# CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

---

## Lecture 1: Basic Definitions and Concepts of Probability Theory

- What is Randomness?
- Random Experiments
- Sample Points, Sample Spaces, Events
- Probability Functions and Probability Spaces
- Classification of Probability Spaces:
  - Discrete (Finite or Countably Infinite)
  - Continuous (Uncountably Infinite)
- Axioms for Probability

# Definition of Randomness

## Wikipedia Definition:

"**Randomness** is the lack of pattern or predictability in events. A random sequence of events, symbols or steps has no order and does not follow an intelligible pattern or combination."

## Information Theory Definition:

"The maximum entropy" OR "the minimum of information"

Example: Flip a coin – before you look at it: Is it heads or tails?

**?**

# Definition of Randomness

## Wikipedia Definition:

"**Randomness** is the lack of **pattern** or **predictability** in events. A random sequence of events, symbols or steps has no order and does not follow an intelligible pattern or combination."

## Information Theory Definition:

"The maximum entropy" OR "the minimum of information"

Example: Flip a coin – before you look at it: Is it heads or tails?


**Heads!**

# Randomness: Life is Uncertain!

We are surrounded by random events! Your chance of

- being audited by IRS: 1 in 175
- finding a pearl in an oyster: 1 in 12,000
- winning $1,000,000 in Powerball: 1 in 11.6 million
- being killed by
  - a champagne cork:
  - a shark
  - a vending machine
  - a cow
  - hot tap water

Which are most and least likely?

# What is Randomness?

We are surrounded by random events! Your chance of

| | |
|---|---|
| ✦ being audited by IRS: | 1 in 175 |
| ✦ finding a pearl in an oyster: | 1 in 12,000 |
| ✦ winning $1,000,000 in Powerball: | 1 in 11.6 million |
| ✦ **being killed by** | |
|   ✦ **hot tap water:** | **1 in 3 million** |
|   ✦ **a champagne cork:** | **1 in 13 million** |
|   ✦ **a cow:** | **1 in 16 million** |
|   ✦ **a vending machine:** | **1 in 146 million** |
|   ✦ **a shark:** | **1 in 319 million** |

# Randomness: "It's Complicated!"

**Our experience is filled with random events, but humans are very bad at analyzing situations involving truly random behavior. (Not surprising, since randomness = lack of information!) We also have a hard time exhibiting random behavior.**

Suppose we attempt to simulate the flipping of a coin by writing down a sequence of 20 symbols from the set { H, T }.

Clearly, these are NOT good examples of random behavior because the patterns are obvious:

H H H H H H H H H H H H H H H H H H

H T H T H T H T H T H T H T H T H T

H T T H H H T T T T H H H T T H

But when is the "lack of patterns" enough?

# Randomness and Non-Randomness

**But of course life is not completely uncertain**, and in the last 350 years or so we have developed mathematical tools for understanding the difference:

> **"Probability Theory is the mathematical study of random phenomena."**
>
> (Encylopedia Britannica)

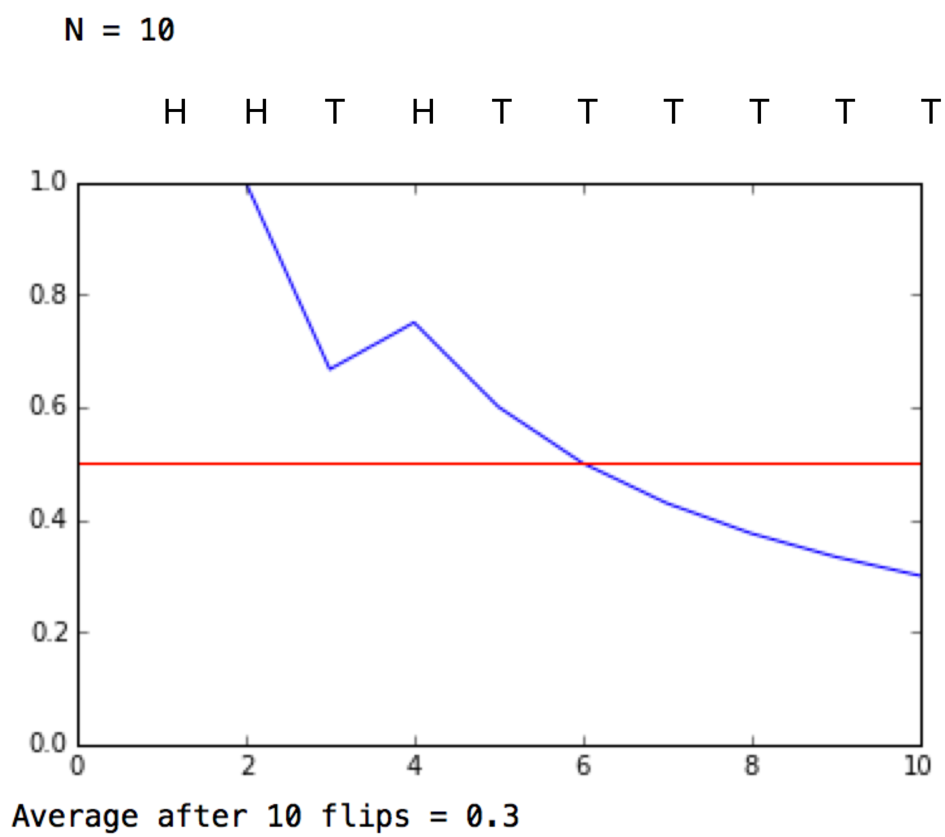
> **"Statistics is the science of learning from data, and of measuring, controlling, and communicating uncertainty…."**
>
> (American Statistical Association)

Much of this work has to do with **discovering structure** within a group or sequence of random events; many random phenomena behave in ways that are unpredictable in the short term, but have non-random characteristics when viewed as a whole – we seek to understand the difference!

# Randomness and Non-Randomness

## Examples of patterns within random events:

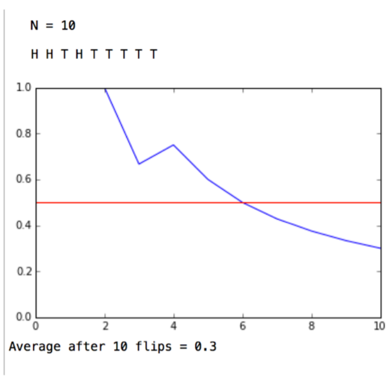
✧ **Example 1:** Flip a coin over and over; what is the average number of heads?

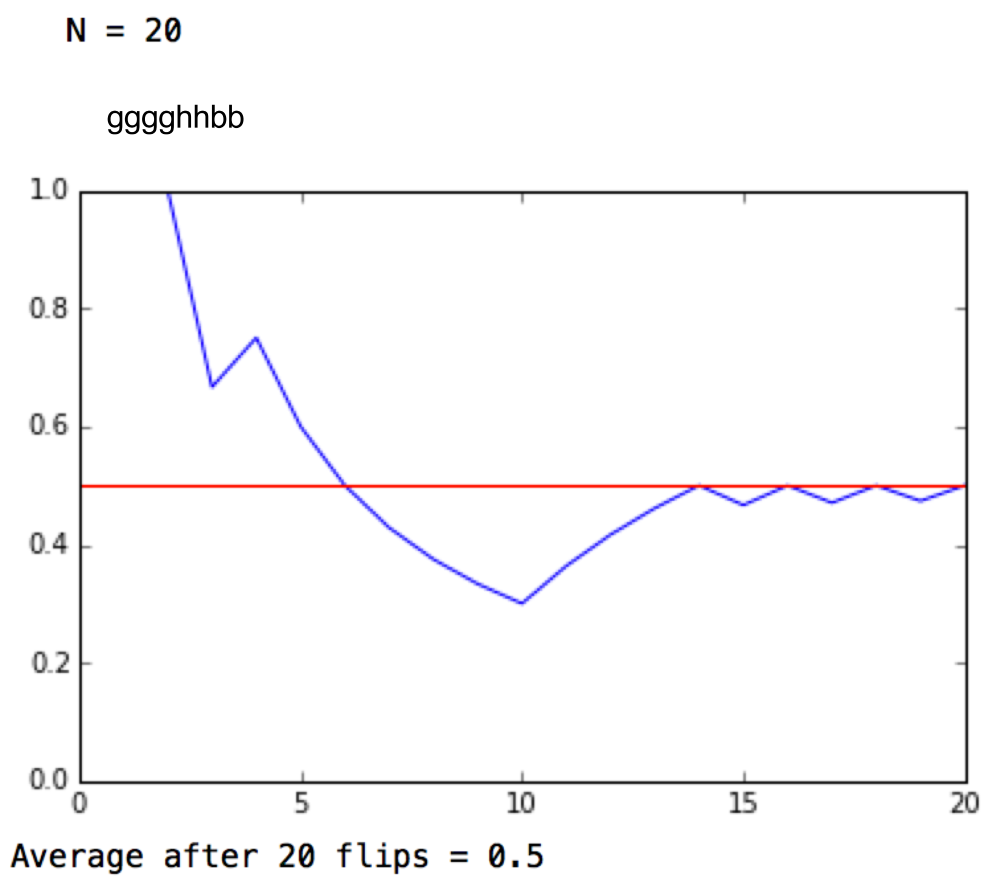


# Randomness and Non-Randomness

## Examples of patterns within random events:

✧ **Example 1:** Flip a coin over and over; what is the average number of heads?

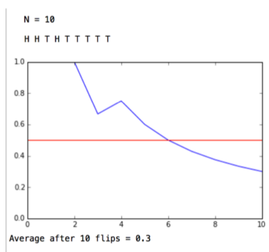


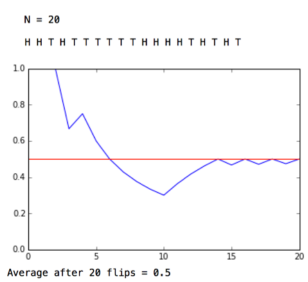



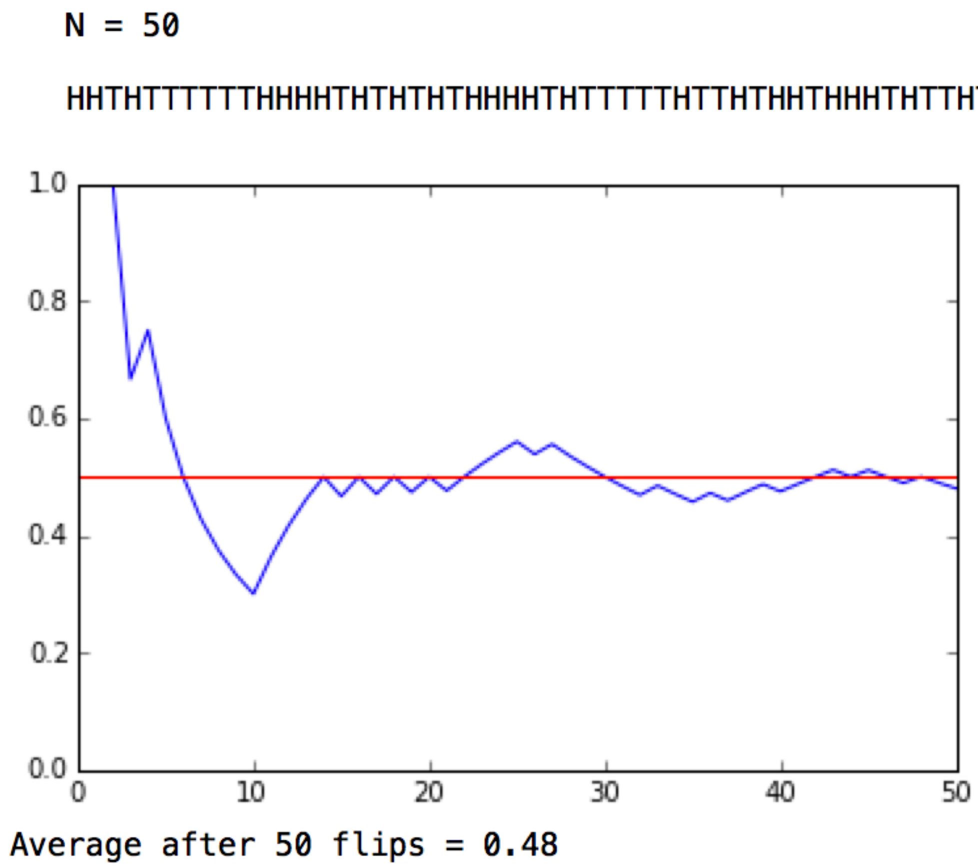
# Randomness and Non-Randomness

## Examples of patterns within random events:

✧ **Example 1:** Flip a coin over and over; what is the average number of heads?



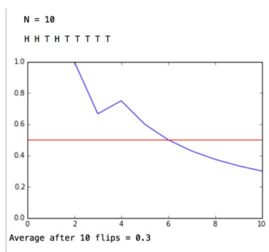



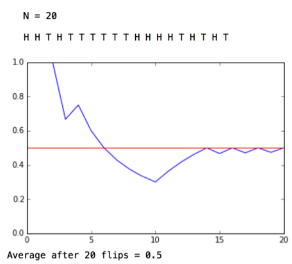



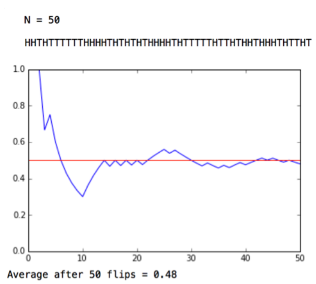
# Randomness and Non-Randomness

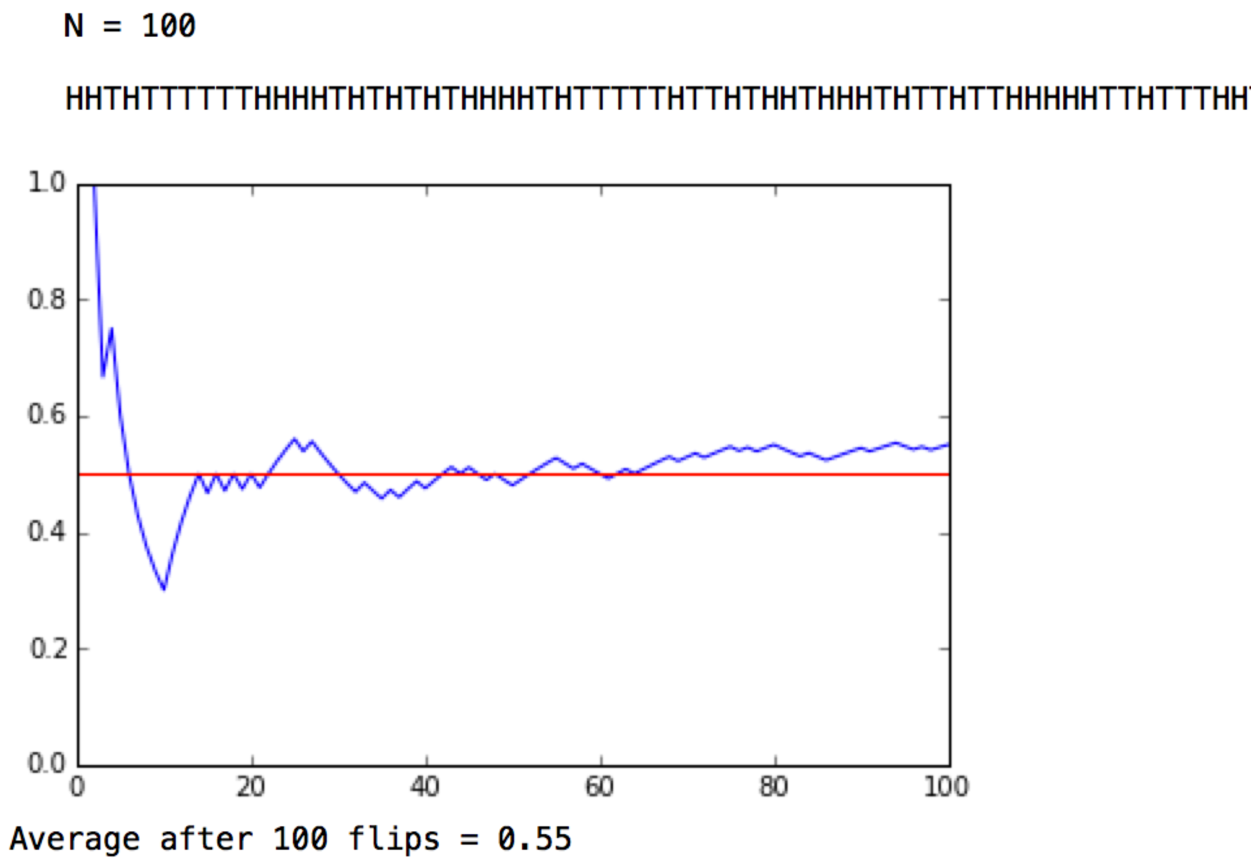
## Examples of patterns within random events:

✧ **Example 1:** Flip a coin over and over; what is the average number of heads?



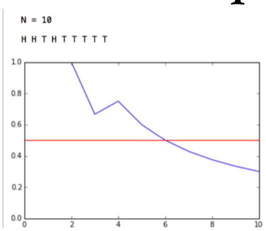



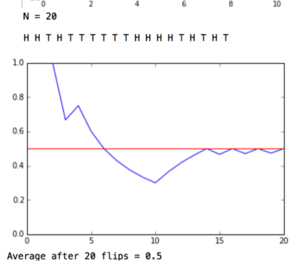



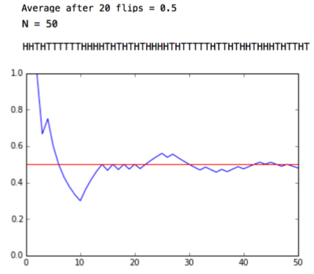
# Randomness and Non-Randomness

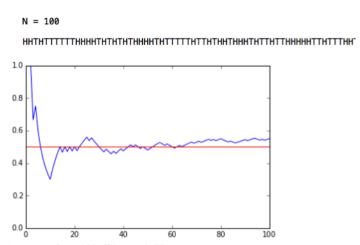
## Examples of patterns within random events:

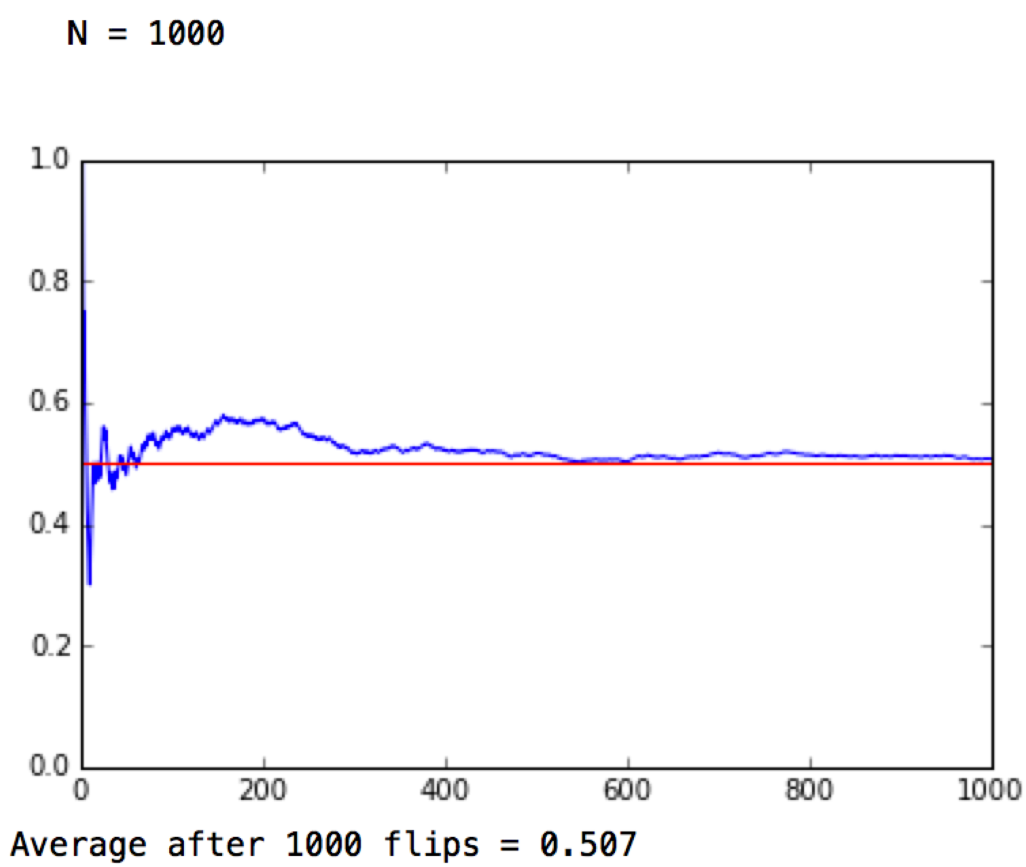
✧ **Example 1:** Flip a coin over and over; what is the average number of heads?



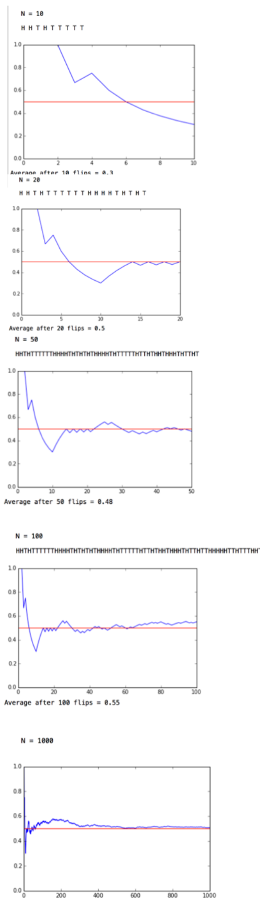



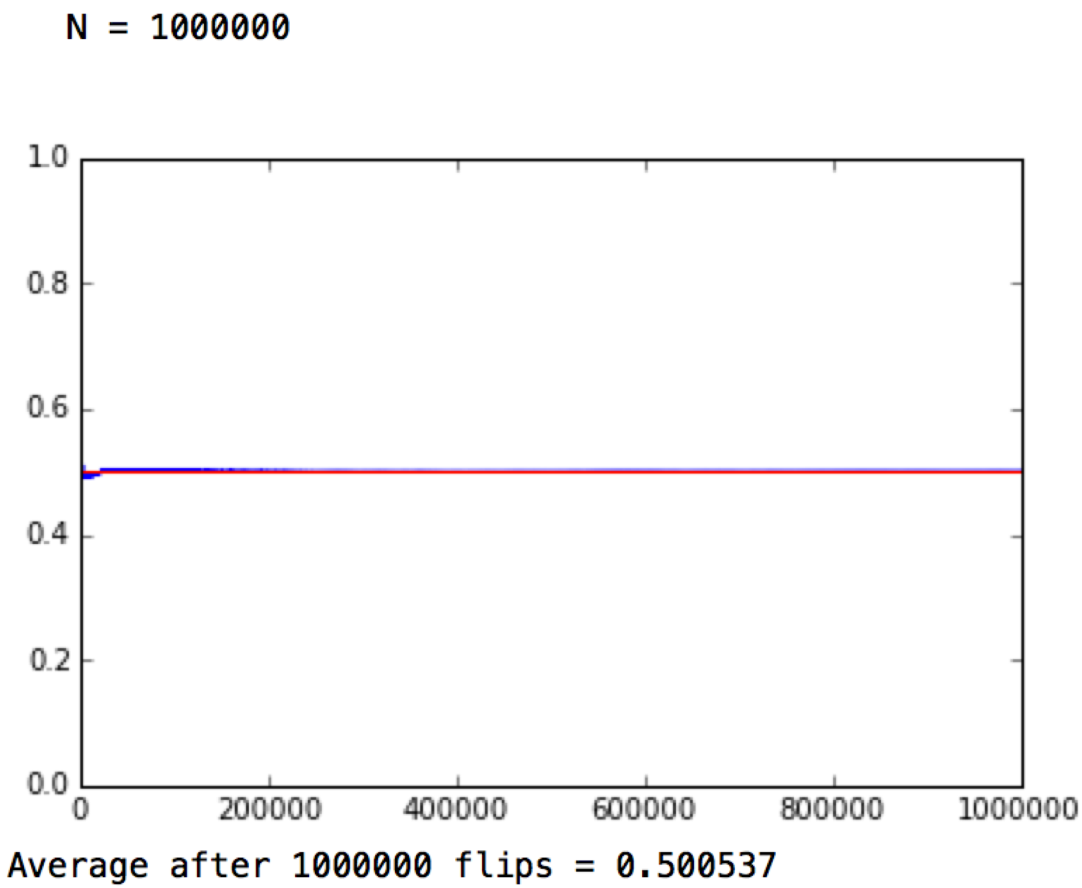



# Randomness and Non-Randomness

## Examples of patterns within random events:

- **Example 1:** Flip a coin over and over; what is the average number of heads?

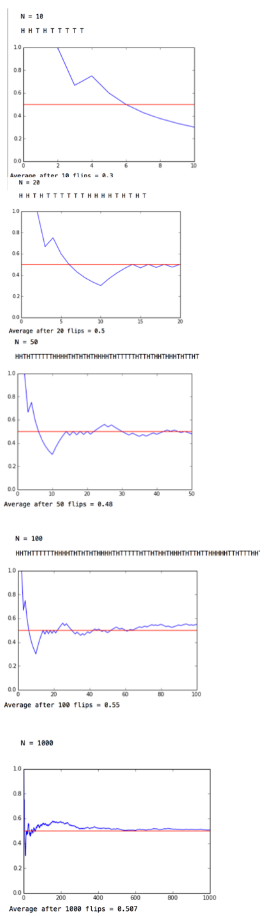


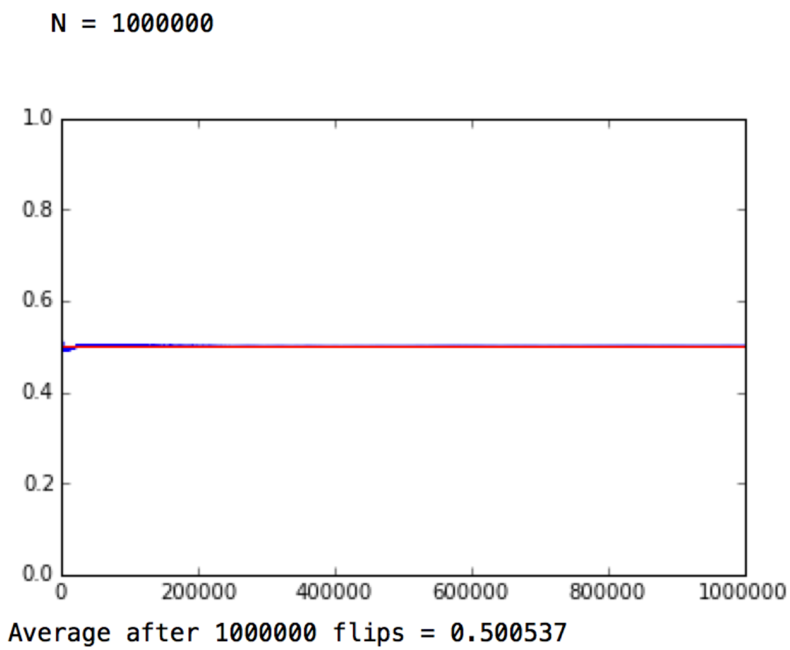


# Randomness and Non-Randomness

## Examples of patterns within random events:

✧ **Example 1:** Flip a coin over and over; what is the average number of heads?




The average number of heads ALWAYS approaches 0.5 as N gets larger!

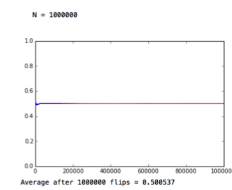
One coin flip is random, but many coin flips are non-random….

Patterns emerge when we repeat random experiments over and over….

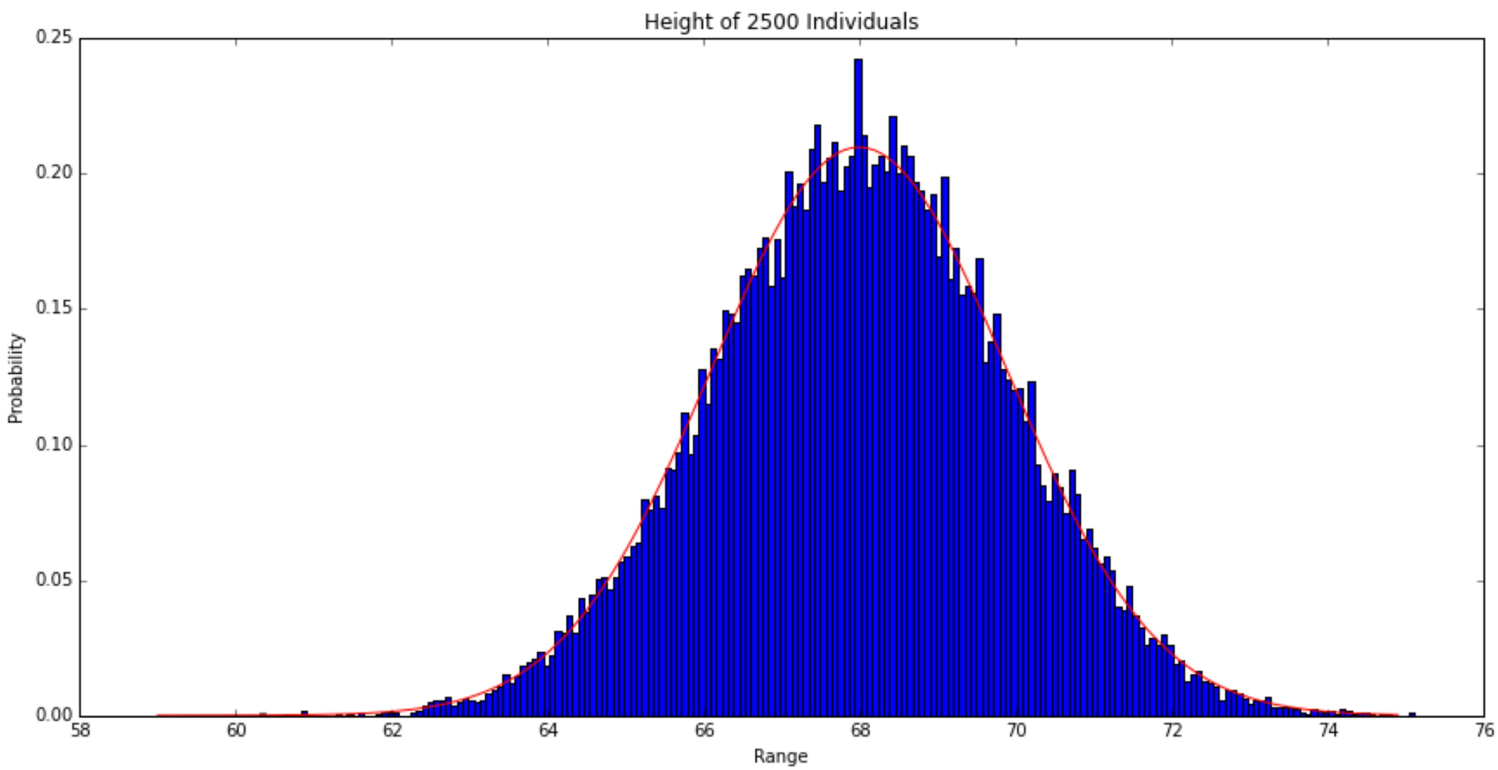
# Randomness and Non-Randomness

## Examples of patterns emerging when we "zoom out" and look at large numbers of seemingly random events:

✧ **Example 1:** Flip a coin over and over; what is the average number of heads?

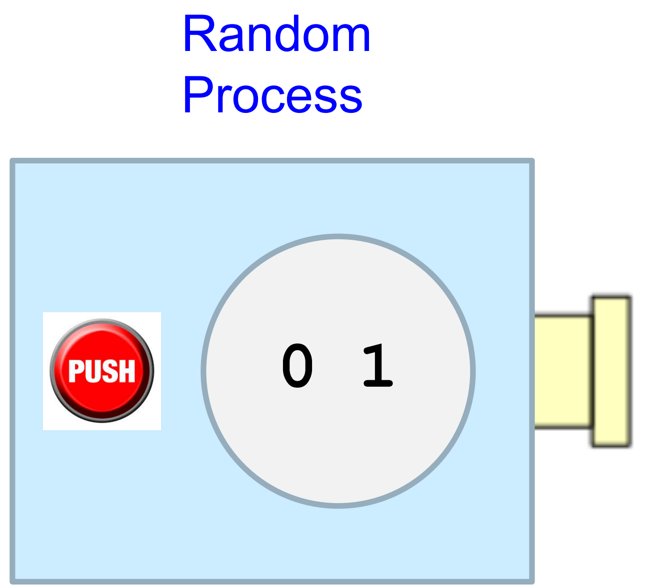



✧ **Example 2:** What is the height of a human being?



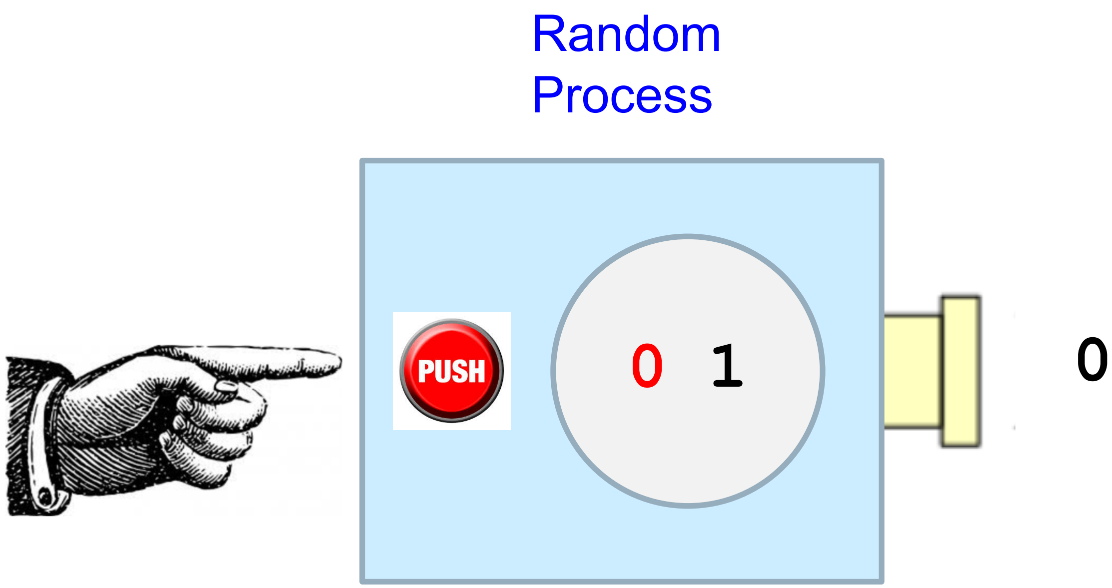
# Random Experiments

A Random Experiment is a process that produces random (as in "unknown") outcomes from a well-defined set of possible outcomes.



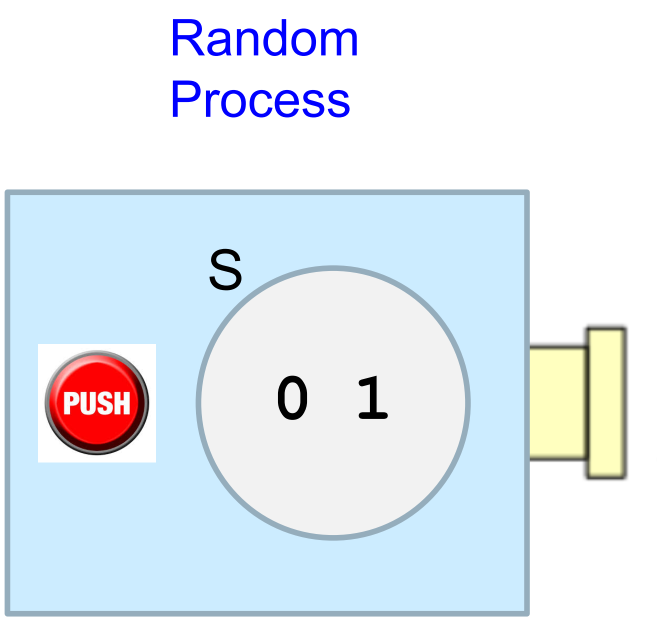
# Random Experiments

A Random Experiment is a process that produces uncertain outcomes from a well-defined set of possible outcomes. Usually there is some kind of physical experiment which is being modeled theoretically.



# Random Experiments, Sample Spaces and Sample Points

The Sample Space (denoted by S or sometimes the Greek letters $\Sigma$ *or* $\Omega$) of the experiment **is the set of possible outcomes,** and any individual outcomes is called a Sample Point. Sample points can be any discrete objects.

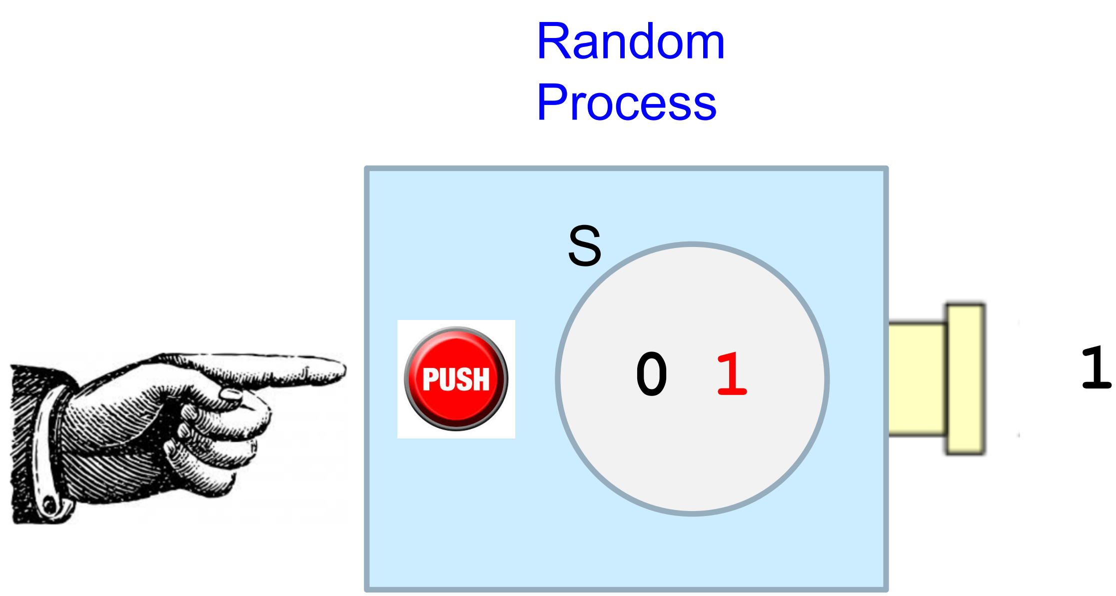



S = { 0, 1 }

# Random Experiments, Sample Spaces and Sample Points

A Random Experiment is a process that produces random (as in "unknown") outcomes from a well-defined set of possible outcomes.

The Sample Space (denoted by S or sometimes the Greek letters $\Sigma$ *or* $\Omega$) of the experiment is the set of possible outcomes, and any individual outcomes is called a Sample Point. Sample points can be any discrete objects.

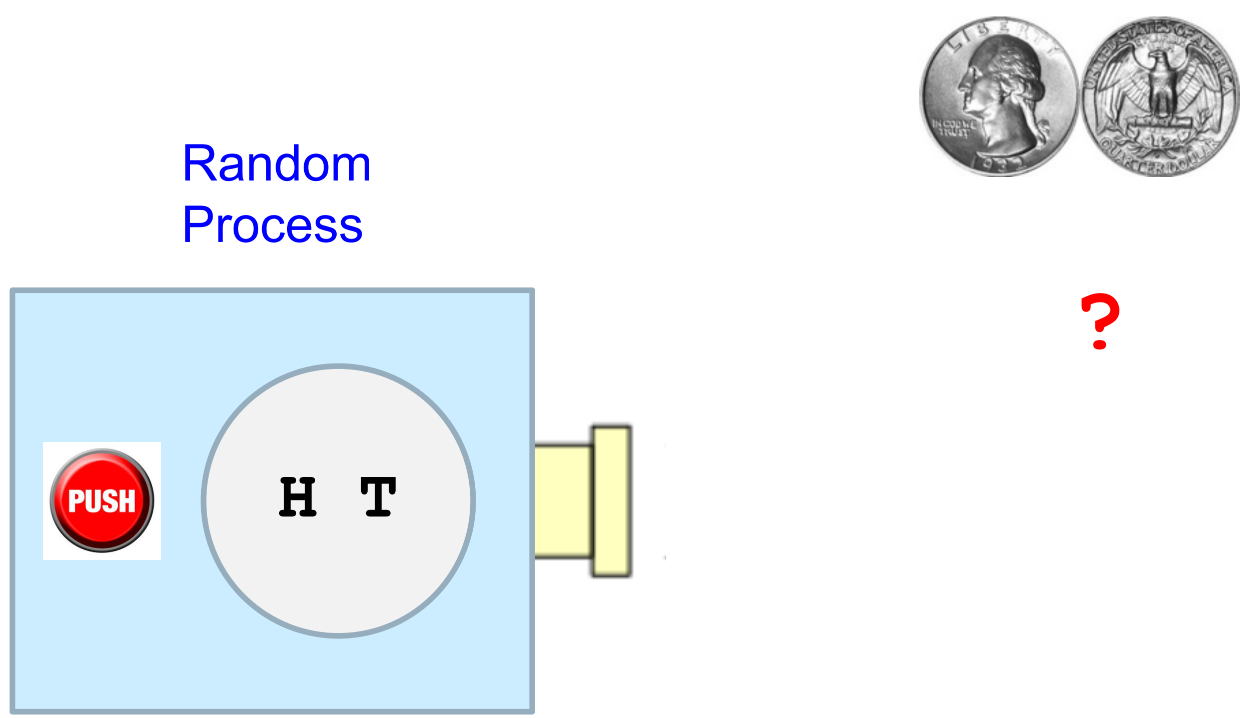



S = { 0, 1 }

# Random Experiments, Sample Spaces and Sample Points

Often such an experiment represents some actual physical experiment.

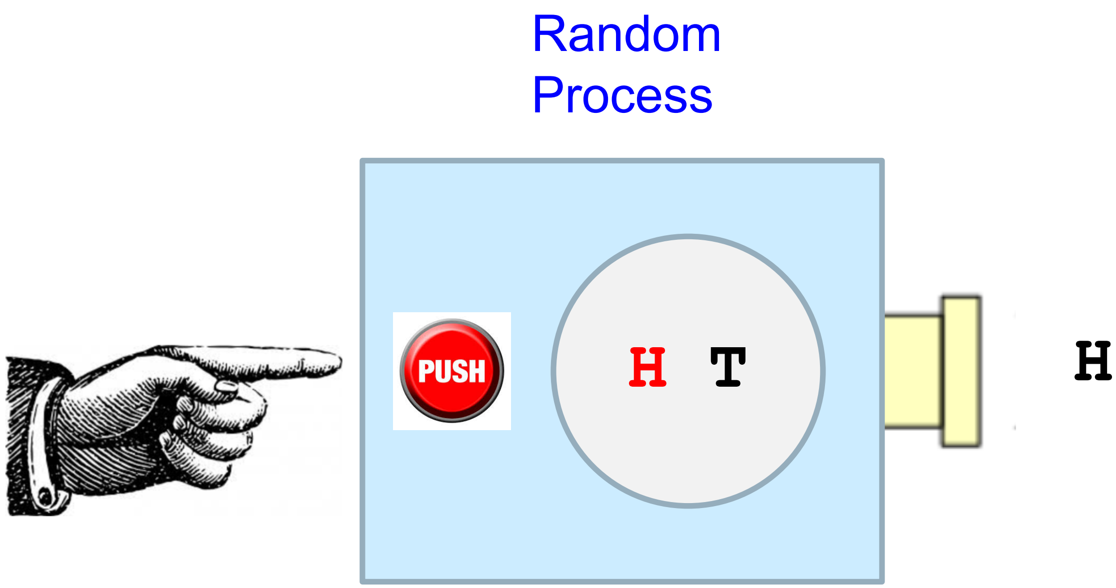
**Example:** Flip a coin, is it heads or tails?



# Random Experiments, Sample Spaces and Sample Points

Often such an experiment represents some actual physical experiment.

**Example:** Flip a coin, is it heads or tails?

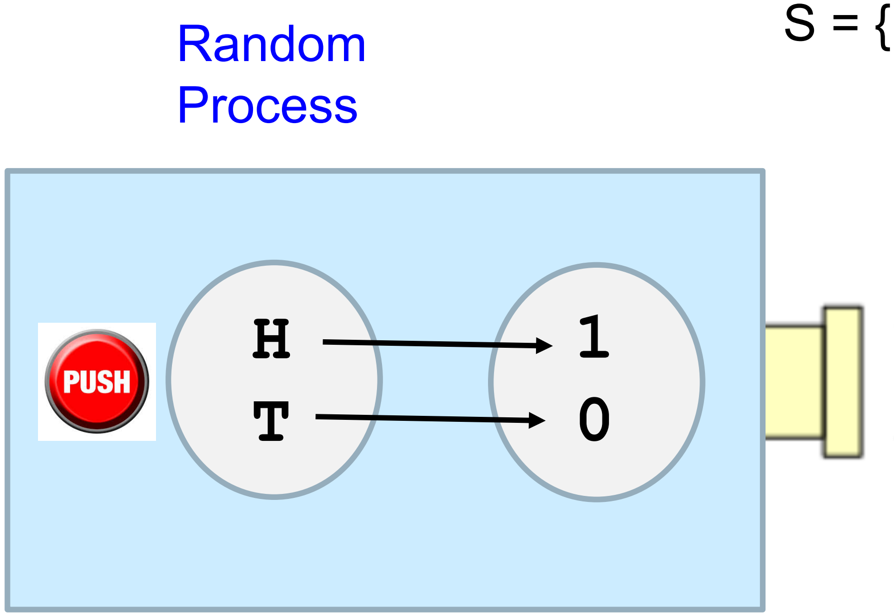


# Random Experiments, Sample Spaces and Sample Points

The description of the random experiment may indicate that there is more than one stage to creating the outcome, for example, a physical analogy may produce results that have to be interpreted as "outcomes."

**Example:** Flip a coin, count how many heads show.

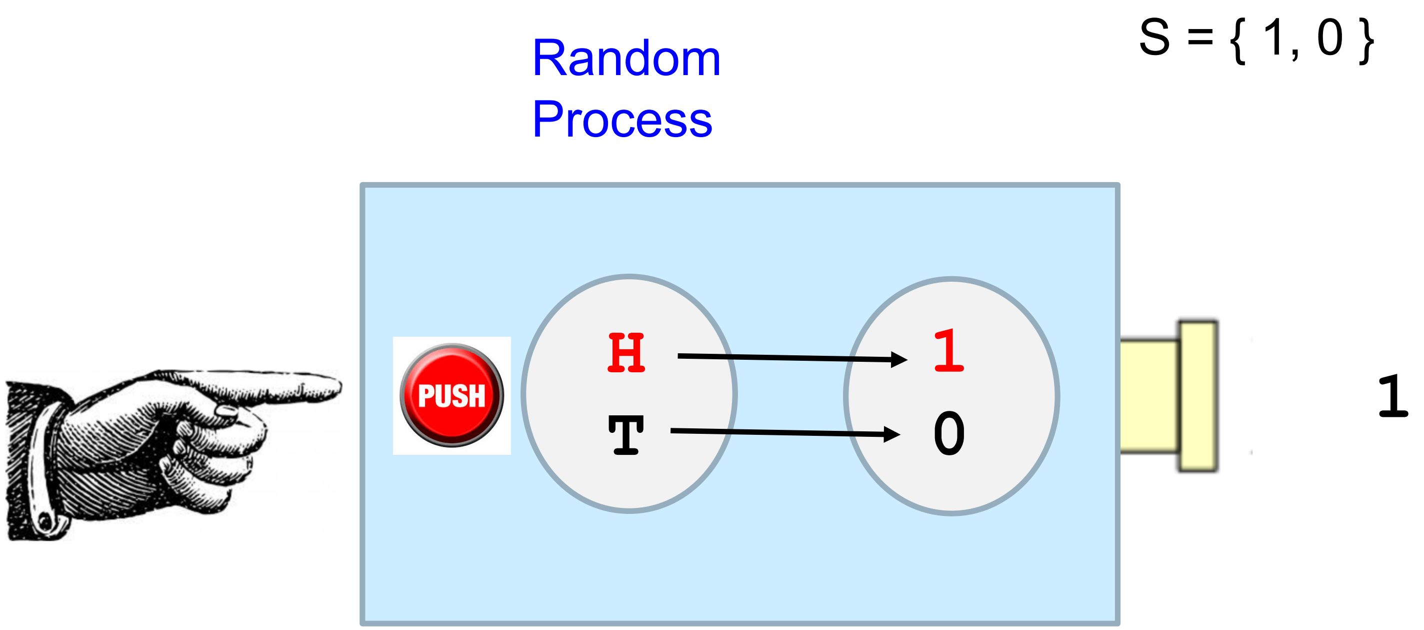
$S = \{ 1, 0 \}$



# Random Experiments, Sample Spaces and Sample Points

The description of the random experiment may indicate that there is more than one stage to creating the outcome, for example, a physical analogy may produce results that have to be interpreted as "outcomes."

**Example:** Flip a coin, count how many heads show.

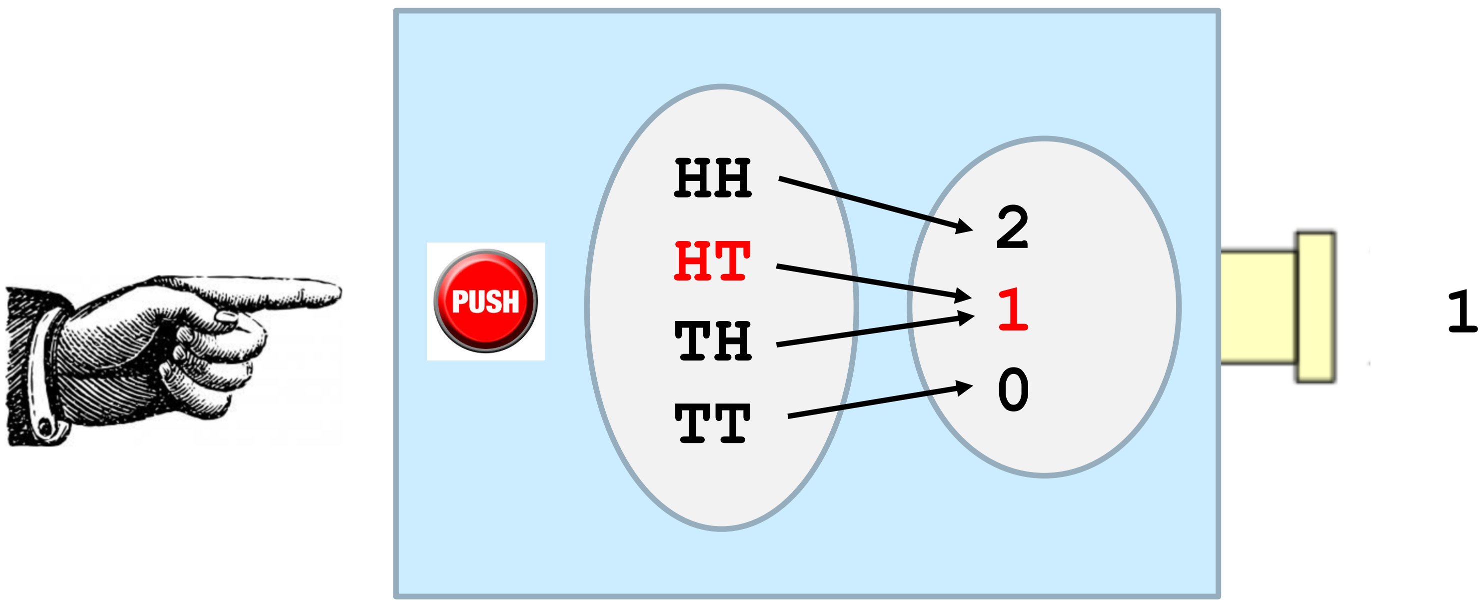


# Random Experiments, Sample Spaces and Sample Points

The description of the random experiment may indicate that there is more than one stage to creating the outcome, for example, a physical analogy may produce results that have to be interpreted as "outcomes."

**Example:** Flip two coins, count how many heads show.

S = { 0, 1, 2 }



# Random Experiments, Sample Spaces and Sample Points

The **Sample Points** can be just about anything (numbers, letters, words, people, etc.) and the Sample Space (= **any** set of sample points) can be

## Finite

**Example:** Flip three coins, and output head if there are two heads showing, and tails otherwise (as if the coins "vote" for the outcome!)

**S= { head, tail }**

## Countably Infinite

**Example:** Flip a coin until heads appears, and report the number of flips

S = { 1, 2, 3, 4, .... }

## Uncountably Infinite

**Example:** Spin a pointer on a circle labelled with real numbers [0..1) and report the number that the pointer stops on.

S = [0 .. 1)

# Random Experiments, Sample Spaces and Sample Points

The Sample Points can be just about anything (numbers, letters, words, people, etc.) and the Sample Space (= **any** set of sample points) can be

Discrete

**Finite**

**Example:** Flip three coins, and output head if there are two heads showing, and tails otherwise (as if the coins "vote" for the outcome!)

**S= { head, tail }**

**Countably Infinite**

**Example:** Flip a coin until heads appears, and report the number of flips

S = { 1, 2, 3, 4, .... }

Continuous

**Uncountably Infinite**

**Example:** Spin a pointer on a circle labelled with real numbers [0..1) and report the number that the pointer stops on.

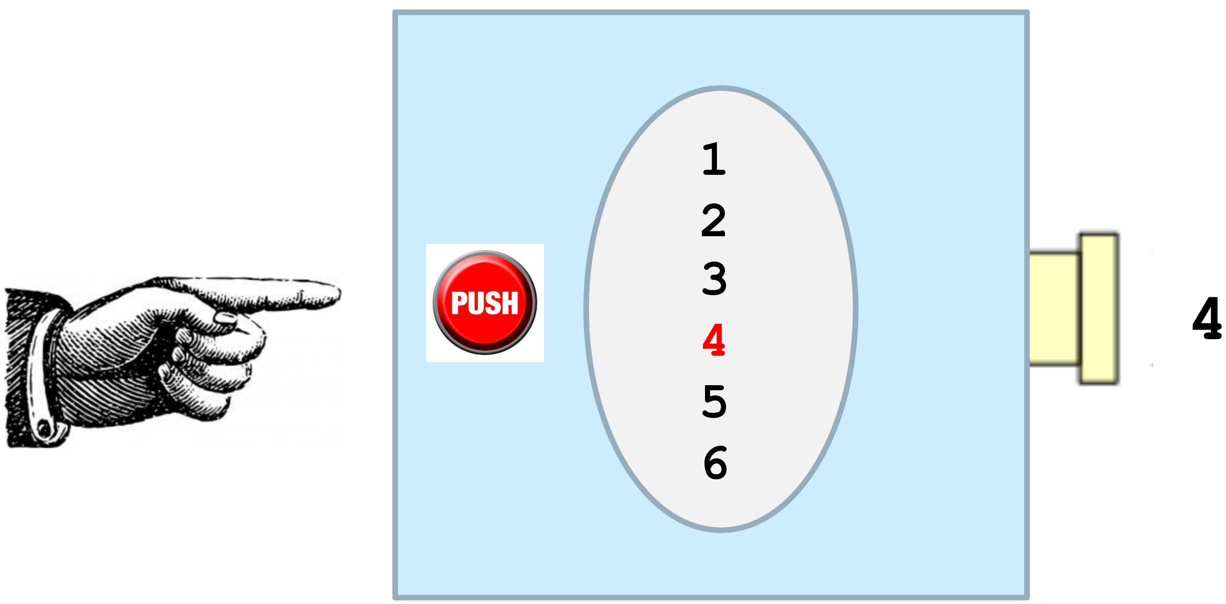
S = [0 .. 1)

# Sample Spaces, Sample Points, and Events

An Event is any subset of the Sample Space. An event A is said to have occurred if the outcome of the random experiment is a member of A. We will be mostly interested specifying a set by some characteristic, and then calculating the probability of that event occurring.

**Example:** Toss a die and output the number of dots showing.
Let A = "there are an even number of dots showing" and
B = "there are at least 5 dots showing."

S = { 1, 2, 3, 4, 5, 6 } A = { 2, 4, 6 } B = { 5, 6 }




The event A occurred, since

$$4 \in \{2,4,6\}$$

The event B did not occur:

$$4 \notin \{ 5, 6 \}$$

# Sample Spaces, Sample Points, and Events

We will be mostly interested in **questions involving the probability of particular events occurring**, so let us pay particular attention to the notion of an event.

**Example:** Toss a die and output the number of dots showing. Let A = "there are an even number of dots showing."

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

The set of possible events is the power-set of S, the set of all subsets,

$$\mathcal{P}(S) = \{ A \mid A \subseteq S \}$$

So for this example we have $2^6 = 64$ possible events, including

- The empty or "impossible event" $\varnothing$ .    ("What is the probability of rolling a 9?")
- The "certain event" S.    ("What is the probability of less than 10 dots?")
- All "elementary events" of one outcome: { 1 }, { 2 } , { 3 }, ..., { 6 }.

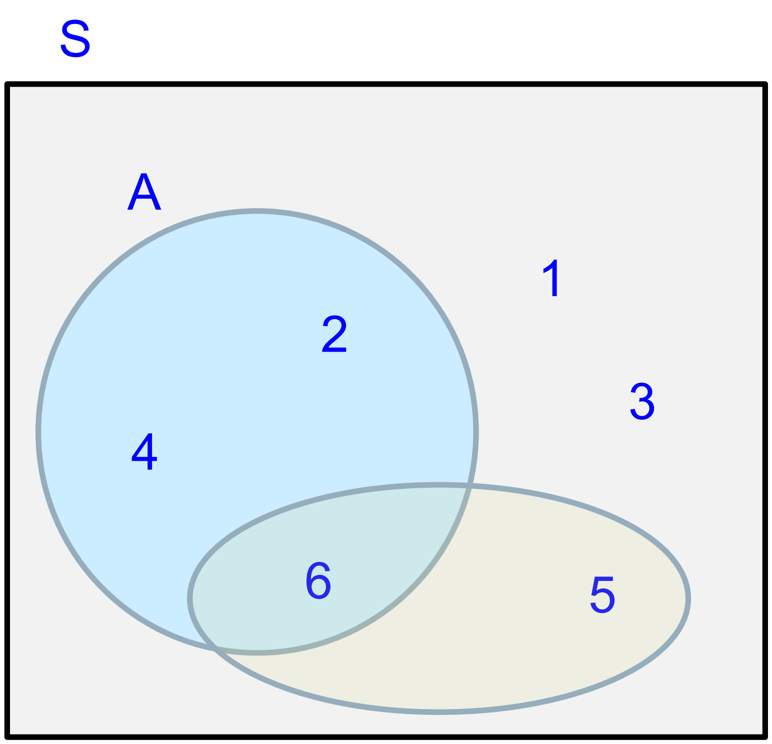
etc.....    This gives you the most flexible way of discussing the results of an experiment....

# Sample Spaces, Sample Points, and Events

Since events are subsets, **we can now use all the machinery of set theory to define and manipulate events.** It will also be useful to visualize events and sample spaces using Venn Diagrams:

**Example:** Toss a die and output the number of dots showing.
Let A = "there are an even number of dots showing" and
B = "there are at least 5 dots showing."

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




For any events A and B, we can define new events using set operations:

$$A \cup B = \{ 2, 4, 5, 6 \}$$

$$A \cap B = \{ 6 \}$$

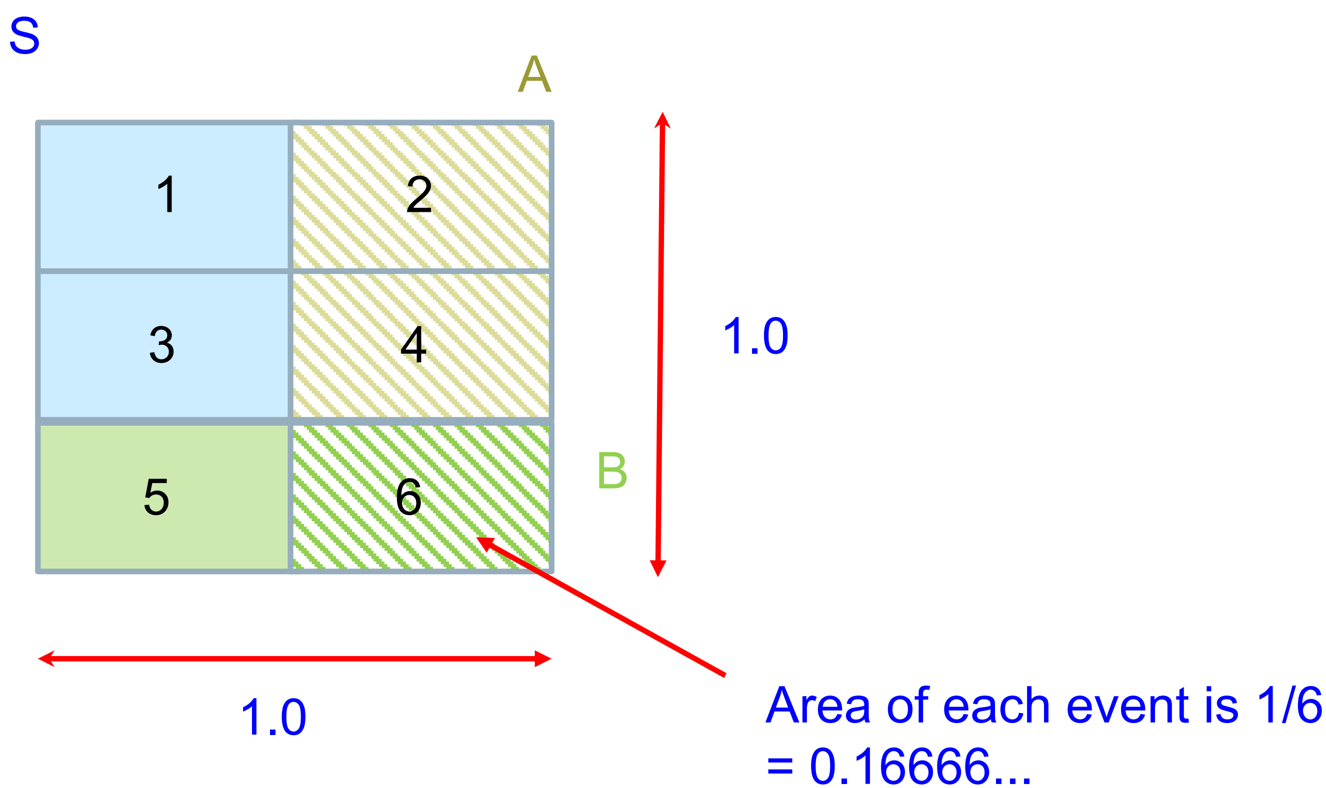
$$A^c = S - A = \{ 1, 3, 5 \}$$

$$A^c \cap B = \{ 5 \}$$

# Probability Spaces and Probability Axioms

These axioms make perfect sense if we consider Venn Diagrams where we use area as indicating probability,  so the area of **S = 1.0**, and the area of an event in the diagram = probability of that event.

**Example:** Toss a die and output the number of dots showing. Let A = "there are an even number of dots showing"  and B = "there are at least 5 dots showing."



# Probability Spaces and Probability Axioms

These axioms make perfect sense if we consider Venn Diagrams where we use area as indicating probability, so the area of an event in the diagram = probability of that event.

**$P_1$:** For any event A, we have $\mathbf{P(A) \geq 0}$.

"The area of each event is non-negative."

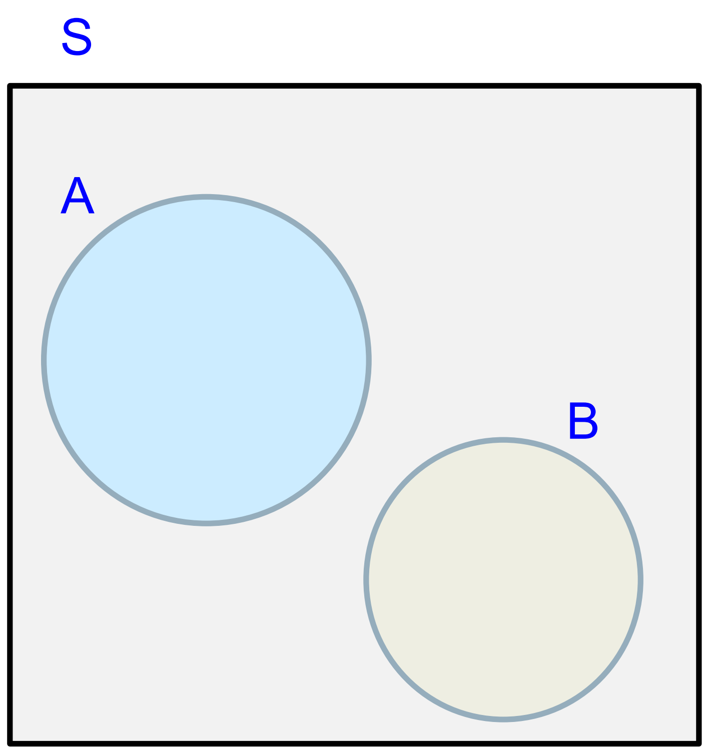
**$P_2$:** For the certain event S, we have $\mathbf{P(S) = 1.0}$.

"The area of the whole sample space is 1.0."

**$P_3$:** For any two disjoint events A and B we have

$$P(A \cup B) = P(A) + P(B)$$

"If two regions of S do not overlap, then the area of the two regions combined is the sum of the area of each region."



# Probability Spaces and Probability Axioms

So we measure the probability of events on a real-number scale from 0 to 1: