4

Probability

This chapter will show you how to
- construct sample space diagrams when considering compound outcomes
- draw Venn and tree diagrams to help work out probabilities in more complex contexts
- work with conditional probabilities.

Before you start

You should know how to:

1 Work with fractions.
   e.g. Calculate
   \[
   \begin{align*}
   a &= \frac{7}{12} + \frac{1}{3} \\
   b &= \frac{1}{6} \times \frac{1}{2}
   \end{align*}
   \]
   \[
   \begin{align*}
   a &= \frac{7}{12} + \frac{1}{3} = \frac{7 + 4}{12} = \frac{11}{12} \\
   b &= \frac{1}{6} \times \frac{1}{2} = \frac{1 \times 1}{6 \times 2} = \frac{1}{12}
   \end{align*}
   \]

2 Identify the basic outcomes of a simple experiment.
   e.g. How many different pairs of letters can be made from the word 'DICE'?
   DI, DC, DE, IC, IE, CE

Check in

1 Calculate
   \[
   \begin{align*}
   a &= \frac{1}{4} \times \frac{2}{3} \\
   b &= \frac{1}{4} + \frac{2}{3}
   \end{align*}
   \]

2 List the possible outcomes when a coin is tossed twice.
A probability experiment has outcomes which occur at random.

Imagine an experiment where you roll a die 500 times, and you see 74 ‘fives’. The relative frequency or experimental probability of throwing a five in the experiment is

\[
\frac{74}{500} = 0.148
\]

However, if the die is fair the theoretical probability of throwing a five is

\[
\frac{1}{6} = 0.166\ldots
\]

Because the outcomes are random, the number of times any particular score appears in an experiment will vary considerably.

The more times you repeat an experiment, the closer the estimated probability is likely to be to the theoretical probability.

An event is a set of possible outcomes from an experiment. So if you throw a die some events could be:

A is the event that a five appears.
B is the event that an even number appears.
C is the event that an odd number appears.

\[A \cup B\] is the union of events A and B. This means A or B or both can happen. 2, 4, 5, 6 are the outcomes which satisfy \(A \cup B\).

\[A \cap C\] is the intersection of events A and C. This means both A and C have to happen. Only the outcome 5 satisfies \(A \cap C\).

\(A'\) means the event ‘A does not happen’. This is the complementary event, and \(P(A') = 1 - P(A)\). 1, 2, 3, 4, 6 are the outcomes which satisfy \(A'\).
Exercise 4.1

1. A letter is chosen at random from the word CAMBRIDGE.
   The events $A$, $B$, $C$ and $D$ are defined as:
   
   $A$: A vowel is chosen.
   $B$: The letter B is chosen.
   $C$: A letter in the first half of the alphabet is chosen.
   $D$: A letter is chosen which has only one letter beside it.

   a. Describe the event $A'$ in words.
   b. For each event $A$, $B$, $C$ and $D$ write down the outcomes which satisfy it.
   c. Give the probability of each event $A$, $B$, $C$, $D$.
   d. List the outcomes which satisfy $A \cap C$.
   e. Write down $P(A \cap C)$.
   f. Find $P(A \cap D)$.
   g. Find $P(A \cup B)$.

Activity

2. a. Throw a coin 20 times and count the number of times it shows a head.
   b. Throw the coin another 20 times and count the number of times it shows a head.
   c. How many times would you ‘expect’ to see a head in 20 throws?
   d. Does this happen in both sets of 20 coin throws?
   e. If you have access to a number of other people's results as well, how often do you see the ‘expected number’ of heads?
   f. How many heads (out of 20 throws) are most commonly seen?

3. a. Throw a die 30 times and count the number of times it shows a five.
   b. How many times would you ‘expect’ to see a five in 30 throws?
   c. Throw the die another 20 times and count the number of times it shows a five.
   d. How many times would you ‘expect’ to see a five in 20 throws?
   e. If you have access to a number of other people's results as well, how often do you see the ‘expected number’ of fives: i) in 30 throws? ii) in 20 throws?
   f. How many fives (out of 20 throws) are most commonly seen?
4.2 Two events

In contexts where two things happen it is often helpful to construct a table showing the possible outcomes. This is sometimes called a possibility space diagram or a sample space diagram.

Two dice are thrown, and the sum of the scores on the two dice is taken. You can represent this in a two-way table:

<table>
<thead>
<tr>
<th>Sum</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

You can use the table to work out the probability of getting a sum of 5.

\[ P(\text{Sum} = 5) = \frac{4}{36} \]

Two dice are thrown, and this time the higher of the scores on the two dice is taken:

<table>
<thead>
<tr>
<th>High</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

You can use this table to work out the probability of the higher score being 5.

\[ P(\text{High} = 5) = \frac{9}{36} \]
Exercise 4.2

1. The numbers 1 to 6 are on cards. Two cards are taken at random.

   Copy and complete the sample space diagram to show the sum of the numbers on the cards.

   Find the probability that the total score is:
   a  5  
   b  4  
   c  2

   

<table>
<thead>
<tr>
<th>Sum</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

2. A coin is tossed and a die is thrown.
   a  List all the possible outcomes in the sample space.

   A head scores 1 and a tail scores 2.

   b  Construct a sample space diagram to show the total score for this experiment.

3. Two dice are thrown.
   a  Construct a sample space diagram to show the product of the scores on the two dice.

   b  Find the probability that the product of scores is:

   i  3  ii  5  iii  6  iv  10

4. Two dice are thrown.
   a  Construct a sample space diagram to show the lower of the scores on the two dice.

   b  Find the probability that the lower score is:

   i  3  ii  5  iii  6

5. Two dice are thrown.
   a  Copy and complete the sample space diagram to show the difference between the scores on the two dice (the unsigned difference).

   b  Find the probability that the difference between the scores is:

   i  3  ii  5  iii  6

In questions 2–5 assume that the dice are fair and six-sided.
A Venn diagram can be a helpful way to visualise the relationship between events.

The diagram shows each event as a circle within a rectangle.

A school has 97 students in Year 12; 55 students take AS Maths and 32 take AS Chemistry; 31 students take neither AS Maths nor AS Chemistry. How many students take both?

The diagram shows $a$, $b$, $c$ and $d$ as the numbers of students in each of those regions.

Given:

- $a + b + c + d = 97$
- $a + b = 55$
- $b + c = 32$
- $d = 31$

$(a + b) + (b + c) + d = 55 + 32 + 31 = 118 \ [= a + 2b + c + d]$

$a + b + c + d = 97$, so $b$ must be 21. Then $a = 24$ and $c = 11$.

21 students take both AS Maths and AS Chemistry.

The events in the previous example can be represented in terms of probabilities. If you do so, you can see an interesting result.

$P(M) = a + b$

$P(C) = b + c$

$P(M \cap C) = b$

$P(M \cup C) = a + b + c$
This leads to a general result for two events:

**Addition rule:** \( P(X \cup Y) = P(X) + P(Y) - P(X \cap Y) \)

In the special case where \( P(X \cap Y) = 0 \) (\( X \) and \( Y \) are **mutually exclusive**), the probability of \( X \) or \( Y \) happening is just the sum of the probabilities of \( X \) and of \( Y \).

For two events \( A \) and \( B \),
\[ P(A) = 0.7, \quad P(B) = 0.4, \quad P(A \cap B) = 0.3 \]

Find

i. \( P(A \cup B) \)

ii. \( P(A' \cap B') \)

iii. \( P(A' \cap B) \)

iv. \( P(A' \cup B) \)

Draw a Venn diagram and label the four regions:

\[(A \cap B) = 0.3 \text{ so } x = 0.4 \text{ from } P(A) \text{ and } y = 0.1 \text{ from } P(B).\]
\[0.3 + 0.4 + 0.1 = 0.8\]
\[\text{so } z = 0.2\]
You sometimes need to extract the necessary information from the context described.

Two-thirds of the pupils in a class have a mobile phone. Half of the pupils have an MP3 player. A quarter of the pupils have neither. What proportion of the class have both?

\[ (a + b) + (b + c) + d = \frac{2}{3} + \frac{1}{2} + \frac{1}{4} = \frac{17}{12} \]

so \( b = P(A \cap B) = \frac{5}{12} \)

so \( \frac{5}{12} \) of the class have both.
Three events can be shown as three intersecting circles, giving eight possible regions.

**Exercise 4.3**

1. For each of the following, draw a copy of the diagram below and shade in the area representing the set.

   a. \( A \cap B \)  
   b. \( A \cup B \)  
   c. \( A' \cap B \)  
   d. \( A \cup B' \)  
   e. \( A' \cap B' \)
2. For each of the following, draw a copy of the diagram below and shade in the area representing the set.

\[\begin{array}{c}
A \cap C \\
\hline
B \\
\hline
C
\end{array}\]

a. \(A \cap C\) 

b. \((A \cup B) \cap C\) 

c. \(A' \cap B \cap C\) 

d. \(A \cup B' \cup C\)

3. There are 80 boys in Year 10; 26 boys played for the rugby team and 17 played for the cricket team. If 12 boys played for both teams, how many played for neither?

4. Once cars are three years old they have to have an annual roadworthiness test called the MOT. 86% of cars at a centre pass the MOT. 9% of the cars are found to have faulty brakes, and 11% have a fault not related to brakes, which means they fail.

   a. What proportion of cars fail only on their brakes?

   b. What proportion fail, but had good brakes?

5. \(P(A) = 0.6, \ P(B) = 0.5, \ P(A \cup B) = 0.8\)

   Calculate

   a. \(P(A \cap B)\) 

   b. \(P(A' \cap B')\) 

   c. \(P(A' \cap B)\)

6. \(P(A) = 0.7, \ P(A \cap B) = 0.5, \ P(A \cup B) = 0.8\)

   Calculate

   a. \(P(B)\) 

   b. \(P(A' \cap B')\) 

   c. \(P(A' \cap B)\)

7. \(P(A) = \frac{2}{3}, \ P(B) = \frac{1}{4}, \ P(A \cup B) = \frac{3}{4}\)

   Calculate

   a. \(P(A \cap B)\) 

   b. \(P(A \cap B')\) 

   c. \(P(A' \cap B)\)

8. \(P(A) = \frac{1}{2}, \ P(A \cap B) = \frac{1}{3}, \ P(A \cup B) = \frac{3}{4}\)

   Calculate

   a. \(P(B)\) 

   b. \(P(A \cap B')\) 

   c. \(P(A' \cap B)\)
9 \[ P(A') = 0.6, \quad P(A' \cap B') = 0.4, \quad P(A \cap B') = 0.1 \]

Calculate
\[ \text{a} \quad P(B) \quad \quad \text{b} \quad P(A \cap B) \quad \quad \text{c} \quad P(A \cup B) \]

10 A survey of a primary school class found that one-third of the pupils had a pet dog, a quarter had a pet hamster and one-sixth of the class had both. A child is chosen at random from the class. What is the probability that the child has:
\[ \text{a} \quad \text{at least one of a pet dog or hamster?} \]
\[ \text{b} \quad \text{a pet dog but not a hamster?} \]

11 \[ P(A) = 0.5, \quad P(B) = 0.5, \quad P(C) = 0.2 \]
\[ P(A \cap B) = 0.2, \quad P(B \cap C) = 0, \quad P(A' \cap B' \cap C') = 0.1 \]

Calculate
\[ \text{a} \quad P(A \cap C) \]
\[ \text{b} \quad P(A \cap B' \cap C') \]

12 A Year 12 group has 112 pupils; 42 take French, 65 take Maths and 32 take Physics. Everyone who takes Physics also takes Maths, but no one takes all three subjects. 12 pupils take Maths and French. How many pupils in the group take none of Maths, French and Physics?

13 A lecturer does an informal survey of his first-year undergraduate politics students. Half the students claim to read the *Daily Telegraph*, half claim to read *The Times* and half claim to read the *Guardian*; \( \frac{1}{12} \) say they read none of these.

No one claims to read all three papers; \( \frac{1}{6} \) claim to read both the *Daily Telegraph* and *The Times*, and \( \frac{1}{6} \) claim to read both the *Daily Telegraph* and the *Guardian*.

\[ \text{a} \quad \text{How many claim to read both *The Times* and the *Guardian*?} \]
\[ \text{b} \quad \text{How many claim to read only the *Guardian*?} \]
\[ \text{c} \quad \text{How many claim to read only the *Daily Telegraph*?} \]
### Sampling without replacement
Consider a bag which has three red and five blue beads in it. If you take a bead out at random, and then take out another without replacing the first, you can represent the possible outcomes in a **probability tree** diagram.

You can put probabilities on the branches to complete the diagram.

You can now scan along the branches to identify any possible outcome. For example,
\[
P(\text{both beads the same colour}) = P(\text{both red}) + P(\text{both blue})
\]
\[
= \left( \frac{3}{8} \times \frac{2}{7} \right) + \left( \frac{5}{8} \times \frac{4}{7} \right)
\]
\[
= \frac{6}{56} + \frac{20}{56}
\]
\[
= \frac{26}{56}
\]
\[
= \frac{13}{28}
\]
Sampling with replacement

If the first bead was returned to the bag before the second one was selected, the probabilities of red or blue would still be $\frac{3}{8}$ and $\frac{5}{8}$ at the second stage:

Now, $P(\text{both beads the same colour}) = P(\text{both red}) + P(\text{both blue})$

\[
\begin{align*}
P(\text{both red}) & = \frac{3}{64} + \frac{25}{64} = \frac{28}{64} = \frac{7}{16} \\
P(\text{both blue}) & = \frac{15}{64} + \frac{15}{64} = \frac{30}{64} = \frac{15}{32}
\end{align*}
\]

A disease is known to affect 1 in 10 000 people. It can be fatal, but it is treatable if it is detected early.

A screening test for the disease shows a positive result for 99% of people with the disease.

The test shows positive for 1% of people who do not have the disease.

For a population of 1 million people:

a how many would you expect to have the disease and test positive?

b how many would you expect to test positive?

First draw a tree diagram:

So there are likely to be about 99 positive tests from people with the disease, and about $99 + 9999 = 10098$ positive tests altogether.
More complex tree diagrams

There are four red, three green and five blue discs in a bag. Find the probability that two discs drawn without replacement are the same colour.

First draw a tree diagram – there are three outcomes at each stage.

\[
P(\text{same colour}) = \frac{12}{132} + \frac{6}{132} + \frac{20}{132} = \frac{38}{132} = \frac{19}{66}
\]

You can use a tree diagram with more than two stages. However, in practice the diagram becomes difficult to work with, and you will not be required to deal with difficult cases.
EXAMPLE 3

A bag has three red beads and five blue beads.
A bead is taken from the bag, its colour is noted, and it is
returned to the bag. This is done three times. Find the
probability that all three beads are the same colour.

Find draw a tree diagram.

From the tree diagram,
\[ P(3 \text{ beads the same colour}) = \frac{27}{512} + \frac{125}{512} \]
\[ = \frac{152}{512} \]
\[ = \frac{19}{64} \]
Exercise 4.4

1. A bag contains five blue and three green balls. A ball is chosen at random, its colour noted, and the ball returned to the bag. A second ball is chosen.
   a. Find the probability that the two balls are different colours.
   b. If the first ball is not returned to the bag before the second ball is chosen, what is the probability the balls are different colours?

2. At a gym, 60% of the members are men. One-third of the men use the gym at least once a week. Three-quarters of the women use the gym at least once a week. A member is chosen at random. Find the probability that
   a. it is a man who does not use the gym at least once a week
   b. it is a person who uses the gym at least once a week.

3. In a certain town, the probability that a person’s car is stolen during a year is 0.06. The probability that a person is assaulted is 0.03. Assuming these events are independent, draw a tree diagram to represent this information. Find the probability that a randomly selected person in the town
   a. is not the victim of either of these crimes during the year
   b. is the victim of exactly one of these crimes during the year
   c. is the victim of both of these crimes during the year

4. A coin is thrown three times. Find the probability that
   a. it shows heads on all three throws
   b. it shows the same face on all three throws
   c. it does not land the same way on two successive throws.
5 Bag A contains five blue and three green balls. A ball is chosen at random, the colour is noted and it is **not** returned to the bag. A second ball is chosen.

   a  Find the probability that the two balls are the same colour.

Bag B contains 50 blue and 30 green balls. Again, a ball is chosen at random, the colour is noted and it is **not** returned to the bag before a second ball is chosen.

   b  Find the probability that the two balls are the same colour.

6 A bag contains four blue, four red and four green balls. Two balls are removed at random, one at a time, and without replacement. Find the probability that:

   a  the second ball drawn is a red

   b  both balls are blue

   c  neither ball is green

   d  at least one ball is green.

7 A bag contains ten counters: four white, three green and three red. Counters are removed at random, one at a time, and without replacement. Find the probability that:

   a  the first counter is red

   b  the first three counters are all white

   c  the first three counters are all different colours.
In the previous section you considered a screening test for a rare disease. Even though the test was remarkably accurate, less than one in 100 of the positive results come from someone with the disease.

Of 10,098 (99 + 9999) positive results 9999 were from healthy people, so the conditional probability that somebody is healthy given that they have a positive result is

\[ P(\text{Healthy} \mid \text{positive test result}) = \frac{9999}{10098} = 0.9902 \]

The conditional probability of an event \( A \) occurring given that an event \( B \) has already occurred can be written as \( P(A \mid B) \).

A GP practice encourages elderly people to have a flu vaccination each year. The doctors say that the vaccination reduces the likelihood of having flu from 40% to 10%. If 45% of the elderly people in the practice have the vaccination, find the probability that an elderly person chosen at random from the practice:

a. gets flu
b. had the vaccination, given that they get flu.

\[
\begin{align*}
\text{Flu} & \quad 0.045 \\
\text{No flu} & \quad 0.955
\end{align*}
\]

\[
\begin{align*}
\text{Vaccination} & \quad 0.45 \\
\text{No vaccination} & \quad 0.55
\end{align*}
\]

\[
\begin{align*}
\text{Flu} & \quad 0.1 \\
\text{No flu} & \quad 0.9
\end{align*}
\]

\[
\begin{align*}
\text{Flu} & \quad 0.4 \\
\text{No flu} & \quad 0.6
\end{align*}
\]

\[
\begin{align*}
\text{Flu} & \quad 0.22 \\
\text{No flu} & \quad 0.33
\end{align*}
\]

\[
\begin{align*}
\text{Flu} & \quad 0.045 + 0.22 = 0.265 \\
\text{No flu} & \quad 0.955
\end{align*}
\]

\[
\begin{align*}
P(V \mid F) = \frac{P(V \cap F)}{P(F)} = \frac{0.045}{0.265} = 0.170 \quad (3 \text{ s.f.})
\end{align*}
\]

In situations like this it is important that the patient is not told that they have the disease on the basis of a positive result from a screening test.
Venn diagrams can be useful when looking at conditional probability. Consider two events $A$ and $B$, with the probabilities shown in the Venn diagram.

$P(A \mid B)$ is the probability of $A$ given that $B$ has occurred.

So $P(A \mid B) = \frac{0.3}{0.7} = \frac{3}{7}$

Generally, for two events $A$ and $B$:

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

Note that generally $P(A \mid B) \neq P(B \mid A)$.

In the diagram above:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{0.3}{0.5} = \frac{3}{5}$$

**Example 2**

$P(A) = 0.75, \ P(B) = 0.35, \ P(A \cup B) = 0.9$

Find $a \ P(A \cap B) \quad b \ P(A \mid B) \quad c \ P(B \mid A')$

---

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

so $P(A \cap B) = 0.75 + 0.35 - 0.9 = 0.2$

then

$b \ P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.35} = \frac{4}{7}$

$c \ P(B \mid A') = \frac{P(B \cap A')}{P(A')} = \frac{0.2}{0.25} = \frac{4}{5}$
Exercise 4.5

1. 95% of drivers wear seat belts. 60% of car drivers involved in serious accidents die if they are not wearing a seat belt, whereas 80% of those who do wear a seat belt survive.
   a. Draw a tree diagram to show this information.
   b. What is the probability that a driver in a serious accident did not wear a seat belt and survived?

2. At an electrical retailer’s, one-third of the light bulbs are from company X, and the rest from Company Y. A report shows that 3% of light bulbs from Company X are faulty, and that 2% from Company Y are faulty.
   a. If the retailer chooses a bulb at random from stock and tests it, what is the probability that it is faulty?
   b. If the bulb is faulty, what is the probability that it came from Company Y?

3. An insurance company classifies drivers in three categories. X is ‘low risk’, and they represent 20% of drivers who are insured. Y is ‘moderate risk’ and they represent 70% of the drivers. Z is ‘high risk’.
   The probability that a category X driver has one or more accidents in a 12-month period is 2%, and the corresponding probabilities for Y and Z are 5% and 9%.
   a. Find the probability that a motorist, chosen at random, is assessed as a category Y risk and has one or more accidents in the year.
   b. Find the probability that a motorist, chosen at random, has one or more accidents in the year.
   c. If a customer has an accident in a 12-month period, what is the probability that the driver was in category Y?

4. \(P(A) = 0.6, \ P(B) = 0.5, \ P(A \cap B) = 0.2\)
   a. Draw a Venn diagram to represent this information.
   b. Find \(\text{i} \ P(A|B)\) \(\text{ii} \ P(B|A)\)

5. \(P(X) = 0.6, \ P(X \cap Y) = 0.3, \ P(X \cup Y) = 0.8\)
   Find \(\text{a} \ P(Y)\) \(\text{b} \ P(Y|X)\) \(\text{c} \ P(X|Y)\) \(\text{d} \ P(X|Y')\)
6 In a school there are 542 pupils, 282 of whom are girls. 364 pupils walk to school, of whom 153 are girls. Find the probability that a pupil chosen at random:

a is a boy
b is a boy who does not walk to school
c does not walk to school given that the pupil is a boy
d is a girl, given that the pupil walks to school.

7 There are 173 pupils in Year 13 in a school. There are 25 prefects in Year 13, of whom 7 are House Captains.

Find the probability that a Year 13 pupil chosen at random:

a is a prefect
b is a House Captain, given that the pupil is a prefect.

8 Of the employees in a large factory one-sixth travel to work by bus, one-third by train, and the rest by car. Those travelling by bus have a probability of \( \frac{1}{4} \) of being late, those by train will be late with probability \( \frac{1}{5} \) and those by car will be late with probability \( \frac{1}{10} \).

Draw and complete a tree diagram and calculate the probability that an employee chosen at random will be late.

9 The homework diaries and homework of two pupils are examined. There is a probability of 0.4 that A does not write in the given homework correctly. She always does the homework if she writes it in, but never checks if it is not written in.

There is a probability of 0.8 that B writes in the homework correctly, and when she does she will do it 90% of the time; if she has nothing written in then she checks with a friend who knows the homework 50% of the time.

She does the homework if she is given it by the friend.

Assume that A and B act independently.

Draw a tree diagram representing this information.

a Find the probability that pupil A does her homework on a particular night.

b Find the probability that both pupils do their homework on a particular night.

c If a homework was checked and was not done, find the probability that it was pupil A.
4.6 Relationships between events

Independence

Two events $A$ and $B$ are **independent** if the outcome of $A$ does not affect the outcome of $B$, and vice versa.

So $P(A \mid B) = P(A) \iff A$ and $B$ are independent.

The probability of $A$ occurring given that $B$ has already occurred will just be the probability of $A$.

The general relationship is:

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

A set of 40 cards shows a number from 1 to 10 of one of four geometrical symbols.
Circles and squares are shown in grey, rectangles and ellipses are blue.
The pack of cards is shuffled and the top card is turned over.
Let $C$ be the event ‘card shows a circle’, $F$ be the event ‘card shows 5 symbols’ and $G$ be the event ‘the card is grey’.

**a** Show that $C$ and $F$ are independent events.

**b** Show that $C$ and $G$ are not independent events.

**a** $P(C) = \frac{10}{40}$ (there are 10 circle cards)
$P(F) = \frac{4}{10}$ (there are 4 cards showing 5 symbols)
$P(C \cap F) = \frac{1}{40}$ (there is only one card with 5 circles)
$P(C\mid F) = \frac{P(C \cap F)}{P(F)} = \frac{\frac{1}{40}}{\frac{4}{40}} = \frac{1}{4} = P(C)$

Since $P(C\mid F) = P(C)$, $C$ and $F$ are independent events.

**b** $P(G\mid C) = 1$, since if you know the card has circles then you know it is grey
$P(G) = 0.5$
So $P(G\mid C) \neq P(G)$, and the events $G$ and $C$ are not independent.

You can test for independence by comparing $P(X \mid Y)$ with $P(X)$, or equivalently you compare $P(X \cap Y)$ with $P(X) \times P(Y)$.

Mutually exclusive events

Two events $A$ and $B$ are **mutually exclusive** if they cannot occur at the same time.

For mutually exclusive events, $P(A \cup B) = P(A) + P(B)$.
Exhaustive events

A set of events is **exhaustive** if it covers all possible outcomes.

Two fair dice are thrown. From the events which are listed below, give two which are:

a  mutually exclusive
b  exhaustive
c  not independent.

A: The two dice show the same number.
B: The sum of the two scores is at least 5.
C: At least one of the two numbers is a 5 or a 6.
D: The sum of the two scores is odd.
E: The largest number shown is a 6.
F: The sum of the two scores is less than 8.

a  A and D are mutually exclusive because if the two dice show the same number, the sum has to be an even number.
b  B and F are exhaustive because the only outcomes not in B are that the sum is 2, 3 or 4 and these are all in F.
c  A and D are not independent (because they are mutually exclusive)
   C and E are not independent since \( P(C \mid E) = 1 \)
   (if \( E \) happens, then you know \( C \) must happen).

Here are some other terms that can be helpful.

**Partition**: a group of sets which are exhaustive and mutually exclusive form a partition. The whole outcome space has been split into disjoint events, so their probabilities total 1, and there is no overlap between any pair.

Compound events can be evaluated simply by going through the group, and seeing whether each set is to be included.

**Complementary event**: this is a two-event partition. If \( A \) and \( B \) are complementary then \( P(B) = 1 - P(A) \). The simplest way of identifying a complementary pair is \( A \) and ‘not \( A’ \).
Over the course of a season, a hockey team play 40 matches, in different conditions, with the following results.

<table>
<thead>
<tr>
<th>Weather</th>
<th>Good</th>
<th>Bad</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Win</td>
<td>13</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>Draw</td>
<td>5</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Lose</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>15</td>
<td>40</td>
</tr>
</tbody>
</table>

For a match chosen at random from the season:
- G is the event ‘Good weather’
- W is the event ‘Team wins’
- D is the event ‘Team draws’
- L is the event ‘Team loses’.

**a** Find the probabilities:
- i  \( P(G) \)
- ii  \( P(G \cap D) \)
- iii  \( P(D \mid G) \)

**b** Are the events \( D \) and \( G \) independent?

**a** i There are 25 good weather matches out of the 40 so
\[
P(G) = \frac{25}{40} = \frac{5}{8}
\]

ii \( G \cap D \) is a draw in good weather, and there are five of those so
\[
P(G \cap D) = \frac{5}{40} = \frac{1}{8}
\]

iii \( P(D \mid G) = \frac{P(G \cap D)}{P(G)} = \frac{\frac{1}{8}}{\frac{5}{8}} = \frac{1}{5} \)

**b** \( P(D) = \frac{8}{40} = \frac{1}{5} \), so since \( P(D \mid G) = P(D) \) the events \( D \) and \( G \) are independent.

---

**Exercise 4.6**

1. \( A \) and \( B \) are independent events. \( P(A) = 0.7, P(B) = 0.4 \).
   Find:  
   - a  \( P(A \cap B) \)
   - b  \( P(A \cup B) \)
   - c  \( P(A' \cap B) \)

2. \( P(A) = 0.7, \ P(B) = 0.4, \ P(A \cup B) = 0.82 \).
   Show that \( A \) and \( B \) are independent.

3. \( P(A) = 0.5, \ P(B \mid A) = 0.6, \ P(B') = 0.7 \).
   Show that \( A' \) and \( B \) are mutually exclusive.
4 X and Y are independent events with \( P(X) = 0.4 \) and \( P(Y) = 0.5 \).
   a Write down \( P(X | Y) \).
   b Write down \( P(Y | X) \).
   c Calculate \( P(X' \cap Y) \).

5 The results of a traffic survey of the colour and type of car are given in the following table:

<table>
<thead>
<tr>
<th></th>
<th>Saloon</th>
<th>Hatchback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silver</td>
<td>65</td>
<td>59</td>
</tr>
<tr>
<td>Black</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>Other</td>
<td>16</td>
<td>19</td>
</tr>
</tbody>
</table>

One car is selected from the group at random. Find the probability that the selected car is:
   i a silver hatchback
   ii a hatchback
   iii a hatchback, given that it is silver.

Show that the type of car is not independent of its colour.

6 Consider the following possible events when a blue and a white die are rolled:
   \( A \): the total is 2 \( B \): the white is a multiple of 2
   \( C \): the total is < 10 \( D \): the white is a multiple of 3
   \( E \): the total is > 7 \( F \): the total is > 9

Which of the following pairs are exhaustive and which are mutually exclusive?
   a \( A, B \)
   b \( A, D \)
   c \( C, E \)
   d \( C, F \)
   e \( B, D \)
   f \( A, E \)

7 Four unbiased coins are tossed together. For the events \( A \) to \( D \) below, say whether the statements a to d are true or false, and give a reason for each answer. (\( X' \) means NOT \( X \))
   \( A \): no heads \( B \): at least one head
   \( C \): no tails \( D \): at least two tails

   a \( A \) and \( B \) are mutually exclusive
   b \( A \) and \( B \) are exhaustive
   c \( B \) and \( D \) are exhaustive
   d \( A' \) and \( C' \) are mutually exclusive
1 A fair die has six faces numbered 1, 2, 2, 3, 3 and 3. The die is rolled twice and the number showing on the uppermost face is recorded each time.

Find the probability that the sum of the two numbers recorded is at least 5. 

2 A bag contains eight purple balls and two pink balls. A ball is selected at random from the bag and its colour is recorded. The ball is not replaced. A second ball is selected at random and its colour is recorded.

a Draw a tree diagram to represent the information.

b the second ball selected is purple

c both balls selected are purple, given that the second ball selected is purple.

3 For the events $A$ and $B$,

$P(A) = 0.5$, $P(A' \cap B) = 0.27$ and $P(A' \cup B) = 0.53$.

a Draw a Venn diagram to illustrate the complete sample space for the events $A$ and $B$.

b Write down the value of $P(B)$.

c Find $P(B|A)$.

d Determine whether or not $A$ and $B$ are independent.

4 The events $A$ and $B$ are such that $P(A) = \frac{5}{12}$, $P(B) = \frac{2}{3}$ and $P(A' \cap B') = \frac{1}{12}$

a Find:

i $P(A \cap B')$

ii $P(A|B)$

iii $P(B|A)$.

b State, with a reason, whether or not $A$ and $B$ are:

i mutually exclusive

ii independent.
5 Two events $A$ and $B$ are mutually exclusive. $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$

a Find $P(A \mid B)$.

b Find $P(A \cup B)$.

c Are events $A$ and $B$ independent? You must provide a reason.

6 Walker’s disease is a rare tropical disease, which is known to be present in only 0.1% of the population. A new screening test has been analysed and shows a 98% probability of showing positive when the person tested has the disease and only 0.2% of showing positive when the person does not have the disease. A person is selected at random from the population and given the screening test.

a What is the probability that the test will show positive?

b What is the probability that the person does not have the disease, given that the test showed positive?

c Jane is a doctor who is unhappy with guidelines which say that patients should be told immediately if the test shows positive. Explain how she could use the answer to part b to argue that these guidelines are not appropriate.

7 A computer-based testing system gives the user a hard question if they got the previous question correct and an easy question if they got previous question wrong. The first question is randomly chosen to be hard or easy.

The probability of Benni getting an easy question right is $\frac{2}{3}$, and the probability he gets a hard question right is $\frac{1}{4}$.

a Draw a tree diagram to represent what can happen in the first two questions Benni has in a test.

b Find the probability that Benni gets his first two questions correct.

c Find the probability that the first question was hard, given that Benni got both of his first two questions correct.
For the events $A$ and $B$:

a. explain in words the meaning of the term $P(B \mid A)$,

b. sketch a Venn diagram to illustrate the relationship $P(B \mid A) = 0$.

Three companies operate a bus service along a busy main road. Amber buses run 50% of the service and 2% of their buses are more than 5 minutes late. Blunder buses run 30% of the service and 10% of their buses are more than 5 minutes late. Clipper buses run the remainder of the service and only 1% of their buses run more than 5 minutes late.

Jean is waiting for a bus on the main road.

c. Find the probability that the first bus to arrive is an Amber bus that is more than 5 minutes late.

Let $A$, $B$ and $C$ denote the events that Jean catches an Amber bus, a Blunder bus and a Clipper bus respectively. Let $L$ denote the event that Jean catches a bus that is more than 5 minutes late.

d. Draw a Venn diagram to represent the events $A$, $B$, $C$ and $L$. Calculate the probabilities associated with each region and write them in the appropriate places on the Venn diagram.

e. Find the probability that Jean catches a bus that is more than 5 minutes late.

A car dealer offers purchasers a three-year warranty on a new car.

He sells two models, the Zippy and the Nifty. For the first 50 cars sold of each model the number of claims under the warranty is shown in the table.

<table>
<thead>
<tr>
<th></th>
<th>Claim</th>
<th>No claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zippy</td>
<td>35</td>
<td>15</td>
</tr>
<tr>
<td>Nifty</td>
<td>40</td>
<td>10</td>
</tr>
</tbody>
</table>

One of the purchasers is chosen at random. Let $A$ be the event that no claim is made by the purchaser under the warranty and $B$ the event that the car purchased is a Nifty.

a. Find $P(A \cap B)$.

b. Find $P(A')$.

Given that the purchaser chosen does not make a claim under the warranty:

c. find the probability that the car purchased is a Zippy

d. show that making a claim is not independent of the make of the car purchased.

Comment on this result.
10. In a factory, machines $X$, $Y$ and $Z$ are all producing metal rods of the same length. Machine $X$ produces 25% of the rods, machine $Y$ produces 45% and the rest are produced by machine $Z$. Of their production of rods, machines $X$, $Y$ and $Z$ produce 4%, 5% and 2% defective rods respectively.

a. Draw a tree diagram to represent this information.

b. Find the probability that a randomly selected rod is
   i. produced by machine $Y$ and is not defective
   ii. is not defective.

c. Given that a randomly selected rod is not defective, find the probability that it was produced by machine $Y$.

11. A golfer enters two tournaments. He estimates the probability that he wins the first tournament is 0.6, that he wins the second tournament is 0.4 and that he wins them both is 0.35.

a. Find the probability that he does not win either tournament.

b. Show, by calculation, that winning the first tournament and winning the second tournament are not independent events.

c. The tournaments are played in successive weeks. Explain why it would be surprising if these were independent events.

12. The events $A$ and $B$ are independent such that $P(A) = \frac{1}{2}$ and $P(B) = \frac{1}{3}$.

Find:

a. $P(A \cap B)$

b. $P(A' \cap B')$

c. $P(A \mid B)$. 

13 A fair die has six faces numbered 4, 4, 4, 5, 6 and 6. The die is rolled twice and the number showing on the uppermost face is recorded each time.

Find the probability that the sum of the two numbers recorded is at least 10.

14 Events $A$ and $B$ are defined in the sample space $S$. The events $A$ and $B$ are independent.

Given that $P(A) = 0.3$, $P(B) = 0.4$ and $P(A \cup B) = 0.65$, find:

a. $P(B)$

b. $P(A \cap B)$

A and $C$ are mutually exclusive and $P(C) = 0.5$.

c. Find $P(A \cup C)$.

15 The events $A$ and $B$ are such that $P(A) = \frac{1}{3}$, $P(B) = \frac{2}{3}$ and $P(A \cap B) = \frac{1}{4}$.

a. Represent these probabilities in a Venn diagram.

Hence, or otherwise, find

b. $P(A' \cup B')$

c. $P(B \mid A)$.

16 a. If $A$ and $B$ are two events which are statistically independent, write down expressions for $P(A \cap B)$ and $P(A \cup B)$ in terms of $P(A)$ and $P(B)$.

b. Anji and Katrina are keen cinema goers, but they decide each Friday independently of one another whether they go to the cinema. On any given Friday, the probability of both going to the cinema is $\frac{1}{3}$, and the probability that at least one of them goes is $\frac{5}{6}$.

Find the possible values for the probability that Anji goes to the cinema on a Friday. (8)
17 Of the pupils who took English in a certain school one year, 60% of them took History, 30% of them took Religious Studies and 10% took both History and Religious Studies.

One of the pupils taking English is chosen at random.

a Find the probability that this pupil took neither History nor Religious Studies.

b Given that the pupil took exactly one of History and Religious Studies, find the probability it was History.

18 Two identical bags each contain 12 discs, which are identical except for colour. One bag (A) contains six red and six blue discs, and the other (B) contains eight red and four blue discs.

a A bag is selected at random and a disc is selected from it. Draw a tree diagram, illustrating this situation, and calculate the probability that the disc drawn will be red.

b The disc selected is now returned to the same bag, along with another two of the same colour, and another disc is chosen from that bag. Find the probability that:

i it is the same colour as the first disc drawn

ii bag A was used, given that two discs of the same colour have been chosen.
Summary

- The relative frequency of an event happening can be used as an estimate of the probability of that event happening. The estimate is more likely to be close to the true probability if the experiment has been carried out a large number of times.  
- A two-way table can be used to show the possible outcomes of a compound event such as throwing two dice.  
- Venn diagrams are useful when you have information about single events and also their union or intersection.  
- Tree diagrams are useful when you know the probabilities of each stage of compound events. You multiply along the branches to get the probability of a pathway, and the probabilities of different pathways can be added.  
- The conditional probability of $V$ given $F$ is $P(V|F) = \frac{P(V \cap F)}{P(F)}$. Be careful to work out the probability of both $V$ and $F$ happening directly and not from the probabilities of $V$ and $F$ happening individually.  
- Events $V$ and $F$ are independent if $P(V|F) = P(V)$ i.e. knowing that $F$ has happened has given no information about the likelihood of $V$ happening.  

Links

Conditional probability reasoning is a fundamental component of using DNA evidence in trials.

Evaluating risk is a fundamental part of our everyday lives – and it is mostly done very informally, so having a good understanding of the way likelihoods of different events combine can help you to make better informed judgements.