Probability in Computing

Lecture 6

CS 237

Last time

- Spinners and Continuous Probability
- Geometric Probability
- Anomalies with Continuous Probability

Today

- Five kinds of sample spaces
- Random Variables

Reminders

- HW 3 due on Thursday

Summary: 5 Types of Sample Spaces

Discrete Sample Spaces are finite or countably infinite

- We assign a probability to each outcome
- The probability of an event is the sum of the probabilities of its members: $A = \{a_1, \dots, a_n\} \rightarrow \Pr = \Pr(a_1) + \cdots + \Pr(a_n)$
- Can be uniform or non-uniform

1. Finite Uniform Sample Spaces: Example

$$\Pr(A) = \frac{|A|}{|\Omega|}$$

- Experiment: Roll a die, count the number of dots showing
- Sample space: $\Omega = \{\ 1,\quad 2,\quad 3,\quad 4,\quad 5,\quad 6\ \}$
- Probability Function:

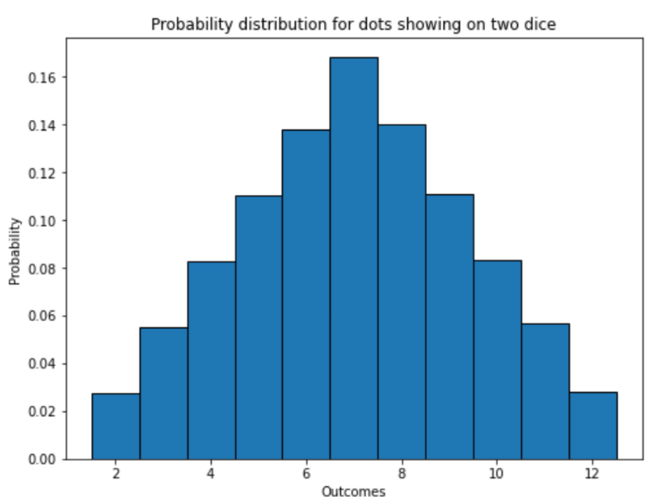
$$P(1) = \frac{1}{6},\quad P(2) = \frac{1}{6},\quad P(3) = \frac{1}{6}$$
$$P(4) = \frac{1}{6},\quad P(5) = \frac{1}{6},\quad P(6) = \frac{1}{6}$$

Summary: 5 Types of Sample Spaces

2. Finite Non-Uniform Sample Spaces: Example

- Experiment: Throw two dice, count the number of dots showing
- Sample space: $\Omega = \{ 2, 3, 4, ..., 12\}$
- Probability Function:

ω :	2	3	4	5	6	7	8	9	10	11	12
Pr(ω) :	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36



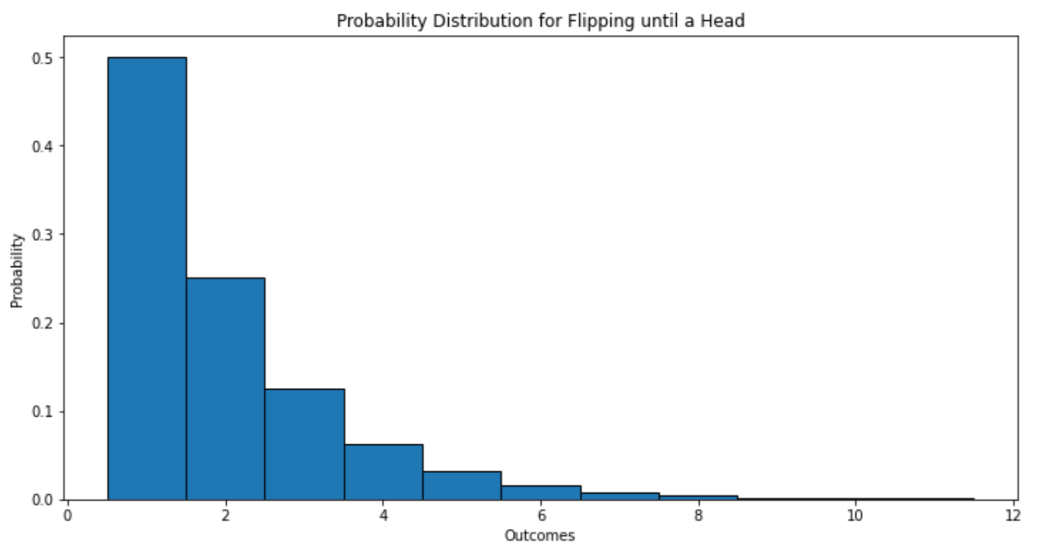
	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Summary: 5 Types of Sample Spaces

3. Countably Infinite Non-Uniform Sample Spaces: Example 1

- Experiment: Flip a coin until a head, count how many flips.
- Sample space: $\Omega = \{ 1, \quad 2, \quad 3, \quad 4, \quad 5, \quad .... \quad \}$
- Probability Function

ω :	1	2	3	4	...	k	...
Pr(ω) :	1/2	1/4	1/8	1/16	...	$1/2^k$	...

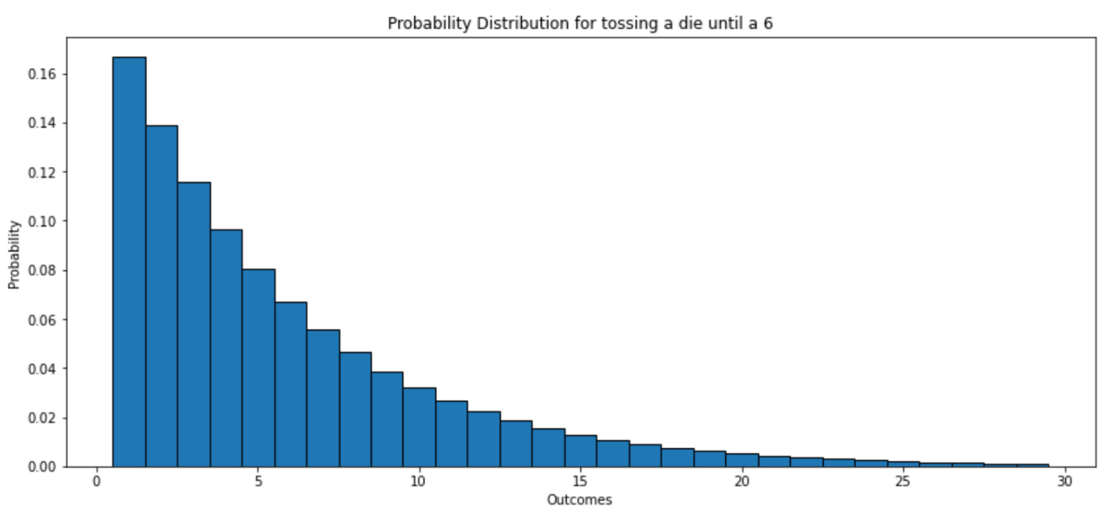


Summary: 5 Types of Sample Spaces

3. Countably Infinite Non-Uniform Sample Spaces: Example 2

- Experiment: Roll a die until a 6 appears, count how many rolls.
- Sample space: $\Omega = \{\ 1,\quad 2,\quad 3,\quad 4,\quad \ldots .\ \}$
- Probability Function:

ω :	1	2	3	.....	k	.....
Pr(ω) :	1/6	5/36	25/216	.....	$5^{k-1}/6^k$	.....



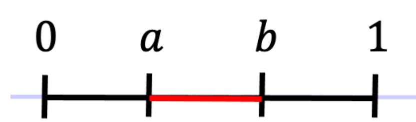
Summary: 5 Types of Sample Spaces

Continuous sample spaces are not countable

- Individual outcomes have probability zero
- We assign probabilities to the extent (length, area, volume, etc.) taken up by the event of interest

4. Uncountably Infinite Uniform Sample Spaces: Example

- Experiment: Uniformly pick an interval from 0 to 1.
- Sample space: $\Omega = [\ 0\ ..\ 1\ ]$
- Event: $\omega \in [a..b]$
- Probability Function: $\Pr(\ [a..b]\ ) = b - a$

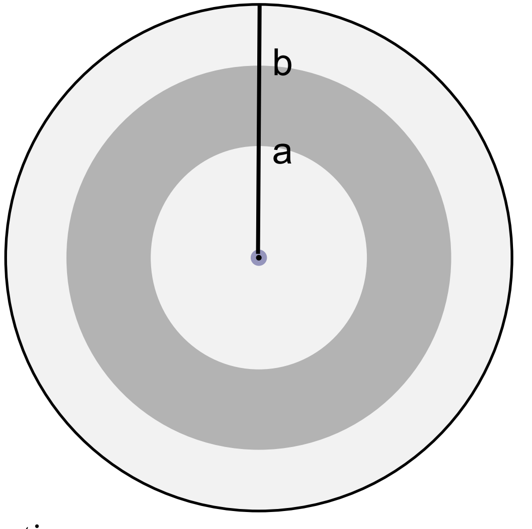


Summary: 5 Types of Sample Spaces

5. Uncountably Infinite Non-Uniform Sample Spaces: Example

- Experiment: Throw a dart at a circular target of radius 1.0, what is the probability of hitting the annulus below?
- Sample space: $\Omega = [\ 0\ ..\ 1\ ]$
- Event: $\omega \in [a..b]$
- Probability Function:

$$\Pr([a\,..\,b]) = \frac{\pi\, b^2 - \pi\, a^2}{\pi\, 1^2} = b^2 - a^2$$



CS 237

Top Hat question (Join Code: 413437)

Experiment: Pick a BU student uniformly at random and measure their height in meters, rounded to 2 decimal places.

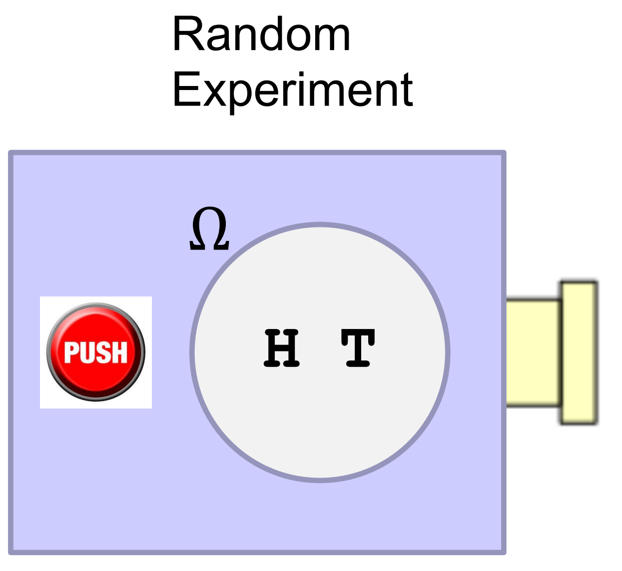
The Sample Space is:

A. Finite and uniform
B. Finite and not uniform
C. Countably infinite
D. Uncountably infinite and uniform
E. Uncountably infinite and not uniform

Random Experiments and Random Variables

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

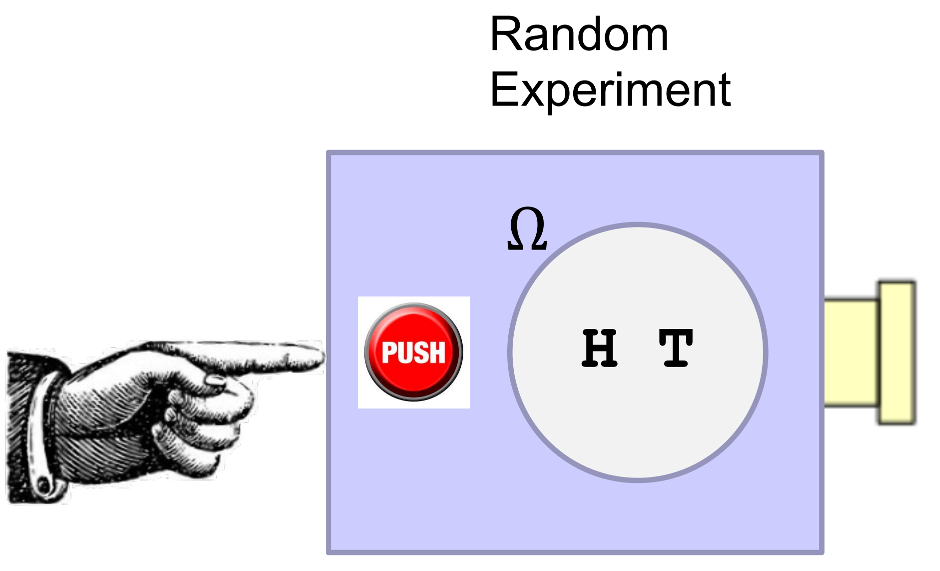
Example: Flip a coin!



Random Experiments and Random Variables

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

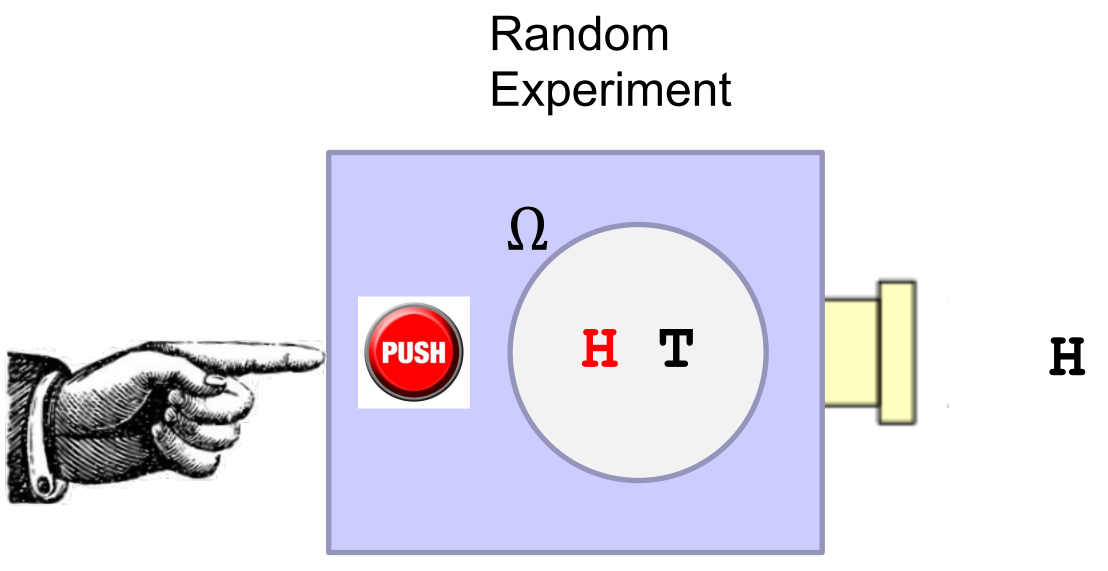
Example: Flip a coin!



Random Experiments and Random Variables

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

Example: Flip a coin!



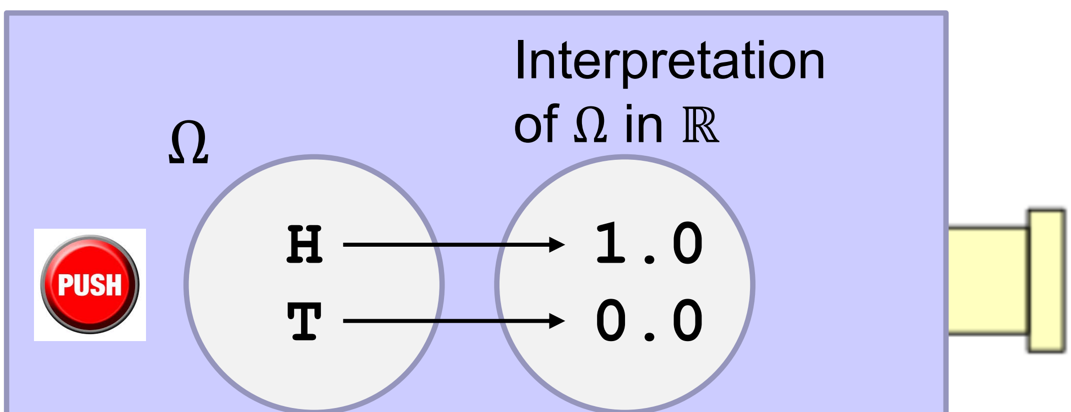
Random Variables

Definition. A Random Variable X is a function from a sample space Ω into the reals:

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

It *interprets* the random experiment as a real number. Now when an outcome is requested, the sample point is translated into a real number:

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



CS 237

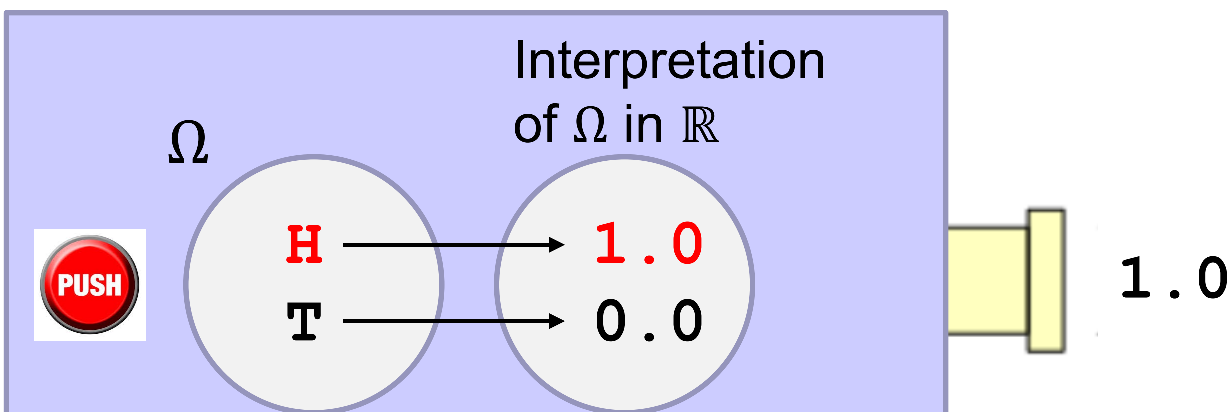
Random Variables

Definition. A Random Variable X is a function from a sample space Ω into the reals:

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

It *interprets* the random experiment as a real number. Now when an outcome is requested, the sample point is translated into a real number:

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

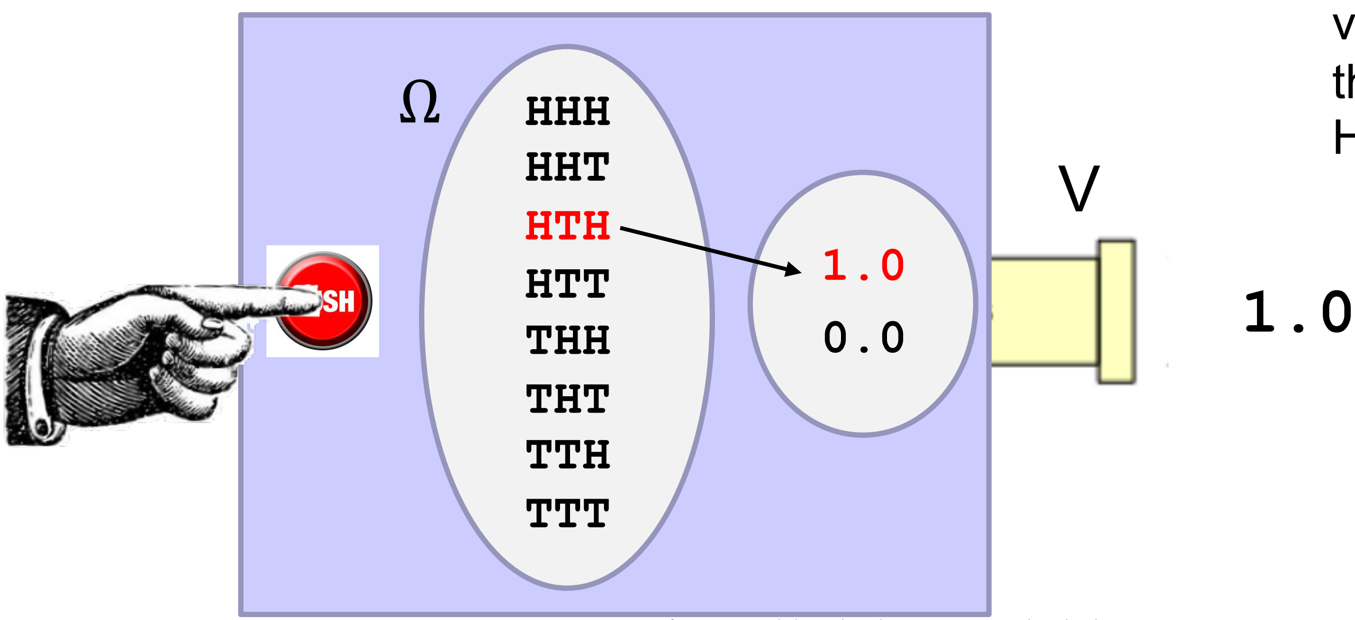


Random Variables

A random variable will not in general simply rename outcomes with real numbers, it will compute some useful information about the outcome.

V = "Flip 3 coins: return 1 if there are more heads showing than tails, 0 if more tails than heads "

V simulates a vote among three coins for H or T!



Range of random variables

- Since a random variable is a function, we can talk about its range.
- The range of a random variable X, shown by Range(X) or R_X, is the set of possible values for X.
- I toss a coin five times. Suppose that I'm interested in the number of heads. Let X be a random variable whose value is the number of heads.

$$\text{Range}(X) = R_X = \{0,1,2,3,4,5\}$$

Top Hat question (Join Code: 033357)

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. $R_Y = \{0,1,2,3,4,5 \dots\}$

C. $R_Y = \{\sum \frac{1}{2^n} : n \in N\}$

D. $R_Y = \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots\}$

E. $R_Y = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \dots\}$

Joint Random Variables

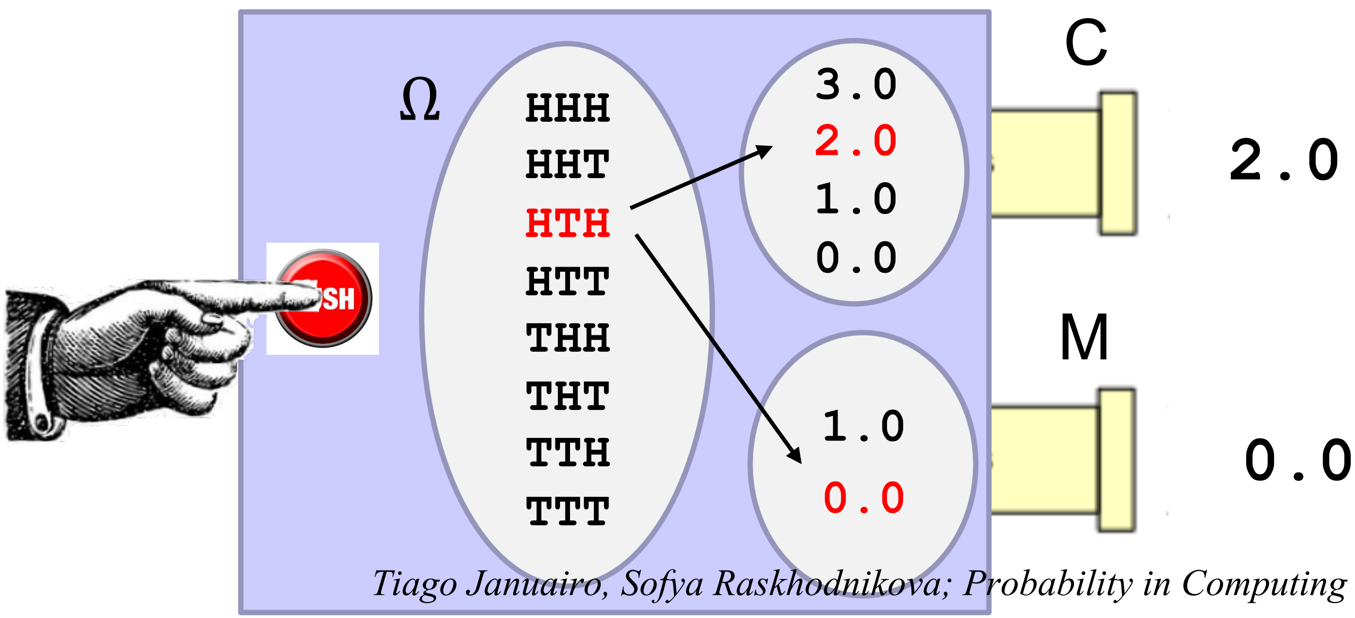
Joint Random Variables are combinations of 2 or more RVs, giving *multiple interpretations* of the same experiment.

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 (Join Code: 033357)

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

Random Variables as Functions

The term "Random Variable" is a misnomer: It's really a function!

Two things to remember about RVs as functions:

1. A random variable X is characterized by its range:

$$\text{Range}(X) = \{ r \mid \exists\, \omega \in \Omega.\ X(\omega) = r \}$$

X is finite/countably infinite/continuous/uniform/non-uniform if it behaves that way on its range.

Example: Choose a student from this class:

X = "Their height rounded to the nearest mm;" (Finite/non-uniform)
Y = "Their age;" (Continuous/non-uniform)
Z = "How many tattoos they have" (Finite/non-uniform)

Random Variables as Functions

Second thing to remember about random variables as functions:

2. Random variables behave just like Python functions which return random values.

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

```
In [1]:
from numpy.random import randint

def X():
    return randint(2)

X()

Out[1]: 0
```

```
In [2]:
X()

Out[2]: 1
```

Random Variables as Functions

"Flip 3 coins and
(C) return the number of heads showing;
(M) return 1 if all heads or all tails, 0 otherwise"

```
In [3]:
def CM():
    omega = (X(), X(), X())
    print("outcome = ", omega)
    C = sum(omega)
    M = (1 if C == 3 or C == 0 else 0)
    return (C,M)

CM()

outcome =  (1, 0, 0)

Out[3]: (1, 0)

In [22]:
CM()

outcome =  (1, 1, 1)

Out[22]: (3, 1)
```

Random Variables: Notation

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

Pr(< some Boolean expression involving X >)

Examples for discrete RVs:

$$\Pr(X = k) \;=\; P_X(k)$$

$$\Pr(X \neq k) \;=\; 1.0 - P_X(k)$$

$$\Pr(X \leq k) \;=\; \sum_{a \leq k} P_X(a)$$

$$\Pr(j < X \leq k) \;=\; \sum_{j < a \leq k} P_X(a)$$

A useful notation in the discrete case from the alternate textbook:

$P_X(k)$ = the probability that X returns k

Discrete Random Variables: Notation

Example: X = “Count the number of flips until the coin shows heads”

Range(X) = { 1, 2, 3, 4, ... }

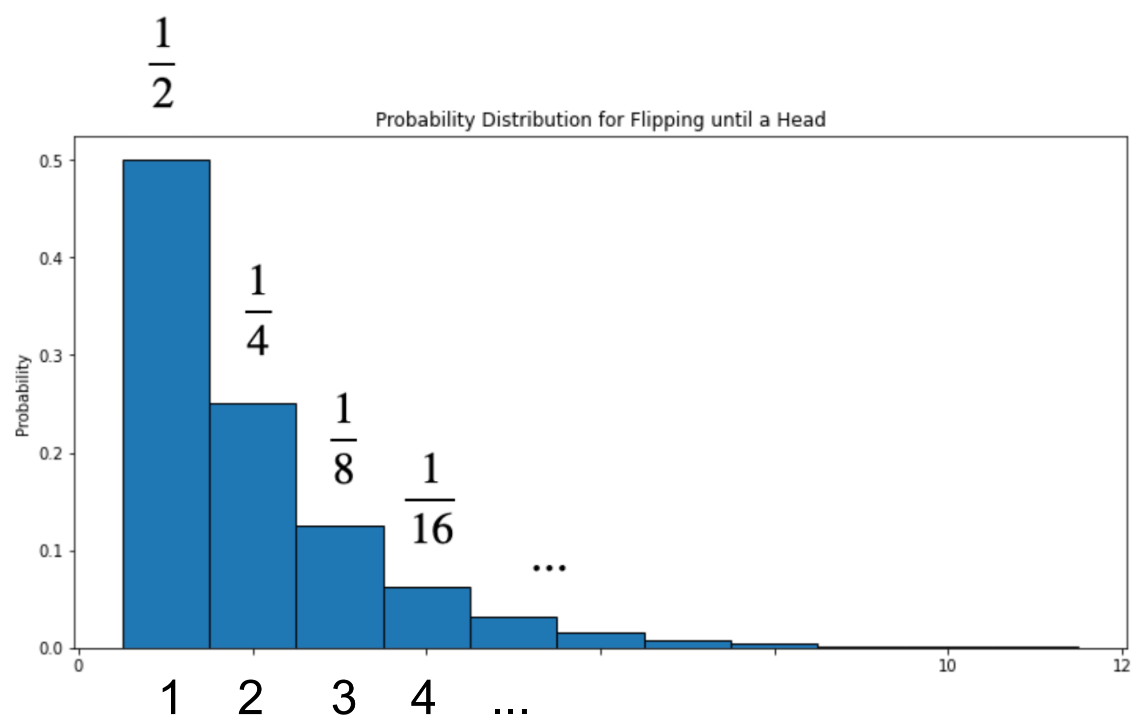
$$Pr(X = k) = \begin{cases} 2^{-k} & k > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Pr(X = 2) = \frac{1}{4}$$

$$Pr(X \leq 4) = \frac{15}{16}$$

$$Pr(2 \leq X < 4) = \frac{3}{8}$$

$$Pr(X \text{ is odd }) = \frac{2}{3}$$



Top Hat question (Join Code: 413437)

For the random variable

X = "Count the number of flips
until the coin shows heads"

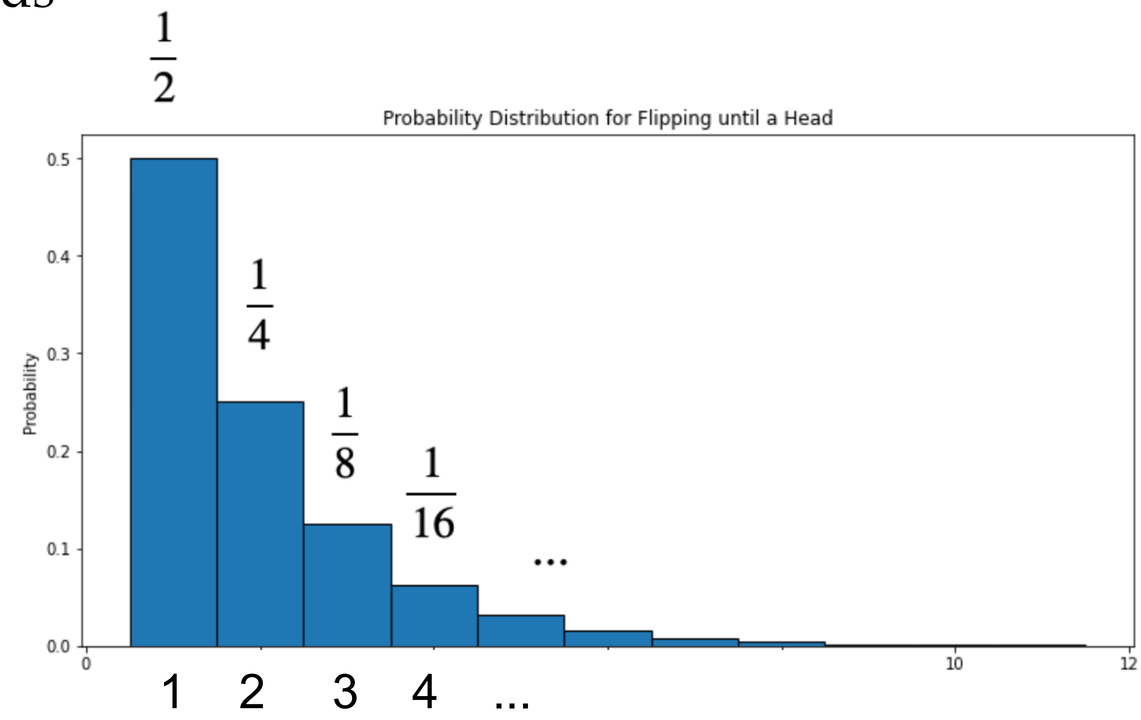
What is Pr(X > k) for k > 0 ?

A. $\frac{1}{k}$

B. $\frac{1}{2^k}$

C. $\frac{1}{2^{k+1}}$

D. $\frac{1}{2^{k-1}}$



Probability Mass Function (PMF)

- *Let* $A = \{X = x_k\}$ be the set of outcomes $s \in S$ for which the corresponding value of X is equal to x_k. In particular,
$$A = \{s \in S \mid X(s) = x_k\}$$
- The probabilities of events $\{X = x_k\}$ are formally shown by the **Probability Mass Function (PMF)** of X, denoted as P_X.
- PMF is a probability measure that gives us probabilities of the possible values for a random variable.

PMF: Example

- I toss a fair coin twice, and let X be defined as the number of heads I observe. Find R_X and P_X
- Sample space: $S=\{HH,HT,TH,TT\}$

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

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

CS 237

Reminders

- **Homework 3 due on Thursday, February 9th**
  - **Analytical problems**
  - **Programming assignment**

- **Ask your questions on Piazza!**

- **This lecture is covered on the second text book, chapters:**
  - P 3.1.1-3.1.3,3.1.6,
  - P 3.2.1, 4.0-4.1

See you on Thursday, Terriers!!!