Introduction to Probability and Statistics

SDS 321

Lecture 1

Overview, Definitions and Axioms of Probability

Course Overview

- This course provides an introduction to probability and statistics.
- The course does not assume any prior knowledge of statistics.
- However, basic calculus will be necessary for the second section of the course.
- The first section will be on fundamentals of probability:
  - Basic concepts of probability
  - Probability trees and special rules
  - Combinatorics
- The second section will be on discrete random variables
  - Probability mass function (PMF) and CDF
  - Discrete random variables and multiple random variables
  - Common discrete distributions
- The third section will be on continuous random variables
  - Probability density function (PDF) and CDF
  - Common continuous distributions
  - Derived distributions
  - Special inequalities and theorems

What is Probability

“If I toss a fair coin, the probability of getting a Head is 0.5”

What does this mean?

We only have two possibilities: Head or Tail

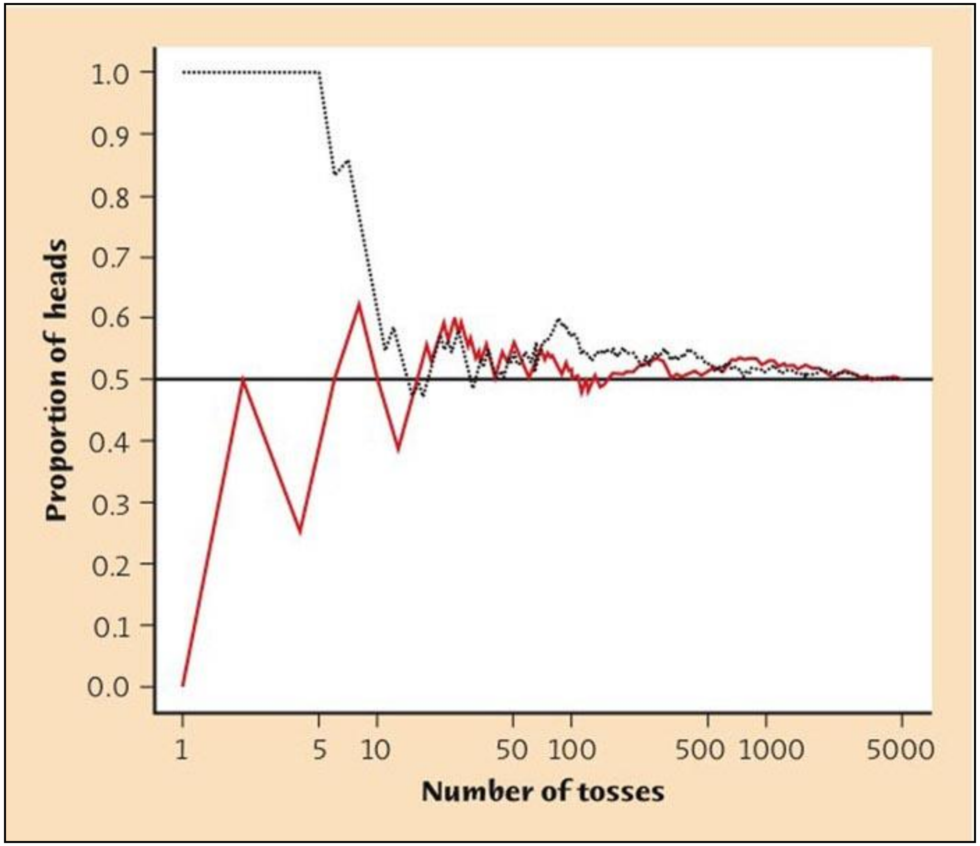
They are equally likely.

So each has a probability = 0.5

So Probability = $\frac{1}{2}$ = $\frac{\text{\# favorable outcomes}}{\text{\# total possible outcomes}}$

What is Probability

Probability: It is also defined as the proportion of times an event will occur if the experiment is repeated over and over again under the same conditions.



Probability: Some coin tossers

Example: Toss a fair coin

(1) Around 1900, the English statistician Karl Pearson heroically tossed a coin 24,000 times. Result: 12,012 heads. A proportion of 0.5005

(1) While imprisoned by the Germans during WW II, the South African mathematician John Kerrich tossed a coin 10,000 times. Result: 5067 heads, a proportion of 0.5067.

> Probability takes values between 0 and 1, where
>
> A certain event has probability = 1 and
> An impossible event has probability = 0.

Experiments and Events

- A Random Experiment is a process that can have more than one possible outcome, where the outcome itself depends on chance and cannot be specified ahead of time.
- An Event is an outcome or a set of outcomes of an experiment.
- Example:

(1) Experiment: Toss a fair coin

Event: ?

(2) Experiment: Throw a fair die

Event: ?

Experiments and Events

- Experiment: Toss a coin twice
  - Event: ?

- Experiment: Throw two dice
  - Event: ?

Sample Space

Definition:

•The Sample Space Ω is the **set** of all possible outcomes of a random experiment.

•The different elements of a sample space must be distinct, mutually exclusive and collectively exhaustive.

•Each outcome in the Sample Space should be in the simplest form, so that it cannot be broken down further.

Mutually Exclusive events:

Events that cannot occur at the same time. If one of them occurs, it excludes the other from occurring.

Example:

event of getting a 1 and event of getting a 2 on a single die roll.

Sample Space

What is the Sample Space for each of the following experiments?

1. Toss one coin
2. Roll one die
3. Toss two coins together
4. Toss three coins together
5. Roll two dice (imagine one is red and the other is green)

Events: Simple and Compound Events

An Event is a collection of some of the possible outcomes in a Sample Space

- Simple event:
  - Your two coin tosses came up *HH*.
  - Your rolled die shows a 6
- Compound event: can be decomposed into simple events
  - Your two coin tosses give two different outcomes
  - The sum of the two rolled dice is six
  - You got two odd faces from rolling two dice.

What are some of the ways you can break these compound events down into simple events?

Events: Simple and Compound Events

- Simple event:
  - Your two coin tosses came up HH.
  - Your rolled die shows a 6
- Compound event: can be decomposed into simple events
  - Your two coin tosses give two different outcomes
    - You got HT or TH.
  - The sum of the two rolled dice is six
    - You got $(1,5)$ or $(2,4)$ or $(3,3)$ or $(4,2)$ or $(5,1)$.
  - You got two odd faces from rolling two dice.
    - You got $(1,1)$, or $(1,3)$ or $\ldots$.

Sets and Sample Spaces

We need to introduce some mathematical concepts to define probability more concretely:

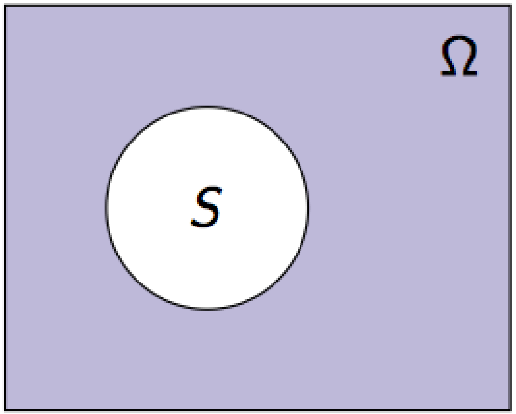
- A **set** is a collection of objects, which are called **elements**
  - The natural numbers are a set, where the elements are individual numbers.
  - This class is the set, where the elements are the professor, the TA and the students.
- If an element x is in a set S, we write $x \in S$.
- If a set contains no elements, we call it the **empty set**, $\emptyset$.
- If a set contains every possible element, we call it the **universal set**, Ω.

Sets

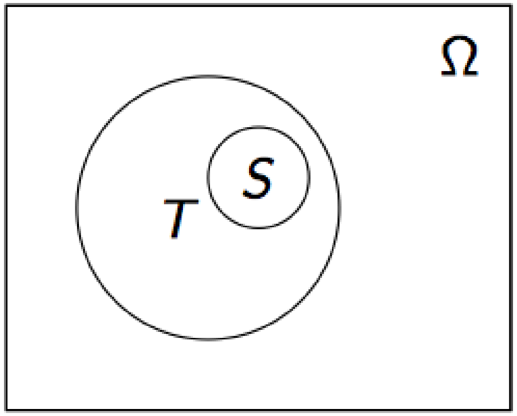
- A set can be **finite** (e.g. the set of people in this class) or **infinite** (e.g. the set of real numbers).
  - Set of primary colors = {red, blue, yellow}.
- If we can enumerate the elements of an infinite set, i.e. arrange the elements in a list, we say it is countable.
  - Set of positive integers $= \{1, 2, \dots\}$
- If we cannot enumerate the elements, we say it is uncountable.
  - the real numbers
  - the set of all subsets of natural numbers, aka the power set
- We can use curly brackets to describe a set in terms of its elements:
  - Sample space of a die roll: $S = \{1, 2, 3, 4, 5, 6\}$
  - Arbitrary set where all the elements meet some criterion C: $S = \{x | x \text{ satisfies } C\}$

Operations on Sets

- Let the **universal set** Ω be the set of all objects we might possibly be interested in.
- The **complement**, S^c, of a set S, w.r.t. Ω, is the set of all elements that are in Ω but not in S. So $\Omega^c = \emptyset$.
- We say $S \subseteq T$, if every element in S is also in T.
- $S \subseteq T$ and $T \subseteq S$ if and only if $S = T$.



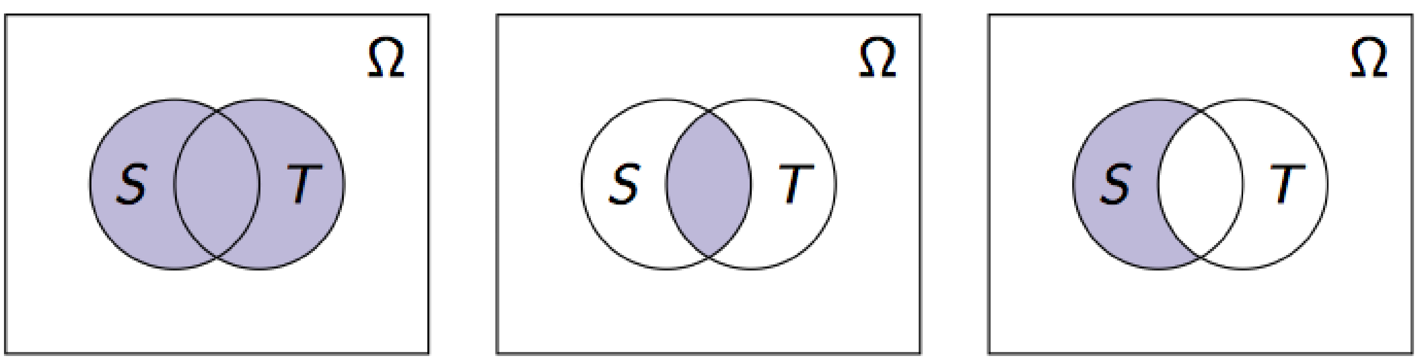
S^c is the shaded region



$S \subset T \subset \Omega$

Operations on Sets: Union, Intersection, Difference

- The **union**, $S \cup T$, of two sets S and T is the set of elements that are in either S or T (or both): $S \cup T = \{x | x \in S \text{ or } x \in T\}$.
- The **intersection**, $S \cap T$, of two sets S and T is the set of elements that are in both S and T: $S \cap T = \{x | x \in S \text{ and } x \in T\}$
- The **difference**, $S \setminus T$, of two sets S and T is the set of elements that are in S, but not in T: $S \setminus T = \{x | x \in S \text{ and } x \notin T\}$



$S \cup T$ $\quad$ $S \cap T$ $\quad$ $S \setminus T = S \cap T^c$

Operations on Sets

- We can extend the notions of union and intersection to multiple (even infinitely many!) sets:

$$\bigcup_{n=1}^{\infty} S_n = S_1 \cup S_2 \cup \cdots = \{x | x \in S_n \text{ for some n}\}$$

$$\bigcap_{n=1}^{\infty} S_n = S_1 \cap S_2 \cap \cdots = \{x | x \in S_n \text{ for all n}\}$$

- We say two sets are **disjoint** if their intersection is empty.
- We say a collection of sets are disjoint if no two sets have any common elements.
- If a collection of disjoint sets have union S, we call them a **partition** of S.

Operations on Sets

A partition of a set *S* is a collection of non-empty subsets of S, such that every element in S occurs in exactly one of these subsets.

Therefore, a collection of sets is a partition of S if and only if:

- This collection doesn't include the empty set Φ
- The union of all these sets is equal to S
- The intersection of any two of the sets is empty (pair-wise disjoint)
- Example: How many non-trivial partitions can you think of for S = {1, 2, 3}? What are they?

{1}, {2}, {3}   {3, 2}, {1}

{1, 2}, {3}

{1, 3}, {2}

{2, 3}, {1}

Probability Laws & Axioms of Probability

- The probability law assigns to an event E a non-negative number $P(E)$ which encodes our belief/knowledge about the "likelihood" of the event E.
- Axioms of probability:
  - **Nonnegativity**: $P(A) \geq 0$, for every event A.
  - **Additivity**: If A and B are two disjoint events, then the probability of their union satisfies $P(A \cup B) = P(A) + P(B)$.
    This extends to the union of infinitely many disjoint events:

    $$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

  - **Normalization**: The probability of the entire sample space Ω is equal to 1, i.e. $P(\Omega) = 1$

Example

You tossed two fair dice together. What is the probability of the event $E = \{\text{sum of the rolls} = 6\}$?

$$\frac{5}{36}$$

$$P(A) = \frac{\text{\# of fav. outcomes}}{\text{outcomes in } \Omega}$$

Properties of Probability Laws

1) If $A \subseteq B$, then $P(A) \leq P(B)$.

2) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. [General Aditivity Rule]

3) $P(A \cup B) \leq P(A) + P(B)$.

4) $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$.

All of the above can be proved by decomposing a set into disjoint partitions and using the additivity and non-negativity rules.

Properties of Probability Laws

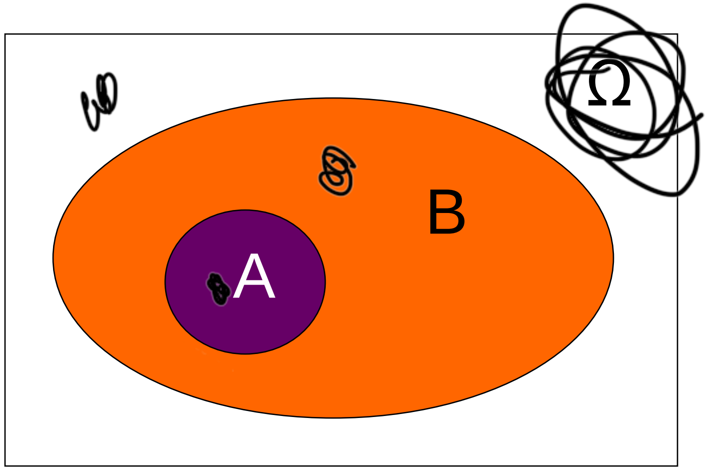
1) **If $A \subseteq B$, then $P(A) \leq P(B)$.**

Let's write B in terms of disjoint partitions:

$B = A \cup (A^c \cap B)$

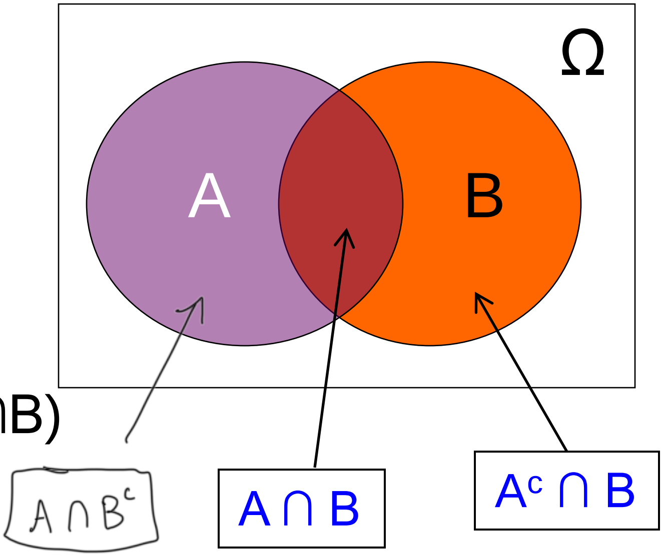
So $P(B) = P(A) + P(A^c \cap B)$

Therefore, $P(B) \geq P(A)$.



Properties of Probability Laws

2) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.



$A \cup B = (B^c \cap A) \cup (A \cap B) \cup (A^c \cap B)$

So $P(A \cup B) = P(B^c \cap A) + P(A \cap B) + P(A^c \cap B)$

Now, $A = (B^c \cap A) \cup (A \cap B)$

So $P(A) = P(B^c \cap A) + P(A \cap B)$

And $B = (A \cap B) \cup (A^c \cap B)$

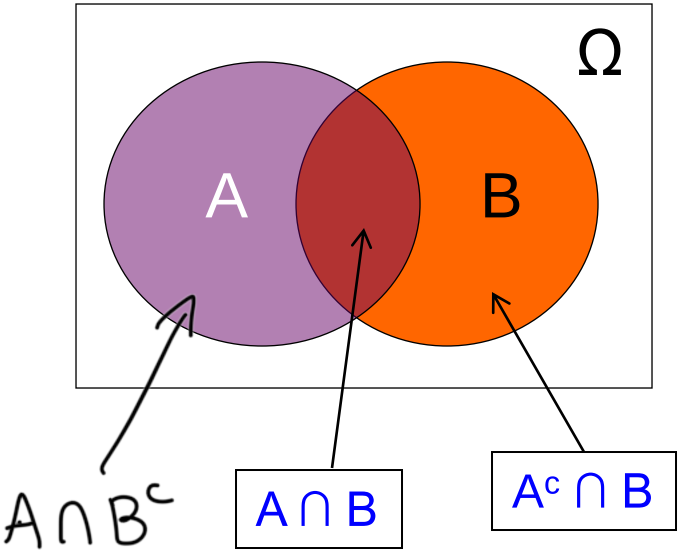
So $P(B) = P(A \cap B) + P(A^c \cap B)$

Can you finish this proof?

Properties of Probability Laws

2) Note that #2 is called the General Addition Rule

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$



If A and B here are Mutually Exclusive events, we have the special addition rule for mutually exclusive events:

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

What is the Venn Diagram for this?

Properties of Probability Laws

3) $P(A \cup B) \leq P(A) + P(B)$

From #2) we know:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Using the non-negativity axiom, $P(A \cap B) \geq 0$.

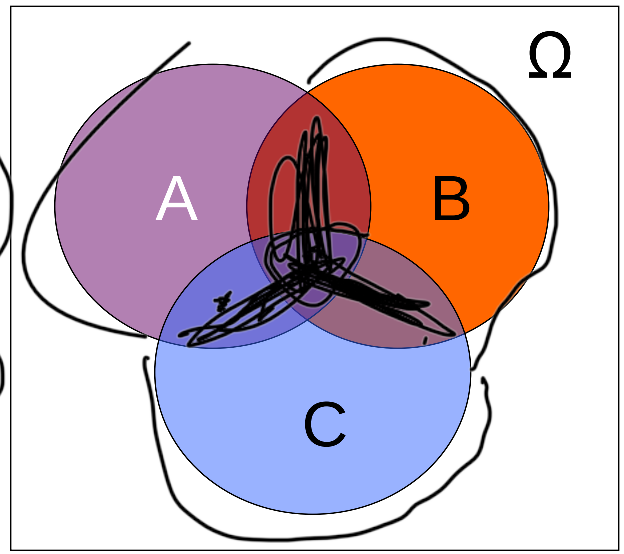
Therefore,

$$P(A \cup B) \leq P(A) + P(B).$$

Properties of Probability Laws

4) $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$.

We can show this by again looking at the disjoint partitions in the following Venn diagram:



Also note from the above figure that:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$