MTH113: Intro. to Probability & Statistics

Lesson 1 - Elementary probability

minerva

Dr. Yinna Ye

Department of Mathematics, XJTLU

Suzhou, PR China

Lesson 1: Elementary probability I

1. Sample space and events
2. Axioms of probability
3. Fundamental formula for probability
4. Some useful results on probability

1. Sample space and events

1.1 Introduction to sample space

Introduction to sample space (1)

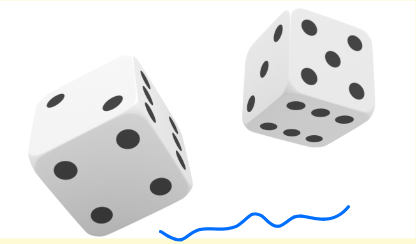
Definition (Experiment and outcome)

A process of observation and trial is called an **experiment**, and the result of this experiment is called **outcome**.

Example (Toss of 2 dice)

A simultaneous throw of 2 dice is an experiment. If we sum up the numbers of dots on the face of those two dice, we may have the possible outcomes: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

$|\Omega| = 6 \times 6 = 36$



$$|\Omega| = \left|\{1, 2, \dots, 6\} \times \{1, 2, \dots, 6\}\right|$$
$$= \left|\{(i, j);\ i, j \in \{1, \dots, 6\}\}\right|$$
$$= 6 \times 6 = 36$$

Introduction to sample space (2)

Definition (Sample space and event)

The set of all possible outcomes of an experiment is called a **sample space**, which is usually denoted by Ω. We will assume from now on that $\Omega \neq \emptyset$. Any subset E of the sample space is called an **event**.

Example (The sex of a newborn child)

$$\Omega = \{g, b\},$$

where the outcome g means that the child is a girl and b that it is a boy.
If $E = \{g\}$, then E is the event that the child is a girl.

Introduction to sample space (3)

Figure: Head and Tail of a UK coin

Introduction to sample space (4)

Example (Flipping two coins)

$|\Omega| = 2 \times 2 = 4$

$$\Omega = \{(H,H),(H,T),(T,H),(T,T)\},$$

where the outcomes

- (H,H) means both coins are **heads** (see the Figure on the last page),
- (H,T) means the first coin is **head** and the second is **tail**
- (T,H) means the first coin is **tail** and the second is **head**
- (T,T) means both coins are **tails**

If $E = \{(H,H),(H,T)\}$, then E is the event that a head appears on the first coin.

Introduction to sample space (5)

Example (Tossing 2 dice)

In the experiment of tossing 2 dice, the sample space

$$\Omega = \{(i,j) : i,j = 1,2,3,4,5,6\} = \{1,\cdots,6\} \times \{1,\cdots,6\}.$$

It consists of $6 \times 6 = 36$ outcomes, where (i,j) is said to occur if i appears on the leftmost die and j on the other die. If $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, then

$$E = \{\text{the sum of the dice equals 7}\}.$$

Introduction to sample space (6)

Example

Weight (in kg) of new-born babies: the sample space is the interval

$$\Omega = (0, 10].$$

uncountable set.

Ω is NOT a countable set.
If $E = \{x : 0 < x \leq 4\}$, then E is the event that the new-born baby is weighted no heavier than 4 kg.

1.2 Set operations

Set operations (1)

Definition (Set operations)

Let Ω be a given sample space and $E, F \subseteq \Omega$ are two events. Define

$$E \cup F := \{x \in \Omega : x \in E \text{ or } x \in F\}, \tag{1}$$

$$E \cap F := \{x \in \Omega : x \in E \text{ and } x \in F\}, \tag{2}$$

$$E^c := \{x \in \Omega : x \notin E\}; \tag{3}$$

where
the operation $\cup$ in (1) is called the **union** of E and F,
the operation $\cap$ in (2) is called the **intersection** of E and F,
the operation ^{c} in (3) is called the **complement** of E.

Set operations (2)

Remark

(1) $E \cup F$ stands for the event that either event E or event F occurs.

(2) $E \cap F$ (or EF) is the event that both E and F occur.

(3) E^c is the **negation** of E ('non E'), the event that E does not occur.

Example (The sex of a new born child)

Define the following events $E := \{g\},\ F := \{b\}$, then

$$E \cup F = \{g, b\} = \Omega.$$

Set operations (3)

Example (Flipping two coins)

- If $E := \{(H,H),(H,T)\}$ and $F := \{(T,H)\}$, then

$$\begin{aligned} E \cup F =& \{(H,H),(H,T),(T,H)\} \\ =& \{\text{an 'H' appear on either coin}\}. \end{aligned}$$

- If $E := \{(H,H),(H,T),(T,H)\}$,
$F := \{(H,T),(T,H),(T,T)\}$, then

$$E = \{\text{ least 1 head occurs}\},$$

$$F = \{\text{at least 1 tail occurs}\},$$

$$\begin{aligned} E \cap F =& \{(H,T),(T,H)\} \\ =& \{\text{exactly 1 'H' and 1 'T' occur}\}. \end{aligned}$$

Set operations (4)

Example (Tossing two dice)

If $E := \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ and $F := \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$, then $E \cap F = \emptyset$ and

$$E = \{\text{the sum of the dice is 7}\},$$

$$F = \{\text{the sum of the dice is 6}\},$$

$$E^c = \{\text{ the sum of the dice does not equal 7}\}.$$

Set operations (5)

Definition

(1) **Null event**, denoted by $\emptyset$, is the event consisting of no outcomes.

(2) If $E \cap F = \emptyset$, then E and F are said to be **mutually exclusive** or **disjoint events**.

(3) More generally, for a given finite collection of events $\{C_1, C_2, \cdots, C_n\}$, if for any $i \neq j$ and $i, j = 1, 2, \ldots, n$, $C_i \cap C_j = \emptyset$, then the events are said to be **pairwise mutually exclusive**.

Set operations (6)

Remark

(1) Note that because the experiment must result in some outcome (i.e. $\Omega \neq \emptyset$), it follows that

$$\Omega^c = \emptyset.$$

(2) For any event $A \subseteq \Omega$,

$$A \cap \Omega = A, \quad A \cup \Omega = \Omega$$
$$A \cap \emptyset = \emptyset, \quad A \cup \emptyset = A.$$

Set operations (7)

Definition (Subsets and equal events)

- For any two events E and F, if all of the outcomes in E are also in F, then we say that E is **contained** in F, or E is a **subset** of F, and write

$$\boxed{E \subseteq F.}$$

  Thus, if $E \subseteq F$, then the occurrence of E implies the occurrence of F.

- If $E \subseteq F$ and $F \subseteq E$, then we say that E and F are **equal** and write

$$E = F.$$

Set operations (8)

Definition (Unions and intersections of more than 2 events)

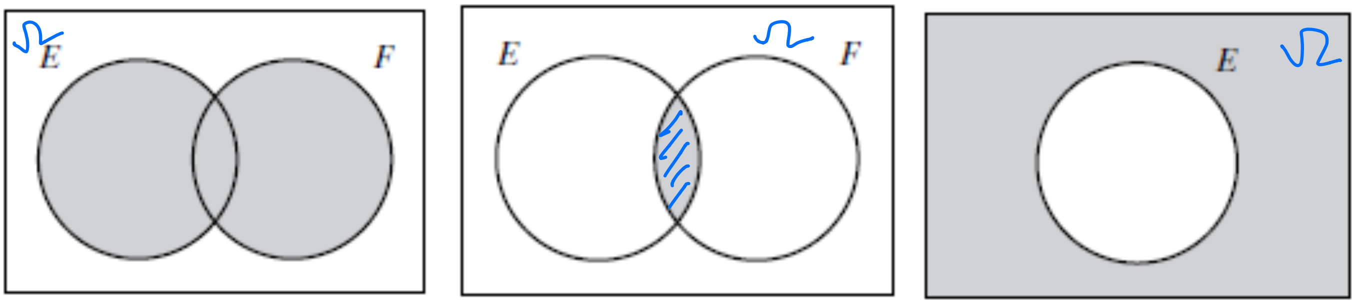
- If $E_1, E_2, \cdots$ are events, then the **union** of these events, denoted by $\bigcup_{n=1}^{\infty} E_n$, is defined to be that event which consists of all outcomes that are in E_n for at least one value of $n = 1, 2, \cdots$.
- Similarly, the **intersection** of the events E_n, denoted by $\bigcap_{n=1}^{\infty} E_n$, is defined to be the event consisting of those outcomes which are in all of the events E_n, $n = 1, 2, \cdots$.

$\bigcup_{n=1}^{\infty} E_n = E_1 \cup E_2 \cup E_3 \cup \cdots \cup \cdots$;

$\bigcap_{n=1} E_n = E_1 \cap E_2 \cap E_3 \cap \cdots \cap \cdots$,

Set operations (9)

The **Venn diagram** is a very useful graphical representation for illustrating logical relations among events.



(a) Shaded region: $E \cup F$

(b) Shaded region: EF

(c) Shaded region: E^c.

Set operations (10)

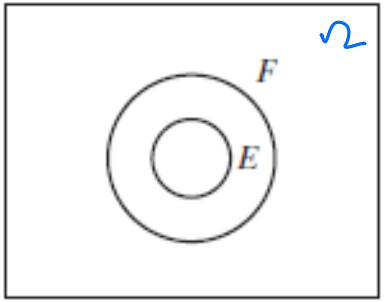


Figure: $E \subset F$

1.3 Properties of set operations

Properties of set operations (1)

Theorem (Properties of set operations)

Let E, F, G be any sets. Then we have the following properties of their unions and intersections:

Commutative laws $E \cup F = F \cup E, \qquad E \cap F = F \cap E$

Associative laws

$(E \cup F) \cup G = E \cup (F \cup G), \qquad (E \cap F) \cap G = E \cap (F \cap G)$

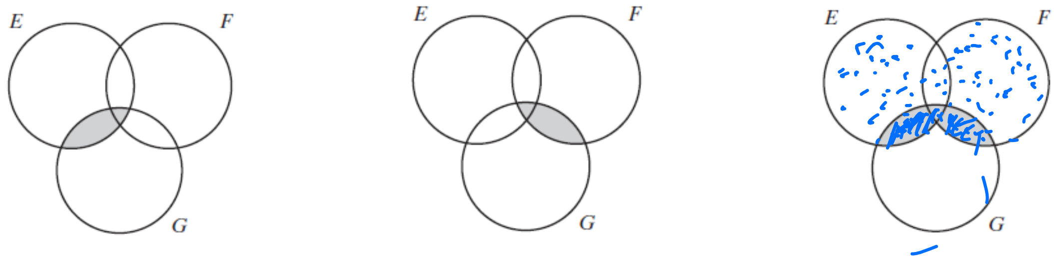
Distributive laws $(E \cup F) \cap G = (E \cap G) \cup (F \cap G),$

$(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$

Idempotence $E \cup E = E, \qquad E \cap E = E.$

$(a+b) \times c = a \times c + b \times c$, if $a, b, c \in \mathbb{R}$.

Properties of set operations (2)



(a) Shaded region: EG.

(b) Shaded region: FG.

(c) Shaded region: $(E \cup F)G$.

Figure: $(E \cup F)G = EG \cup FG$

More generally, the properties of set operations above are also true for finite number of sets.

Properties of set operations (3)

Theorem (DeMorgan's laws)

Let $E_1, E_2, \ldots, E_n$ be a collection of n events in the sample space Ω.

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c,$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c.$$

We need to prove

1) $\left(\bigcup_{i=1}^{n} E_i\right)^c \subseteq \bigcap_{i=1}^{n} E_i^c$

and at the same time

2) $\bigcap_{i=1}^{n} E_i^c \subseteq \left(\bigcup_{i=1}^{n} E_i\right)^c$.

Properties of set operations (4)

Proof: We only prove the first equality here, the method for proving the second one is similar.

Suppose that $x \in \left(\bigcup_{i=1}^{n} E_i\right)^c$. Then

$$x \notin \bigcup_{i=1}^{n} E_i \Longrightarrow x \notin E_i, \text{ for any } i = 1, 2, \ldots, n.$$
$$\Longrightarrow x \in E_i^c, \text{ for any } i = 1, 2, \ldots, n.$$
$$\Longrightarrow x \in \bigcap_{i=1}^{n} E_i^c.$$

So we have found $\left(\bigcup_{i=1}^{n} E_i\right)^c \subseteq \bigcap_{i=1}^{n} E_i^c$.

Properties of set operations (5)

Conversely, suppose now that $x \in \bigcap_{i=1}^{n} E_i^c$. Then

$$
\begin{aligned}
& x \in E_i^c, \text{ for any } i = 1, 2, \ldots, n \\
\Longrightarrow & x \notin E_i, \text{ for any } i = 1, 2, \ldots, n \\
\Longrightarrow & x \notin \bigcup_{i=1}^{n} E_i \\
\Longrightarrow & x \in \left(\bigcup_{i=1}^{n} E_i \right)^c .
\end{aligned}
$$

So we have $\bigcap_{i=1}^{n} E_i^c \subseteq \left(\bigcup_{i=1}^{n} E_i \right)^c$.

In conclusion, $\left(\bigcup_{i=1}^{n} E_i \right)^c = \bigcap_{i=1}^{n} E_i^c$.

Properties of set operations (6)

Exercise

Try to prove the second of DeMorgan's laws by taking the complements and the first law.

2. Axioms of probability

Axioms of probability (1)

$2^{\Omega} = \{$all the subsets of $\Omega\}$.

$\mathbb{P} : 2^{\Omega} \longrightarrow [0,1]$.

Definition

Consider an experiment whose sample space is Ω. The probability $\mathbb{P}$ is a function which assigns a set in 2^{Ω} to a value in $[0, 1]$. More precisely, for each event E $(\subseteq \Omega)$, the **probability of the event** E, denoted by $\mathbb{P}(E)$, is defined and satisfies the following three axioms:

Axiom 1 $0 \leq \mathbb{P}(E) \leq 1$;

Axiom 2 $\mathbb{P}(\Omega) = 1$ and Ω is the unique set in 2^{Ω} s.t. this equality holds;

Axiom 3 For any sequence of pairwise mutually exclusive events $E_1, E_2, \ldots$ (that is, events for which $E_i \cap E_j = \emptyset$ when $i \neq j$),

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

The Axioms 1-3 are called **Kolmogorov's axioms**.

Axioms of probability (2)

Remark

In words,

- Axiom 1 states that the probability that the outcome of the experiment is an outcome in E is some number between 0 and 1.
- Axiom 2 states that, with probability 1, the outcome will be a point in the sample space Ω.

  $\mathbb{P}(\Omega) = 1$
- Axiom 3 states that, for any sequence of pairwise mutually exclusive events, the probability of at least one of these events occurring is just the sum of their respective probabilities.

3. Fundamental formula for probability

Fundamental formula for probability (1)

As an example, let's see one of the most classical probabilities defined by the following 'fundamental formula'.

Theorem (Fundamental formula for probability)

Let Ω be a finite sample space, then for any event $A \subseteq \Omega$,

$$\mathbb{P}(A) = \boxed{\frac{|A|}{|\Omega|}}.$$

$|A| \to$ cardinality of the set A

$|A|$ = no. of members in A.

Proof: In order to prove this theorem, we need to check the Kolmogorov's axioms hold for the given $\mathbb{P}$.

(1) Since $\emptyset \subseteq A \subseteq \Omega$, we have

$$0 = |\emptyset| \leq |A| \leq |\Omega|.$$

From this, we obtain

$$0 = \frac{|\emptyset|}{|\Omega|} \leq \frac{|A|}{|\Omega|} = \mathbb{P}(A) \leq \frac{|\Omega|}{|\Omega|} = 1.$$

Fundamental formula for probability (2)

(2) Let $A = \Omega$, then the fundamental formula for probability gives

$$\mathbb{P}(\Omega) = \frac{|\Omega|}{|\Omega|} = 1.$$

And Ω is the only event with probability 1.

(3) Suppose that $E_1, E_2, \cdots$ is a sequence of pairwise mutually exclusive events in the sample space Ω. Since $E_i \cap E_j = \emptyset$, for any positive integers $i \neq j$,

$$\left| \bigcup_{i=1}^{\infty} E_i \right| = \sum_{i=1}^{\infty} |E_i|.$$

So the fundamental formula for probability gives

$= \mathbb{P}(E_i)$

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i \right) = \frac{|\bigcup_{i=1}^{\infty} E_i|}{|\Omega|} = \frac{\sum_{i=1}^{\infty} |E_i|}{|\Omega|} = \sum_{i=1}^{\infty} \boxed{\frac{|E_i|}{|\Omega|}} = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

By now we have shown that $\mathbb{P}$ defined in this theorem is a

Fundamental formula for probability (3)

Remark

- The Fundamental formula for probability can be interpreted as the probability of an event which is given by the ratio of the number of outcomes favorable to the total number of possible outcomes.
- In practice, we shall use this formula when each of the outcome in an experiment is equally probable to occur.

Example

If our experiment consists of tossing a coin and if we assume that a head is as likely to appear as a tail, then we could have

$$\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = 1/2.$$

From FFP, $\mathbb{P}(\{H\}) = \frac{|\{H\}|}{|\Omega|} = \frac{|\{H\}|}{|\{\{H\},\{T\}\}|} = \frac{1}{2}$

Fundamental formula for probability (4)

Example

If a die is rolled and we suppose that all six sides are equally likely to appear, i.e.

$$\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{1}{6}.$$

From Axiom 3, it would thus follow that the probability of rolling an even number would equal

$$\mathbb{P}(\{2,4,6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

because $\{2\}, \{4\}, \{6\}$ is a collection of m.e. events, i.e.

$$\{2\} \cap \{4\} = \emptyset = \{2\} \cap \{6\} = \{4\} \cap \{6\}.$$

Fundamental formula for probability (5)

Example

On a toss of two fair dice, we record the sum total of spots on the uppermost sides of the dice. What is the probability of having this number equal to 7?

Solution: We have seen before all the possible outcomes are as follows

$|\Omega| = 6 \times 6 = 36$

$$\Omega = \{(i,j) : i,j = 1,2,\cdots 6\}$$

$$= \begin{Bmatrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{Bmatrix}$$

Fundamental formula for probability (6)

This gives a total number of $|\Omega| = 6^2 = 36$ outcomes, each of which is equally likely to occur.
Denote by (ω_1, ω_2) an arbitrary outcome. The event, say A, we are interested is then represented by

$$\begin{aligned} A =& \{(\omega_1, \omega_2) \in \Omega : \omega_1 + \omega_2 = 7\} \\ =& \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}. \end{aligned}$$

Therefore, by Fundamental formula,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}.$$

4. Some useful results on probability

Some useful results on probability (1)

Theorem (Rule of complement)

If $A \subseteq \Omega$ is an event in sample space Ω, then

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A).$$

Proof.

Note that $A^c \cup A = \Omega$ and $A^c \cap A = \emptyset$. By using Kolmogorov's axioms, we have

$$1 \overset{\text{Ax.2}}{=} \mathbb{P}(\Omega) = \mathbb{P}(A^c \cup A) \overset{\text{Ax.3}}{=} \mathbb{P}(A^c) + \mathbb{P}(A),$$

which yields the desired result. □

Some useful results on probability (2)

Theorem (Rule of total probability)

If $A, B \subseteq \Omega$ are events in the sample space Ω, then

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c). \tag{4}$$

Proof.

By using distributive law we can rewrite A as

$$A = A \cap \Omega = A \cap (B \cup B^c) \overset{\text{distr.}}{=} (A \cap B) \cup (A \cap B^c). \tag{5}$$

This is a disjoint union (why? Exercise). Then by Kolmogorov's axiom 3, we have

$$\mathbb{P}(A) = \mathbb{P}((A \cap B) \cup (A \cap B^c)) \overset{\text{Ax.3}}{=} \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c).$$

□

Some useful results on probability (3)

Remark (Generalization of the rule of total probability)

More generally, if a finite collection of pairwise mutually exclusive events $\{C_1, C_2, \cdots, C_n\}$ satisfy

$$\boxed{C_1 \cup C_2 \cup \cdots \cup C_n = \Omega}$$

(we will see this kind of events is called a **partition** of Ω in the next chapter), then for any event $A \subseteq \Omega$,

$$\mathbb{P}(A) = \mathbb{P}(A \cap C_1) + \mathbb{P}(A \cap C_2) + \cdots + \mathbb{P}(A \cap C_n).$$

(As exercise, prove this generalization.)

Some useful results on probability (4)

Theorem (Monotonicity)

If $A, B \subseteq \Omega$ are events in the sample space Ω and $\boxed{A \subseteq B}$, then

$$\boxed{\mathbb{P}(A) \leq \mathbb{P}(B).}$$

Proof.

By exchanging A and B in (4), we have

$$\mathbb{P}(B) = \mathbb{P}(\boxed{A \cap B}) + \mathbb{P}(B \cap A^c).$$

Since $A \subseteq B$, $A \cap B = A$. So

$$\mathbb{P}(B) = \mathbb{P}(A) + \boxed{\underbrace{\mathbb{P}(B \cap A^c)}_{\geq 0 \text{ by Ax.1}}},$$

that is $\mathbb{P}(B) - \mathbb{P}(A) = \mathbb{P}(B \cap A^c) \geq 0$. $\square$

Some useful results on probability (5)

Theorem (Rule of addition)

If $A, B \subseteq \Omega$ *are events in the sample space* Ω, *then*

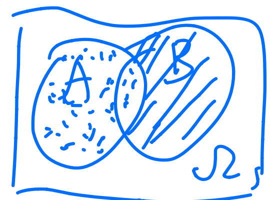
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Some useful results on probability (6)

Proof.

Note that $A \cup B$ can be written as the union of two disjoint events A and $B \cap A^c$, that is

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



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

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

$= \phi \cap B = \phi$

So by Kolmogrov's axiom 3, we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (B \cap A^c)) \overset{\text{Ax.3}}{=} \mathbb{P}(A) + \mathbb{P}(B \cap A^c). \quad (6)$$

Using rule of total probability (4), we have

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Substituting $\mathbb{P}(B \cap A^c)$ into (6), we thus obtain the desired result. □

Some useful results on probability (7)

Exercise

Try to use the Venn diagram to prove the Rule of addition.

The end of Lesson 2