ST2334

Probability and Statistics

Academic Year 2025/2026
Semester I

One

Basic Concepts of Probability

1 Probability Concepts and Definitions

In this section we introduce the basic terminology of probability theory: experiment, outcomes, sample space, events.

DEFINITION 1 (EXPERIMENT, SAMPLE SPACE, EVENT)
*A **statistical experiment** is any procedure that produces data or observations.*

*The **sample space**, denoted by S, is the set of all possible outcomes of a statistical experiment.* The sample space depends on the problem of interest!

*A **sample point** is an outcome (element) in the sample space.*

*An **event** is a subset of the sample space.*

EXAMPLE 1.1
Consider the experiment of **rolling a die**.

(i) If the problem of interest is "the number that shows on the top face", then

- Sample space: $S = \{1,2,3,4,5,6\}$.
- Sample point: 1 or 2 or 3 or 4 or 5 or 6.
- Some possible events are:
  - an event where an odd number occurs = $\{1,3,5\}$;
  - an event where a number greater than 4 occurs = $\{5,6\}$.

(ii) If the problem of interest is “whether the number is even or odd", then

- Sample space: $S = \{\text{even, odd}\}$.
- Sample point: “even" or “odd".
- A possible event is:
  - an event where an odd number occurs = $\{\text{odd}\}$.

Remark

The sample space is itself an event and is called a **sure event**.

An event that contains no element is the empty set, denoted by $\emptyset$, and is called a **null event**.

2 Event Operations & Relationships

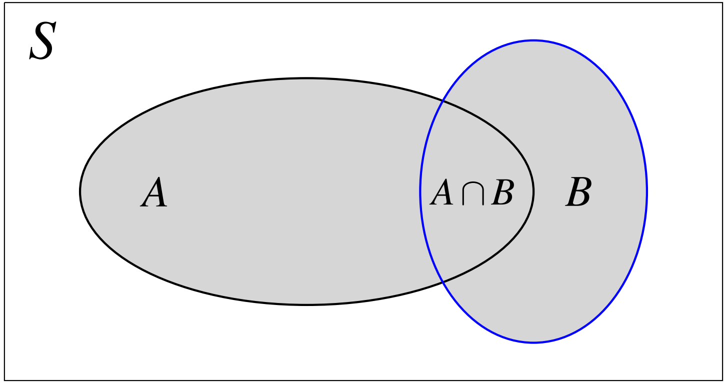
Let A and B be two events in the sample space S. We shall go through some event operations and relationships involving A and B.

- Event operations:
  (i) Union; (ii) Intersection; (iii) Complement.
- Event relationships:
  (i) Contained; (ii) Equivalent; (iii) Mutually exclusive.

Union

The **union** of events A and B, denoted by $A \cup B$, is the event containing all elements that belong to A or B or both. That is

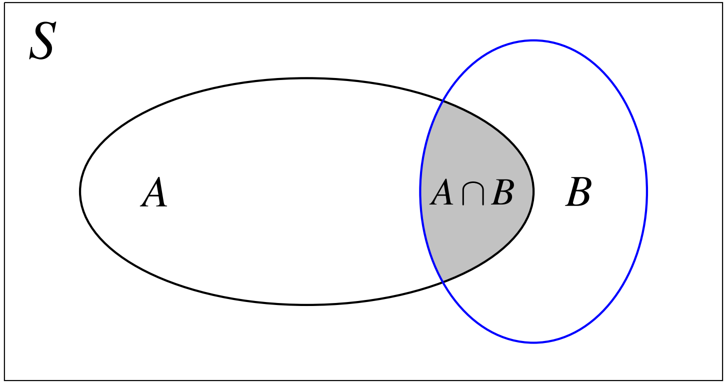
$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$



Intersection

The **intersection** of events A and B, denoted by $A \cap B$ or simply AB, is the event containing elements that belong to both A and B. That is

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$



We can also consider the **union** and **intersection** of n events: $A_1, A_2, \dots, A_n$.

- **Union**:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \ldots \cup A_n = \{x : x \in A_1 \text{ or } x \in A_2 \text{ or } \ldots \text{ or } x \in A_n\},$$

  comprises of elements that belong to one or more of $A_1, \dots, A_n$.

- **Intersection**:

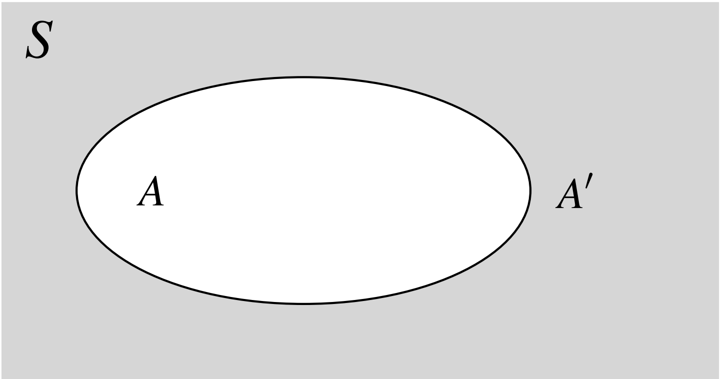
$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \ldots \cap A_n = \{x : x \in A_1 \text{ and } x \in A_2 \text{ and } \ldots \text{ and } x \in A_n\},$$

  comprises of elements that belong to every $A_1, \dots, A_n$.

Complement

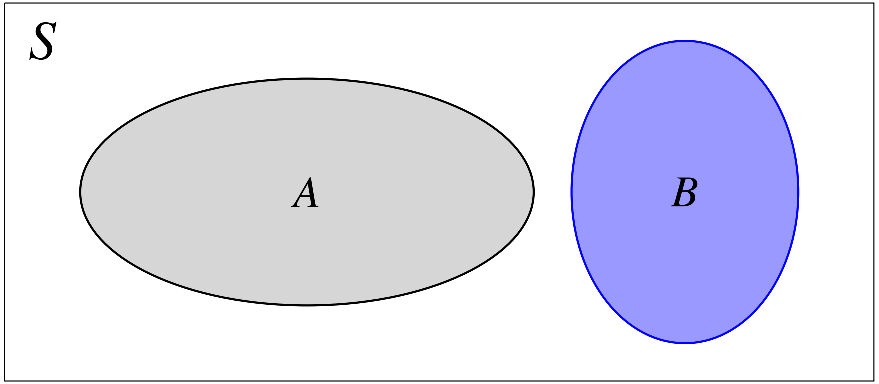
The **complement** of the event A with respect to S, denoted by A', is the event with elements in S, which are not in A. That is

$$A' = \{x : x \in S \text{ but } x \notin A\}.$$



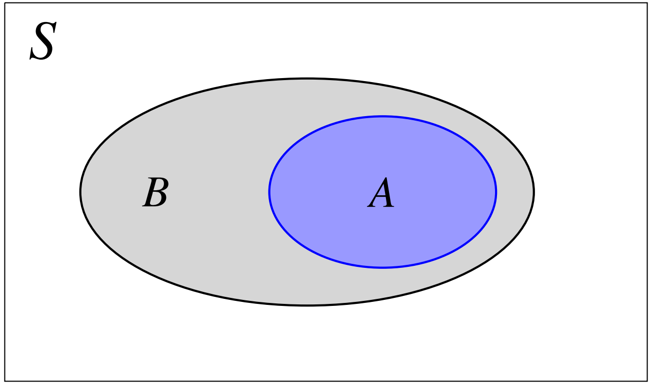
Mutually Exclusive

Events A and B are said to be **mutually exclusive** or **disjoint**, if $A \cap B = \emptyset$. That is, A and B have no element in common.



Contained and Equivalent

If all elements in A are also elements in B, then we say A is **contained** in B, denoted by $A \subset B$, or equivalently $B \supset A$.



If $A \subset B$ and $B \subset A$, then $A = B$. That is, set A and B are **equivalent**.

EXAMPLE 1.2
Consider the sample space and events:

$$S = \{1,2,3,4,5,6\}, \quad A = \{1,2,3\}, \quad B = \{1,3,5\}, \quad C = \{2,4,6\}.$$

Then

(i) $A \cup B = \{1,2,3,5\}; \quad A \cup C = \{1,2,3,4,6\}; \quad B \cup C = S.$

(ii) $A \cap B = \{1,3\}; \quad A \cap C = \{2\}; \quad B \cap C = \emptyset.$

(iii) $A \cup B \cup C = S; \quad A \cap B \cap C = \emptyset.$

(iv) $A' = \{4,5,6\}; \quad B' = \{2,4,6\} = C.$

Note that B and C are mutually exclusive, since $B \cap C = \emptyset$. On the other hand, A and B are not mutually exclusive as $A \cap B = \{1,3\} \neq \emptyset$.

More Event Operations

(a) $A \cap A' = \emptyset$

(b) $A \cap \emptyset = \emptyset$

(c) $A \cup A' = S$

(d) $(A')' = A$

(e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(f) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(g) $A \cup B = A \cup (B \cap A')$

(h) $A = (A \cap B) \cup (A \cap B')$

De Morgan's Law

For any n events $A_1, A_2, \ldots, A_n$,

(i) $(A_1 \cup A_2 \cup \ldots \cup A_n)' = A_1' \cap A_2' \cap \ldots \cap A_n'$.

A special case: $(A \cup B)' = A' \cap B'$.

(j) $(A_1 \cap A_2 \cap \ldots \cap A_n)' = A_1' \cup A_2' \cup \ldots \cup A_n'$.

A special case: $(A \cap B)' = A' \cup B'$.

Example 1.3
We return to Example 1.2 where

$$S = \{1,2,3,4,5,6\}, \quad A = \{1,2,3\}, \quad B = \{1,3,5\}, \quad C = \{2,4,6\}.$$

We have

$$A' = \{4,5,6\}, \quad B' = \{2,4,6\}, \quad C' = \{1,3,5\}.$$

We check that

$$(A \cup B)' = \{1,2,3,5\}' = \{4,6\}; \quad A' \cap B' = \{4,5,6\} \cap \{2,4,6\} = \{4,6\}.$$

This agrees with $(A \cup B)' = A' \cap B'$.

Also,

$$(A \cap B)' = \{1,3\}' = \{2,4,5,6\}; \quad A' \cup B' = \{4,5,6\} \cap \{2,4,6\} = \{2,4,5,6\}.$$

This agrees with $(A \cap B)' = A' \cup B'$.

Similarly, we can check that

$$(A \cup B \cup C)' = \emptyset = A' \cap B' \cap C' \quad \text{and} \quad (A \cap B \cap C)' = S = A' \cup B' \cup C'.$$

3 Counting Methods

In many instances, we need to count the number of ways that some operations can be carried out or that certain situations can happen.

There are two fundamental principles in counting:

Multiplication principle

Addition principle

They can be applied to derive some important counting methods: **permutation** and **combination**.

Multiplication Principle

Suppose that r different experiments are to be performed sequentially. Suppose

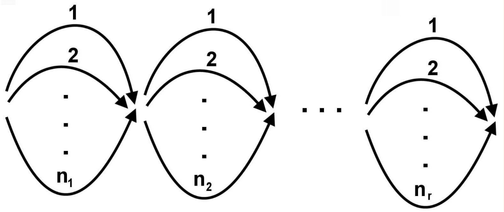
experiment 1 results in n_1 possible outcomes;

for each outcome above, experiment 2 results in n_2 possible outcomes;

$\cdots \ \cdots \ \cdots \ \cdots \ \cdots$

for each outcome above, experiment r results in n_r possible outcomes.

Then there are $n_1 n_2 \cdots n_r$ possible outcomes for the r experiments.



EXAMPLE 1.4

How many possible outcomes are there when a die and a coin are thrown together?

Solution:
Note that for

- experiment 1: throwing a die, there are 6 possible outcomes: {1,2,3,4,5,6}.
- experiment 2: throwing a coin, with each outcome of experiment 1, there are 2 possible outcomes: $\{H, T\}$.

So altogether there are $6 \times 2 = 12$ possible outcomes.

In fact, the sample space is given by

$$S = \{(x, y) : x = 1, \ldots, 6; y = H \text{ or } T\}.$$

Addition Principle

Suppose that an experiment can be performed by k different procedures.

Procedure 1 can be carried out in n_1 ways;

Procedure 2 can be carried out in n_2 ways;

$\ldots \ldots \ldots \ldots \ldots$

Procedure k can be carried out in n_k ways.

Suppose that the “ways" under different procedures *do not overlap*. Then the total number of ways we can perform the experiment is

$$n_1 + n_2 + \ldots + n_k.$$

EXAMPLE 1.5 (ORCHARD ROAD)
We can take the MRT or bus from home to Orchard road. Suppose there are three bus routes and two MRT routes. How many ways can we go from home to Orchard road?

Solution:
Consider the trip from home to Orchard road as an experiment. Two procedures can used to complete the experiment:

Procedure 1: take MRT – 2 ways.

Procedure 2: take bus – 3 ways.

These ways do not overlap. So the total number of ways we can go from home to Orchard road is $2+3=5$.

PERMUTATION

A **permutation** is a selection and arrangement of r objects out of n. In this case, *order* is taken into consideration.

The number of ways to choose and arrange r objects out of n, where $r \leq n$, is denoted by P_r^n, where

$$P_r^n = \frac{n!}{(n-r)!} = n(n-1)(n-2)\ldots(n-(r-1)).$$

obj 1	obj 2	obj 3	$\cdots$	obj r
n ways	$(n-1)$ ways	$(n-2)$ ways	$\cdots$	$(n-(r-1))$ ways

Remark

When $r = n$, $P_n^n = n!$.

Essentially, it is the number of ways to arrange n objects in order.

EXAMPLE 1.6

Find the number of possible four-letter code words in which all letters are different.

Solution:
Note that there are $n = 26$ alphabets, and $r = 4$ in our case.

So the number of possible four-letter code words is

$$P_4^{26} = (26)(25)(24)(23) = 358800.$$

COMBINATION

A **combination** is a selection of r objects out of n, *without regard to the order*.

The number of combinations of choosing r objects out of n, denoted by C_r^n or $\binom{n}{r}$, is given by as

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Note that this formula immediately implies $\binom{n}{r} = \binom{n}{n-r}$.

The derivation is as follows.

(A) Thinking in terms of permutation, *the number of ways to choose and arrange r objects out of n* is P_r^n.

(B) On the other hand, the same permutation task can be achieved by conducting the following two experiments sequentially:

(1) Select r objects out of n without regard to the order: $\binom{n}{r}$ ways.

(2) For each such combination, permute its r objects: P_r^r ways.

(C) Therefore, by the multiplication rule, *the number of ways to choose and arrange r objects out of n* is $\binom{n}{r} \times P_r^r$.

(D) As a consequence, $\binom{n}{r} \times P_r^r = P_r^n$, and so we obtain

$$\binom{n}{r} = \frac{P_r^n}{P_r^r} = \frac{n!/(n-r)!}{r!} = \frac{n!}{r!(n-r)!}.$$

EXAMPLE 1.7

From 4 women and 3 men, find the number of committees of size 3 that can be formed with 2 women and 1 man.

Solution:

The number of ways to select 2 women from 4 is $\binom{4}{2} = 6$.

The number of ways to select 1 man from 3 is $\binom{3}{1} = 3$.

By the multiplication rule, the number of committees formed with 2 women and 1 man is

$$\binom{4}{2} \times \binom{3}{1} = 6 \times 3 = 18.$$

4 Probability

Intuitively, the term probability is understood as the chance or how likely a certain event may occur.

More specifically, let A be an event in an experiment. We typically associate a number, called probability, to quantify how likely the event A occurs. This is denoted as $P(A)$.

Let us now investigate how we can obtain such a number.

You will discover that the fundamental concept of probability is extended from an idea based on intuition to a rigorous, abstract, and advanced mathematical theory. It has also wide applications in various scientific disciplines.

Interpretation of Probability: Relative Frequency

Suppose that we repeat an experiment E for a total of n times.

Let n_A be the number of times that the event A occurs.

Then $f_A = n_A/n$ is called the **relative frequency** of the event A in the n repetitions of E.

Clearly, f_A may not equal to $P(A)$ exactly. However when n grows large, we expect f_A to be close to $P(A)$; in the sense that $f_A \approx P(A)$. Or mathematically,

$$f_A \to P(A), \quad \text{as } n \to \infty.$$

Thus f_A “mimics" $P(A)$, and has the following properties:

(a) $0 \leq f_A \leq 1$.

(b) $f_A = 1$ if A occurs in every repetition.

(c) If A and B are mutually exclusive events, $f_{A \cup B} = f_A + f_B$.

Extending this idea, we can define **probability on a sample space** mathematically.

Axioms of Probability

Probability, denoted by $P(\cdot)$, is a **function** on the collection of events of the sample space S, satisfying:

Axiom 1. For any event A,

$$0 \leq P(A) \leq 1.$$

Axiom 2. For the sample space,

$$P(S) = 1.$$

Axiom 3. For any two mutually exclusive events A and B, that is, $A \cap B = \emptyset$,

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

EXAMPLE 1.8

Let H denote the event of getting a head when a coin is tossed. Find $P(H)$, if

(i) the coin is fair;

(ii) the coin is biased and a head is twice as likely to appear as a tail.

Solution:
The sample space is $S = \{H, T\}$.

(i) "The coin is fair" means that $P(H) = P(T)$.

The events $\{H\}$ and $\{T\}$ are mutually exclusive. Thus based on Axioms 2 and 3, we have

$$1 = P(S) = P(\{H\} \cup \{T\}) = P(H) + P(T) = 2P(H).$$

This gives $P(H) = 1/2$.

(ii) “A head is twice likely to appear as a tail" means $P(H) = 2P(T)$; therefore

$$1 = P(S) = P(\{H\} \cup \{T\}) = P(H) + P(T) = 3P(T).$$

This gives $P(T) = 1/3$ and $P(H) = 2/3$.

Basic Properties of Probability

Using the axioms, we can derive the following propositions.

Proposition 2
The probability of the empty set $\emptyset$ *is* $P(\emptyset) = 0$.

Proof Since $\emptyset \cap \emptyset = \emptyset$ and $\emptyset = \emptyset \cup \emptyset$, applying Axiom 3 leads to

$$P(\emptyset) = P(\emptyset \cup \emptyset) = P(\emptyset) + P(\emptyset) = 2P(\emptyset).$$

This implies that $P(\emptyset) = 0$. ✠

PROPOSITION 3

If $A_1, A_2, \ldots, A_n$ are mutually exclusive events, that is $A_i \cap A_j = \emptyset$ for any $i \neq j$, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Proof This proposition can be established easily using induction and Axiom 3. ✠

PROPOSITION 4

For any event A, we have

$$P(A') = 1 - P(A).$$

Proof Since $S = A \cup A'$ and $A \cap A' = \emptyset$, based on Axioms 2 and 3, we have

$$1 = P(S) = P(A \cup A') = P(A) + P(A').$$

The result follows. ✠

PROPOSITION 5

For any two events A and B,

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

Proof Based on the properties

$$A = (A \cap B) \cup (A \cap B') \quad \text{and} \quad (A \cap B) \cap (A \cap B') = \emptyset,$$

we have

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

✠

Proposition 6

For any two events A and B,

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

Proof Based on the event operations

$$A \cup B = B \cup (A \cap B') \quad \text{and} \quad B \cap (A \cap B') = \emptyset,$$

and Proposition 5 which states

$$P(A \cap B') = P(A) - P(A \cap B),$$

we have

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

✠

Proposition 7

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

Proof Since $A \subset B$, we have $A \cup B = B$. Also, we have

$$A \cup B = A \cup (B \cap A') \quad \text{and} \quad A \cap (B \cap A') = \emptyset.$$

Thus we obtain

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

✠

EXAMPLE 1.9

A retail establishment accepts either the American Express or the VISA credit card.

A total of 24% of its customers carry an American Express card, 61% carry a VISA card, and 11% carry both.

What is the probability that a customer carries a credit card that the establishment will accept?

Solution:
Let

$$A = \{\text{the customer carries an American Express Card}\}$$

and

$$V = \{\text{the customer carries an VISA Card}\}.$$

Then

$$P(A) = 0.24, \quad P(V) = 0.61, \quad P(A \cap V) = 0.11.$$

The question asked for $P(A \cup V)$, which is given as

$$P(A \cup V) = P(A) + P(V) - P(A \cap V) = 0.24 + 0.61 - 0.11 = 0.74.$$

FINITE SAMPLE SPACE WITH EQUALLY LIKELY OUTCOMES

Consider a sample space $S = \{a_1, a_2, \ldots, a_k\}$.

Assume that all outcomes in the sample space are **equally likely** to occur, i.e.,

$$P(a_1) = P(a_2) = \cdots = P(a_k).$$

Then for any event $A \subset S$,

$$P(A) = \frac{\text{number of sample points in } A}{\text{number of sample points in } S}.$$

Example 1.10

A box contains 50 bolts and 150 nuts. Half of the bolts and half of the nuts are rusted.

If one item is chosen at random, what is the probability that it is rusted or is a bolt?

Solution:
We define the following events

$$A = \{\text{the item is rusted}\}, \quad B = \{\text{the item is a bolt}\}, \quad S = \{\text{all the items}\}.$$

Since the item is selected at random, each of the 200 elements in S is equally likely to be chosen.

- A consists of $25 + 75 = 100$ elements;
- B consists of 50 elements; and
- $A \cap B$ consists of 25 elements.

These give

$$P(A) = 100/200 \quad , \quad P(B) = 50/200 \quad , \quad P(A \cap B) = 25/200.$$

Therefore the required probability is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 5/8.$$

5 Conditional Probability

Sometimes we need to compute the probability of some events when some **partial information** is available.

Specifically, we might need to compute the probability of an event B, given that we have the information “an event A has occurred".

Mathematically, we denote

$$P(B|A)$$

as the conditional probability of the event B, given that event A has occurred.

Definition 8 (Conditional Probability)
For any two events A and B with $P(A) > 0$, the ***conditional probability*** *of B given that A has occurred is defined by*

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

EXAMPLE 1.11

A fair die is rolled twice.

(i) What is the probability that the sum of the 2 rolls is even?

(ii) Given that the first roll is a 5, what is the (conditional) probability that the sum of the 2 rolls is even?

Solution:

We define the following events:

$$\begin{aligned} B &= \{\text{the sum of the 2 rolls is even}\}, \\ A &= \{\text{the first roll is a } 5\}. \end{aligned}$$

(i) The sample space is given by

		2nd roll					
		1	2	3	4	5	6
	1	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(1,6)$
	2	$(2,1)$	$(2,2)$	$(2,3)$	$(2,4)$	$(2,5)$	$(2,6)$
1st roll	3	$(3,1)$	$(3,2)$	$(3,3)$	$(3,4)$	$(3,5)$	$(3,6)$
	4	$(4,1)$	$(4,2)$	$(4,3)$	$(4,4)$	$(4,5)$	$(4,6)$
	5	$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$
	6	$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$

It is easy to see that $P(B) = {}^{18}\!/_{36}$.

(ii) Since we know that A has already happened, we can just look at the fifth row:

		2nd roll					
		1	2	3	4	5	6
1st roll	5	$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$

We are interested to look for instances along this row that gives an even sum. So $P(B|A) = {}^{3}/_{6}$.

Alternatively, we can use the formula:

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{3}{36}}{\frac{6}{36}}.$$

Remark (Reduced Sample Space)

$P(B|A)$ is read as:

"the conditional probability that B occurs given that A has occurred."

Since we know that A has occurred, regard A as our new, or **reduced sample space**.

The conditional probability that the event B given A will equal the probability of $A \cap B$ relative to the probability of A.

Multiplication Rule

Starting from the definition of conditional probability, and rearranging the terms, we have

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A), \quad \text{if } P(A) \neq 0 \\ \text{or} \quad P(A \cap B) &= P(B)P(A|B), \quad \text{if } P(B) \neq 0. \end{aligned}$$

This is known as the **Multiplication Rule**.

Inverse Probability Formula

The multiplication rule together with the definition of the conditional probability gives us:

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

This is known as the **Inverse Probability Formula**.

EXAMPLE 1.12

Deal 2 cards from a regular playing deck without replacement. What is the probability that both cards are aces?

Solution:

$$\begin{aligned}
P(\text{both aces}) &= P(\text{1st card is ace and 2nd card is ace})\\
&= P(\text{1st card ace}) \cdot P(\text{2nd card ace}|\text{1st card ace})\\
&= \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}.
\end{aligned}$$

6 Independence

We saw several examples where conditioning on one event changes our beliefs about the probability of another event.

In this section, we discuss the important concept of independence, where learning that the event B occurred gives us no information that would change our probabilities for another event A occurring.

Definition 9 (Independence)

*Two events A and B are **independent** if and only if*

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

We denote this by $A \perp B$.

*If A and B are not independent, they are said to be **dependent**, denoted by $A \not\perp B$.*

Remark

If $P(A) \neq 0$, $A \perp B$ if and only if $P(B|A) = P(B)$.

This follows from the definition of conditional probability –

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

Intuitively, this is the same as saying that A and B are independent if the knowledge of A does not change the probability of B.

Likewise, if $P(B) \neq 0$, $A \perp B$ if and only if $P(A|B) = P(A)$.

EXAMPLE 1.13
Suppose we roll 2 fair dice.

(i) Let

$$A_6 = \{\text{the sum of two dice is 6}\}, \quad B = \{\text{the first die equals 4}\}.$$

Thus

$$P(A_6) = {}^{5}\!/_{36}, \quad P(B) = {}^{6}\!/_{36} = {}^{1}\!/_{6} \quad \text{and} \quad P(A_6 \cap B) = {}^{1}\!/_{36}.$$

As $P(A_6 \cap B) \neq P(A_6)P(B)$, we say that A_6 and B are **dependent**.

(ii) Let

$$A_7 = \{\text{the sum of two dice is } 7\}.$$

Then

$$P(A_7 \cap B) = {}^1\!/_{36}, \quad P(A_7) = {}^1\!/_6 \quad \text{and} \quad P(B) = {}^1\!/_6.$$

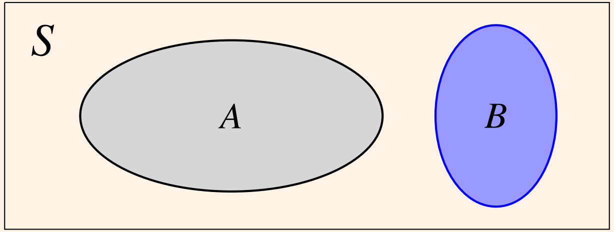
As $P(A_7 \cap B) = P(A_7)P(B)$, we say that A_7 and B are **independent**.

Independent vs Mutually Exclusive

Independence and **mutually exclusivity** are totally different concepts:

$$\begin{aligned} A, B \text{ independent } &\Leftrightarrow P(A \cap B) = P(A)P(B) \\ A, B \text{ mutually exclusive } &\Leftrightarrow A \cap B = \emptyset \end{aligned}$$

“Mutually exclusivity" can be illustrated by a Venn diagram (like below), but we can not do that for “independence".



7 The Law of Total Probability

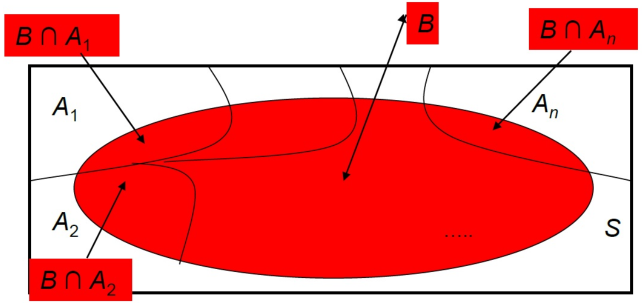
The definition of conditional probability has far-reaching consequences.

In this section we look at the Law of Total Probability (LOTP), which relates conditional probability to unconditional probability.

Definition 10 (Partition)
*If $A_1, A_2, \ldots, A_n$ are mutually exclusive events and $\cup_{i=1}^{n} A_i = S$, we call $A_1, A_2, \ldots, A_n$ a **partition** of S.*

Theorem 11 (Law of Total Probability)
Suppose $A_1, A_2, \ldots, A_n$ is a partition of S. Then for any event B, we have

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i) P(B|A_i).$$



Special case: Law of Total Probability

For any events A and B, we have

$$P(B) = P(A)P(B|A) + P(A')P(B|A').$$

EXAMPLE 1.14 (FRYING FISH)

At a nasi lemak stall, the chef and his assistant take turns to fry fish.

The chef burns his fish with probability 0.1, his assistant burns his fish with probability 0.23.

If the chef is frying fish 80% of the time, what is the probability that the fish you order is burnt?

Solution:
Let

$$
\begin{aligned}
B &= \{\text{the fish is burnt}\}, \\
C &= \{\text{the fish is fried by the chef}\}.
\end{aligned}
$$

We then need to compute $P(B)$. Using the Law of Total Probability,

$$P(B) = P(C)P(B|C) + P(C')P(B|C') = 0.8 \times 0.1 + 0.2 \times 0.23.$$

8 Bayes' Theorem

We now discuss Bayes' Theorem (or Bayes' Rule), which will allow us to relate $P(A|B)$ to $P(B|A)$ and compute conditional probabilities in a wide range of problems.

Theorem 12 (Bayes' Theorem)
Let $A_1, A_2, \ldots, A_n$ be a partition of S, then for any event B and $k = 1, 2, \ldots, n$,

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_k)}.$$

Proof Bayes' Theorem can be derived based on the definition of conditional probability, the Multiplication Rule, and the Law of Total Probability.

In particular,

$$\begin{aligned} P(A_k|B) &= \frac{P(A_k \cap B)}{P(B)} = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^n P(B \cap A_i)} \\ &= \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^n P(A_i)P(B|A_i)}. \end{aligned}$$

✠

Special case: Bayes' Theorem

Let us consider a special case of Bayes' Theorem when $n = 2$.

$\{A, A'\}$ becomes a partition of S, and we have

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

EXAMPLE 1.15
The previous formula is practically meaningful.

For example, consider the events

$$A = \text{disease status of a person}, \quad B = \text{symptom observed}.$$

Then

- $P(A)$: probability of a disease in general;
- $P(B|A)$: if diseased, probability of observing symptom;
- $P(A|B)$: if symptom observed, probability of diseased.

EXAMPLE 1.16

Historically, we observe the collapse of some newly constructed house.

The chance that the design of the house is faulty is 1%. If the design is faulty, the chance that the house collapses is 75%; otherwise, the chance is 0.01%.

We observe that a newly constructed house collapsed, what is the probability that the design is faulty?

Solution:

Let

$$B = \{\text{The design is faulty}\}, \quad A = \{\text{The house collapses}\}.$$

We then have

$$P(B) = 0.01, \quad P(A|B) = 0.75, \quad \text{and} \quad P(A|B') = 0.0001.$$

The question asked for $P(B|A)$. We will compute it using Bayes' Theorem.

The denominator can be computed using the Law of Total Probability:

$$\begin{aligned} P(A) &= P(B)P(A|B) + P(B')P(A|B') \\ &= (0.01)(0.75) + (0.99)(0.0001) = 0.007599. \end{aligned}$$

The numerator is

$$P(A|B)P(B) = 0.75(0.01) = 0.0075.$$

As a consequence,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = 0.9870.$$