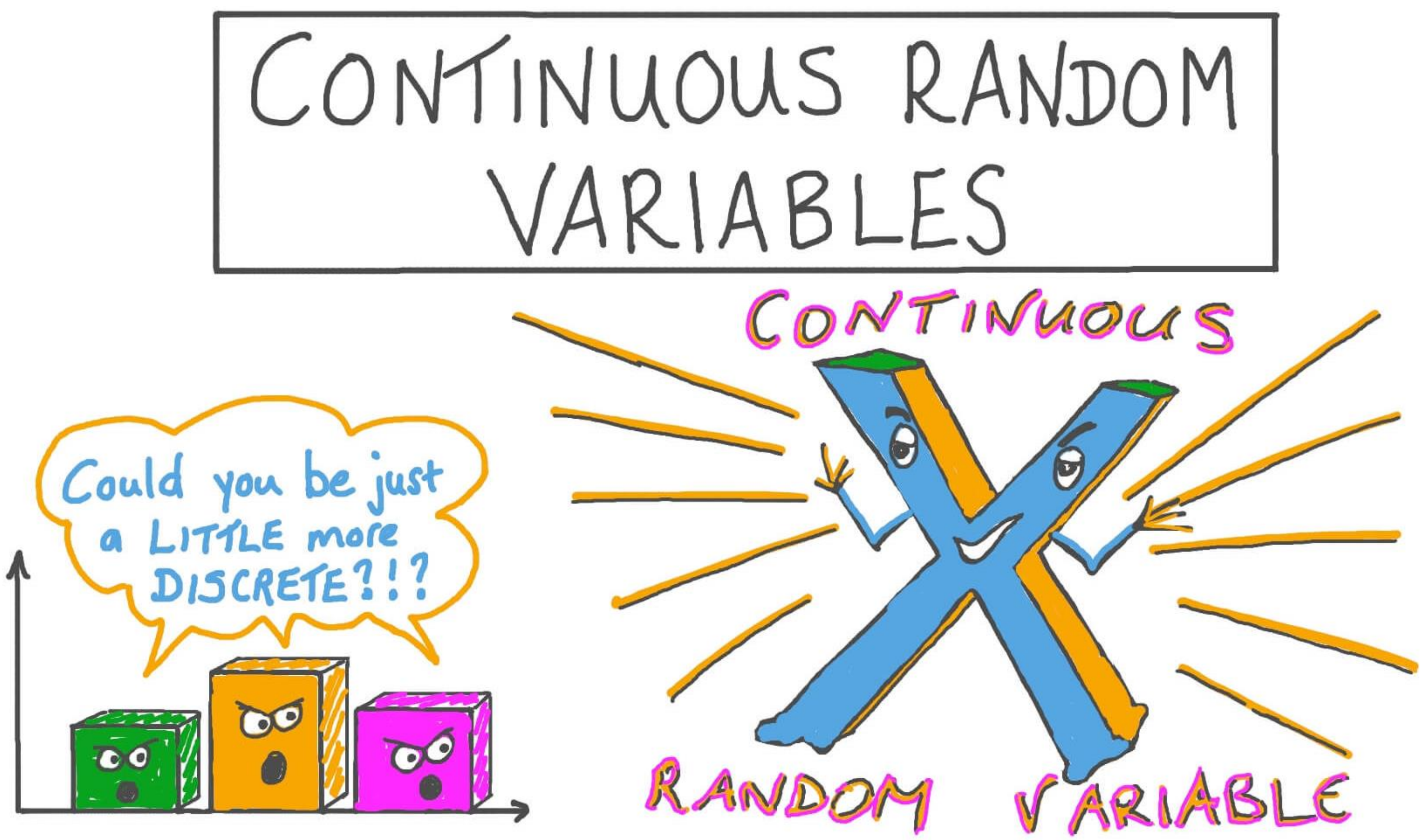
CS 237: PROBABILITY IN COMPUTING

INSTRUCTORS: TIAGO JANUARIO, SOFYA RASKHODNIKOVA

BOSTON UNIVERSITY



LECTURE 6

Last time

- Continuous Probability Spaces
- Anomalies with Continuous Probability

Today

- Random variables
- Probability Mass Function

GET READY FOR VOTING!

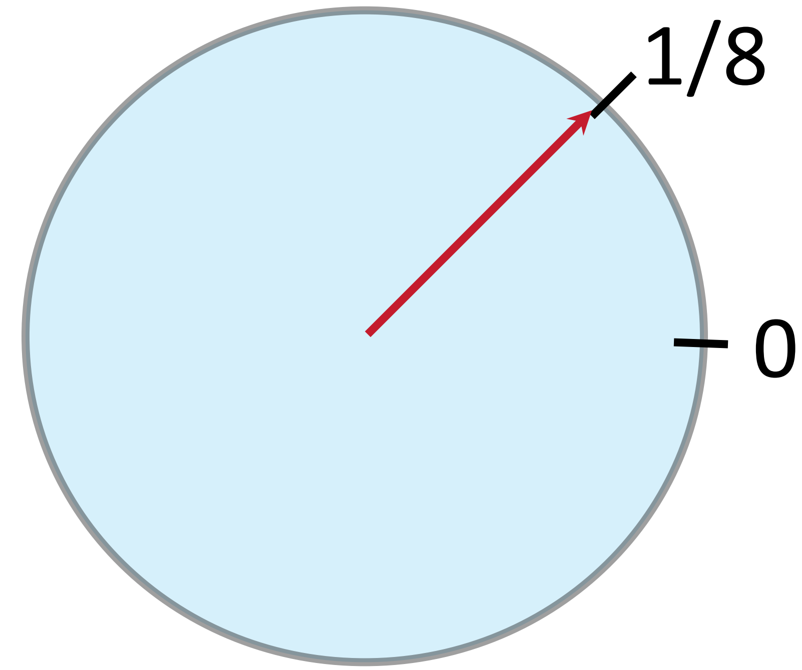
- Students from A1 section:
  1. Enter this URL: **https://app.tophat.com/e/**037447
  2. Login as needed using your Top Hat account info.
  3. If asked about enrolling, click *Enroll*.
- Students from A2 section:
  1. Enter this URL: **https://app.tophat.com/e/**606363
  2. Login as needed using your Top Hat account info.
  3. If asked about enrolling, click *Enroll*.

TOP HAT QUESTION

- What is the probability that the dial lands on 1/8?

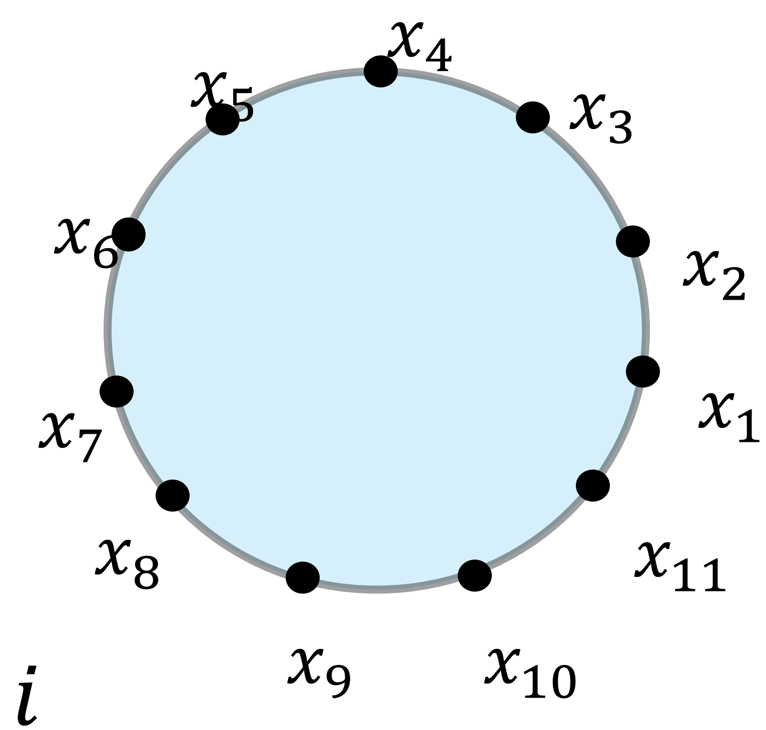
(A) $^1/_8$ = 0.125 (B) 0

(C) 0.01 (D) none of the above



SPINNERS

- What is the probability that the dial lands on 1/8?
- Suppose we suspect that it is, say, 0.1
- Pick 11 distinct points on the circle



by symmetry: $Pr(\text{dial lands on } x_i) = 0.1 \text{ for all } i$

by additivity: $Pr(\text{dial lands on one of } x_1, \dots, x_{11}) = 11 \cdot 0.1 = 1.1$

contradiction!

SPINNERS

- Let p be the probability of an outcome, e.g., 1/8
- We can show that p = 0 by contradiction
- If $p \neq 0$: pick distinct points $x_1, \ldots, x_k$, where $k > 1/p$

by symmetry: $Pr(\text{dial lands on } x_i) = p \; \forall i \in \{1, 2, \ldots, k\}$

by additivity: $Pr(\text{dial lands on one of } x_1, \ldots, x_k) = k \cdot p > 1$

contradiction!

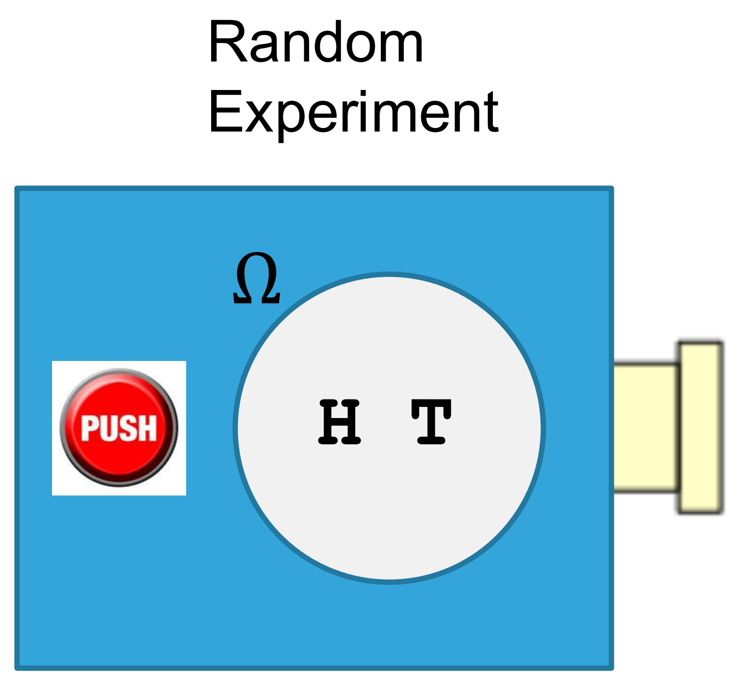
BEYOND EVENTS

- In many settings, we would like to know more than the probabilities of events
  - How many days will a system run before it fails?
  - How many people support candidate A?
  - How much money will I make if I invest in this stock?
  - How many heads in n coin tosses?
  - How many die rolls until we see a 6?
  - How much time will it take me to a mine a Bitcoin on my laptop?

RANDOM EXPERIMENT

- A Random Experiment is a repeatable procedure that produces uncertain outcomes from a well-defined sample space.

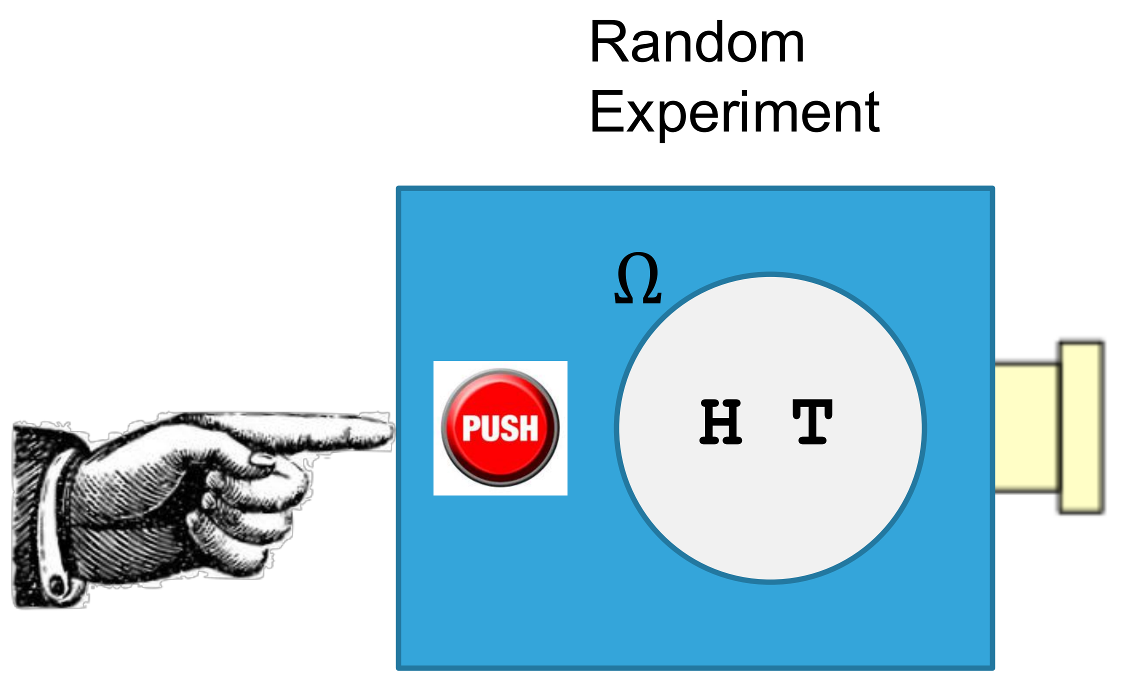
- Example: Flip a coin!



RANDOM EXPERIMENT

- A Random Experiment is a repeatable procedure that produces uncertain outcomes from a well-defined sample space.

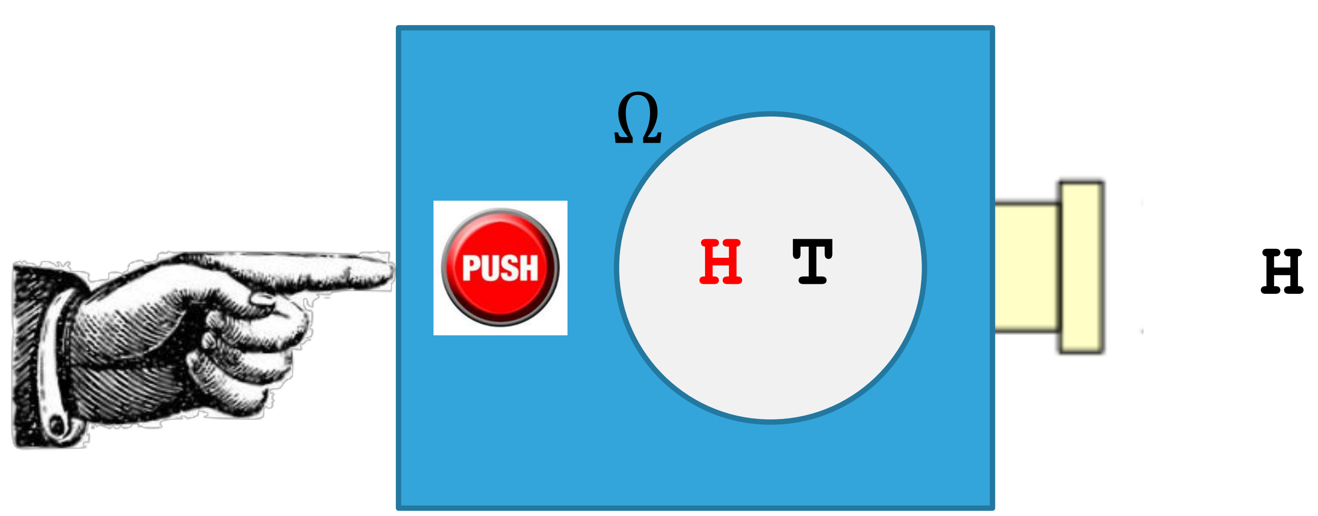
- Example: Flip a coin!



RANDOM EXPERIMENT

- A Random Experiment is a repeatable procedure that produces uncertain outcomes from a well-defined sample space.

- Example: Flip a coin!

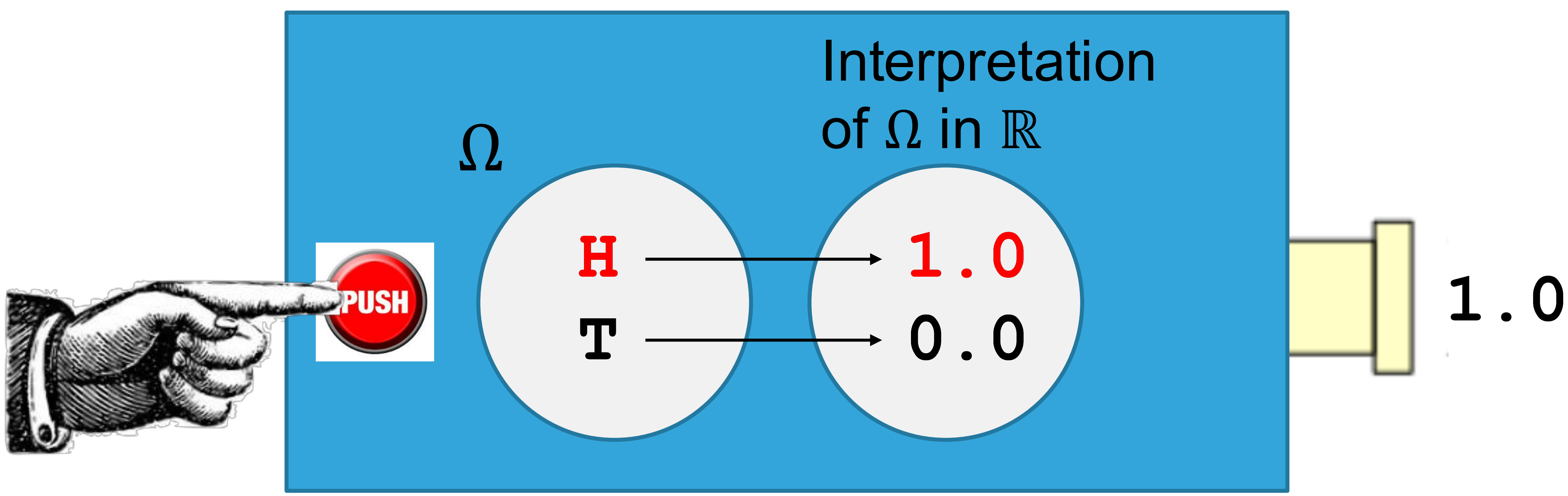


A RANDOM VARIABLE X IS A FUNCTION

- It maps a sample space Ω into $\mathbb{R}$, *i.e., interprets* the random experiment as a real number.
- Now when an outcome is requested, the sample point is translated into a real number:

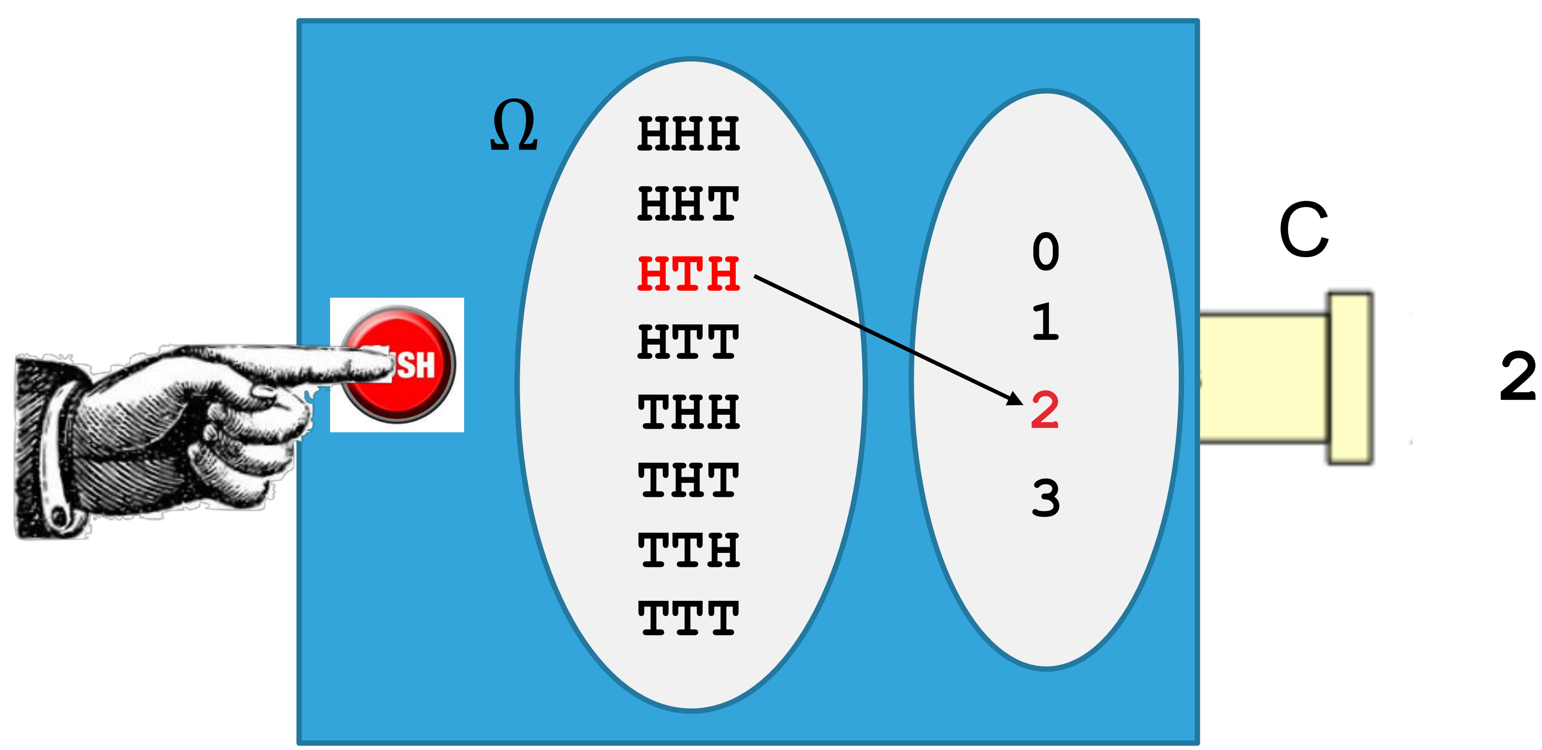
$$X : \Omega \longrightarrow \mathbb{R}$$

- X = "Flip a coin and count the number of heads showing"



RANDOM VARIABLES

- A random variable will not in general simply rename outcomes with real numbers, it will compute some useful information about the outcome.
- C = "Flip 3 coins and return the number of heads showing"



RANDOM VARIABLES: DEFINITION

- Definition: A random variable X is a function whose domain is the sample space and co-domain is a subset of $\mathbb{R}$:

$$X : \Omega \to \mathbb{R}$$

- X maps each outcome $\omega \in \Omega$ to a value $X(\omega) \in \mathbb{R}$
- Once an outcome is requested it is translated into a real number
- The name "random variable" is a misnomer, random variables are actually functions

RANDOM VARIABLES: DEFINITION

- Definition: A random variable X is a function from the sample space Ω to numbers.
- For a given value x_k ,the event

$$\{X = x_k\} = \{\, \omega \in \Omega : X(\omega) = x_k \,\}$$

- is the subset of outcomes in the sample space that make X equal to x_k.

$C : \Omega \to \{0, 1, 2, 3\}$

RANDOM VARIABLES - EXAMPLE

- Experiment: toss a fair coin 3 times
- Let C be the number of heads that appear
- How should we think about C?

$\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

$C(HHH) = 3$, $C(HHT) = C(HTH) = C(THH) = 2$

$C(TTT) = 0$, $C(TTH) = C(THT) = C(HTT) = 1$

$\{C = 2\} = \{\omega \in \Omega \mid C(\omega) = 2\} = \{HHT, HTH, THH\}$

RANDOM VARIABLES - EXAMPLE

- Experiment: toss a fair coin 3 times
- Let C be the number of heads that appear
- How should we think about C?
- C maps each outcome to a value

$$C(HHH) = 3 \qquad C(HHT) = C(HTH) = C(THH) = 2$$

$$C(TTT) = 0 \qquad C(TTH) = C(THT) = C(HTT) = 1$$

RANDOM VARIABLES - EXAMPLE

- Experiment: toss a fair coin 3 times
- Let C be the number of heads that appear
- How should we think about C?
- C maps each outcome to a value

$$C(HHH) = 3 \qquad C(HHT) = C(HTH) = C(THH) = 2$$

$$C(TTT) = 0 \qquad C(TTH) = C(THT) = C(HTT) = 1$$

- C maps each outcome to a value , i.e., C is a function from the sample space to values

$$C: \Omega \rightarrow \{0,1,2,3\}$$

TOP HAT QUESTION

Consider the following experiment:
"I toss a coin until the first heads appears. Let Y be the total number of coin tosses."

- What is the range of the random variable Y?:

A. $R_Y = \{H, T\}$
B. $R_Y = \{1,2,3,4,5 \ldots\}$
C. $R_Y = \{\sum \frac{1}{2^n} : n \in N\}$
D. $R_Y = \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \ldots\right\}$
E. $R_Y = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \ldots\right\}$

Random variables and Events

C = # of heads after tossing a fair coin 3 times

RANDOM VARIABLES: NOTATION

- A very typical notion for probabilities involving a random variable X is:

Pr(< some Boolean expression involving X >)

$Pr(C \neq 2) = 1 - Pr(C = 2) = Pr(\{C=0\} \cup \{C=1\} \cup \{C=3\}) = Pr(C=0) + Pr(C=1) + Pr(C=3)$

$Pr(C=1) = Pr(\{C=1\})$

$$\Pr(X = k) = Pr_X(k) = Pr(\{\omega \in \Omega : X(\omega) = k\})$$

$$\Pr(X \neq k) = 1 - Pr_X(k)$$

$$\Pr(X \leq k) = \sum_{j \leq a \leq k} Pr_X(a)$$

- Here, $Pr_X(k)$ means the probability that X returns k

$Pr(C \leq 1) = 1 - Pr(C > 1) = Pr(C=0) + Pr(C=1)$

$Pr(\{C=1\}) = Pr(\{\omega \in \Omega : C(\omega) = 1\}) = Pr(\{TTH, THT, HTT\})$

RANDOM VARIABLES AND EVENTS

- The set of outcomes that map to specific values of an r.v. also defines an event.
- C = number of heads in three tosses

$$\underbrace{TTT}_{C=0},\underbrace{HTT,THT,\ TTH}_{C=1},\underbrace{HHT,\ HTH,\ THH}_{C=2},\underbrace{HHH}_{C=3}$$

C(HHT) = C(HTH) = C(THH) = 2

$$Pr(C = 2) = Pr(\{HHT, HTH, THH\}) = 3/8$$

Pr(HHT) + Pr(HTH) + Pr(THH) = 3/8

$$Pr(C \leq 1) = Pr(\{TTT, HTT, THT, TTH\}) = 1/2$$

Examples of Random Variables

RANDOM VARIABLES

- Discrete random variables = discrete range
  - The possible values (outputs) are listable/countable
  - Number of H in 3 tosses, sum of two rolls of a die, number of tosses until H, …
- Continuous random variables = continuous range
  - The possible values (outputs) form an interval of $\mathbb{R}$
  - Spinners, darts, …

RANDOM VARIABLES: EXAMPLES

- Experiment: toss a fair coin 3 times

$$M = \begin{cases} 1 & \text{if all three tosses match (all H or all T)} \\ 0 & \text{otherwise} \end{cases}$$

$$M: \Omega \rightarrow \{0,1\}$$

$$M(HHH) = M(TTT) = 1$$

$$M(HHT) = M(HTH) = M(HTT) = 0$$
$$M(THH) = M(THT) = M(TTH) = 0$$

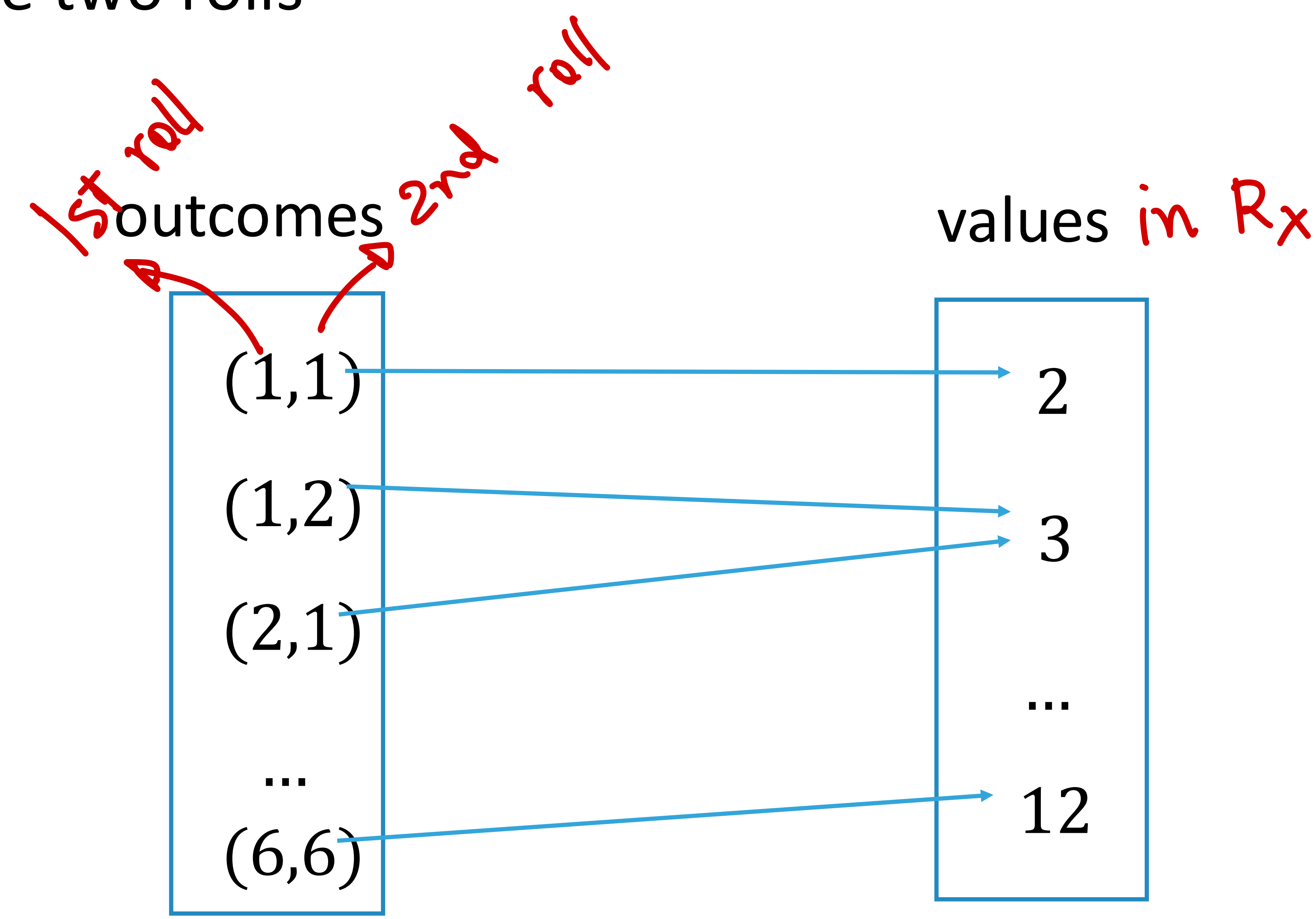
- An R.V. whose codomain is {0, 1} is an *indicator* variable.

This R.V. is super important!

- Indicator variables are another way to define *events.*

RANDOM VARIABLES: EXAMPLES

- Experiment: roll a fair die twice
- X = sum of the two rolls

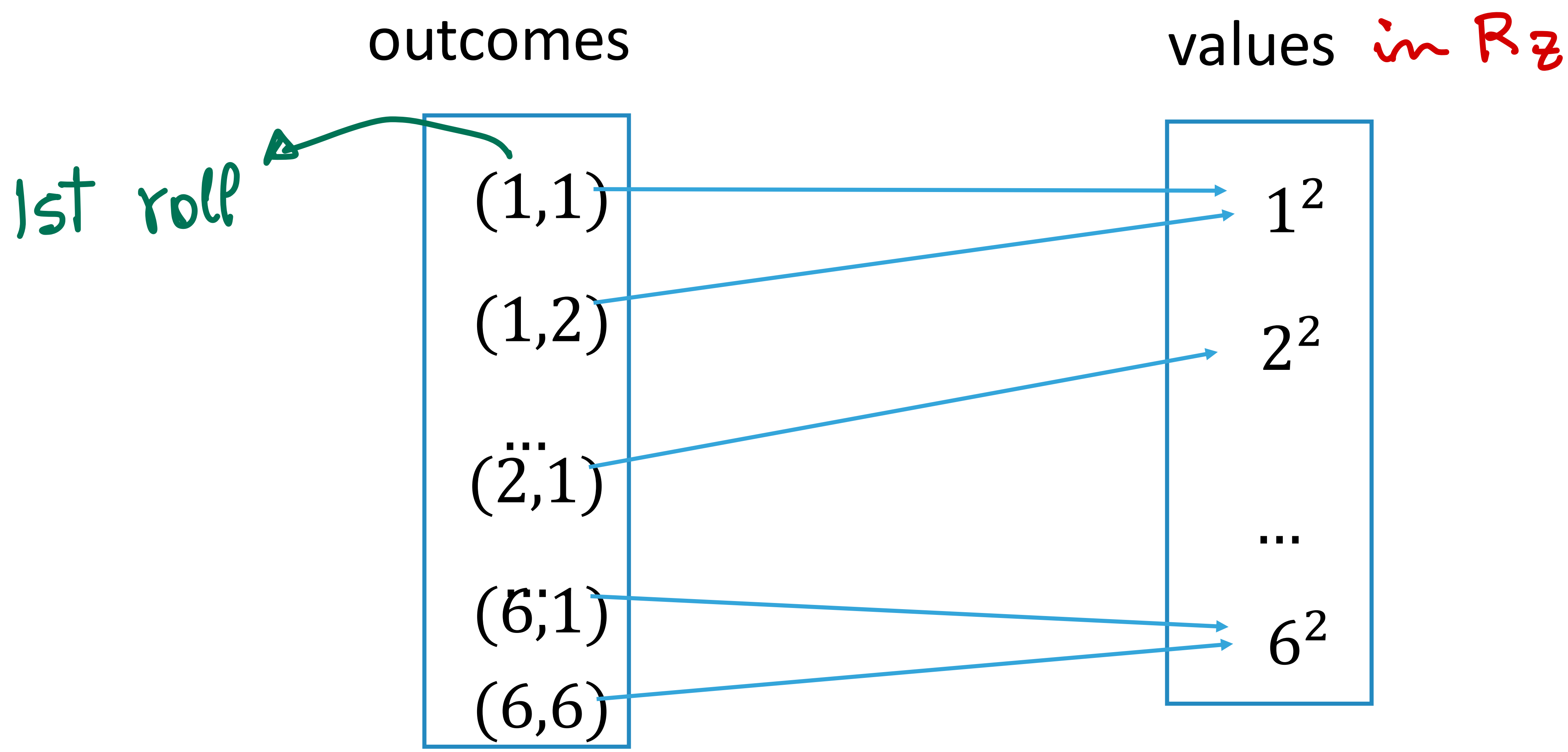


$$X: \Omega \to \{2,3,\dots,12\}$$

RANDOM VARIABLES: EXAMPLES

- Experiment: roll a fair die twice
- Z = first roll squared ⇒ 1st-roll to the power of 2

$R_Z = \{1^2, 2^2, 3^2, 4^2, 5^6, 6^2\}$



RANDOM VARIABLES: EXAMPLES

- Experiment: sample a value uniformly from [-1, 1]

$$X: [-1,1] \to \mathbb{R}, X(r) = r \quad \forall r \in [-1,1]$$

domain (inputs)	values (outputs)
$[-1,1]$	$[-1,1]$

RANDOM VARIABLES: EXAMPLES

- Experiment: sample a value uniformly from [-1, 1]

$$Y: [-1,1] \to \mathbb{R}, Y(r) = r^2 \quad \forall r \in [-1,1]$$

domain (inputs)	values (outputs)
$[-1,1]$	$[0,1]$

RANDOM VARIABLES: EXAMPLES

- Experiment: sample a value x uniformly from [-1, 1]

$$Z: [-1,1] \to \mathbb{R}, Z(r) = \begin{cases} 1 & \text{if } r > 0 \\ 0 & \text{if } r = 0 \\ -1 & \text{if } r < 0 \end{cases}$$

, w.p. 0.5
, w.p. 0
, w.p. 0.5

domain (inputs)

$[-1,1]$

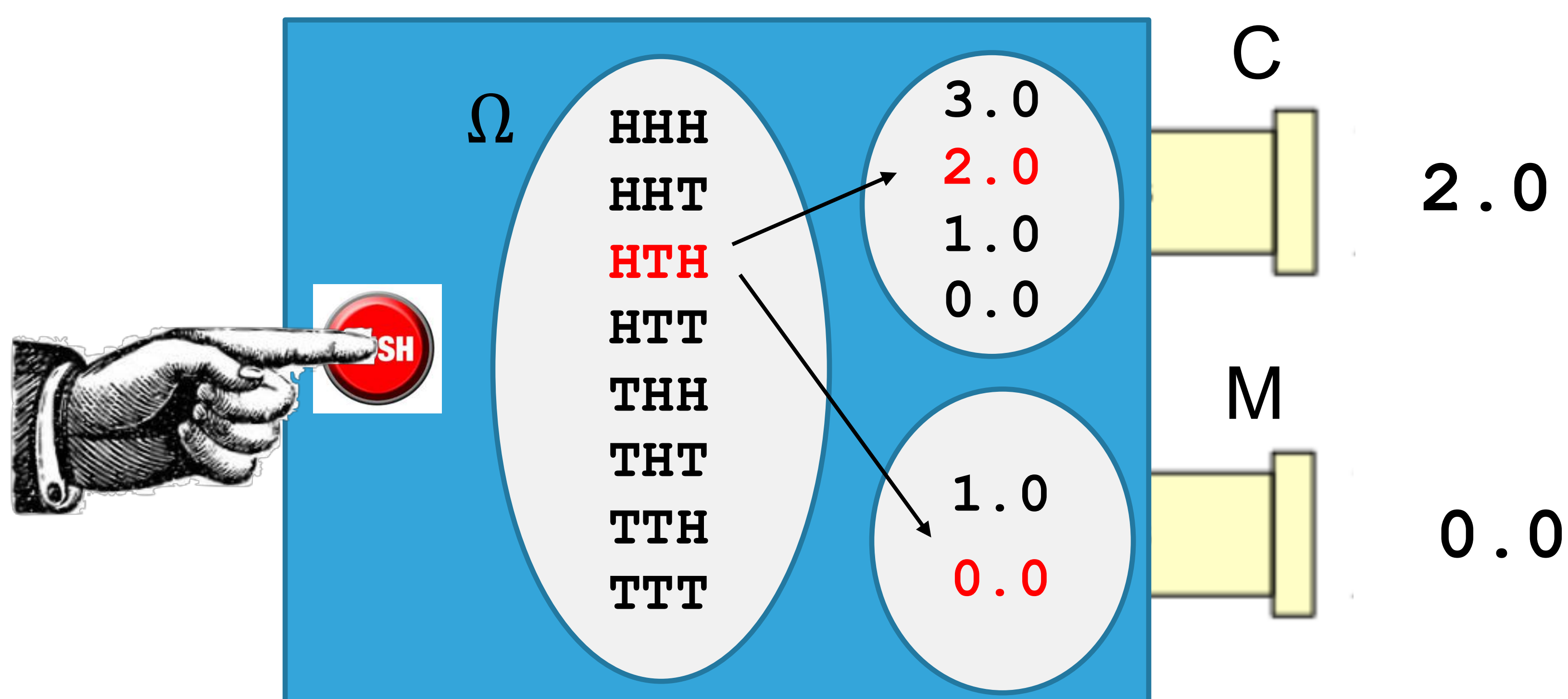
values (outputs)

$\{1,0,-1\}$

discrete Range

JOINT RANDOM VARIABLES

- Joint Random Variables are combinations of 2 or more RVs, giving *multiple interpretations* of the same experiment.
- Example (from LLM):
- “Flip 3 coins and
  (C) return the number of heads showing;
  (M) return 1 if all heads or all tails, 0 otherwise”



TOP HAT QUESTION

Consider the following experiment:
"Flip 3 coins and
(T) return the number of tails showing;
(V) return 1 if there are more heads showing than tails, 0 if more tails than heads

- Suppose the random experiment chooses HHT from Ω then the result is:

A. T = 1 V = 0
B. T = 2 V = 0
C. T = 2 V = 1
D. None of the above

T(HHT) = 1

V(HHT) = 1

PROBABILITY MASS FUNCTION (PMF)

$x_k \in R_X$

- For a given value x_k ,the event

$$\{X = x_k\} = \{\, \omega \in \Omega : X(\omega) = x_k \,\}$$

- is the subset of outcomes in the sample space that make X equal to x_k.
- The **Probability Mass Function (PMF)** of X, denoted by P_X ,assigns probabilities to these values:

$$P_X(x_k) = P(X = x_k) = \Pr(\{\omega \in \Omega : X(\omega) = x_k\})$$

- In other words, the PMF gives the probability that the random variable takes each of its possible values.

PROBABILITY MASS FUNCTION (PMF)

- A discrete random variable X takes values in a set $R_X = \{x_1, x_2, \dots\}$.
- The probability that X equals a particular value $x \in R_X$ is given by the **PMF**:

$$P_X(x) = P(X = x).$$

- The **PMF** assigns a probability to each possible value in the range of X.
- These probabilities must satisfy:

$$P_X(x) \geq 0 \; \textit{for all } x, \textit{and} \sum_x P_X(x) = 1.$$

PMF: EXAMPLE

- Let X be the number of heads obtained when a fair coin is tossed twice. Find R_X and P_X → PMF

PMF: EXAMPLE

- Let X be the number of heads obtained when a fair coin is tossed twice. Find R_X and P_X

Range ← ; P.M.F.

- Sample space: $S=\{HH,HT,TH,TT\}$

$$R_X = \{0,1,2\}$$

- $P_X(k) = P(X = k)$ for $k = 0,1,2$
- $P_X(0) = P(X = 0) = P(TT) = \frac{1}{4}$

$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

- $P_X(1) = P(X = 1) = P(\{HT, TH\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
- $P_X(2) = P(X = 2) = P(HH) = \frac{1}{4}$

PROBABILITY MASS FUNCTION (PMF)

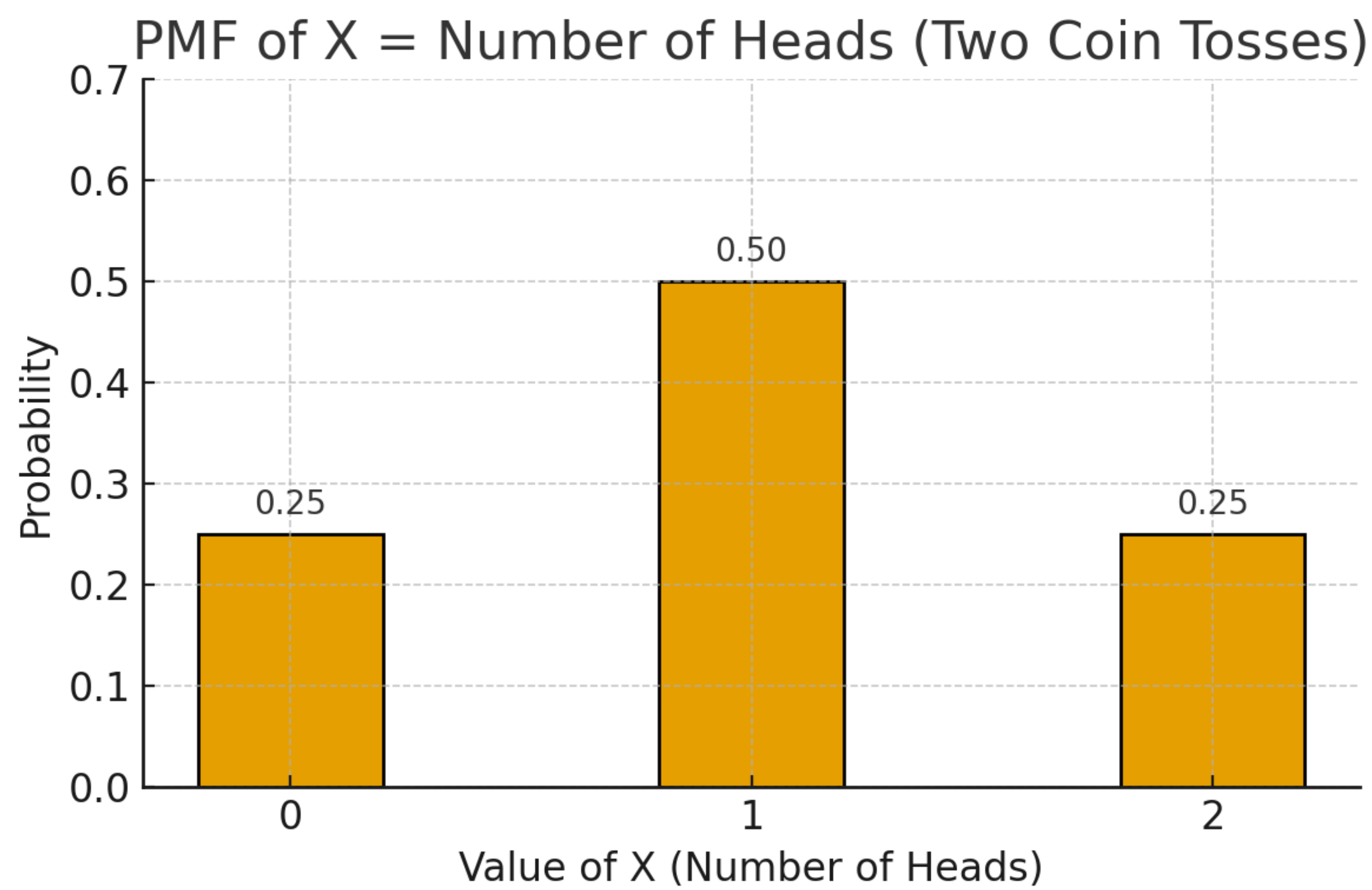
- Example: Toss two fair coins. Let $X =$ number of heads.

PROBABILITY MASS FUNCTION (PMF)

- Example: Toss two fair coins. Let X = number of heads.

$$P_X(0) = \frac{1}{4}, P_X(1) = \frac{1}{2}, P_X(2) = \frac{1}{4}.$$

- The PMF shows how probability is distributed across the values of X.

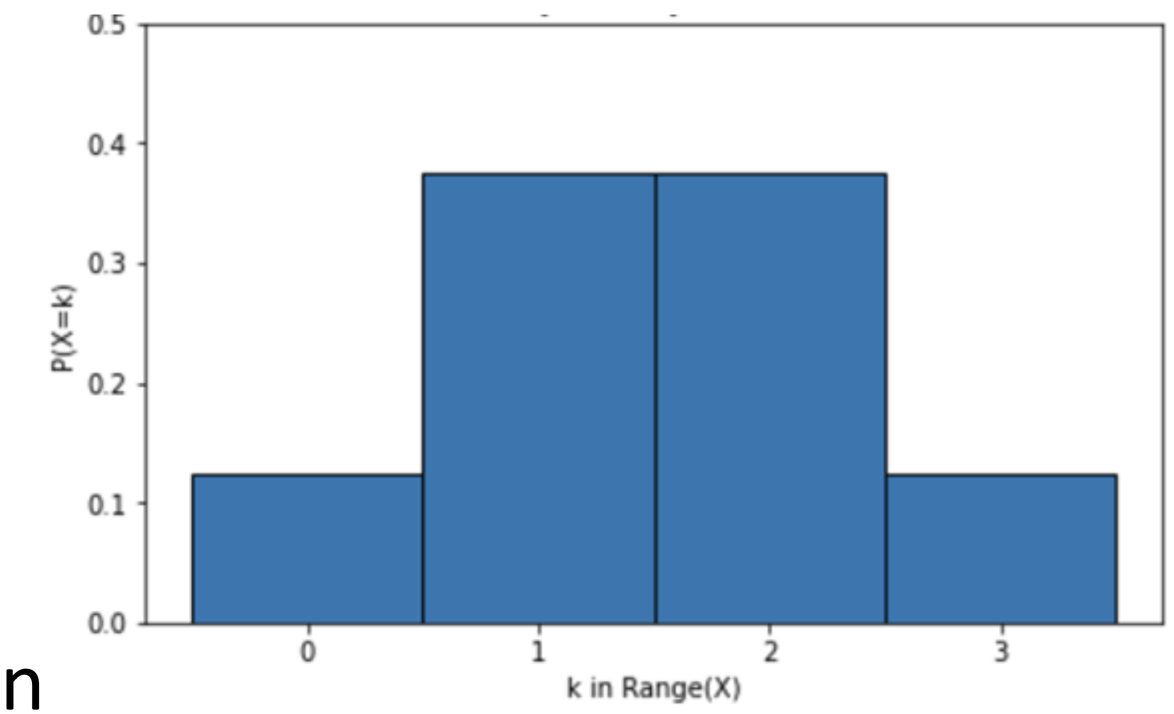


PROBABILITY MASS FUNCTION

- For discrete random variables, the **Probability Mass Function** (PMF) is also called the probability distribution. Thus, when asked to find the probability distribution of a discrete random variable X, we can do this by finding its PMF.
- C = "Number of heads after three coin-flips"
- Range(C) = {0, 1, 2, 3}
- Probability Mass Function P_C

$Pr(C=0) = Pr(C=3) = 1/8$

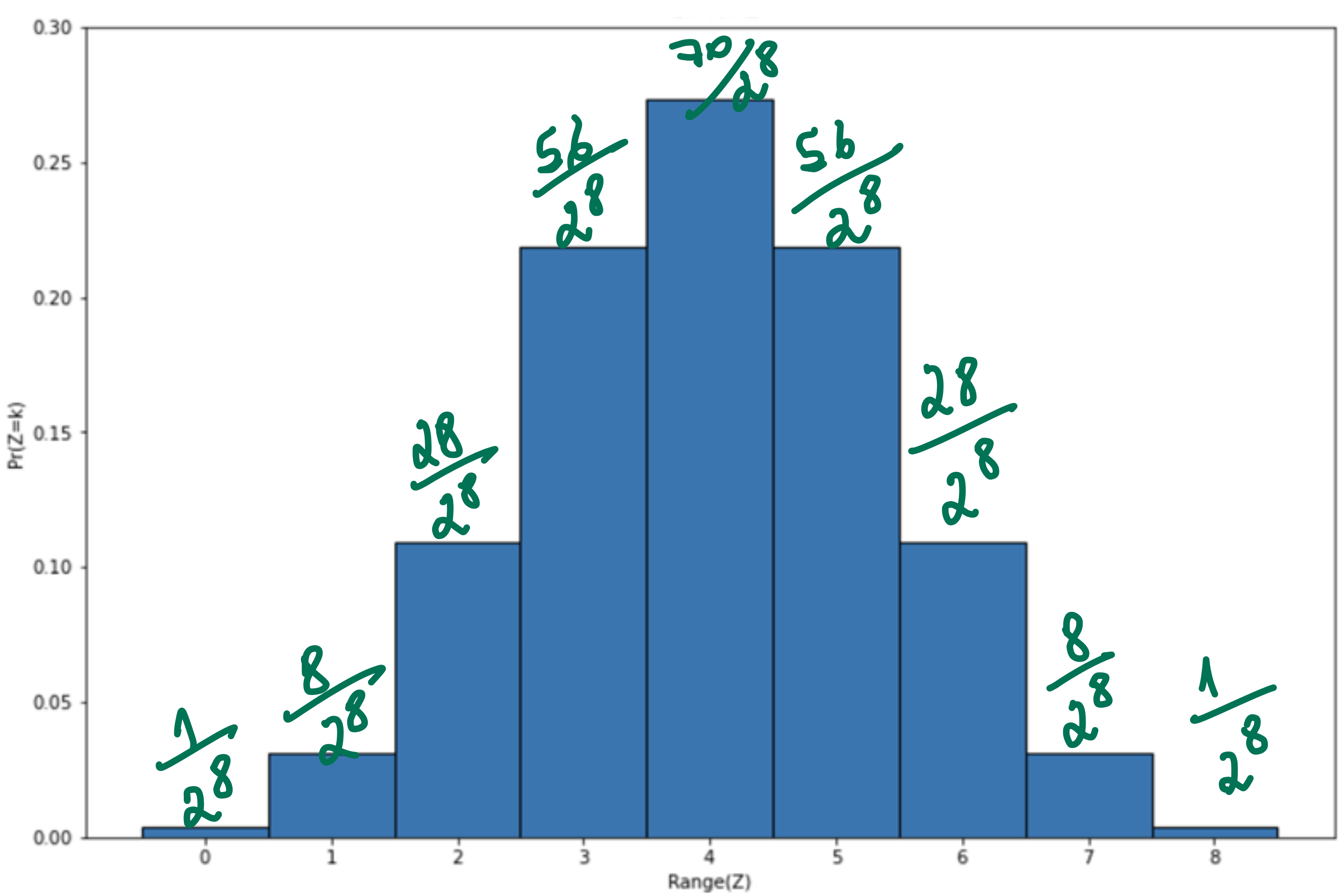
$Pr(C=1) = Pr(C=2) = 3/8$



- A PMF contains all the information you need to kn is often illustrated as a histogram.

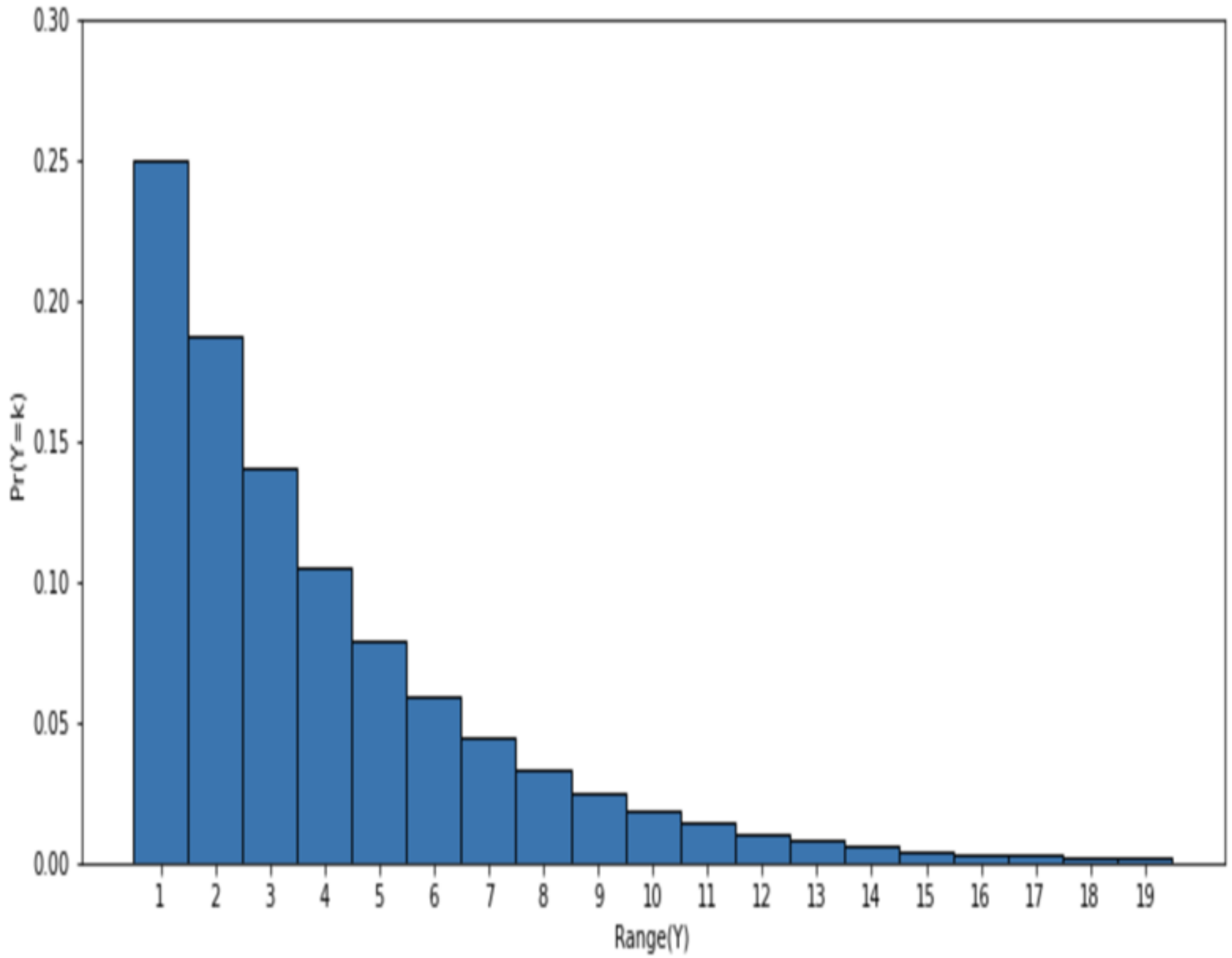
PROBABILITY MASS FUNCTION

- Z = "Toss 8 fair coins and count the number of heads"
- Range(Z) = {0, 1, 2, 3, 4, 5, 6, 7, 8}



PROBABILITY MASS FUNCTION

- Y = "Flip a 4-sided die and count the number of flips until it lands a 4."



$R_Y = \{1, 2, \ldots\}$

$Pr(Y=1) = 1/4$

$Pr(Y=2) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$

$Pr(Y=3) = \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{4}$