ST2334

Probability and Statistics

Academic Year 2022/2023
Semester I

Chapter 1: Basic Concepts of Probability

1 Basic Probability Concepts and Definitions

- **Statistical Experiment**: Any procedure that obtains data (observations).
- **Sample Space** (denoted by S): The set of all possible outcomes of a statistical experiment.

  It depends on the problem of interest!
- **Sample Point**: Every outcome (element) in a sample space.
- **Events**: Subset of a sample space.

Example 1.1
Consider an experiment of **tossing a die**.

- If the problem of interest is “the number 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.
  - Events: (1) An event that an odd number occurs = $\{1,3,5\}$; (2) An event that a number greater than 4 occurs = $\{5,6\}$.
- If the problem of interest is “whether the number is even or odd", then

- Sample space: $S = \{\text{even}, \text{odd}\}$.
- Sample point: “even" or “odd".
- Events: An event that an odd number occurs = {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

- Denote by S the sample space; let A and B be two events. Event operations and possible relationships are summarized below.

- Event operations include:

  (1). Union: $A \cup B$; (2). Intersection: $A \cap B$; (3). Complement: A'.

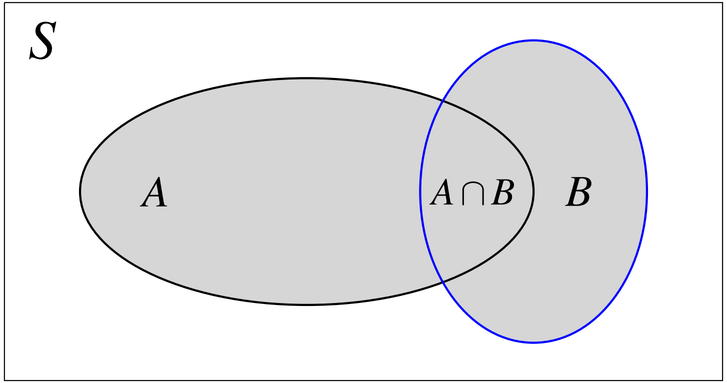
- Possible event relationships:

  (1). Contained: $A \subset B$; (2). Equivalent: $A = B$; (3) Mutually exclusive: $A \cap B = \emptyset$; (4) Independent: $A \perp B$ (postponed).

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

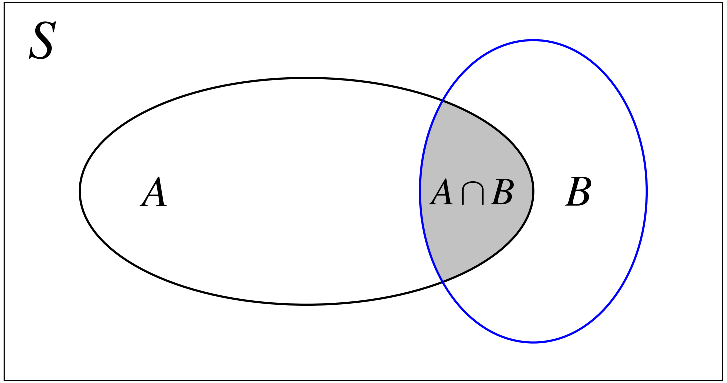
$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$



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

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



Union and **intersection** can be extended to n events: $A_1, A_2, \ldots, A_n$.

- **Union**:

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

  composed of elements that belong to one or more of $A_1, \ldots, A_n$.

- **Intersection**:

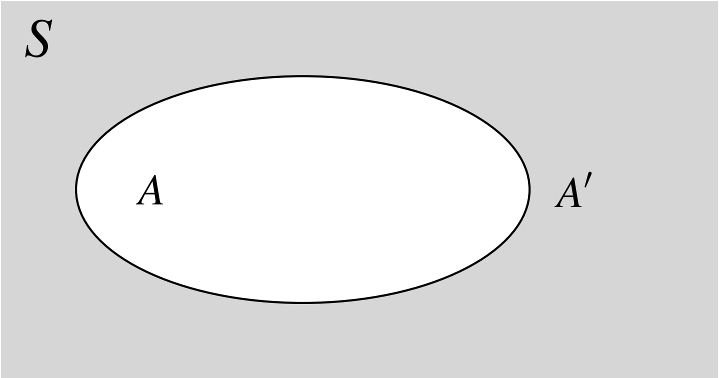
$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \ldots \cap A_n = \{x : x \in A_1 \text{ and } x \in A_2 \text{ and } \ldots \text{ and } x \in A_n\},$$

  composed of elements that belong every $A_1, \ldots, A_n$.

Complement

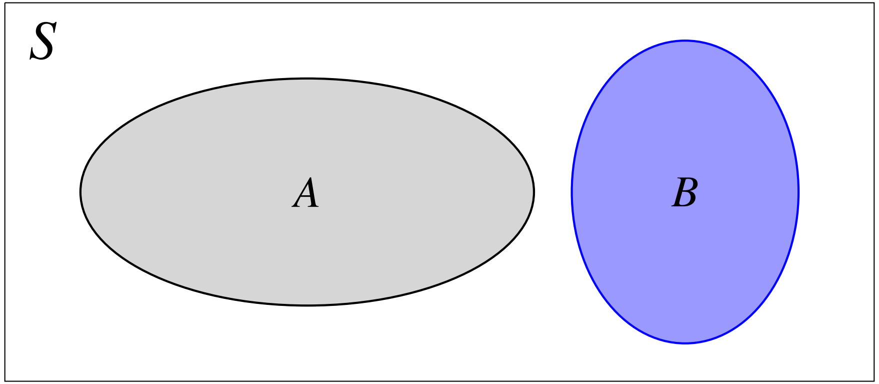
The **complement** of the event A (with respect to S), denoted by A', is the event with elements in S, which are not in A. That is

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



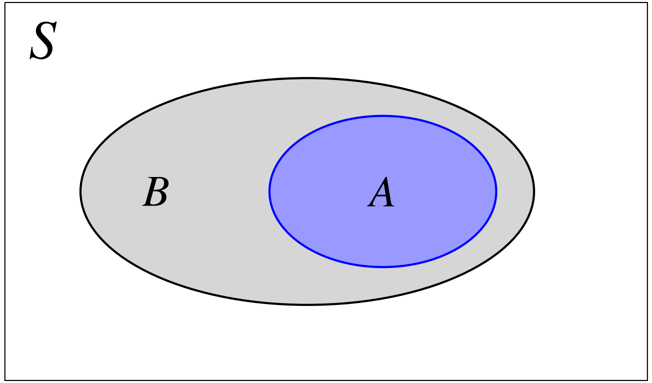
Mutually Exclusive

Events A and B are said to be mutually exclusive or disjoint, if $A \cap B = \emptyset$, i.e., 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 $A = B$, i.e., set A and B are **equivalent**.

Example 1.2 Consider the sample space and events: $S = \{1,2,3,4,5,6\}$, $A = \{1,2,3\}$, $B = \{1,3,5\}$, $C = \{2,4,6\}$. Then

- $A \cup B = \{1,2,3,5\}$; $A \cup C = \{1,2,3,4,6\}$; $B \cup C = S$.
- $A \cap B = \{1,3\}$; $A \cap C = \{2\}$; $B \cap C = \emptyset$.
- $A \cup B \cup C = S$; $A \cap B \cap C = \emptyset$.
- $(A \cap B) \cup C = \{1,3\} \cup \{2,4,6\} = \{1,2,3,4,6\}$.
- $A' = \{4,5,6\}$; $B' = \{2,4,6\} = C$.
- B and C are mutually exclusive, since $B \cap C = \emptyset$; A and B are not mutually exclusive since $A \cap B = \{1,3\} \neq \emptyset$.

Some Basic Properties of Event Operations

(1). $A \cap A' = \emptyset$ (2). $A \cap \emptyset = \emptyset$

(3). $A \cup A' = S$ (4). $(A')' = A$

(5). $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(6). $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(7). $A \cup B = A \cup (B \cap A')$

(8). $A = (A \cap B) \cup (A \cap B')$

De Morgan's Law

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

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

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

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

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

Example 1.3 Adopt the setting of Example 1.2: $S = \{1,2,3,4,5,6\}$, $A = \{1,2,3\}$, $B = \{1,3,5\}$, $C = \{2,4,6\}$. We have

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

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

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

- Similarly, we can check $(A \cup B \cup C)' = \emptyset = A' \cap B' \cap C'$; and $(A \cap B \cap C)' = S = A' \cup B' \cup C'$.

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 obtain some important counting methods: **permutation** and **combination**.

Multiplication Principle

Suppose that r different experiments are to be performed sequentially.

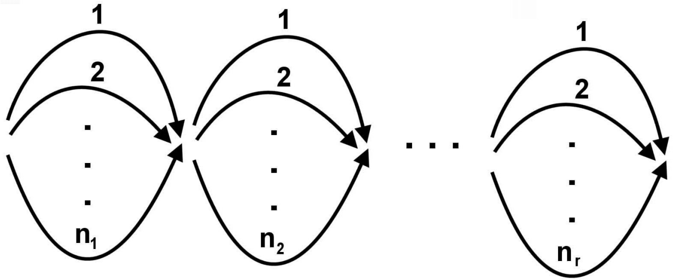
experiment 1 results in n_1 possible outcomes;

for each of the above result, experiment 2 results in n_2 possible outcomes;

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

for each of the above result, experiment r results in n_r possible outcomes.

Together there are $n_1 n_2 \cdots n_r$ possible outcomes of the r experiments.



Example 1.4 How many possible outcomes are there when a die and a coin are thrown together?

Solution:

- experiment 1: throwing a die; it has 6 possible outcomes: {1,2,3,4,5,6}.
- experiment 2: throwing a coin; for each outcome of experiment 1, it has 2 possible outcomes: $\{H, T\}$

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

In fact, the sample space is given by

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

Additional 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.

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

Procedure k can be carried out in n_k ways.

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

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

Example 1.5 We can take MRT or bus from home to Orchard road. If there are three bus routes and two MRT routes. How many ways we can go from home to Orchard road?

Solution: Consider that we go 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.

The ways are not overlapped. So the total number of ways that 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 ($r \leq n$) is denoted by P_r^n, which has the value:

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

ob1	ob2	ob3	$\cdots$	ob$r$
n ways	$(n-1)$ ways	$(n-2)$ ways	$\cdots$	$(n-(r-1))$ ways

- A special case: when $r = n$, $P_n^n = n!$. Practically, it is the number of ways to arrange n objects in order.

Example 1.6 Find the number of all possible four-letter code words in which all letters are different.

Solution: $n = 26$, $r = 4$. So the number of all 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

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

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

- The derivation is as follows:
  - Based on permutation, **the number of ways to choose and arrange r objects out n** is P_r^n.
  - On the other hand, this permutation task can be achieved by conducting two experiments sequentially:

    Exp. 1 get a combination, i.e., select r objects out n without regard to the order; there are $\binom{n}{r}$ ways.

    Exp. 2 for each combination, get a permutation on its r objects; there are P_r^r ways.

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

– As a consequent, $\binom{n}{r} \times P_r^r = P_r^n$, and 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 3 that can be formed with 2 women and 1 man.

<u>Solution</u>:

- 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, “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. We denote “$P(A)$".

- But..., how could we obtain such a number?

- Even more, as a fundamental concept, it has been extended from the intuition to more rigorous, abstract, and advanced mathematical theory and has wide applications in scientific disciplines.

Interpretation of Probability by Relative Frequency

- Suppose that we repeat an experiment E for 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 event A in the n repetition of E.
- Clearly, f_A may not equal to $P(A)$ exactly. But when n grows large, we may expect that f_A may "approach" it; in a sense $f_A \approx P(A)$. Or more mathematically,

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

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

  (1) $0 \leq f_A \leq 1$;

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

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

- Extending this, we 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:

(i) For any event A,
$$0 \leq P(A) \leq 1.$$

(ii) For the sample space,
$$P(S) = 1.$$

(iii) For any two mutually exclusive events A and B, i.e., $A \cap B = \emptyset$,
$$P(A \cup B) = P(A) + P(B).$$

Example 1.8
Let H denote the event of getting head when tossing a coin. Find $P(H)$, if

(a) the coin is fair;

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

Solution:

(a) The sample space is $S = \{H, T\}$.

- "Fair" means $P(H) = P(T)$.
- The events $\{H\}$ and $\{T\}$ are mutually exclusive.
- Based on Axioms 2 and 3, we have

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

which implies $P(H) = 1/2$.

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

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

which leads to $P(T) = 1/3$ and $P(H) = 2/3$.

Basic Properties of Probability

Using the axioms, we can derive the following propositions.

PROPOSITION 1
The probability of the empty set is $P(\emptyset) = 0$.

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

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

which implies $P(\emptyset) = 0$.

Proposition 2

If $A_1, A_2, \ldots, A_n$ *are mutually exclusive events* ($A_i \cap A_j = \emptyset$ *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 is immediately observed based on the induction.

PROPOSITION 3
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 4

For any two events A and B,

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

Proof Based on property (8) of event operations, i.e., $A = (A \cap B) \cup (A \cap B')$, and $(A \cap B) \cap (A \cap B') = \emptyset$, we have

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

PROPOSITION 5

For any two events A and B,

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

Proof Based on property (7) of event operations, i.e., $A \cup B = B \cup (A \cap B')$, and $B \cap (A \cap B') = \emptyset$; and proposition 4 that $P(A \cap B') = P(A) - P(A \cap B)$, we have

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

PROPOSITION 6

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

Proof Since $A \subset B$, we have $A \cup B = B$; based on property (7) of the event operations, i.e, $A \cup B = A \cup (B \cap A')$; and based on $A \cap (B \cap A') = A \cap B \cap A' = \emptyset$, we have

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

Example 1.9

- A retail establishment accepts either the American Express or the VISA credit card.
- A total of 24% of it customer 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 $A = \{\text{the customer carries an American Express Card}\}$; $V = \{\text{the customer carries an VISA Card}\}$.
- Then $P(A) = 0.24$; $P(V) = 0.61$; and $P(A \cap V) = 0.11$.
- The question is asking $P(A \cup V)$, and

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

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.,
$$P(a_1) = P(a_2) = \ldots = P(a_k).$$
- 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:

- Let $A = \{\text{the item is rusted}\}$, $B = \{\text{the item is a bolt}\}$,

  $S = \{\text{all the items}\}$.

- Since the item is selected at random, the elements in S are equally likely. S contains 200 elements. A contain $25 + 75 = 100$ elements, B contains 50, and $A \cap B$ contains 25.

- $P(A) = 100/200$, $P(B) = 50/200$, $P(A \cap B) = 25/200$;

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

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 7 (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.

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

(b) 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:

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

(a) The sample space is given by

$$\begin{array}{cc|cccccc} & & & & \text{2nd roll} & & & \\ & & 1 & 2 & 3 & 4 & 5 & 6 \\ & 1 & (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ & 2 & (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ \text{1st roll} & 3 & (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ & 4 & (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ & 5 & (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ & 6 & (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array}$$

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

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

$$\begin{array}{cccccccc} & & & & \text{2nd roll} & & & \\ & & 1 & 2 & 3 & 4 & 5 & 6 \\ \text{1st roll} & 5 & (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \end{array}$$

We are interested to look instances among 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)

- $P(B|A)$ can also be 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

Rearranging the definition of the conditional probability, we have

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

This together with the definition of the conditional probability, we have the inverse probability formula:

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

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

Independence is one of the most important concepts in probability.

Definition 8 (Independence)
*Two events A and B are **independent** if and only if*

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

We denote it as $A \perp B$.

If A and B are not independent, they are **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 is observed from the definition of conditional probability:

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

- Intuitively, this is stating: A and B 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.

(a) Let

$$\begin{aligned} A_6 &= \{\text{the sum of two dice is } 6\} \\ B &= \{\text{the first die equals } 4\}. \end{aligned}$$

Hence, $P(A_6) = {}^5\!/_{36}$, $P(B) = {}^6\!/_{36} = {}^1\!/_6$ and $P(A_6 \cap B) = {}^1\!/_{36}$. Since

$$P(A_6 \cap B) \neq P(A_6)P(B),$$

we say that A_6 and B are **dependent**.

(b) Let $A_7 = \{\text{the sum of two dice is } 7\}$. Then $P(A_7 \cap B) = 1/36$, $P(A_7) = 1/6$ and $P(B) = 1/6$. Hence

$$P(A_7 \cap B) = P(A_7)P(B),$$

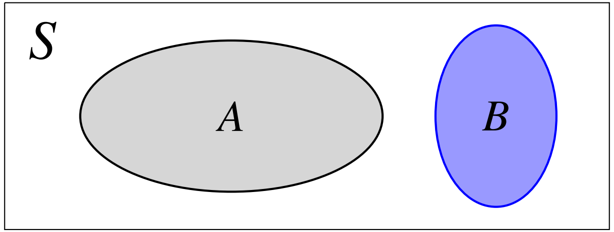
and we say that A_7 and B are **independent**.

Remark (Independent vs Mutually Exclusive)

Independent and **mutually exclusive** are totally different concepts:

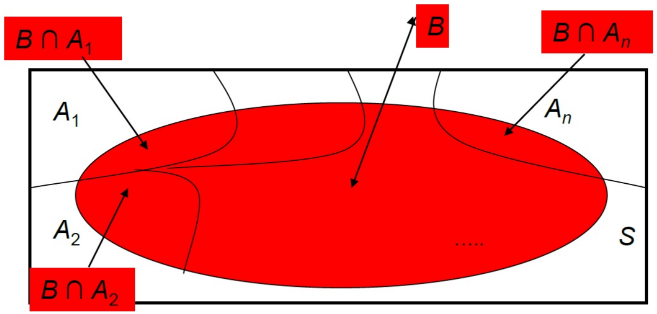
$$\begin{aligned} A, B \text{ independent } &\Leftrightarrow P(A \cap B) = P(A)P(B) \\ A, B \text{ mutually exclusive } &\Leftrightarrow A \cap B = \emptyset \end{aligned}$$

“Mutually exclusive" can be illustrated by the Venn diagram, but “independent" can not.



7 The Law of Total Probability

Definition 9 (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 10 (The law of total probability)

If $A_1, A_2, \ldots, A_n$ is a partition of S, then for any event B, we have

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

A special case: 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 $P(B)$.

- Use 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

Theorem 11 (Bayes' Theorem)
Let $A_1, A_2, \ldots, A_n$ be a partition of S, then for any event B and $k = 1, 2, \ldots, n$,

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

- A special case of Bayes' theorem: take $n = 2$. A and A' become a partition of S. We have

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

- This formula is practically meaningful. For example:
  - $A =$ disease status of a person;
  - $B =$ symptom observed;
  - $P(A)$: the probability of a disease in general;
  - $P(B|A)$: if diseased, probability of observing symptom;
  - $P(A|B)$: if symptom observed, probability of diseased.

- Bayes' theorem can be derived based on the conditional probability, multiplication rule, and the law of the total probability.
- In particular,

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

Example 1.15

- Historically we observe some newly constructed house to collapse.
- The chance that the design is faulty is 1%.
- If the design is faulty, the chance that the house is to collapse is 75%, otherwise, the chance is 0.01%.
- If we observed that a newly constructed house collapsed, what is the probability that the design is faulty?

Solution:

- Let

$$
\begin{aligned}
B &= \{\text{The design is faulty}\}, \\
A &= \{\text{The house collapses}\}.
\end{aligned}
$$

- We have $P(B) = 0.01$, $P(A|B) = 0.75$, and $P(A|B') = 0.0001$.
- The question is asking to compute $P(B|A)$.

- We compute it based on Bayes′ theorem. The denominator can be computed based on the law of total probability:

$$\begin{aligned} P(A) &= P(B)P(A|B) + P(B')P(A|B') \\ &= (0.01)(0.75) + (0.99)(0.00001) = 0.007599. \end{aligned}$$

- The numerator is

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

- As a consequence $P(B|A) = P(A|B)P(B)/P(A) = 0.9870.$