4 Probability

Tests for diseases can return false positives (a positive result when the patient does not have the disease) as well as positive results for patients who do have the disease. Conditional probability allows doctors to have some idea of the likelihood the patient really has the disease when a positive results occurs. Misunderstanding of conditional probabilities has caused serious miscarriages of justice; for example, in the UK, Sally Clark was wrongly convicted of murdering her two sons who died from sudden death infant syndrome.

Objectives

- Evaluate probabilities in simple cases by means of enumeration of equiprobable elementary events (e.g. for the total score when two fair dice are thrown).
- Use addition and multiplication of probabilities, as appropriate, in simple cases.
- Understand the meaning of exclusive and independent events, and calculate and use conditional probabilities in simple cases, e.g. situations that can be represented by means of a tree diagram.

Before you start

You should know how to:

1. Work with fractions, e.g. calculate

a) $\frac{7}{12}+\frac{1}{3}$ **b)** $\frac{1}{6}\times\frac{1}{2}$

a) $\frac{7}{12}+\frac{1}{3}=\frac{7+4}{12}=\frac{11}{12}$ **b)** $\frac{1}{6}\times\frac{1}{2}=\frac{1}{12}$

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

Skills check:

1. Calculate

a) $\frac{1}{4}+\frac{2}{3}$ **b)** $\frac{1}{4}\times\frac{2}{3}$

2. List the possible outcomes when a coin is tossed twice.

4.1 Basic concepts and language of probability

A **probability experiment** has outcomes which occur unpredictably.

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, provided the die is fair, the **theoretical probability** of throwing a five is $\frac{1}{6} = 0.166....$ So, if the die is fair, 'on average' you would see $\frac{1}{6}$ of $500 \approx 83$ fives.

The **probability** is $\frac{1}{6}$ because there are six possible scores when rolling a die.

It might be argued that the die used in the experiment is slightly biased against rolling a five, but the evidence for this is not very strong. When the experiment is recreated, the number of fives varies considerably.

To be more confident, you should roll the die more times: 7416 fives in 50 000 rolls would be stronger evidence that the die was not fair than 74 fives in 500 rolls, even though the proportion of fives is almost the same.

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.

An **event** is a set of possible outcomes from an experiment.

So for rolling a die, you could say:

A is the event that a five is rolled.

B is the event that an even number is rolled.

C is the event that an odd number is rolled.

$A \cup B$ is the **union** of events A and B.
This means A or B or both can happen.
Rolling 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.
Rolling 5 is the only outcome that satisfies $A \cap C$.

A' means the event 'A does not happen'.

This is the **complementary event**, and $P(A') = 1 - P(A)$.

Rolling 1, 2, 3, 4, 6 are the outcomes which satisfy A'.

Sometimes the complementary probability is much easier to work out directly.

Note that $B \cap C$ has no outcomes satisfying it – there are no numbers which are both even and odd. This can be written as $B \cap C = \varnothing$ or $B \cap C = \{\}$ and is referred to as the **null set** or **empty set**.

Exercise 4.1

1. A normal die is thrown. The events A, B, C and D are defined as:

   A: A factor of 4 is seen. B: A square number is seen.

   C: A prime number is seen. D: A multiple of 3 is seen.

   a) For each event, A, B, C and D, write down the outcomes which satisfy it.

   b) Give the probability of each event A, B, C and D.

   c) List the outcomes which satisfy $A \cap C$.

   d) Write down $P(A \cap C)$.

   e) Find $P(A \cap D)$.

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

$P(A \cap C)$ stands for 'the probability that A and C *both* happen.'

2. 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 and 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 $P(A \cup B)$.

3. **a)** Toss a coin 20 times and count the number of times it shows heads.

   b) Toss the coin another 20 times and count the number of times it shows heads.

   c) How many times would you 'expect' to see heads in 20 tosses?

   d) Did you get this in both sets of 20 coin tosses?

   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?

4. **a)** Roll 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 rolls?

   c) Roll 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 rolls?

   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 rolls **ii)** in 20 rolls?

4.2 Two (or more) events

There are many contexts in which you are interested in the outcome of more than one thing at a time – that is, of a **compound event**. In simple cases it may be possible to make a list of all the possible outcomes. In order to make it easier to keep track, it is common to write the outcomes as ordered pairs (or larger groups, if more terms are needed) inside brackets, ( ), and to put the whole list inside curly brackets, { }. This list is called the **sample space** of the experiment.

So for tossing a coin twice the sample space will be: {(H, H); (H, T); (T, H); (T, T)}

In a context where two things happen and you combine them according to some rule, it is often helpful to construct a table showing the possible outcomes. This is sometimes called a **sample space diagram** (also referred to as a **possibility space diagram**).

For example, a situation in which two dice are thrown and the sum of the scores is taken is represented in the *two-way table* shown here.

		Score on second die					
Score on first die	**Sum**	**1**	**2**	**3**	**4**	**5**	**6**
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

It helps to write a sample space list in a logical, methodical order. For tossing a coin three times the sample space can be written as: {(H, H, H); (H, H, T); (H, T, H); (H, T, T); (T, H, H); (T, H, T); (T, T, H); (T, T, T)}.

Within this sample space, the list for tossing a coin twice appears two times – first in red with 'H' in front and then in blue with 'T' in front.

This list of eight outcomes for tossing a coin three times can then form the basis of the sixteen outcomes for tossing a coin four times. This is created by putting an extra 'H' in front and then an extra 'T' in front; it can in turn can be used for the list for tossing a coin five times, and so on.

The blue cell represents the ordered pair (1, 5) – as indicated by the row and column headings – but you can then insert the **value** of interest inside the cell; in this case you can insert the sum of the two scores, which is 6.

Since all 36 of the cells in the table are equally likely outcomes, you can use the table to work out the probability of getting a sum of 5 (the instances of a sum of 5 are in grey cells):

$$P(\text{sum} = 5) = \frac{4}{36}$$

The advantage of this approach is that the (row, column) pair identifies what is actually seen in the experiment, while the contents of the cell show the result of this ordered pair of outcomes, according to the particular rule to be applied.

This rule can change. For example, two dice are again thrown, but this time the higher of the scores on the two dice is taken:

High	**1**	**2**	**3**	**4**	**5**	**6**
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

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

$$\text{P(high} = 5) = \frac{9}{36}$$

> A two-way table can be used to show the possible outcomes of a compound event such as throwing two dice.

Exercise 4.2

1. There are three starters on a restaurant menu: vegetable pakora (V), onion bhaji (O) and chicken tikka (C). Jen and Kay each order a starter.

   List the sample space of possible orders.

2. In question 1, if there is only one of each of the starters left, list the sample space of possible orders.

3. A set of six cards show the numbers 1 to 6. Two cards are taken at random.

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

Sum	**1**	**2**	**3**	**4**	**5**	**6**
1			4			
2						
3						
4						
5						
6		8				

   Find the probability that the total score is

   a) 5 **b)** 4 **c)** 2.

4. A coin is tossed and a die is thrown. A head scores 1 and a tail scores 2.

   a) List all the possible outcomes in the sample space.

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

5. 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 score is

      i) 3 ii) 5 iii) 6 iv) 10.

6. 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 score is

      i) 3 ii) 5 iii) 6.

7. Two dice are thrown.

   a) Construct a sample space diagram to show the (unsigned) difference between the scores on the two dice.

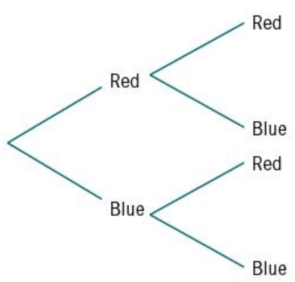
   b) Find the probability that the difference is

      i) 3 ii) 5 iii) 6.

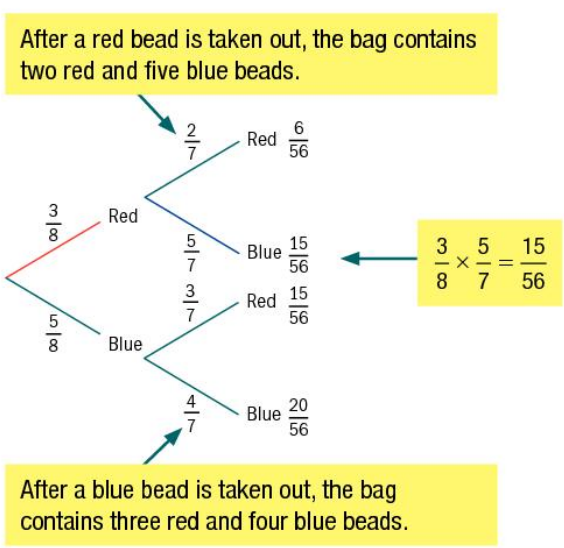
4.3 Tree diagrams

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 possibility **tree diagram**, as shown on the right.



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



P(both beads the same colour) = P(both beads red) + P(both beads blue)

$$= \left(\frac{3}{8} \times \frac{2}{7}\right) + \left(\frac{5}{8} \times \frac{4}{7}\right)$$

$$= \frac{6}{56} + \frac{20}{56}$$

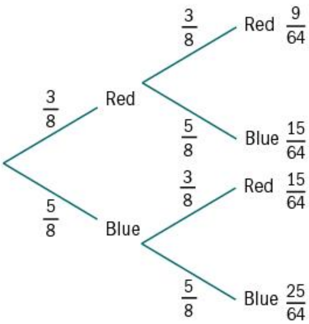
$$= \frac{26}{56}$$

$$= \frac{13}{28}$$

Think × for 'and', and + for 'or'.
So (red *and* red) OR (blue *and* blue)
= (red × red) + (blue × blue)

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 be $\frac{3}{8}$ and $\frac{5}{8}$ at the second stage, as well as the first.



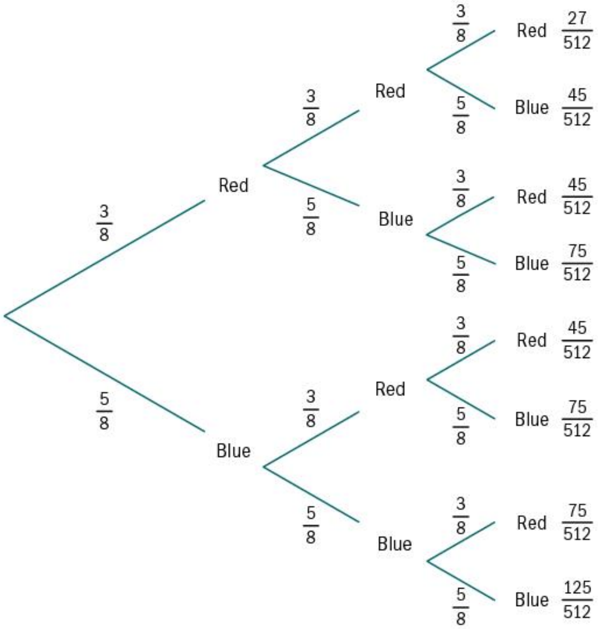
Now,

P(both beads the same colour)
= P(both beads red) + P(both beads blue)

$$= \frac{9}{64} + \frac{25}{64}$$

$$= \frac{34}{64}$$

If the first and second bead are returned to the bag, and a third bead is taken out, the tree diagram will look like this:



Example 1

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 2% of people who do not have the disease.

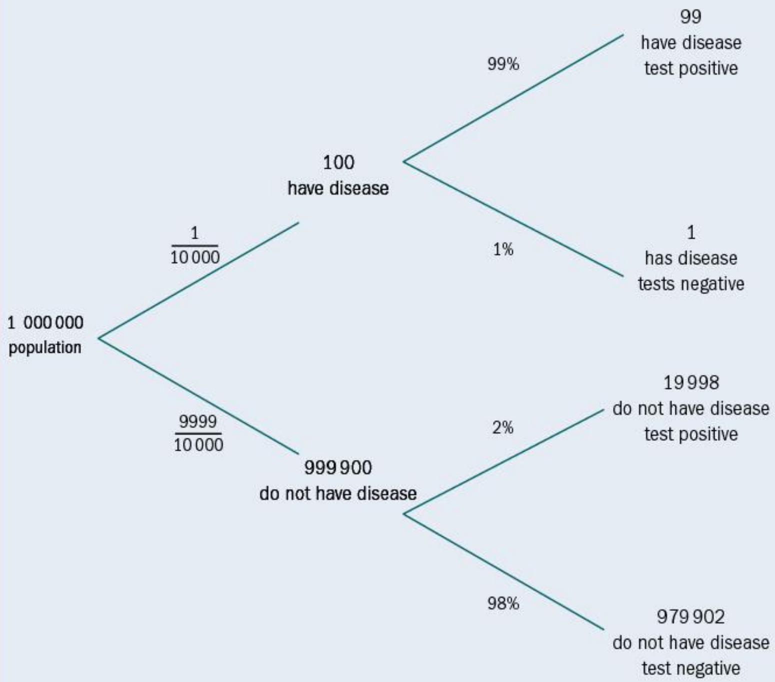
For a population of one million people,

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

b) how many would you expect to test positive?

A screening test will detect the presence in people's bodies of substances which are almost always present with the disease, but which also occur naturally in a small proportion of people.

First draw a tree diagram:



So there are likely to be about 99 positive tests from people with the disease, and about 99 + 19 998 = 20 097 positive tests altogether. In this scenario, fewer than 1 in 200 positive tests is from a person with the disease, despite the test being very accurate.

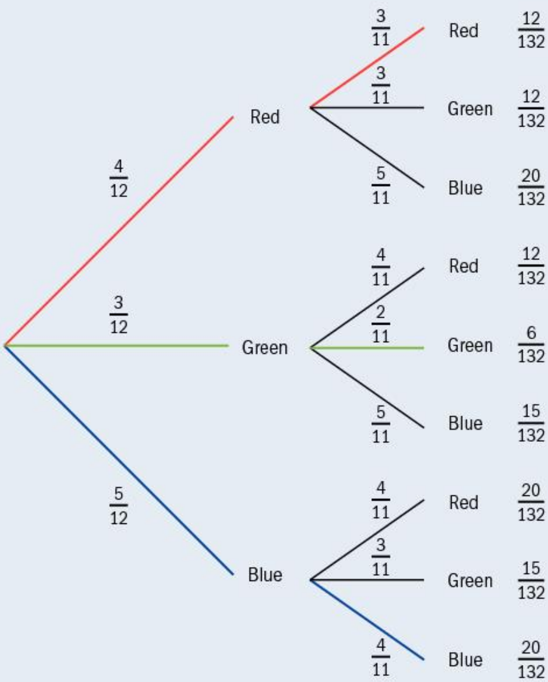
More complex tree diagrams

Example 2

There are four red, three green and five blue discs in a bag. Two discs are drawn out.

Find the probability that two discs the same colour are drawn.

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



Even though some fractions may be simplified, it is much easier to keep the same denominator for the probabilities at each stage.

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

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.

Exercise 4.3

1. A bag contains five blue and three green balls.

A ball is chosen at random and the colour noted, then the ball is 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, what is the probability the two 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. For a person living in a particular town, the probability that during a period of one year they will start a new job is 0.06. The probability that they will be fired from a job is 0.03.

   Assuming these events are independent, draw a tree diagram to represent this information.

   Find the probability that during one year a randomly selected person living in the town has

   a) neither of these events happen

   b) exactly one of these events happen

   c) both of these events happen.

4. A coin is tossed three times. Find the probability that

   a) it shows heads on all three tosses

   b) it shows the same on all three tosses

   c) it does not show the same on two successive tosses.

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

   a) the first counter drawn is red

   b) the first three counters drawn are all white

   c) the first three counters drawn are all different colours.

4.4 Conditional probability

In situations like this it is important that a person testing positive is not told that they have the disease based only on the test result.

In Example 1 you considered a screening test for a disease and found that, even though the test is very accurate, the fact that the disease is rare – there are many more people who do not suffer from it than people who do – means that fewer than 1 in 200 of those testing positive actually have the disease.

19 998 out of 20 097 positive test results were from healthy people, so the **conditional probability** that somebody is healthy *given that* they have a positive result is

$$\text{P(healthy} \mid \text{positive test result)} = \frac{19\,998}{20\,097} \approx 0.9951$$

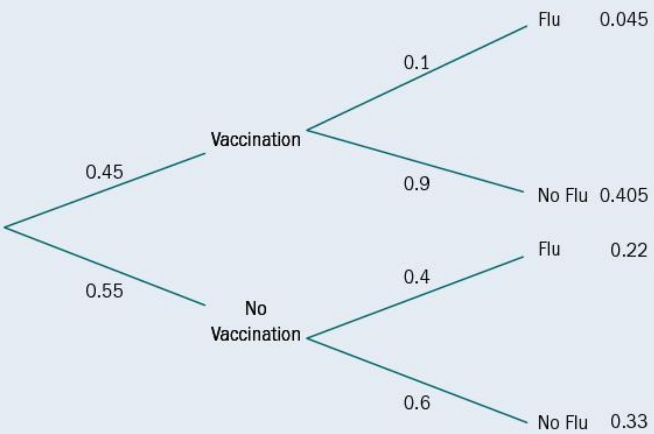
The conditional probability of an event A occurring given that an event B has already occurred can be written as $\mathrm{P}(A|B)$.

Example 3

A medical centre encourages elderly people to have a flu vaccination each year. The vaccination reduces the likelihood of getting flu from 40% to 10%.

If 45% of the elderly people visiting the centre have the vaccination, find the probability that an elderly person chosen at random

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



The symbol | stands for 'given that', i.e. $P(V \mid F)$ stands for the probability of V, given that F has already happened.

a) $P(F) = 0.045 + 0.22 = 0.265$ **b)** $P(V \mid F) = \dfrac{P(V \cap F)}{P(F)} = \dfrac{0.045}{0.265} = 0.170$ (3 s.f.)

The conditional probability of V given F is $P(V \mid F) = \dfrac{P(V \cap F)}{P(F)}$.

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.

Exercise 4.4

1. From a sample taken, 95% of car drivers wear seat belts, 60% of car drivers involved in serious accidents die if they are not wearing a seat belt, and 80% of those that 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. A shopkeeper buys one third of his stock of light bulbs from Company X, and the rest from Company Y.

An independent report states that 3% of light bulbs from Company X are faulty and that 2% from Company Y are faulty.

a) If the shopkeeper chooses a bulb at random from his 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. The homework diaries and completed homework of two students, A and B, are examined.

There is a probability of 0.4 that Student A does not make a note in her diary of the homework set. She always does the homework if it is written in her diary, but never does the homework if it is not written in.

There is a probability of 0.8 that Student B writes the homework given in her diary. When she does this, she will do the homework 90% of the time. If she has nothing written in her diary then she checks with a friend, who knows what the homework is 50% of the time; Student B always does the homework if she is told what it is by her friend.

Draw tree diagrams representing this information.

a) Find the probability that Student B does her homework on a particular night.

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

c) If one piece of homework was given to a student but wasn't done, find the probability that it was given to Student A.

4. In a school there are 542 students, of whom 282 are girls. Of the 542 students, 364 walk to school, and 153 of those are girls. Find the probability that a student 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 they are a boy

d) is a girl, given that they walk to school.

5. There are 173 students in one year in a school. From the year, 25 students play hockey, and of these 7 are in the school's hockey team.

Find the probability that a student chosen from the year at random

a) plays hockey

b) plays in the school's team, given that they play hockey.

6. 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 to show this information. **Calculate** the probability that an employee chosen at random will be late.

7. An insurance company separates car drivers into three categories: Category X is 'low risk', and this category represents 20% of drivers who are insured with the company; Y is 'moderate risk' and represents 70% of drivers insured; 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 drivers in Categories Y and Z are 5% and 9%, respectively.

   a) Find the probability that a driver insured with the company, chosen at random, is assessed as a Category Y risk and has one or more accidents in a 12-month period.

   b) Find the probability that a driver insured with the company, chosen at random, has one or more accidents in a 12-month period.

   c) If a customer has an accident in a 12-month period, what is the probability that they are a Category Y driver?

8. Two identical bags each contain 12 discs, which are identical except for colour. Bag A contains 6 red and 6 blue discs. Bag B contains 8 red and 4 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 returned to the same bag, along with another two 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 the same colour have been chosen.

4.5 Relationships between events

Events are often connected to each other by some type of relationship. The following are the main types.

Note: $\Leftrightarrow$ means 'implies and is implied by', i.e. $p \Leftrightarrow q$ means if p, then q; and if q, then p (p is equivalent to q).

Independence

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

$$\mathrm{P}(A \mid B) = \mathrm{P}(A) \Leftrightarrow A \text{ and } B \text{ are independent, and}$$
$$\mathrm{P}(B \mid A) = \mathrm{P}(B) \Leftrightarrow B \text{ and } A \text{ are independent.}$$

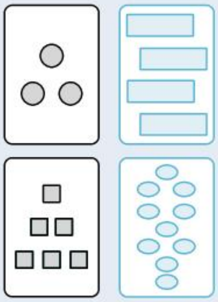
The probability of A occurring given that B has already occurred will just be the probability of A.

Together with the conditional probability definition given on page 72, this gives the multiplication law for independent events:

If A and B are independent, $\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B)$.

Example 4

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 'the card shows a circle', F be the event 'the 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.

Knowing that F has happened has given no additional information concerning whether C is likely to happen or not.

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.

Addition law for probabilities

If we take all the outcomes that satisfy A, and then all the outcomes that satisfy B, then any outcomes which satisfy both will be double counted, so

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

That is, the probability of A or B can be found by adding the probabilities of A and B and then subtracting what has been double counted.

You can check this works in simple examples like events based on throwing a die, but logic says it is always going to be true, even in situations where events are not equally likely or where you cannot list all possible outcomes.

If A is the event a fair die shows a factor of 6 and B is the event that the fair die shows a square number, then 1, 2, 3 and 6 satisfy A, while 1 and 4 satisfy B. A and B is satisfied by 1, and A or B is satisfied by 1, 2, 3, 4 and 6.

$$P(A \cup B) = \frac{5}{6};\ P(A) = \frac{4}{6};\ P(B) = \frac{2}{6};\ P(A|B) = \frac{1}{6}.$$

$$\frac{5}{6} = \frac{4}{6} + \frac{2}{6} - \frac{1}{6}$$

The addition law for probabilities is:

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

Mutually exclusive

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

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

But remember that the general relationship is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Exhaustive

A set of events is **exhaustive** if they cover all possible outcomes.

Example 5

Two fair dice are thrown. From the events listed, give two which, when taken together, 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: if the two dice show the same number the sum has to be an even number.

b) B and F are exhaustive: 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 $P(C|E) = 1$ (if E happens, then you know C must happen).

There are other possibilities for part **(c)**.

Here are some other terms that you need to be aware of:

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 events are a special case of a partition; they are a two-event partition. If A and B are complementary then $P(B) = 1 - P(A)$. The simplest way of representing a complementary pair is as 'A and "not A"'.

Example 6

Over the course of a season, a hockey team play 40 matches, in different conditions, with the following results.

		Weather		Total
		Good	Poor	
Result	Win	13	6	19
	Draw	5	3	8
	Loss	7	6	13
Total		25	15	40

This is a two-way 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 probability in each case.

i) $P(G)$ **ii)** $P(G \cap D)$ **iii)** $P(D|G)$

b) Are the events D and G independent?

a) i) During the season, 25 matches are played in good weather, so $P(G) = \frac{25}{40} = \frac{5}{8}$

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

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

b) $P(D) = \frac{8}{40} = \frac{1}{5}$. Since $P(D|G) = P(D)$ the events D and G are independent.

However, note that W & G are not independent and L & G are not independent either.

Exercise 4.5

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. $P(A) = 0.5$, $P(B|A) = 0.6$, $P(B') = 0.7$

   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 survey of colours and types of cars are shown in the table.

	Saloon	**Hatchback**
Silver	65	59
Black	27	22
Other	16	19

   One car is selected from the group at random.

   a) Find the probability that the selected car is

   i) a silver hatchback

   ii) a hatchback

   iii) a hatchback, given that it is silver.

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

6. Consider the following possible outcomes of rolling a blue die and a white die:

   A: The total is 2.
   B: The white die shows a multiple of 2.
   C: The total is less than 10.
   D: The white die shows a multiple of 3.
   E: The total is greater than 7.
   F: The total is greater than 9.

   Which of the following pairs of events are exhaustive? Which are mutually exclusive?

   a) A, B **b)** A, D

   c) C, E **d)** C, F

   e) B, D **f)** A, E

7. The two sides of a coin are known as 'head' and 'tail'. Four unbiased coins are tossed together. Possible events are:

A: No heads.

B: At least one head.

C: No tails.

D: At least two tails.

Say whether each statement is true or false, giving a reason for your answer.

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.

Recall that X' means 'not X'.

Summary exercise 4

EXAM-STYLE QUESTION

1. 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 this information.

Find the probability that

b) the second ball selected is purple

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

2. Walker's disease is a rare tropical disease, known to be present in only 0.1% of the population. A new screening test has been analysed, and there is a 98% probability of testing positive when the person tested has the disease, and only a 0.2% probability of testing 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 person will test positive?

b) What is the probability that the person does not have the disease, given that they test 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.

EXAM-STYLE QUESTION

3. 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? Provide a reason for your answer.

EXAM-STYLE QUESTIONS

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) $P(A \mid B)$

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

   b) **State**, giving a reason, whether or not A and B are

      i) mutually exclusive

      ii) independent.

5. 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 the 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 for the first two questions Benni gets in a test.

   b) Find the probability 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.

6. In a factory, machines X, Y and Z are all producing identical metal rods. Machine X produces 25% of the rods, machine Y produces 45% and the rest are produced by machine Z. The production of rods from machines X, Y and Z are 4%, 5% and 2% defective, 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 not defective

      ii) not defective.

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

7. 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 both tournaments 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.

EXAM-STYLE QUESTION

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

9. 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 14.

EXAM-STYLE QUESTIONS

10. Events A, B and C are defined in the sample space S.

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

a) $P(A \cap B)$

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

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

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

11. 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) Each Friday, Anji and Katrina decide independently of one another whether to go to the cinema. On any given Friday, the probability of them 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 particular Friday.

12. Of the students who took English in a certain school one year, 60% of them took History, 30% of them took Biology, and 10% took both History and Biology.

One of the students taking English is chosen at random.

a) Find the probability that the student took neither History nor Biology.

b) Given that the student took exactly one of History and Biology, find the probability it was History.

13. Agneska has a spare ticket for a concert tomorrow and phones her friend Venus. It is equally likely that Venus will answer immediately or the phone will go to voicemail. If the phone goes to voicemail there is a probability of 0.7 that Venus calls Agneska back today. If Venus speaks to Agneska today she will attend the concert.

i) Find the probability that Venus attends the concert.

ii) Find the conditional probability that Agneska left a voicemail given that Venus attended the concert. Give your answer correct to 2 decimal places.

14. Jamie tosses a biased coin and throws two fair dice. The probability that the coin shows a head is $\frac{1}{3}$. Each of the dice has six faces, numbered 1, 1, 2, 3, 3 and 4. Jamie's score is calculated from the numbers on the faces that the dice land on, as follows:

- if the coin shows a head, the two numbers from the dice are added together;
- if the coin shows a tail, the two numbers from the dice are multiplied together.

Find the probability that the coin showed a head given that Jamie scores 6.

15. Data about age of males and females in a small rural area are shown in the table.

	Under 35	**35 and over**
Male	345	380
Female	362	472

A person from this area is chosen at random. Let M be the event that the person is male and let Y be the event that the person is under 35.

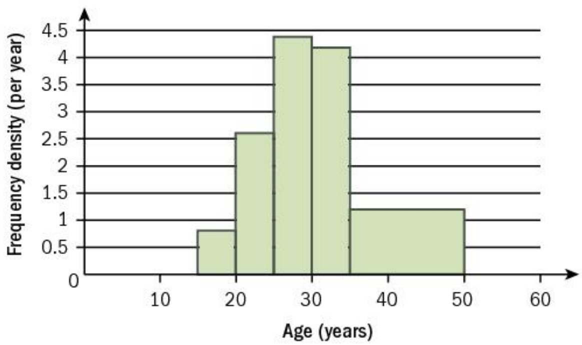
i) Find $P(M)$.

ii) Find $P(M \text{ and } Y)$.

iii) Are M and Y independent events? Justify your answer.

iv) Given that the person chosen is under 35, find the probability that the person is female.

16. Each mother in a random sample of mothers was asked how old she was when her first child was born. The following histogram represents the information.



i) What is the modal age group?

ii) How many mothers were between 20 and 25 years old when their first child was born?

iii) How many mothers were in the sample?

iv) Find the probability that a mother, chosen at random from the group, was between 20 and 25 years old when her first child was born, give that she was not older than 25 years.

Chapter 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 the total score when throwing two dice.
- 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 A given B is $\mathrm{P}(A|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$.

  Be careful to work out the probability of both A and B happening directly and not from the probabilities of A and B happening individually.
- Two events, A and B, are independent if $\mathrm{P}(A|B) = \mathrm{P}(A)$, but if A and B are independent events then $\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B)$. That is, knowing that B has happened has given no information about the likelihood of A happening, and vice versa.
- The addition law for probabilities is:
  $$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$$
- Two events A and B are mutually exclusive if they cannot occur at the same time:
  $$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$$
- A set of events is exhaustive if they cover all possible outcomes.