

## Lecture 2 : Probability (§2.1, 2.2)

Before we delve into the vast area of probability, let us first introduce the relevant terminology needed from set theory.

### I. Sets

Def. A set is a collection of objects which are called elements of the set. For example, if  $x$  is an element in a set  $S$ , we write  $x \in S$ . If  $x$  is not in  $S$ , we write  $x \notin S$ . An empty set is a set that has no elements, and is denoted by the Greek letter  $\emptyset$ .

Consider a set  $S$  that has a finite number of elements  $x_1, x_2, \dots, x_n$ . We write:

$$S = \{x_1, x_2, \dots, x_n\}$$

If  $S$  has infinitely many elements, we write:

$$S = \{x_1, x_2, \dots\}$$

If every element of  $S$  is also an element of a set  $T$ , we say that  $S$  is a subset of  $T$ , and write  $S \subset T$ . If  $S \subset T$  and  $T \subset S$ , then  $S = T$ .

Consider a universal set  $\Omega$ , such that  $\Omega$  contains all sets of interest in a particular context.

- \* The complement of a set  $S$  is denoted by  $\bar{S}$ , and it contains everything in  $\Omega$  that is not in  $S$ . Note that  $\bar{\Omega} = \emptyset$ .
- \* The union of two sets  $S$  and  $T$  is the set of all elements in  $S$  or  $T$  (or both), and is denoted  $S \cup T$ .
- \* The intersection of two sets  $S$  and  $T$  is the set of all elements that belong to both  $S$  and  $T$ , and is denoted  $S \cap T$ .

\* Consider an infinite number of sets  $S_1, S_2, \dots$ . Then, the union of all sets is denoted by

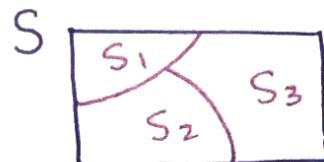
$$\bigcup_{n=1}^{\infty} S_n = S_1 \cup S_2 \cup \dots,$$

and the intersection is denoted by

$$\bigcap_{n=1}^{\infty} S_n = S_1 \cap S_2 \cap \dots$$

\* Two sets are said to be disjoint if there are no common elements between them, i.e., sets  $S$  and  $T$  are disjoint if  $S \cap T = \emptyset$ .

\* A collection of sets is said to be a partition of  $S$  if the sets in the collection are disjoint and their union is  $S$ .



$S_1, S_2$ , and  $S_3$  are a partition of  $S$

## Algebra of sets

$$S \cup T = T \cup S$$

$$S \cap (T \cup W) = (S \cap T) \cup (S \cap W)$$

$$S \cup (T \cup W) = (S \cup T) \cup W$$

$$S \cup (T \cap W) = (S \cup T) \cap (S \cup W)$$

## II. Probability Models

A probabilistic model is a mathematical description of an uncertain situation. Every probabilistic model involves an underlying process, called the experiment, which will produce one of several outcomes.

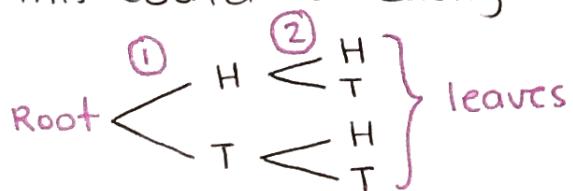
- \* Sample space : The set of all possible outcomes of the experiment. It is denoted by  $\Omega$ .
- \* Event: A subset of the sample space.

Ex. Consider the experiment of tossing a coin once. The possible outcomes are : Heads (H) or Tails (T). Therefore,  $\Omega = \{ H, T \}$ .

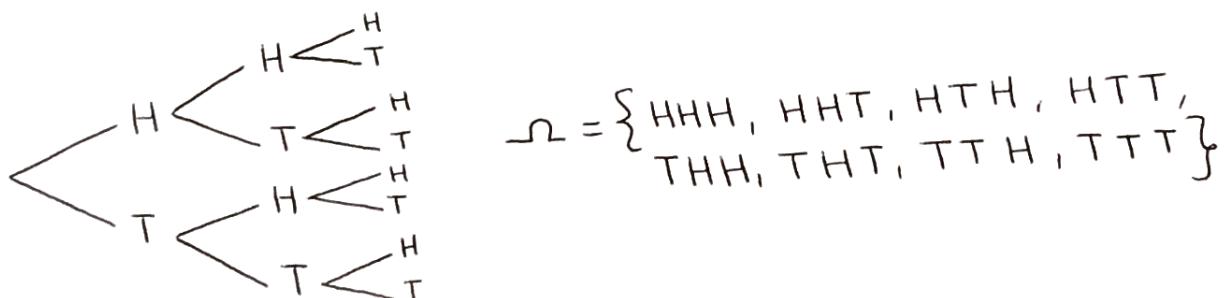
Ex. Consider the experiment of rolling a die once. The sample space is  $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ .

\* sequential probability models: Many experiments are sequential in nature. In this case , a tree representation is often useful. For example consider tossing a coin twice. The sample space is  $\Omega = \{ HH, HT, TH, TT \}$ .

This could be easily visualized using a tree :



Similarly, tossing a coin three times



### III. Probability Laws

Suppose we performed an experiment which gave us the sample space  $\Omega$ . To complete the probabilistic model, we must introduce a probability law. The probability law will tell us something about the likelihood of any given event in the sample space. Let  $A \subset \Omega$  be an event in the sample space of our experiment, then the probability law assigns a number  $P(A)$  to  $A$ , called the probability of  $A$ , satisfying the following axioms:

1. (Nonnegativity) For any event  $A$ ,  $P(A) \geq 0$ .
2. (Additivity) If  $A$  and  $B$  are two disjoint events, then the probability of their union satisfies

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

If the sample space has an infinite number of elements, and  $A_1, A_2, \dots$  is an infinite collection of disjoint events, then

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

3. (Normalization)  $P(\Omega) = 1$

Several properties can be derived from these axioms, e.g.,  $P(\emptyset) = 0$ .

Ex. Suppose you toss a coin once, and you are told that  $P(H) = p$ . Since  $\Omega = \{H, T\}$ , and  $P(\Omega) = P(H) + P(T) = 1$ , then  $P(T) = 1 - p$ .

List all possible events in  $\Omega = \{H, T\}$

$\{H, T\}$ ,  $\{H\}$ ,  $\{T\}$ ,  $\emptyset$  occurs with probability 0  
occurs with probability 1      occurs with probability  $p$       occurs with probability  $(1-p)$

Ex. Consider tossing a coin three times. Then, the sample space is  $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$ . Suppose the coin is fair and each of these outcomes is equally likely. What is the event that exactly two heads occur?

$$A = \text{exactly two heads occur} = \{HHT, HTH, THH\}$$

$$P(A) = P(\{HHT\}) + P(\{HTH\}) + P(\{THH\})$$

$$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

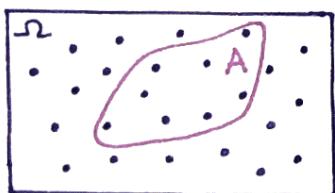
### Discrete Probability Law:

If  $\Omega$  consists of a finite number of outcomes, then the probability law is specified by the probabilities of the events that consist of a single element. Consider an event  $\{s_1, s_2, \dots, s_n\}$  with probability  $P(\{s_1, s_2, \dots, s_n\}) = P(s_1) + \dots + P(s_n)$ .

In the special case of  $P(s_1) = P(s_2) = \dots = P(s_n) = p$ , then  $p = \frac{1}{n}$  if the sample space consists of  $n$  possible outcomes,

i.e.,  $\Omega = \{s_1, s_2, \dots, s_n\}$ . If  $A$  is an event in  $\Omega$ ,

then  $P(A) = \frac{\# \text{ of elements in } A}{n}$



$$P(A) = \frac{8}{25}$$

## Properties of Probability Laws:

Consider an event A, which occurs with probability  $P(A)$ , then:

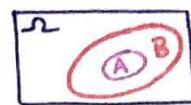
- \*  $\bar{A}$  is the complement of A, with  $P(\bar{A}) = 1 - P(A)$

- \*  $P(A) \leq 1$

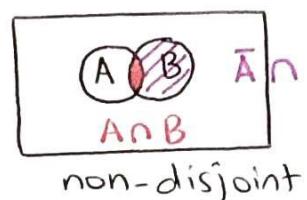
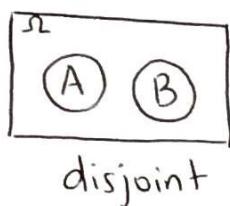
- \* If  $A \subset B$ , then  $P(A) \leq P(B)$

- \* If A and B are two events which are not mutually exclusive (i.e. not disjoint), then

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



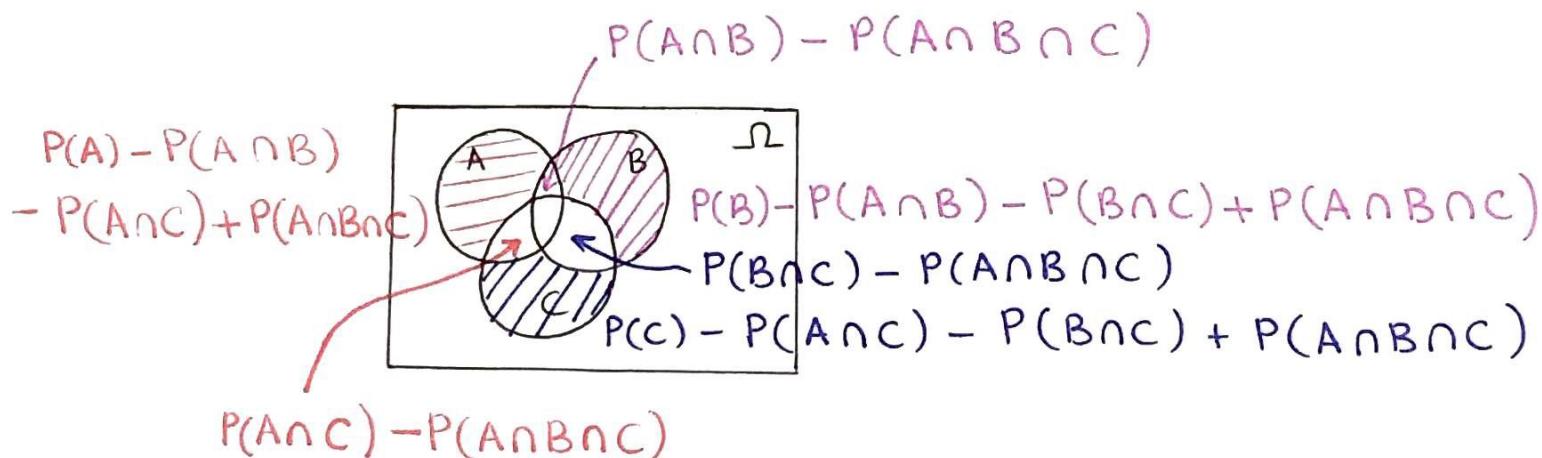
$$\begin{aligned} P(B) &= P(A) + P(\bar{A} \cap B) \\ &\geq P(A) \end{aligned}$$



Proof:  $P(A \cup B) = P(A) + P(\bar{A} \cap B) = P(A) + [P(B) - P(A \cap B)]$

- \*  $P(A \cup B) \leq P(A) + P(B)$  equal only when A and B are disjoint

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



$P(A \cap B \cap C)$  is added three times and subtracted three times  $\Rightarrow$  must be added again to account for it.