
Lecture 02: Probability Theory

INF1004 Mathematics II

Name: Timothy Chia

Date: 12/01/2025

1 Types of Probability

Probability is defined as a number in the range $[0, 1]$ that is assigned to an event associated with a random experiment. There are two interpretations of probability theory that guide how such values are assigned:

1.1 Classical Probability

In classical probability, all outcomes in the sample space are assumed to be equally likely. The probability of a random event A occurring is given as

$$P(A) = \frac{\text{Number of different outcomes in } A}{\text{Total number of possible outcomes}}$$

Note: This definition is not very practical for everyday life because it assumes the outcomes are equally likely.

1.2 Empirical Probability

In empirical probability theory, the probability of an event A occurring approximated by the relative frequency of its occurrence over a large number of trials such that

$$P(A) = \frac{\text{Number of times an outcome occurs}}{\text{Total number of observations}}$$

This interpretation relies on collecting data over the years or existing historical data to propose such relative frequencies.

1.3 Definitions

- We define a sample space S as the set of all possible outcomes of a random phenomenon.
- An event A is defined as the set of outcomes under investigation.
- By definition of probability, the following statements must hold to be valid:

$$0 \leq P(A) \leq 1, \quad P(S) = 1$$

- If A_1, A_2, \dots are mutually exclusive events, then

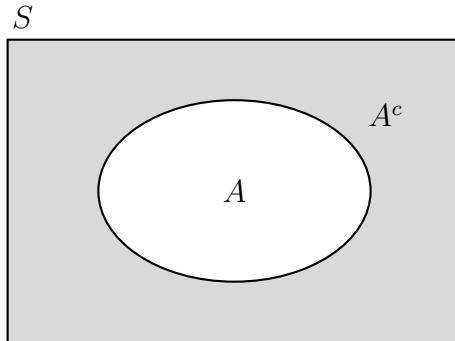
$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

2 Complementary Events

The complement of an event A denoted $P(A^c)$ and is given by

$$P(A^c) = P(S) - P(A)$$

Conceptually, the complement of an event can be thought of as the probability that it does not occur. Visually, this can be represented as the area in a venn diagram that is not in A .

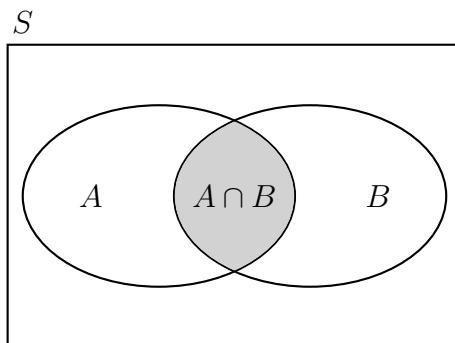


3 Intersection of Events

Given two events A and B , the probability of their intersection denoted

$$P(A \cap B)$$

is the probability that both events occur. In a venn diagram, it is the area where both events overlap.

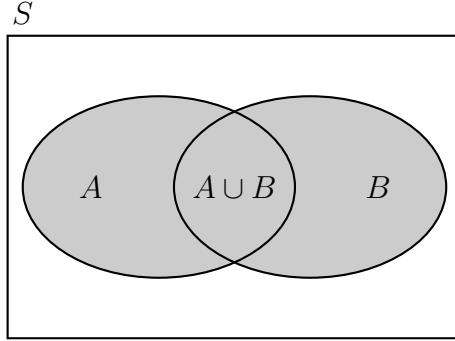


If two events A and B are mutually exclusive, then $P(A \cap B) = 0$ because both events never overlap.

4 Union of Events

Given two events A and B , the *union* of these events, denoted $A \cup B$, is the event that *at least one* of the events occurs. That is, $A \cup B$ occurs if event A occurs, or event B occurs, or both occur.

In terms of probability, $P(A \cup B)$ represents the probability that at least one of the two events occurs. Visually, in a Venn diagram, the union corresponds to the total area covered by both events A and B .



In general, the probability of the union of two events is given by

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

This formula accounts for the fact that the overlapping region $A \cap B$ is included in both $P(A)$ and $P(B)$, and therefore must be subtracted once to avoid double counting.

4.1 Mutually Exclusive Events

If events A and B are *mutually exclusive*, then they cannot occur at the same time. In this case,

$$P(A \cap B) = 0,$$

and the formula simplifies to

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

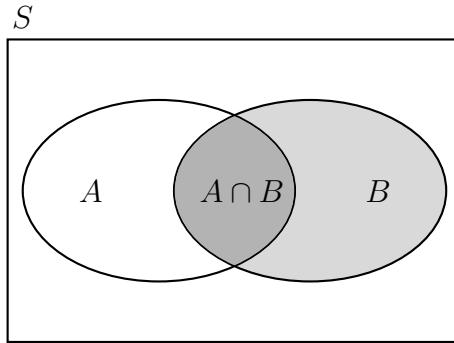
5 Conditional Probabilities

In practice, it happens naturally that probability of an event A can be changed if additional information is available, which leads to the discovery/creation of *conditional probability*.

5.1 Definition

The conditional probability of event A given that event B has occurred, denoted $P(A | B)$, is defined (for $P(B) > 0$) as

$$P(A | B) = \frac{P(A \cap B)}{P(B)}.$$



Note: Conditional probability is *not symmetric*. In general,

$$P(A | B) \neq P(B | A),$$

because the events A and B may have different probabilities and provide different information when conditioned upon.

Additionally, when working with conditional probability $P(A | B)$, the sample space is no longer the original sample space S . Instead, the sample space is restricted to event B . All probabilities are therefore measured *relative to B*.

6 Multiplicative Rule of Probability

The definition of conditional probability provides a way to compute the intersection of events A and B . Since,

$$\begin{aligned} P(A | B) &= \frac{P(A \cap B)}{P(B)}, \\ P(B | A) &= \frac{P(A \cap B)}{P(A)}. \end{aligned}$$

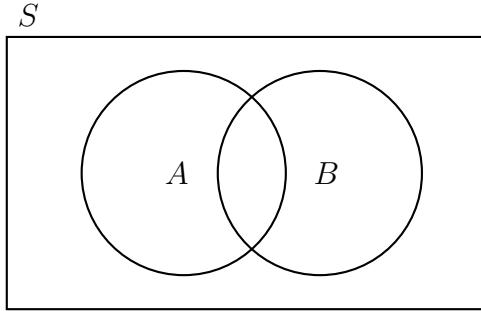
We have that,

$$\begin{aligned} P(A \cap B) &= P(A | B) \cdot P(B), \\ P(A \cap B) &= P(B | A) \cdot P(A). \end{aligned}$$

7 Independent and Mutually Exclusive Events

Events may be related in different ways depending on how the occurrence of one affects another. Two important relationships are *independence* and *mutual exclusivity*. These concepts are fundamentally different and should not be confused.

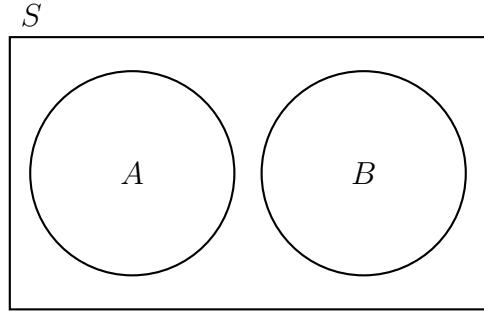
Independent Events



Both events may occur, but the occurrence of one does not affect the probability of the other.

Note: Independent events may overlap, whereas mutually exclusive events never do.

Mutually Exclusive Events



Events cannot occur at the same time, so $A \cap B = \emptyset$.

7.1 Independent Events

Two events A and B are said to be **independent** if

$$P(A \cap B) = P(A) P(B).$$

Equivalently, provided the relevant probabilities are non-zero,

$$P(A | B) = P(A) \quad \text{or} \quad P(B | A) = P(B).$$

That is, the occurrence of one event does not affect the probability of the other. Using the definition of conditional probability,

$$P(A \cap B) = P(A | B) P(B).$$

If A and B are independent, then $P(A | B) = P(A)$, and hence

$$P(A \cap B) = P(A) P(B).$$

This result is known as the **multiplication rule for independent events**.

More generally, if A_1, A_2, \dots, A_n are mutually independent events, then the probability that all events occur is

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i).$$

8 Law of Total Probability

The **Law of Total Probability** allows us to compute the probability of an event by decomposing the sample space into several mutually exclusive and exhaustive cases.

8.1 Definition

Suppose S_1, S_2, \dots, S_n form a **partition** of the sample space S , meaning:

- $S_i \cap S_j = \emptyset$ for $i \neq j$ (mutually exclusive),
- $\bigcup_{i=1}^n S_i = S$ (exhaustive),
- $P(S_i) > 0$ for all i .

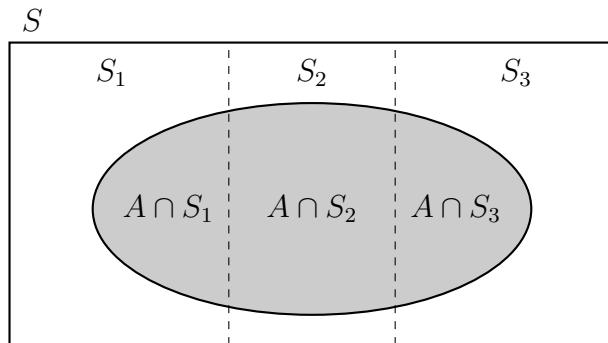
Then for any event $A \subseteq S$, we may write A as the union of its intersections with each S_i :

$$A = (A \cap S_1) \cup (A \cap S_2) \cup \dots \cup (A \cap S_n).$$

Since these intersections are mutually exclusive, their probabilities add, giving

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \dots + P(A \cap S_n).$$

This concept can be visualised with the following Venn diagram.



Using the definition of conditional probability,

$$P(A \cap S_i) = P(A | S_i) P(S_i).$$

Substituting this into the previous expression yields the **Law of Total Probability**:

$$P(A) = \sum_{i=1}^n P(A | S_i) P(S_i).$$

9 Bayes' Theorem

Bayes' Theorem provides a way to reverse conditional probabilities. In particular, it allows us to compute $P(B | A)$ in terms of $P(A | B)$, together with prior information about B . From the definition of conditional probability, we have

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad \text{and} \quad P(B | A) = \frac{P(A \cap B)}{P(A)}.$$

Rearranging the first expression gives

$$P(A \cap B) = P(A | B) P(B).$$

Substituting this into the second expression yields **Bayes' Theorem**:

$$P(B | A) = \frac{P(A | B) P(B)}{P(A)}.$$

In practice, the probability $P(A)$ in the denominator is often unknown directly. It is commonly computed using the **Law of Total Probability**.

Suppose B and B^c form a partition of the sample space, meaning they are mutually exclusive and exhaustive. Then

$$P(A) = P(A | B) P(B) + P(A | B^c) P(B^c).$$

Substituting this into Bayes' Theorem gives the expanded form:

$$P(B | A) = \frac{P(A | B) P(B)}{P(A | B) P(B) + P(A | B^c) P(B^c)}.$$

This formulation highlights how Bayes' Theorem updates the probability of an event using new information, weighted by prior probabilities.