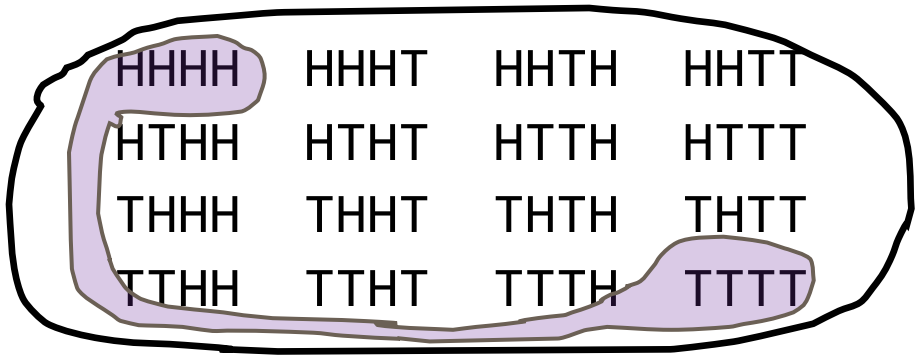
$\Omega =$



$\Omega =$



$\Omega =$



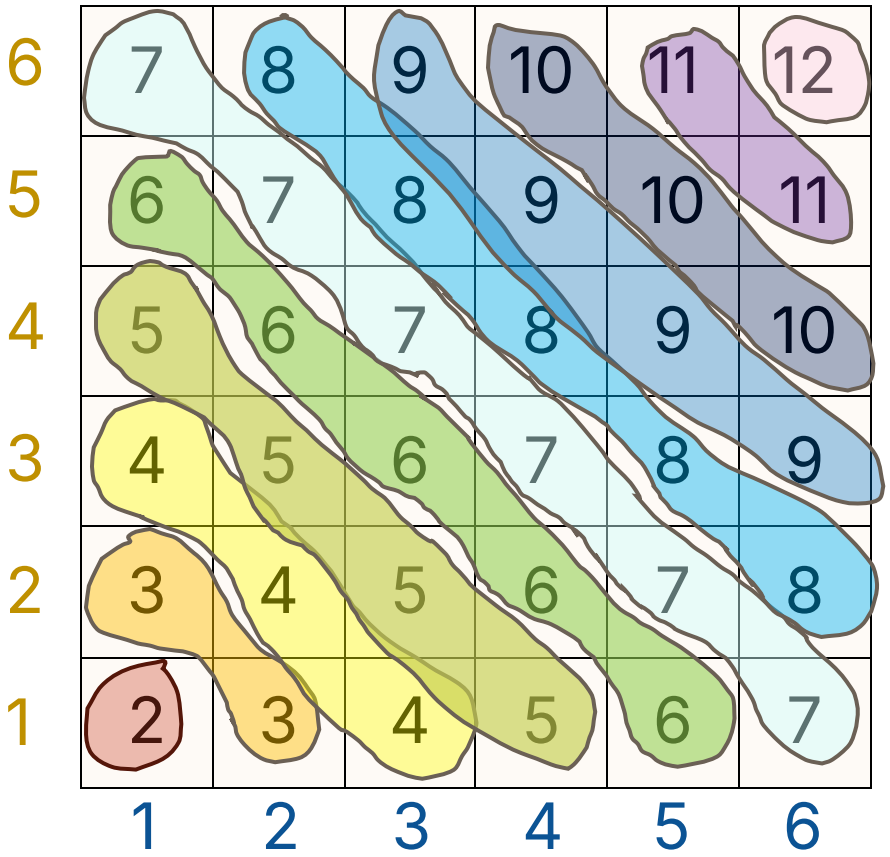
CC BY NC SA

Distribution of a Random Variable

The **distribution** of a random variable X is the set of value/probability pairs:

$$\{(a, P(X = a)) : a \in \text{range}(X)\}$$

- Other texts call $a \mapsto P(X = a)$ the probability mass function (PMF).



Example: If X is the sum of two six sided dice, the distribution of X is:

$$\left\{ \begin{array}{c} \left(2, \frac{1}{36}\right), \left(3, \frac{2}{36}\right), \left(4, \frac{3}{36}\right), \left(5, \frac{4}{36}\right), \left(6, \frac{5}{36}\right), \left(7, \frac{6}{36}\right), \\ \left(8, \frac{5}{36}\right), \left(9, \frac{4}{36}\right), \left(10, \frac{3}{36}\right), \left(11, \frac{2}{36}\right), \left(12, \frac{1}{36}\right) \end{array} \right\}$$

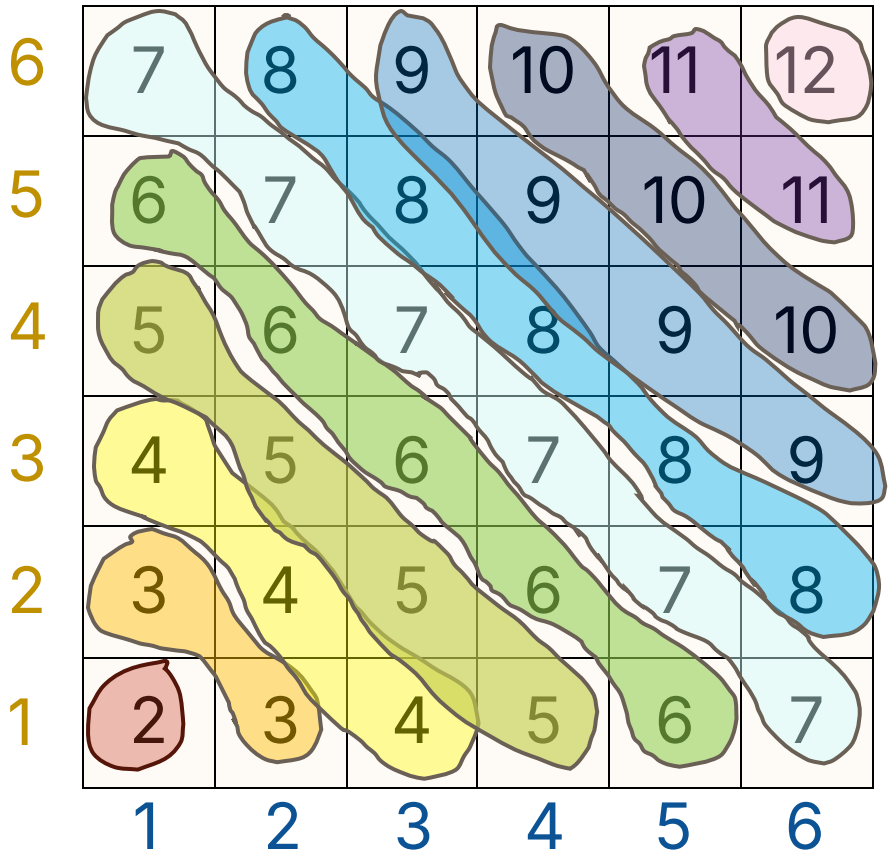
cc BY NC SA

Distribution of a Random Variable

The distribution of a random variable X is the set of value/probability pairs:

$$\{(a, P(X = a)) : a \in \text{range}(X)\}$$

- Other texts call $a \mapsto P(X = a)$ the probability mass function (PMF).



Can also show the distribution of X as a table:

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

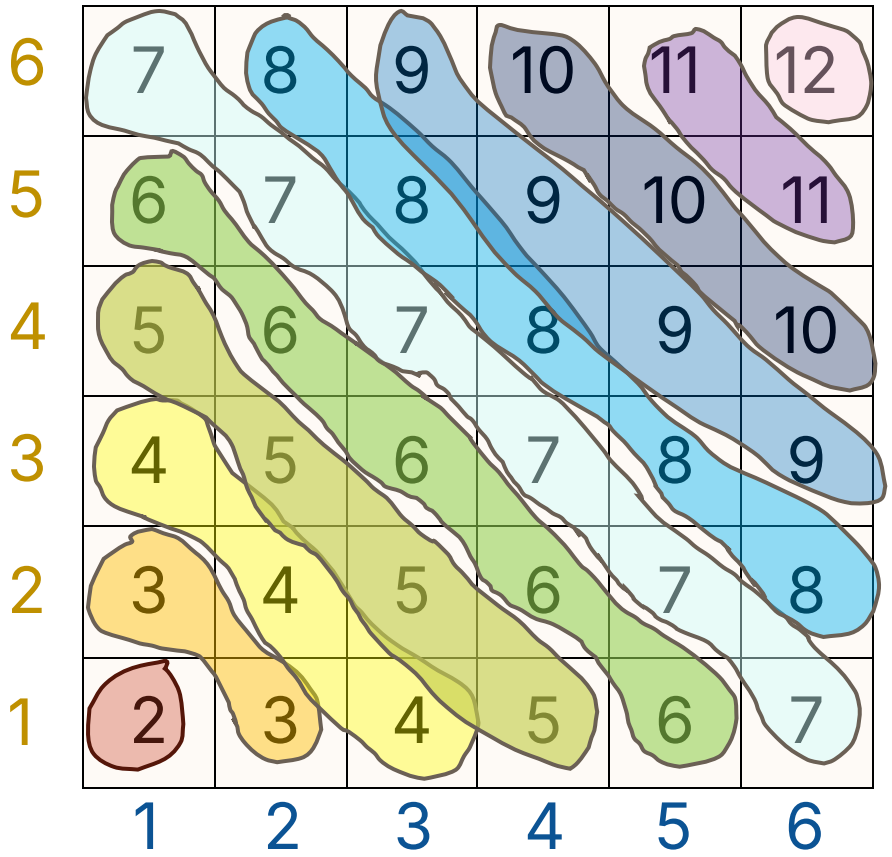
CC BY NC SA

Distribution of a Random Variable

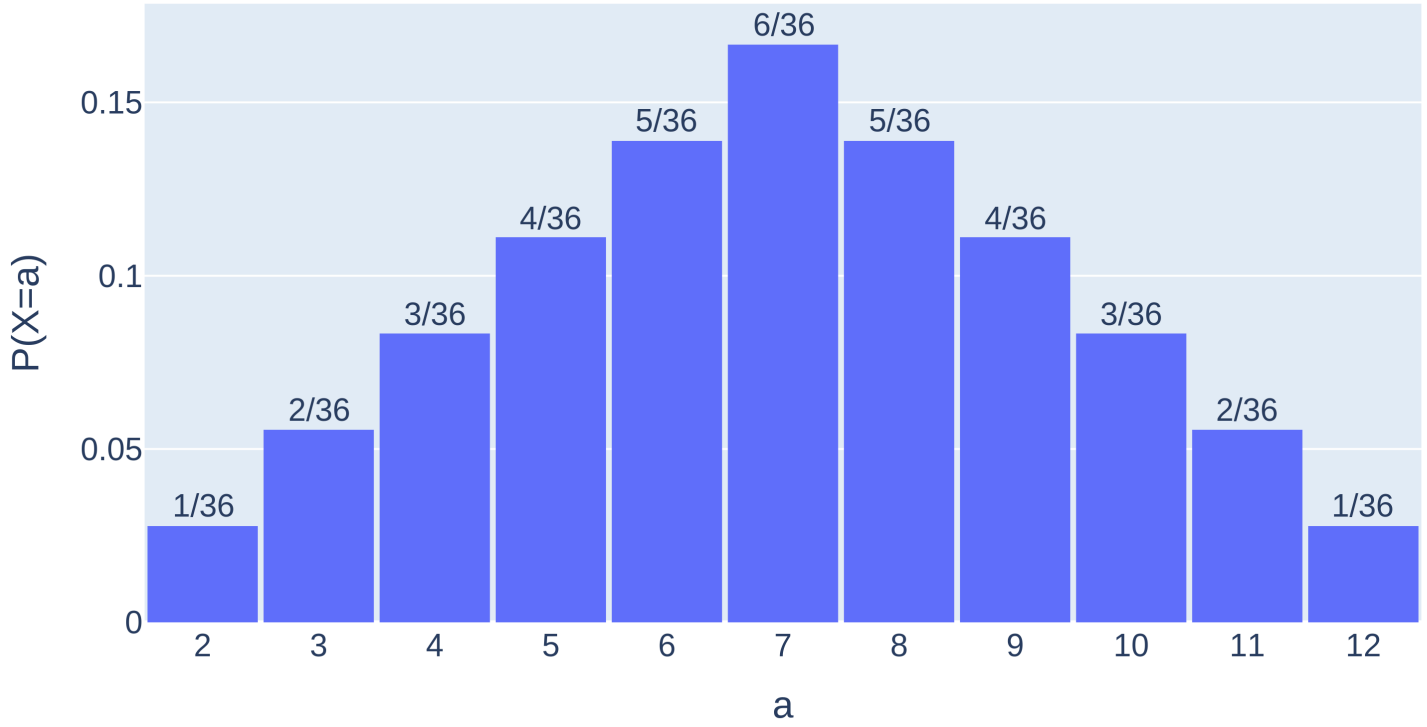
The distribution of a random variable X is the set of value/probability pairs:

$$\{(a, P(X = a)) : a \in \text{range}(X)\}$$

- Other texts call $a \mapsto P(X = a)$ the probability mass function (PMF).



Can also show the distribution of X as a plot:

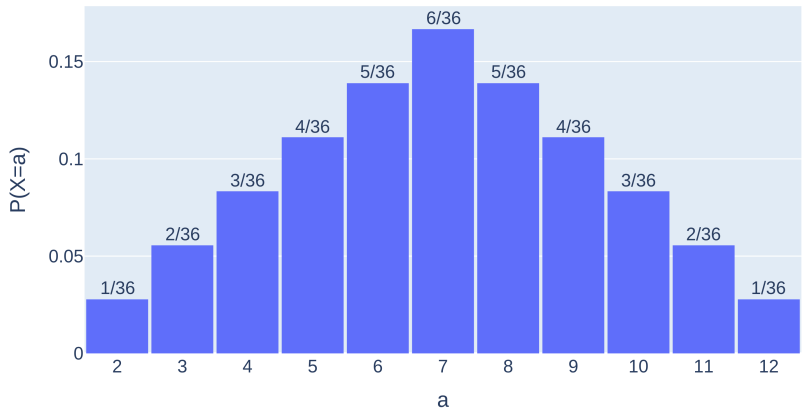


Note

Summing random variables can yield radically different distributions!

- Let D be a random variable that models rolling a single six sided die.
- The sum of two independent rolls, $D_1 + D_2$, looks totally different.

x	$P(D = x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6



x	$P(D_1 + D_2 = x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

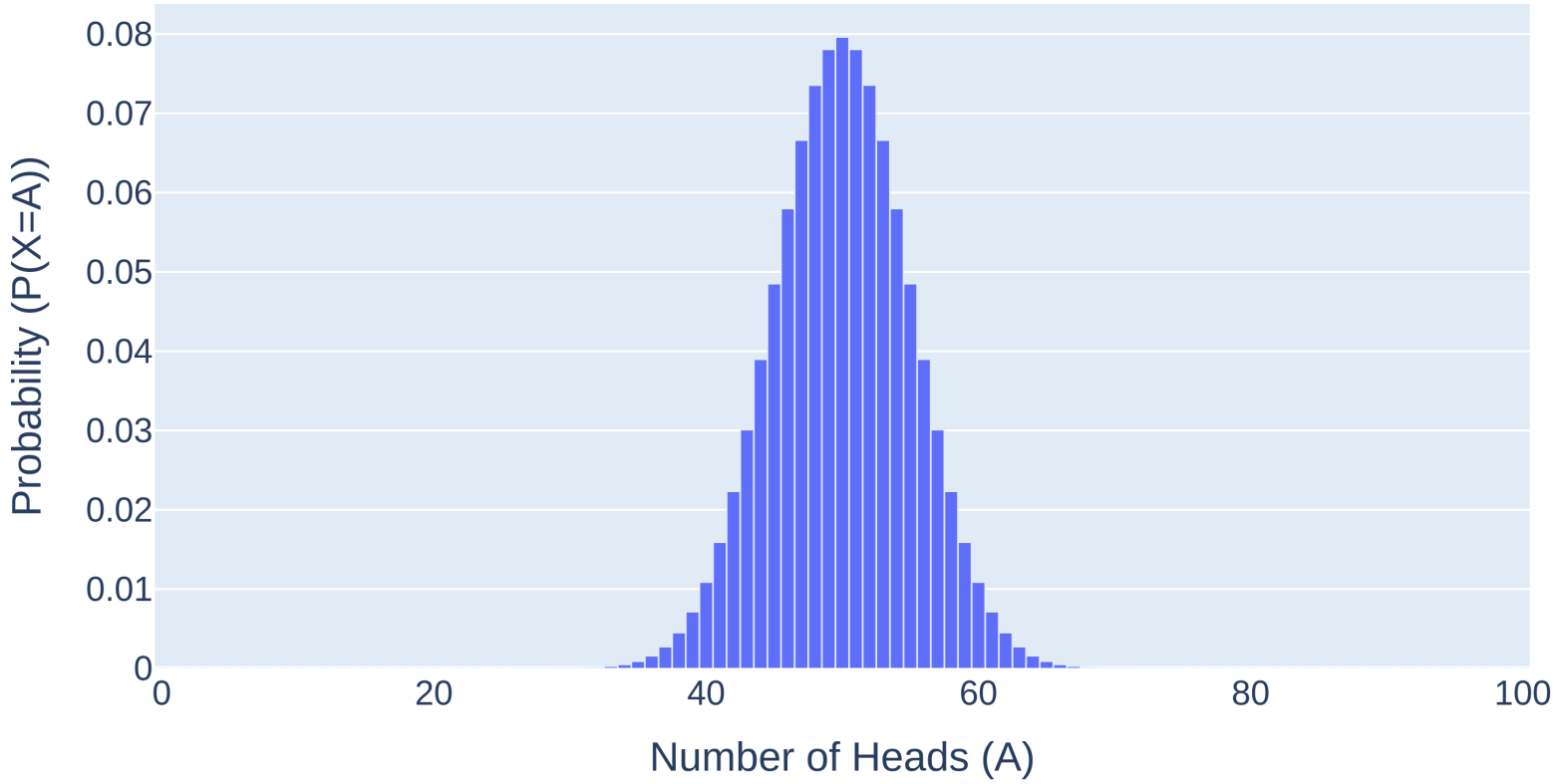
CC BY NC SA

Distribution Example 2: Tossing 100 Coins

Toss 100 coins. Let X be the number of heads.

- Reminder: X is a function that maps samples to reals, e.g. $X(HHH\ldots H) = 100$

Distribution of X: $$P(X = A) = \frac{\#\text{ outcomes with A heads}}{|\Omega|} = \frac{\binom{100}{A}}{2^{100}}$$



Distribution Example 3: Handing Back Assignments

Random Experiment: Suppose we take homework from 3 students, and randomly give each student one homework. Let X be the number of students who get their own homework.

$$\Omega = \{123,\ 132,\ 213,\ 231,\ 312,\ 321\}$$

What is the distribution of X?

https://joshh.ug/C

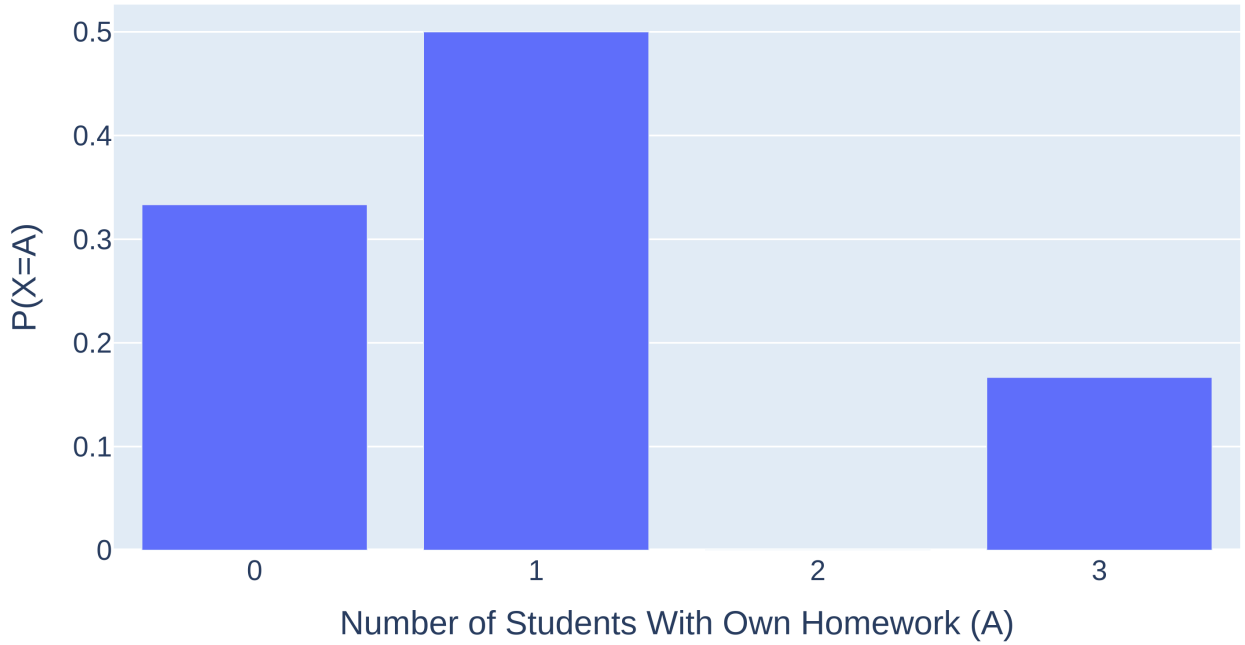
CC BY NC SA

Distribution Example 3: Handing Back Assignments

Random Experiment: Suppose we take homework from 3 students, and randomly give each student one homework. Let X be the number of students who get their own homework.

X	3	1	1	0	0	1

$$\Omega = \{123,\ 132,\ 213,\ 231,\ 312,\ 321\}$$



A	$P(X = A)$
0	$2/6$
1	$3/6$
2	0
3	$1/6$

CC BY NC SA

The Expectation of a Random Variable

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- Distribution of a Random Variable
- **Expectation of a Random Variable**

Example Distributions

- Bernoulli
- Binomial
- Uniform
- Geometric

CC BY NC SA

Distributions and Expectation

Random Experiment: Pick a random student from CS70. X is their grade on the first midterm.

What are some interesting questions we might ask about X?

- What is the distribution of X?
- What is the average of X? We'll call this the **expectation** of X.
- What is the standard deviation of X?

CC BY NC SA

Expectation Informally

Let's do a little experiment and flip 4 coins. Let X be the number of heads. We'll count:

- How many times we get each value a from $\{0, 1, 2, 3, 4\}$.
- The average of the values we get.

Simulate | Pause

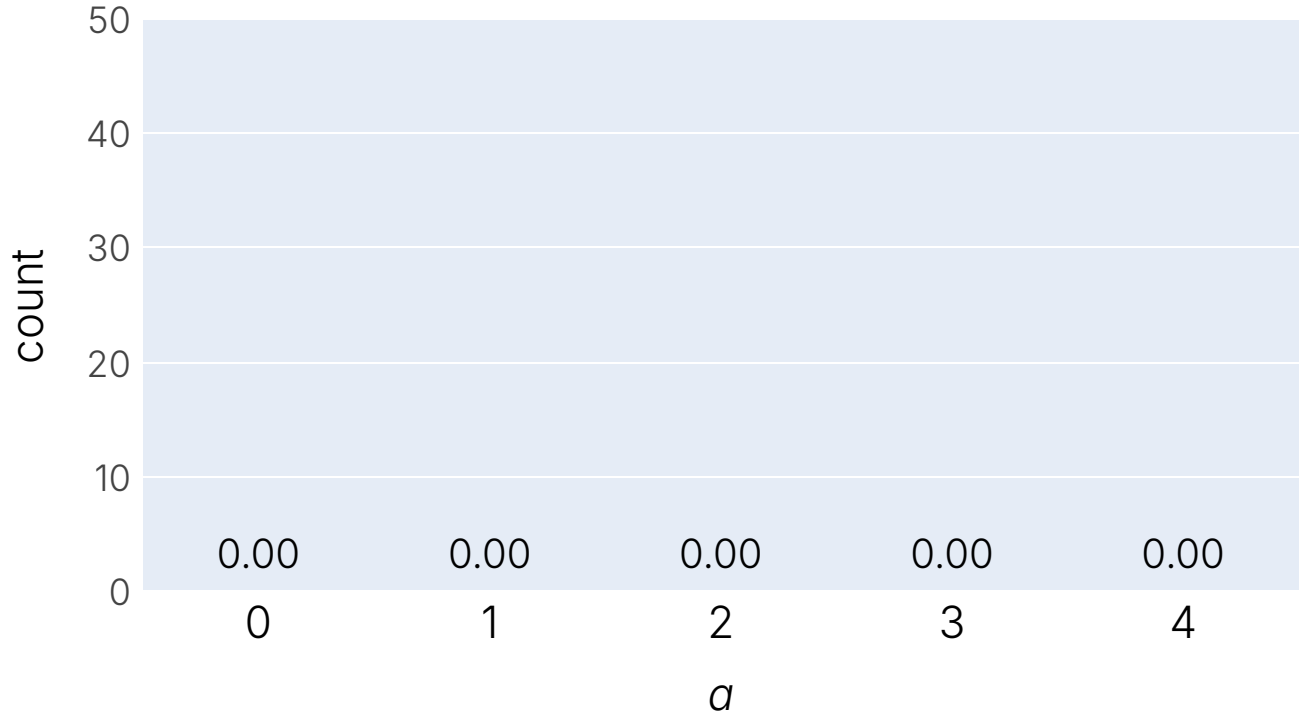
Flip Time (ms): 300 ms

Current flip: N/A

Total number of heads across all flips: N/A

Total number of flips: N/A

Average a: N/A



https://joshh.ug/cs70/four_coin_flips_expectation_simulator.html

CC BY NC SA

Simulation Results (My Experiment - Done Before Lecture)

Out of 20 flips, the average number of heads per flip was 1.7.

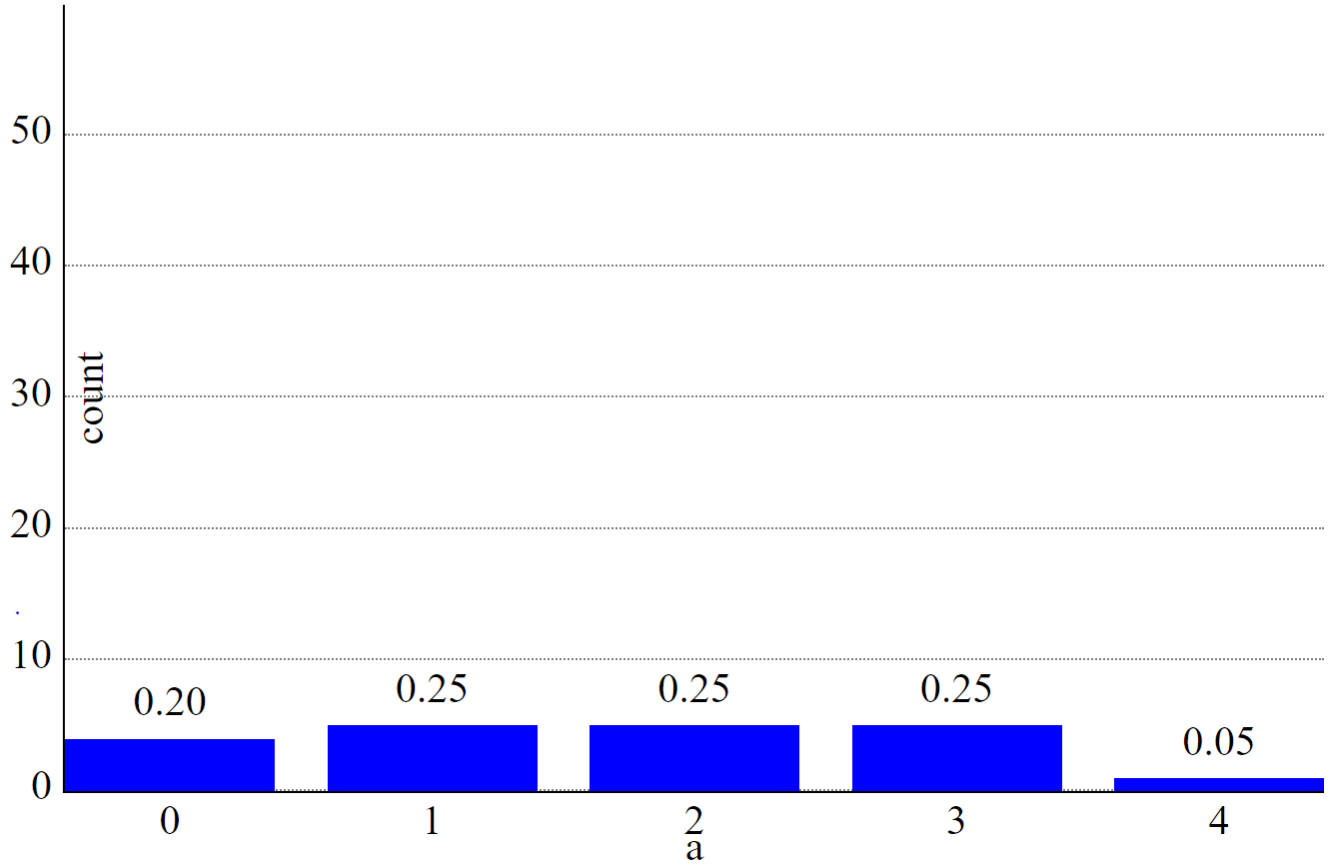
$$\frac{0+2+3+1+2+2+\cdots+2}{20} = 1.7$$

1, 2, and 3 were the most common value, occurring 25% of the time.

Total number of heads across all flips: 34

Total number of flips: 20

Average a: 34/20 = 1.70



CC BY NC SA

Simulation Results (My Experiment - Done Before Lecture)

Out of 150 flips, the average number of heads per flip was 1.99.

$$\frac{0+2+3+1+2+2+\cdots+1}{150} = 1.99$$

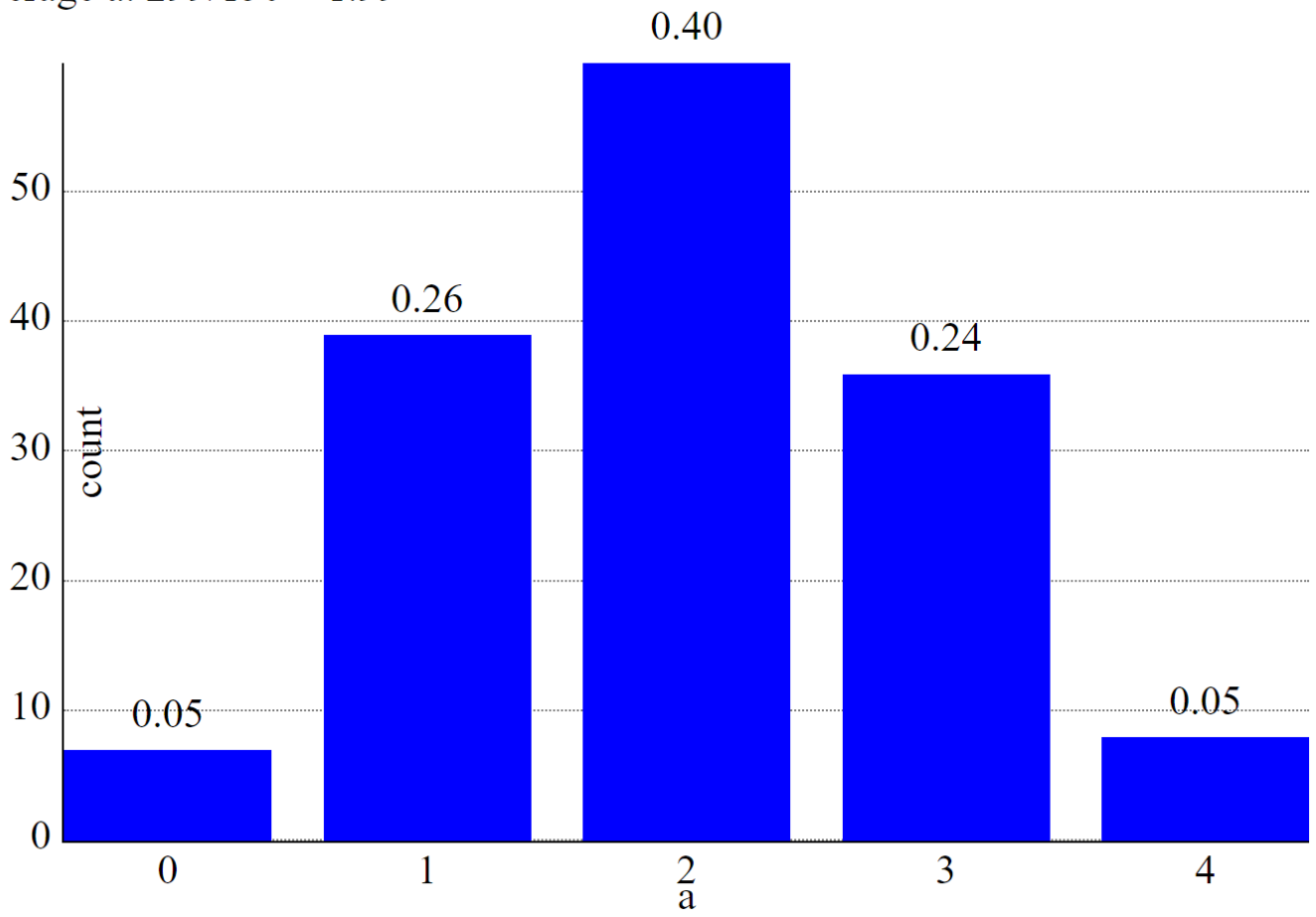
2 was the most common value, occurring 40% of the time.

Note: As the number of trials grows, the histogram looks increasingly like the distribution for the random variable X.

Total number of heads across all flips: 299

Total number of flips: 150

Average a: 299/150 = 1.99



Simulation Results (My Experiment - Done Before Lecture)

Note: As the number of trials grows, the histogram looks increasingly like the distribution for the random variable X.

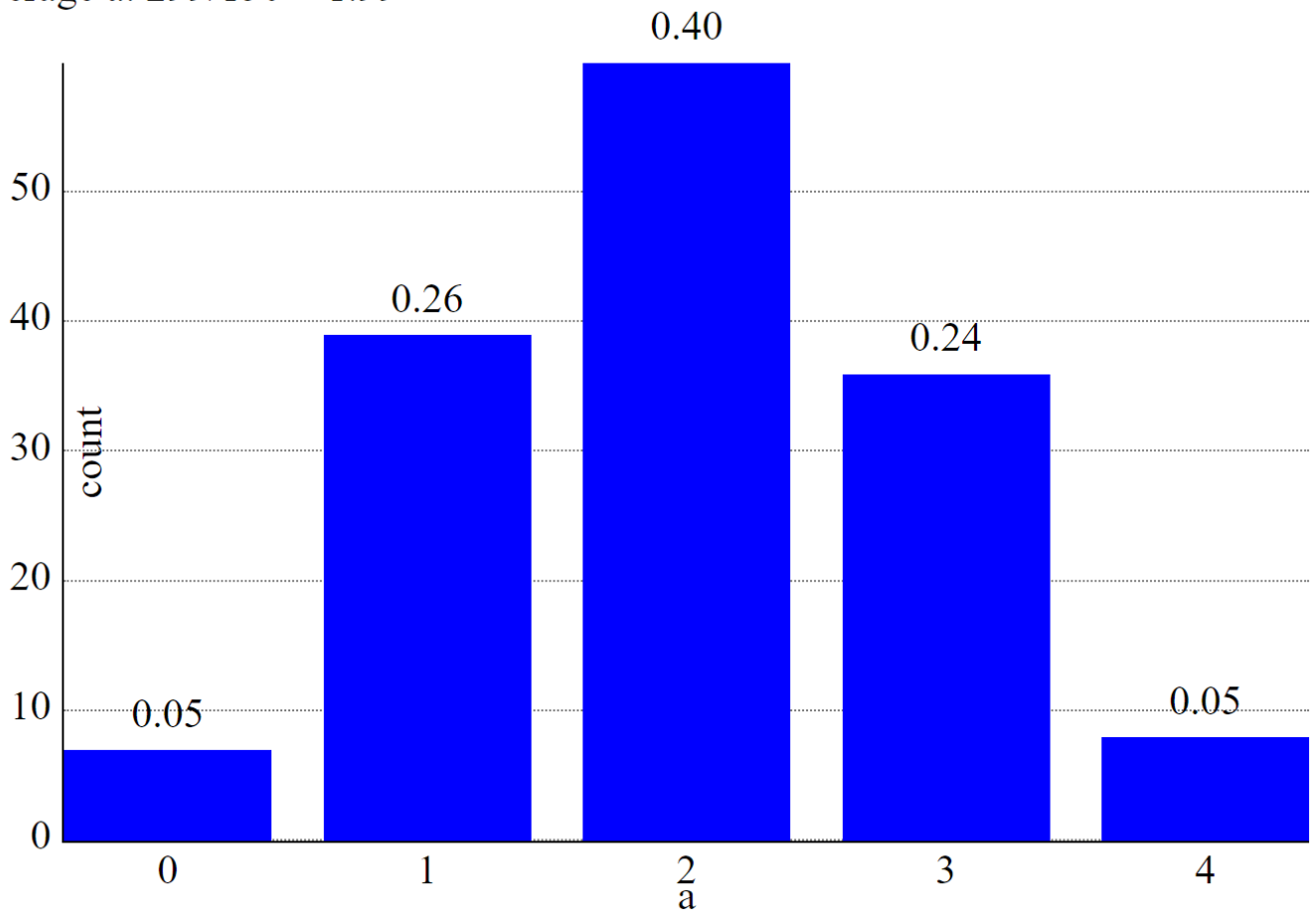
Question: If we know the exact distribution for X, how can we compute the average or **expectation** of X?

- This is the average value that the simulation will eventually converge to (given enough samples).

Total number of heads across all flips: 299

Total number of flips: 150

Average a: 299/150 = 1.99



Expectation (Formally)

The expectation of a random variable is defined as:

CC BY NC SA

Expectation (Formally)

The expectation of a random variable is defined as:

$$E[X] = \sum_{a \in \text{range}(X)} a \times P(X = a)$$

Example: Flipping four coins.

ω	a	$P(X = a)$
TTTT	0	$1/16$
HTTT, THTT, TTHT, TTTH	1	$4/16$
HHTT, HTHT, HTTH, THHT, THTH, TTHH	2	$6/16$
HHHT, HHTH, HTHH, THHH	3	$4/16$
HHHH	4	$1/16$

CC BY NC SA

Expectation (Formally)

The expectation of a random variable is defined as:

$$E[X] = \sum_{a \in \text{range}(X)} a \times P(X = a)$$

Example: Flipping four coins.

- $E[X] = 0 \times 1/16 + 1 \times 4/16 + 2 \times 6/16 + 3 \times 4/16 + 4 \times 1/16 = 32/16 = 2$

ω	a	$P(X = a)$
TTTT	0	$1/16$
HTTT, THTT, TTHT, TTTH	1	$4/16$
HHTT, HTHT, HTTH, THHT, THTH, TTHH	2	$6/16$
HHHT, HHTH, HTHH, THHH	3	$4/16$
HHHH	4	$1/16$

CC BY NC SA

Tail Sum Formula

There's also an alternate formula for expectation called the "tail sum formula". Let X be a random variable whose range is a subset of $\mathbb{N}$, i.e. $\{0, 1, 2, \ldots\}$.

$$E[X] = \sum_{i=1}^{\infty} \boldsymbol{P}(X \geq i)$$

This formula is only valid for X whose range is a subset of $\mathbb{N}$. It's occasionally useful.

Example: Flipping four coins.

- $E[X] = \boldsymbol{P}(X \geq 1) + \boldsymbol{P}(X \geq 2) + \boldsymbol{P}(X \geq 3) + \boldsymbol{P}(X \geq 4)$
  $= 15/16 + 11/16 + 5/16 + 1/16 = 32/16 = 2$

a	$\boldsymbol{P}(X = a)$	$\boldsymbol{P}(X \geq a)$
0	$1/16$	$16/16$
1	$4/16$	$15/16$
2	$6/16$	$11/16$
3	$4/16$	$5/16$
4	$1/16$	$1/16$

CC BY NC SA

Proof of Tail Sum Formula

Let X be random variable whose range is a subset of $\mathbb{N}$, i.e. $\{0, 1, 2, \ldots\}$

$$E[X] = \sum_{i=1}^{\infty} \boldsymbol{P}(X \geq i)$$

Proof: Let $p_i = \boldsymbol{P}(X = i)$, just for the sake of notational conciseness in our proof.

$$E[X] = \sum_{a \in \text{range}(X)} a \times \boldsymbol{P}(X = a)$$

$$= (0 \times p_0) + (1 \times p_1) + (2 \times p_2) + (3 \times p_3) + (4 \times p_4) + \cdots$$

$$= p_1 + (p_2 + p_2) + (p_3 + p_3 + p_3) + (p_4 + p_4 + p_4 + p_4) + \cdots$$

$$= (p_1 + p_2 + p_3 + p_4 + \cdots) + (p_2 + p_3 + p_4 + \cdots) + (p_3 + p_4 + \cdots) + (p_4 + \cdots) + \cdots$$

$$= \boldsymbol{P}(X \geq 1) + \boldsymbol{P}(X \geq 2) + \boldsymbol{P}(X \geq 3) + \boldsymbol{P}(X \geq 4) + \cdots$$

CC BY NC SA

Bernoulli Random Variables

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- Distribution of a Random Variable
- Expectation of a Random Variable

Example Distributions

- **Bernoulli**
- Binomial
- Uniform
- Geometric

CC BY NC SA

Bernoulli Random Variable

A Bernoulli random variable has a distribution given by the equation below, where $0 \leq p \leq 1$.

$$P(X = a) = \begin{cases} p, & \text{if } a = 1 \\ 1 - p, & \text{if } a = 0 \end{cases}$$

Example: Suppose we flip a coin that comes up heads 75% of the time. Let X be 1 if the coin comes up heads, and 0 if it comes up tails.

- We'd say "X is distributed as a Bernoulli random variable with $p = 0.75$"
- Or in writing we can say "$X \sim \text{Bernoulli}(0.75)$".
  - This is just shorthand for saying "X is distributed as a Bernoulli random variable with $p = 0.75$".

a	$P(X = a)$
0	0.25
1	0.75

CC BY NC SA

Bernoulli Random Variables and Expectation

Suppose $X \sim \text{Bernoulli}(0.5)$. What is $E[X]$?

$$E[X] = \sum_{a \in \text{range}(X)} a \times P(X = a)$$

$$P(X = a) = \begin{cases} p, & \text{if } a = 1 \\ 1 - p, & \text{if } a = 0 \end{cases}$$

We have that $E[X] = 0 \times P(X = 0) + 1 \times P(X = 1)$.

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

So: $E[X] = 0 \times 0.5 + 1 \times 0.5 = 0.5$

- Note: The expected value of X is a value that X can never equal.

CC BY NC SA

Bernoulli Random Variables and Expectation

Suppose $X \sim \text{Bernoulli}(p)$. What is $E[X]$?

$$E[X] = \sum_{a \in \text{range}(X)} a \times P(X = a)$$

$$P(X = a) = \begin{cases} p, & \text{if } a = 1 \\ 1 - p, & \text{if } a = 0 \end{cases}$$

We have that $E[X] = 0 \times P(X = 0) + 1 \times P(X = 1)$.

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

So: $E[X] = 0 \times (1 - p) + 1 \times p = p$

- That is, the expected value of a Bernoulli random variable is just the probability that it is 1 (i.e. that the "coin flip" comes up "heads").

CC BY NC SA

The Sum of Two Bernoulli Random Variables

Suppose $X_1 \sim \text{Bernoulli}(p)$ and $X_2 \sim \text{Bernoulli}(p)$. Suppose they are independent.

Let $Y = X_1 + X_2$.

- What is the distribution of Y?
- What is $E[Y]$?

CC BY NC SA

The Sum of Two Bernoulli Random Variables

Suppose $X_1 \sim \text{Bernoulli}(p)$ and $X_2 \sim \text{Bernoulli}(p)$. Suppose they are independent.

Let $Y = X_1 + X_2$.

$$P(Y = a) = \begin{cases} p^2, & \text{if } a = 2 \\ 2p(1-p), & \text{if } a = 1 \\ (1-p)^2, & \text{if } a = 0 \end{cases}$$

- What is the distribution of Y?
  - To get $a = 2$, both Bernoulli RVs must be 1.
  - To get $a = 1$, one Bernoulli RV is 1, and the other is 0.
  - To get $a = 0$, both Bernoulli RVs must be 0.
- What is $E[Y]$?

CC BY NC SA

The Sum of Two Bernoulli Random Variables

Suppose $X_1 \sim \text{Bernoulli}(p)$ and $X_2 \sim \text{Bernoulli}(p)$. Suppose they are independent.

Let $Y = X_1 + X_2$.

$$P(Y = a) = \begin{cases} p^2, & \text{if } a = 2 \\ 2p(1-p), & \text{if } a = 1 \\ (1-p)^2, & \text{if } a = 0 \end{cases}$$

- What is the distribution of Y?
  - To get $a = 2$, both Bernoulli RVs must be 1.
  - To get $a = 1$, one Bernoulli RV is 1, and the other is 0.
  - To get $a = 0$, both Bernoulli RVs must be 0.
- What is $E[Y]$?

$$\begin{aligned} E[Y] &= 2 \times p^2 + 1 \times 2p(1-p) + 0 \times (1-p)^2 \\ &= 2p^2 + 2p - 2p^2 \\ &= 2p \end{aligned}$$

CC BY NC SA

Binomial Random Variables

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- Distribution of a Random Variable
- Expectation of a Random Variable

Example Distributions

- Bernoulli
- **Binomial**
- Uniform
- Geometric

CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Let's work out the distribution by considering a specific example. Suppose that $X \sim \text{Binomial}(n = 5, p)$

- $P(HHTHT) = p \cdot p \cdot (1-p) \cdot p \cdot (1-p) = p^3(1-p)^2$
- $P(TTHHH) = (1-p) \cdot (1-p) \cdot p \cdot p \cdot p = p^3(1-p)^2$

CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Let's work out the distribution by considering a specific example. Suppose that $X \sim \text{Binomial}(n = 5, p)$

- $P(HHTHT) = p \cdot p \cdot (1-p) \cdot p \cdot (1-p) = p^3(1-p)^2$
- $P(TTHHH) = (1-p) \cdot (1-p) \cdot p \cdot p \cdot p = p^3(1-p)^2$
- $P(X = 3) =$???

https://joshh.ug/D

CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Let's work out the distribution by considering a specific example. Suppose that $X \sim \text{Binomial}(n = 5, p)$

- $P(HHTHT) = p \cdot p \cdot (1-p) \cdot p \cdot (1-p) = p^3(1-p)^2$
- $P(TTHHH) = (1-p) \cdot (1-p) \cdot p \cdot p \cdot p = p^3(1-p)^2$
- $P(X = 3) = (\#\text{ of sequences with } 3 \text{ heads})p^3(1-p)^2$

$$= \binom{5}{3} p^3(1-p)^2$$

CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Let's work out the distribution by considering a specific example. Suppose that $X \sim \text{Binomial}(n = 5, p)$

- $P(X = i) = (\#\text{ of sequences with i heads})p^i(1-p)^{5-i}$

$$= \binom{5}{i} p^i (1-p)^{5-i} \quad \text{for } i = 0,\ 1,\ 2,\ 3,\ 4,\ 5$$

CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Let's work out the distribution by considering a specific example. Suppose that $X \sim \text{Binomial}(n, p)$

- $P(X = i) = (\#\text{ of sequences with i heads}) p^i (1-p)^{n-i}$

$$= \binom{n}{i} p^i (1-p)^{n-i} \quad \text{for } i = 0, 1, 2, \ldots, n$$

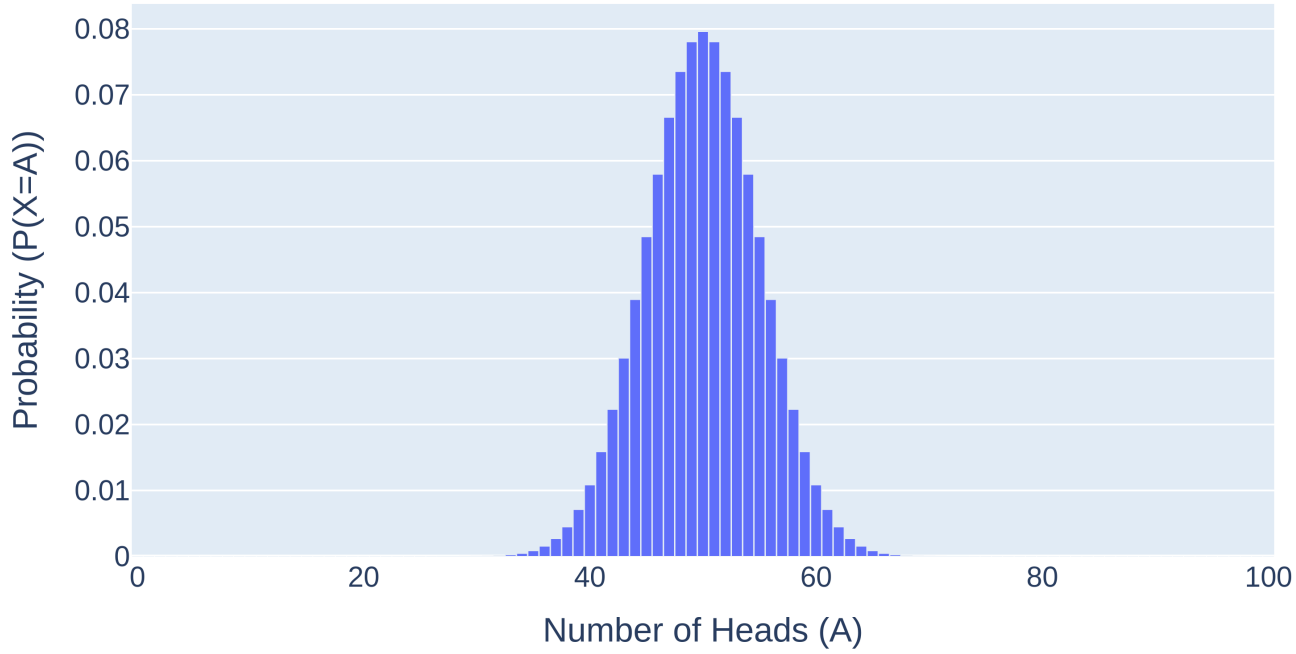
CC BY NC SA

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

- $P(X = i) = (\#\text{ of sequences with i heads})p^i(1-p)^{n-i}$

$$= \binom{n}{i} p^i (1-p)^{n-i} \qquad \text{for } i = 0, 1, 2, \ldots, n$$



$\text{Binomial}(100,\ 0.5)$

Binomial Distribution

The Binomial distribution models counting the number of heads if we flip n coins that come up heads with probability p.

Suppose that $X \sim \textbf{Binomial}(n, p)$

- $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ for $i = 0, 1, 2, \ldots, n$

Exercise: Verify that $\sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = 1$.

We'll cover the expectation of a Binomial RV next lecture.

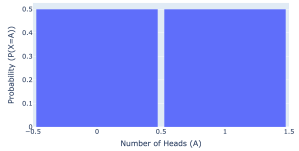
- You can do it using the usual sum, but that's much messier than the nicer approach from the next lecture.

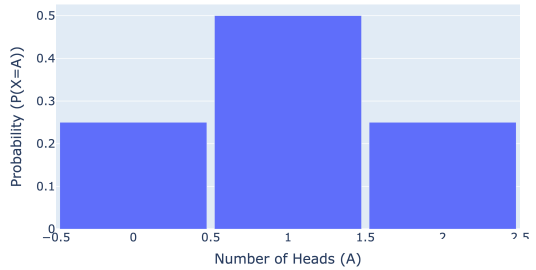
CC BY NC SA

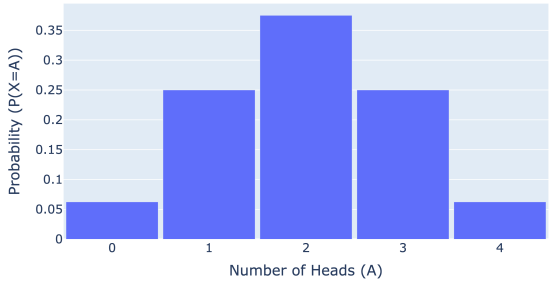
Binomial vs. Bernoulli Distribution

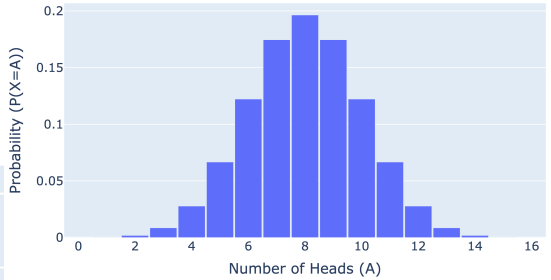
Note that a binomial random variable $X \sim \text{Binomial}(n, p)$ is simply the sum of n independent random variables that are $\text{Bernoulli}(p)$.

- We have not yet defined independent random variables.









Uniform Random Variables

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- Distribution of a Random Variable
- Expectation of a Random Variable

Example Distributions

- Bernoulli
- Binomial
- **Uniform**
- Geometric

CC BY NC SA

Uniform Random Variable

A uniform random variable is equally likely to be any of $1, 2, \ldots, n$:

$$P(X = i) = \frac{1}{n}, \quad \text{for } i = 1, 2, \ldots, n$$

Example: Roll a fair six-sided die and let X be the number that comes up.

- We'd say "X is uniformly distributed on $\{1, \ldots, 6\}$"
- Or in writing we can say "$X \sim \text{Uniform}\{1, \ldots, 6\}$".

Suppose $X \sim \text{Uniform}\{1, \ldots, n\}$. What is $E[X]$?

- $E[X] = \sum_{i=1}^{n} i \times \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$
- For our die, $E[X] = \frac{7}{2} = 3.5$.

CC BY NC SA

Geometric Random Variables

Lecture 19, CS70 Fall 2026

Definitions

- Definition of a Random Variable
- Distribution of a Random Variable
- Expectation of a Random Variable

Example Distributions

- Bernoulli
- Binomial
- Uniform
- **Geometric**

CC BY NC SA

Geometric Distribution

Imagine a random experiment where we repeatedly flip a coin with a probability of heads equal to p until we get heads. The flips are independent.

- Let X be the number of flips.

This distribution is called the geometric distribution, i.e. $X \sim \text{Geometric}(p)$

In the coin flipping example, what are the domain and range of X?

- Domain: Any sequence of 0 or more tails, followed by a heads.
- Range: Any integer greater than or equal to 1, i.e. $\mathbb{N}^+$

What is $P(X = 1)$?

- p

CC BY NC SA

Geometric Distribution

Imagine a random experiment where we repeatedly flip a coin with a probability of heads equal to p until we get heads. The flips are independent.

- Let X be the number of flips.

What is $P(X = 3)$?

A. p^3

B. $(1-p)^2 p$

C. $(1-p)p^2$

D. $3(1-p)^2 p$

E. $(1-p)^3$

https://joshh.ug/E



CC BY NC SA

Geometric Distribution

Imagine a random experiment where we repeatedly flip a coin with a probability of heads equal to p until we get heads. The flips are independent.

- Let X be the number of flips.

What is $P(X = 3)$?

- $X = 3$ exactly when the flips are TTH: two tails, then a head.
- The flips are independent, so $P(X = 3) = (1 - p) \cdot (1 - p) \cdot p = (1 - p)^2 \cdot p$.

cc BY NC SA

Geometric Distribution

Imagine a random experiment where we repeatedly flip a coin with a probability of heads equal to p until we get heads. The flips are independent.

- Let X be the number of flips.

This distribution is called the geometric distribution, i.e. $X \sim \text{Geometric}(p)$

What is $P(X = i)$?

- $P(X = 2) = (1 - p) \cdot p$
- $P(X = 3) = (1 - p)^2 \cdot p$
- $P(X = i) = (1 - p)^{i-1} \cdot p$

CC BY NC SA

Verifying that the Probabilities Sum to 1

Imagine a random experiment where we repeatedly flip a coin with a probability of heads equal to p until we get heads. The flips are independent.

- Let X be the number of flips.

This distribution is called the geometric distribution, i.e. $X \sim \text{Geometric}(p)$

- $P(X = i) = (1-p)^{i-1}p$ for $i = 1, 2, \ldots$

Let's verify that the probabilities sum to 1.

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$\sum_{i=1}^{\infty} P(X = i) = \sum_{i=1}^{\infty} (1-p)^{i-1}p = p\sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

cc BY NC SA

Geometric Distribution

Suppose $X \sim \text{Geometric}(0.2)$.

- That is, we flip a coin that comes up heads with probability 0.2 until we get heads, and X is the number of flips.

What is $P(X > 6)$?

A. $0.8^5 \cdot 0.2$

B. 0.8^5

C. 0.8^6

D. 0.8^7

E. $1 - 0.8^6$

https://joshh.ug/F

CC BY NC SA

Geometric Distribution

Suppose $X \sim \text{Geometric}(0.2)$. What is $P(X > 6)$?

- $X > 6$ exactly when the first 6 flips are all tails.
- So $P(X > 6) = 0.8^6 \approx 0.262$.

More generally, if $X \sim \text{Geometric}(p)$:

$$P(X > a) = (1-p)^a \text{ for } a = 0, 1, 2, \ldots$$

- $X > a$ exactly when the first a flips are all tails.

CC BY NC SA

Geometric Random Variables and the Memoryless Property

Geometric random variables (and processes modeled by them) are memoryless:

- It basically means: Even if you've waited a long time, the chance of the event occurring soon isn't any better.
- Example: You're flipping a coin and it came up tails 10 times in a row. The chance of the next heads isn't better just because you had 10 tails.
- Example: You've lost all your money on your last 5 purchases of stock options. The chance of getting a good return is equally bad, and it doesn't matter that you've made 5 bad trades in a row.

CC BY NC SA

Geometric Random Variables and the Memoryless Property

Geometric random variables (and processes modeled by them) are **memoryless**:

- It basically means: Even if you've waited a long time, the chance of the event occurring soon isn't any better.

In terms of probability, we can state this as "If we've already tossed a coin m times without getting a head, the probability that we need n additional tosses is the same as if we hadn't tossed the first m coins at all."

Or symbolically we have:

$$P(X > n + m \mid X > m) = P(X > n)$$

CC BY NC SA

Geometric Random Variables and the Memoryless Property

Geometric random variables (and processes modeled by them) are memoryless:

- It basically means: Even if you've waited a long time, the chance of the event occurring soon isn't any better.

Proof of $\mathrm{P}(X > n+m \mid X > m) = P(X > n)$

$$\mathrm{P}(X > n+m \mid X > m) = \frac{P(X > n+m)}{P(X > m)}$$

$$= \frac{(1-p)^{n+m}}{(1-p)^{m}}$$

$$= (1-p)^{n}$$

$$= P(X > n)$$

CC BY NC SA

Without Proof: The Expectation of a Geometric Random Variable

Suppose we have a coin where the probability of heads is $p = 0.2$. How many flips do you think we need to wait on average before we get our first heads?

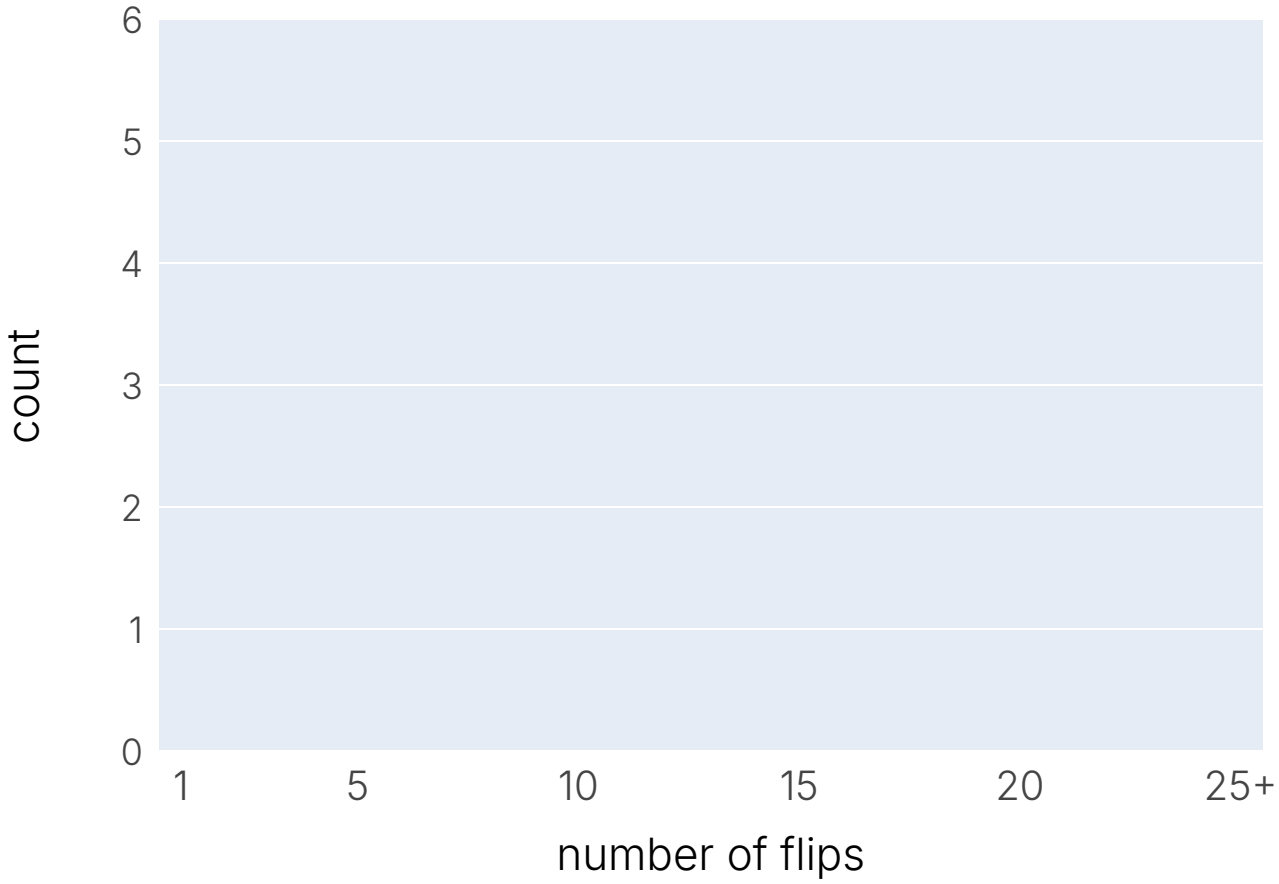
Run once | Run 100× | Run 10,000× | Reset

Last run: N/A

Number of runs: N/A

Total flips across all runs: N/A

Average flips per run: N/A



CC BY NC SA

Without Proof: The Expectation of a Geometric Random Variable

Suppose we have a coin where the probability of heads is $p = 0.2$. How many flips do you think we need to wait on average before we get our first heads?

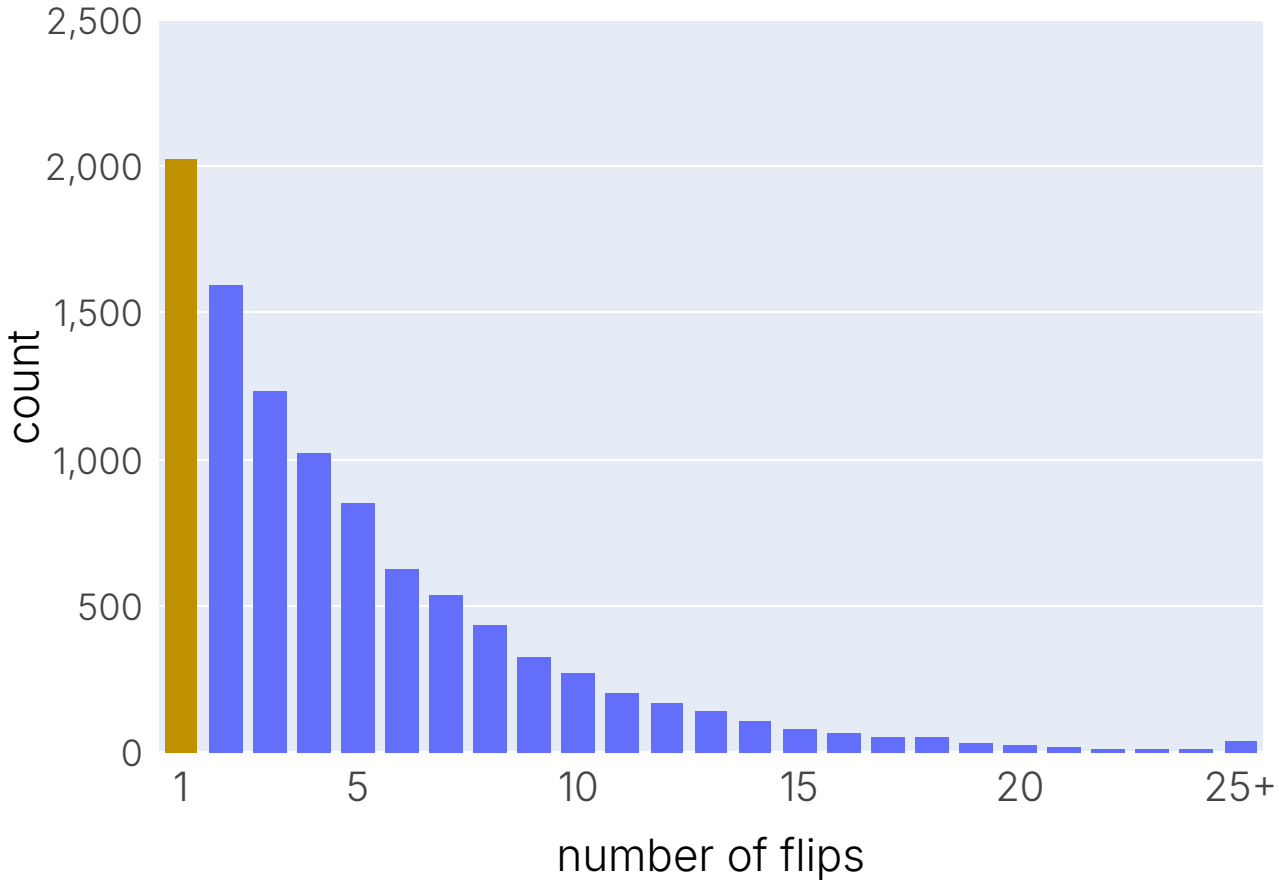
Run once | Run 100× | Run 10,000× | Reset

Last run: H (1 flip)

Number of runs: 10,000

Total flips across all runs: 49,796

Average flips per run: 49,796/10,000 = 4.98



If $X \sim \text{Geometric}(0.2)$, then $E[X] = 5$

Expectation of a Geometric RV

Suppose $X \sim \mathbf{Geometric}(p)$. Since the range of X is a subset of $\mathbb{N}$, i.e. $\{0, 1, 2, \ldots\}$, its expectation is given by the tail sum formula:

$$E[X] = \sum_{i=1}^{\infty} \boldsymbol{P}(X \geq i)$$

So we need to know $\boldsymbol{P}(X \geq i)$. What is it?

- Hint: What's the probability that the first $i - 1$ flips are all tails?
- Hint 2: You don't need to worry about the flips after the first $i - 1$.

CC BY NC SA

Expectation of a Geometric RV

Suppose $X \sim \text{Geometric}(p)$. Since the range of X is a subset of $\mathbb{N}$, i.e. $\{0, 1, 2, \ldots\}$, its expectation is given by the tail sum formula:

$$E[X] = \sum_{i=1}^{\infty} \boldsymbol{P}(X \geq i)$$

$\boldsymbol{P}(X \geq i)$ is the probability that the first $i-1$ flips are tails, which is given by:

$$\boldsymbol{P}(X \geq i) = (1-p)^{i-1}$$

Plugging this into the tail sum formula, we have:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$E[X] = \sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{1}{1-(1-p)} = \frac{1}{p}$$

CC BY NC SA

Summary

- A random variable X is a function $\Omega \to \mathbb{R}$, i.e. $X(\omega)$ is a real number for every $\omega \in \Omega$.
- The distribution of X is the set of value/probability pairs $\{(a, P(X = a)) : a \in \mathbf{range}(X)\}$.
- The expectation of X is $E[X] = \sum_{a \in \mathbf{range}(X)} a \times P(X = a)$.
  - If the range of X is a subset of $\mathbb{N}$, the tail sum formula also works: $E[X] = \sum_{i=1}^{\infty} \boldsymbol{P}(X \geq i)$
- Example distributions:
  - $X \sim \mathbf{Bernoulli}(p)$: $P(X = 1) = p$ and $P(X = 0) = 1 - p$. $E[X] = p$.
  - $X \sim \mathbf{Binomial}(n, p)$: the number of heads in n flips. $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$.
  - $X \sim \mathbf{Uniform}\{1, \ldots, n\}$: $P(X = i) = 1/n$. $E[X] = (n+1)/2$.
  - $X \sim \mathbf{Geometric}(p)$: the number of flips until the first heads. $P(X = i) = (1-p)^{i-1} p$ and $P(X > a) = (1-p)^a$. $E[X] = 1/p$.
  - Geometric random variables are memoryless: $P(X > n + m \mid X > m) = P(X > n)$.

CC BY NC SA