# 5 Quantifying uncertainty: probability

Probability enables us to quantify the likelihood of events occurring and evaluate risk. This chapter looks at the language of probability, how to quantify probability and the basic tools you need to solve problems involving probability.

**Concepts**

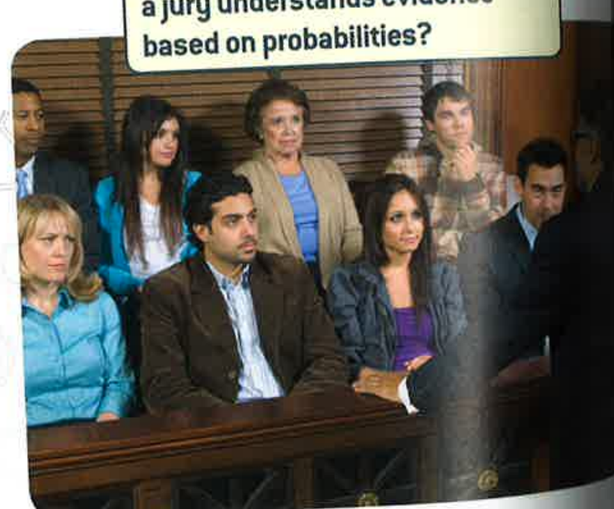
- Representation
- Quantity

**Microconcepts**

- Uncertainty and random behaviour
- Trial, outcome, equally likely outcomes and relative frequency
- Sample space and event
- Theoretical probability
- Venn diagrams, tree diagrams, sample space diagrams and tables of outcome
- Compound events
- Exclusive, independent and dependent events

**How can a geneticist quantify the chance that a child may inherit the same colour of eyes as his father?**

**How can a lawyer make sure that a jury understands evidence based on probabilities?**



**"37% chance of rain tomorrow." How can you reach a common agreement on how to interpret and apply statements like this?**



Daniel Kahneman (Nobel Prize winner in economic science) and Amos Tversky (cognitive and mathematical psychologist) spent decades collaborating and researching together. Below is an adaptation of one of the questions they set to students:

Two taxi companies operate in Mathcity: Blackcabs and Yellowrides.

85% of the cabs in the city work for Blackcabs and are coloured black.

The rest of the cabs in the city work for Yellowrides and are coloured yellow.

A taxi was involved in a hit and run accident at night. A witness told police that the taxi involved was yellow. The court carried out a series of tests on the reliability of the witness, asking her to identify the colour of a random sequence of taxis. The witness correctly identified each one of the two colours 80% of the time and failed 20% of the time.

1 What is the probability that the taxi involved in the accident was yellow?

2 Karolina carries out a traffic survey in Mathcity. She sits at an interchange and notes the colour of the first six cabs that pass her. What number of yellow cabs is she most likely to observe?

- What types of diagram can help represent the problem?
- What assumptions did you make?

## Developing inquiry skills

Write down any similar inquiry questions you might ask if you were asked to predict the reliability of the witness if 50% of the cars in the city were Yellowrides or if another taxi company Blue Taxis also operated in Mathcity.

What questions might you need to ask in these scenarios?

Think about the questions in this opening problem and answer any you can. As you work through the chapter, you will gain mathematical knowledge and skills that will help you to answer them all.

## Before you start

**You should know how to:**

1 Find simple probabilities.

eg A number is chosen at random from the set of natural numbers {1, 2, 3, ..., 100}.

Find the probability of choosing a cube number.

The cube numbers in the set are 1, 8, 27 and 64. Probability of a cube number

$= \frac{4}{100} = 0.04$

**Skills check**

Click here for help with this skills check

1 A number is chosen at random from the set {1, 2, 4, 5, 9, 10, 11, 16, 17, 25, 26, 27}

Find the probability that the number is:

**a** prime **b** odd **c** a square number.

2 A student collects this data:

| | Smoker | Non-smoker |
|---|---|---|
| Male | 12 | 47 |
| Female | 6 | 51 |

A person is chosen at random from the survey. Find the probability that they are:

**a** female **b** a male smoker

**c** a non-smoker.

# 5.1 Reflecting on experiences in the world of chance. First steps in the quantification of probabilities

Probability is synonymous with uncertainty, likelihood, chance and possibility. You can quantify probability through three main approaches: subjective, experimental and theoretical.

You may judge that you are more likely to get to school on time if you take a particular route, based on your experience with traffic. Subjective probabilities are based on past experiences and opinions rather than formal calculations.

### Investigation 1

We use subjective probability in everyday life every time we make a judgement of how likely something is to occur. For example, you may judge that you are more likely to get to school on time if you take a particular route, based on your experience with traffic. We justify subjective probabilities to ourselves through a mixture of past experiences, opinions and intuition.

How likely do you judge these outcomes to be?

| | | |
|---|---|---|
| A: There will be a financial crisis in Europe during the next 10 years. | B: It will rain tomorrow. | C: There will be a financial crisis in Asia during the next 10 years. |
| D: Choosing one digit at random from the decimal expansion of $\frac{1}{6}$, you get 6. | E: The world will be free of all dictators within the next 10 years. | F: The sequence 999999 is found somewhere in the first 1000 digits of pi. |
| G: The team winning the FIFA World Cup in 2030 will be from the Americas. | H: Humans will land on Mars by 2050. | I: If you cut a strip of paper into three lengths at random, they can form a triangle. |

Display your answers by plotting them on this probability scale:

0 (impossible to occur) — 0.5 (just as likely to occur as not) — 1 (certain to occur)

**Compare, contrast, discuss** and **justify** your answers within a small group.

You may find disagreements with others, based on your opinions, experience or beliefs.

When is it easier to reach a common agreement on the value of a subjective probability?

**TOK**

Do you rely on intuition to help you make decisions?





## Experimental probability

You should use these terms when discussing and quantifying probabilities:

**Experiment:** A process by which you obtain an observation.

**Trials:** Repeating an experiment a number of times.

**Outcome:** A possible result of an experiment.

**Event:** An outcome or set of outcomes.

**Sample space:** The set of all possible outcomes of an experiment, always denoted by $U$.

These terms are illustrated in the following example:

Erin wants to explore the probability of throwing a prime number with an octahedral die. She designs an **experiment** that she feels is efficient and bias-free. Erin places the die in a cup, shakes it, turns the cup upside down, then reads and records the number thrown.

Erin repeats her experiment until she has completed 50 **trials**. She knows that the **outcome** of each trial can be any number from $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and that the **event** she is exploring can be described as a statement: "throw a prime with an octahedral die" or a set of outcomes that make the statement true: $\{2, 3, 5, 7\}$.

Erin can either write **P(throw a prime)** to represent the probability of her event occurring or **P($A$)** if $A$ denotes the set $\{2, 3, 5, 7\}$.

**International-mindedness**

Probability theory was first studied to increase the chances of winning when gambling. The earliest work on the subject was by Italian mathematician Girolamo Cardano in the 16th century.

**HINT**

A crucial assumption in many problems is that of **equally likely outcomes**.

A consequence of the geometry of the shapes shown here is that they form *fair dice*. Each outcome on a fair die is equally likely as any other.

One way to quantify probability is with relative frequency, also known as experimental probability. The general formula for the relative frequency of an event $A$ after $n$ trials is:

$$\text{Relative frequency of } A = \frac{\text{Frequency of occurrence of event } A \text{ in } n \text{ trials}}{n}$$

This is also known as the experimental probability of the event $A$.

Theoretical probability gives you a way to quantify probability that does not require carrying out a large number of trials.

Statistics and probability

The formula for the theoretical probability $\mathrm{P}(A)$ of an event $A$ is:

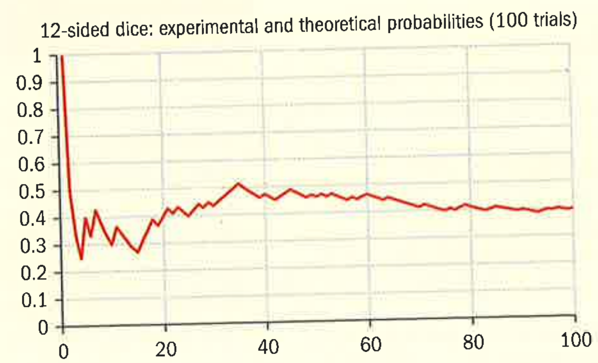
$\mathrm{P}(A) = \frac{n(A)}{n(U)}$ where $n(A)$ is the number of outcomes that make $A$ happen and $n(U)$ is the number of outcomes in the sample space.

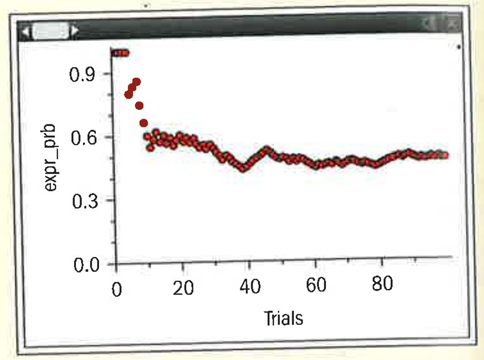
Whenever $\mathrm{P}(A)$ represents a subjective, experimental or theoretical probability, then $0 \leq \mathrm{P}(A) \leq 1$.

## Investigation 2

1 Imagine throwing a fair 12-sided die 15 times. Let $A$ be the event "throw a prime number".

2 Use technology to show the sequence of experimental probabilities of $A$ after $1, 2, 3, \ldots, 100$ throws. You should be able to create a graph like one of these:




| Number of trial ($n$) | Outcome | Event | Frequency of occurrence in $n$ trials | Relative frequency in $n$ trials |
|---|---|---|---|---|
| 1 | 2 | Prime | 1 | 1 |
| 2 | 11 | Prime | 2 | 1 |
| 3 | 5 | Prime | 3 | 1 |
| 4 | 9 | Not prime | 3 | 0.75 |
| 5 | 4 | Not prime | 3 | 0.6 |
| 6 | 2 | Prime | 4 | 0.666667 |
| 7 | 4 | Not prime | 4 | 0.571429 |
| 8 | 8 | Not prime | 4 | 0.5 |
| 9 | 2 | Prime | 5 | 0.555556 |
| 10 | 12 | Not prime | 5 | 0.5 |
| 11 | 6 | Not prime | 5 | 0.454545 |
| 12 | 1 | Not prime | 5 | 0.416667 |
| 12 | 1 | Not prime | 5 | 0.416667 |
| 14 | 3 | Prime | 6 | 0.428571 |
| 15 | 11 | Prime | 7 | 0.466667 |





3 Show that for this experiment, $\mathrm{P}(A) = \frac{5}{12} \approx 0.417$. Add a horizontal line with equation $y = 0.417$ to your graph. Press F9 (spreadsheet) or Ctrl+R (TiNspire) to carry out another 100 trials.

4 Repeat until you have seen each of these three scenarios:

- The experimental probability is always greater than the theoretical probability.
- The experimental probability is always less than the theoretical probability.
- The experimental probability is often equal to the theoretical probability.

You may wish to adapt your spreadsheet so that it carries out 1000 trials. Examine the columns in your spreadsheet and the features on your graph.

5 **Factual** What is the set of *all* possible values of theoretical probabilities?

6 What relationship does your graph have with the line $y = \frac{5}{12}$?

7 **Factual** What is the relationship between relative frequency and theoretical probability in the short term?

8 **Factual** What is the relationship between relative frequency and theoretical probability in the long term?

9 **Conceptual** In the short term, does random behaviour involve predictability or unpredictability?

10 **Conceptual** In the long term, does random behaviour involve predictability or unpredictability?

11 **Conceptual** How may we interpret and apply the number quantified by the formula for the theoretical probability of an event?

## Example 1

Find the probability of each event and determine which event is least likely.

$T$: throw a factor of 24 on a four-sided die.
$O$: throw a prime on an eight-sided die.
$D$: throw at least 11 on a 12-sided die.
$C$: throw at most 3 on a three-sided die.
$I$: throw a multiple of 5 on a 20-sided die.

All the dice are fair and are numbered from 1 up to the number of sides on the die.

| | |
|---|---|
| $n(T) = n(\{1, 2, 3, 4\}) = 4 = n(U)$<br>so $\mathrm{P}(T) = 1$ | Every element of $\{1, 2, 3, 4\}$ is a factor of 24.<br>$\mathrm{P}(T) = \frac{n(T)}{n(U)} = \frac{4}{4} = 1$, so $T$ is certain to happen. |
| $n(O) = n(\{2, 3, 5, 7\}) = 4$, $n(U) = 8$<br>so $\mathrm{P}(O) = \frac{4}{8} = 0.5$ | |
| $n(D) = n(\{11, 12\}) = 2$, $n(U) = 12$ | "at least 11" means "11 or more" |

Continued on next page

Statistics and probability

| | |
|---|---|
| so $P(D) = \frac{2}{12} = \frac{1}{6} = 0.1\dot{6}$<br>$n(C) = n(\{1, 2, 3\}) = 3$, $n(U) = 6$<br>so $P(C) = \frac{3}{6} = 0.5$<br>$n(I) = n(\{5, 10, 15, 20\}) = 4$, $n(I) = 20$<br>so $P(I) = \frac{4}{20} = 0.25$<br>Hence, $D$ is the least likely event. | "at most 3" means "3 or less" |

Just as theoretical probability gives you a way to predict long-term behaviour of relative frequency, a simple rearrangement gives you a way to predict how many times an event is likely to occur in a given number of trials.

## Example 2

**a** A fair coin is flipped 14 times. Predict the average number of times you expect a head to be face up.

**b** Statistical data built up over 5 years shows that the probability of a student being absent at a school is 0.05. There are 531 students in the school.

Predict the number of students that you expect to be absent on any given day and interpret your answer.

**c** State the assumptions supporting your answer for part **b**.

| | |
|---|---|
| **a** $14 \times 0.5 = 7$.<br>Seven heads are expected. | The expected number of occurrences is $nP(A)$. |
| **b** $531 \times 0.05 = 26.55$. So, around 26 or 27 students are expected to be absent. | Note that 26.55 students cannot actually be absent. |
| **c** This assumes that absences on all days of the year are equally likely. | |

**TOK**

Play the game of the St Petersburg Paradox and decide how much you would pay to play the game.

## Exercise 5A

1 A letter is picked at random from the letters of RANDOM. Calculate the probability that it is a letter from MATHS.

2 This dartboard has 20 sectors each of equal area.

If a dart lands in a numbered sector at random, find the probability that the number is:

**a** at least 4 **b** more than 6

**c** less than 30 **d** no more than 14

**e** prime **f** square

**g** a solution to the equation $x^2 = 3$.

3 A survey was carried out in a small city centre street one Saturday afternoon. Shoppers were asked about how they travelled that day. The results are shown in the table below.

| Mode of transport | Car | Bus | Foot |
|---|---|---|---|
| Male | 40 | 59 | 37 |
| Female | 33 | 41 | 29 |

One shopper is randomly selected.

**a** Find the probability that this shopper travelled by car.

One male shopper is randomly selected.

**b** Find the probability that this male shopper travelled by foot.

**c** 1300 shoppers visit the town in one week. Estimate the number of shoppers who travelled by bus.

4 A personal identification number (PIN) consists of four digits. Consider the PIN 0005 equal to the number 5, etc. Find the probability that a PIN number is:

**a** equal to 0000

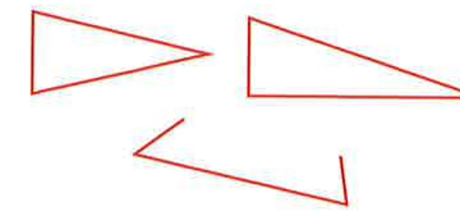
**b** less than 8000 and more than 7900

**c** divisible by 10

**d** at least 13.

5 Take a narrow strip of paper 20 cm long. Use your calculator to generate a random decimal length between 0 and 20 and cut the strip into two strips at this length. Label the two strips H and T. Toss a coin. If the coin shows heads choose strip H. Measure its length and use your calculator to find a place to cut it at random into two strips.

Can you make a triangle with your three pieces?



Make a guess on the probability scale as to how likely it is that a triangle can be formed following this process. Use your classmates' results to quantify the experimental probability.

6 A health professional is investigating the theoretical probability that a randomly chosen female smokes is 0.17. She organizes a survey and asks 11 278 females if they smoke or not. Using her theoretical probability, determine the number of females she would predict to be smokers.

7 A multiple-choice test consists of 10 questions. Each question has five answers. Only one of the answers is correct. For each question, Jose randomly chooses one of the five answers. Predict the expected number of questions Jose answers correctly.

## Developing inquiry skills

There are four outcomes in the first opening scenario:

- A taxi is yellow and is identified as yellow.
- A taxi is yellow and is identified as black.
- A taxi is black and is identified as yellow.
- A taxi is black and is identified as black.

Are these equally likely outcomes?

In 1000 trials, how many occurrences of each outcome would you expect?

Statistics and probability





# 5.2 Representing combined probabilities with diagrams

You have taken the first steps in the quantification of probabilities, experienced random experiments and investigated how to make predictions in the world of chance by application of formulae.

Probability situations themselves have a structure that you can represent in different ways, for example in problems where two or more sets are combined in some way.

## Investigation 3

1 For each situation, think about how best to represent the situation with a diagram.
Compare and contrast your diagrams with others in your class **then** solve the problems.

**Situation 1**

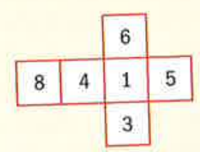
In a class survey on subject choices, Isabel, Clara, Coco, Anastasiia and Fangyu all state that they study biology. Isabel, Clara, Fangyu and Tomas all study chemistry whereas Barbora, Coco and Achille study neither biology nor chemistry.

2 Find the probability that a student chosen randomly from this class studies both biology and chemistry.

3 Create your own probability question using your representation of the situation and have another student answer it.

4 **Situation 2**

One example of a Sicherman die is a fair cubical die with this net:




It is thrown together with a fair octahedral die whose faces are numbered 1, 2, 3, 4, 5, 6, 7 and 8.

5 Find the probability that the number obtained by adding the two numbers thrown on each die is prime.

6 Find and describe a pattern in your representation of this situation and acquire some knowledge from your pattern.

7 **Conceptual** What advantages are there in using a diagram in problem-solving with combined probabilities?

Two frequently used representations of probability problems are Venn diagrams and sample space diagrams.

A **Venn diagram** represents the sample space with a rectangle. Within the rectangle, each event is represented by a set of outcomes in a circle or an oval shape and is labelled accordingly.

A **sample space diagram** is a useful way to represent the whole sample space and often takes the form of a table.

**International-mindedness**

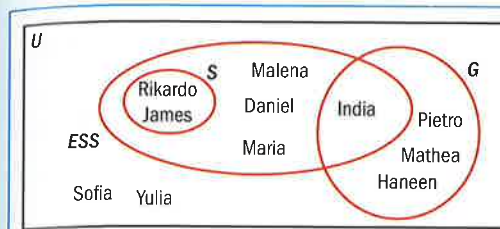
A well-known French gambler, Chevalier de Méré, consulted Blaise Pascal in Paris in 1650 with questions about some games of chance. Pascal began to correspond with his friend Pierre de Fermat about these problems, which began their study of probability.





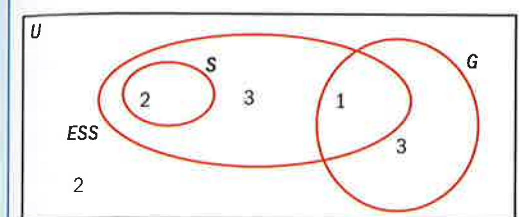
## Example 3

In a class survey, Rikardo, Malena, Daniel, Maria, India and James reported that they study environmental systems and societies *(ESS)*. India, Pietro, Mathea and Haneen said that they study geography *(G)*. Rikardo and James were the only ones who reported that they studied Spanish *(S)* whereas Sofia and Yulia studied none of the subjects mentioned in the survey. Draw the data in a Venn diagram.




Each set is represented by a italic capital letter.

*U* represents the entire sample space. In set terminology, this is called the universal set.




This diagram can be simplified to show the number of students in each region.

## Example 4

Use the Venn diagram in Example 3 to find the probabilities that a student chosen randomly from this class:

**a** studies ESS
**b** studies ESS but not Spanish
**c** studies all three subjects
**d** studies exactly two of the subjects.

**a** $P(ESS) = \frac{n(ESS)}{n(U)} = \frac{2+3+1}{11} = \frac{6}{11}$

*P(ESS)* represents the "Probability of choosing a student at random from the set *ESS*".

**b** $\frac{4}{11}$

There are a total of four students within the *ESS* oval but outside the Spanish oval.

**c** $0$

The diagram clearly shows that there are no students who study all three subjects.

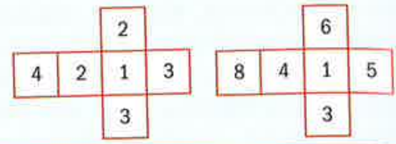
**d** $\frac{3}{11}$

The diagram clearly shows that two students Rikardo and James study Spanish and ESS, whereas one student India – studies both geography and ESS. These are the only three students who study exactly two of the subjects surveyed.

Statistics and probability

## Example 5

It is claimed that when this pair of Sicherman dice is thrown and the two numbers obtained added together, the probability of each total is just the same as if the two dice were numbered with 1, 2, 3, 4, 5 and 6. Verify this claim.




Sample space diagram for the total of two die numbered 1, 2, 3, 4, 5 and 6:

| | 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 |

Sample space diagram for the two Sicherman dice:

| | 1 | 2 | 2 | 3 | 3 | 4 |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 3 | 4 | 4 | 5 |
| 3 | 4 | 5 | 5 | 6 | 6 | 7 |
| 4 | 5 | 6 | 6 | 7 | 7 | 8 |
| 5 | 6 | 7 | 7 | 8 | 8 | 9 |
| 6 | 7 | 8 | 8 | 9 | 9 | 10 |
| 8 | 9 | 10 | 10 | 11 | 11 | 12 |

In both tables, $P(T = 2) = P(T = 12) = \frac{1}{36}$

$P(T = 3) = P(T = 11) = \frac{2}{36} = \frac{1}{18}$

$P(T = 4) = P(T = 10) = \frac{3}{36} = \frac{1}{12}$

$P(T = 5) = P(T = 9) = \frac{4}{36} = \frac{1}{9}$

$P(T = 6) = P(T = 8) = \frac{5}{36}$

and $P(T = 7) = \frac{6}{36} = \frac{1}{6}$

The probability of each total is the same for each pair of dice, so the claim is true.

Form a sample space diagram for each experiment. Enter each total in the table as shown.

Then find the probability of each outcome in the sample space, representing the total as $T$.

State your conclusion.

Once time has been invested in drawing a diagram, it can be used to quantify many different probabilities.

**TOK**

Do ethics play a role in the use of mathematics?





## Exercise 5B

1 Alex throws a fair tetrahedral (four-sided) dice and a fair octahedral (eight-sided) dice. He defines $M$ as the product of his two numbers. Find:

**a** P($M$ is odd) **b** P($M$ is prime)

**c** P($M$ is both odd and prime)

Bethany has two fair six-sided dice, which she throws. She defines $N$ as the product of her numbers. Find:

**d** P($N$ is odd) **e** P($N$ is more than 13)

**f** P($N$ is a factor of 36)

Bethany and Alex can see that the probability that $M$ is odd equals the probability that $N$ is odd. Try to find more events that have the same probabilities for each of their experiments. Find at least one such event.

2 A survey of 127 consumers found that 81 had a tablet computer, 70 had a smartphone and 29 had both a smartphone and a tablet computer.

**a** Find the number of consumers surveyed who had neither a smartphone nor a tablet.

**b** Find the probability that when choosing one of the consumers surveyed at random, a consumer who has only a smartphone is chosen.

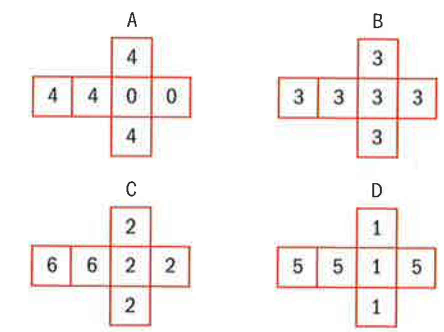
**c** In a population of 10 000 consumers, predict how many would have only a tablet computer.

3 In a class of 20 students, 12 study biology, 15 study history and 2 students study neither biology nor history.

**a** Find the probability that a student selected at random from this class studies both biology and history.

**b** Given that a randomly selected student studies biology, find the probability that this student also studies history.

4 These dice compete in the "Dice World Cup". A pair of dice is thrown and the highest number wins. The semi-finals are A vs B and C vs D. The winners of each semi-final go in to the final.




Construct sample space diagrams to find the probabilities of the outcomes of each semi-final.

5 The dice in the previous question are called non-transitive dice. Show that A is likely to beat B, that B is likely to beat C and that C is likely to beat A. You may wish to explore the meaning of the term *transitive* and try to design your own non-transitive dice.

6 Two cubical dice are rolled in a game. The score is the greater of the two numbers. If the same number appears on both dice, then the score is that number. Find the probability that the score is at most 4.

Statistics and probability

## Developing inquiry skills

In the first opening scenario, imagine 100 trials. How many outcomes would you expect in each area shown on this diagram?

| | | Cab yellow? | |
|---|---|---|---|
| | | Yes | No |
| Witness correct? | Yes | ?? | ?? |
| | No | ?? | ?? |

# 5.3 Representing combined probabilities with diagrams and formulae

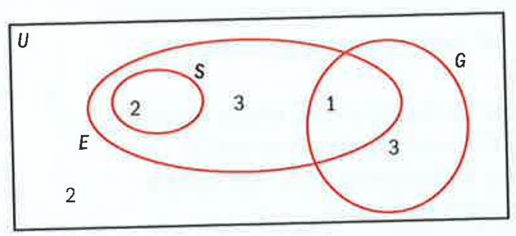
In Section 5.2, you found probabilities by representing combined events in a sample space diagram or a Venn diagram. There are other ways to find probabilities of combined events, which can add to your problem-solving skills.

In this section, you will use Venn diagrams to investigate and represent laws of probability and you will use these symbols, language and definitions:

| Name | Symbol applied to events | Informal language | Formal definition |
|---|---|---|---|
| Intersection | $A \cap B$ | $A$ and $B$ | Events $A$ and $B$ **both occur** |
| Union | $A \cup B$ | $A$ or $B$ | Events $A$ **or** $B$ **or both** occur |
| Complement | $A'$ | Not $A$ | Event $A$ **does not** occur |
| Conditional | $A\|B$ | $A$ given $B$ | Event $A$ **given that** event $B$ **has** occurred |

## Example 6

A student is chosen at random from this class. If $E$ is the event "the student takes ESS" and $G$ is the event "the student takes geography", then find these probabilities and interpret what they mean:




**a** $P(E \cap G)$ and $P(G \cap E)$ **b** $P(E \cup G)$ and $P(G \cup E)$

**c** $P(E')$ **d** $P(E|G)$ and $P(G|E)$

**a** $P(E \cap G) = P(G \cap E) = \frac{1}{11}$ is the probability that a randomly chosen student studies **both** ESS and geography.

Only one student takes both ESS and geography. This example illustrates that $E \cap G$ means exactly the same as $G \cap E$. In fact this is always true.

**b** $P(E \cup G) = \frac{2+3+1+3}{11} = \frac{9}{11}$ is the probability that a randomly chosen student studies ESS **or** geography **or both**.

Similarly, $E \cup G$ means the same as $G \cup E$ hence $P(E \cup G) = P(G \cup E)$ is always true.

**c** $P(E') = \frac{5}{11}$ is the probability that a randomly chosen student **does not** study ESS

There are 5 students outside the ESS oval. $P(E') = 1 - \frac{6}{11} = \frac{5}{11}$ is another way to find the probability required.



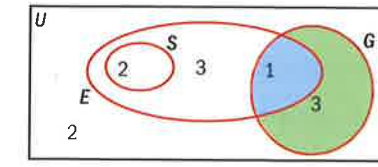


**d** $P(E|G) = \frac{1}{1+3} = \frac{1}{4}$, the probability that a randomly chosen student studies ESS **given** that he/she studies geography.

However, $P(G|E) = \frac{1}{2+3+1} = \frac{1}{6}$.

These are not equal since the information **given** changes the sample space. This example shows that $P(E|G) = P(G|E)$ is not generally true.

Since it is **given** that $G$ has occurred, the sample space is now $G$, not $U$.




Only 1 student studies ESS and geography, hence $P(E|G) = \frac{1}{4}$. Notice how this contrasts with $P(E) = \frac{6}{11}$.

**International-mindedness**

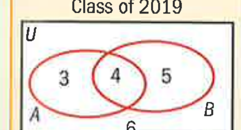
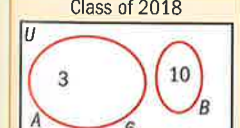
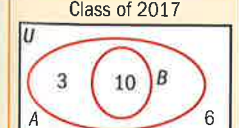
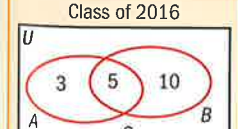
The Dutch scientist Christiaan Huygens, a teacher of Leibniz, published the first book on probability in 1657.

Just as areas of mathematics like trigonometry or sequences have formulae, so does probability. In this investigation, you will consider some relationships that you can generalize as laws of probability.

## Investigation 4

The following Venn diagrams represent how many students study art or biology in four different classes, using the sets $A$ and $B$.

Fill in the probabilities for each Venn diagram and investigate your answers.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Venn Diagram | $P(A)$ | $P(A')$ | $P(B)$ | $P(A \cap B)$ | $P(A \cup B)$ | $P(A\|B)$ | $P(A)+P(B)$ | $P(A) \times P(B)$ | $P(A)+P(B) - P(A \cap B)$ | $\frac{P(A \cap B)}{P(B)}$ |
| Class of 2019  | | | | | | | | | | |
| Class of 2018  | | | | | | | | | | |
| Class of 2017  | | | | | | | | | | |
| Class of 2016  | | | | | | | | | | |

Continued on next page

Statistics and probability

Examine your results.

Answer these questions and discuss your answers in a group.

1 What relationship exists between the probabilities in columns 1 and 2?

2 This relationship is true in general. Why?

3 What relationship exists between the probabilities in columns 5 and 9?

4 This relationship is true in general. Why?

5 What relationship exists between the probabilities in columns 6 and 10?

6 **Factual** Which Venn diagram shows "mutually exclusive events", ie ones that cannot occur together?

7 **Factual** Which Venn diagram shows "independent events", ie events for which the outcome of one is unaffected by the outcome of the other?

8 **Conceptual** What is the difference between mutually exclusive and independent events? Can mutually exclusive events be independent? Why or why not?

For events $A$ and $B$, the following laws apply:

- $P(A \cup B) = P(A) + P(A) - P(A \cap B)$
- $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
- This can be rearranged to give the multiplicative law of probabilities: $P(A \cap B) = P(A|B)P(B)$.

Two events $A$ and $B$ are **mutually exclusive** if they cannot **both** occur. Hence:

- Knowing that $A$ has occurred means you know that $B$ cannot, and that knowing that $B$ occurs means you know that $A$ cannot.
- $P(A \cap B) = P(B \cap A) = 0$
- $P(A \cup B) = P(A) + P(B)$

Two events $A$ and $B$ are **independent** if the occurrence of each event does **not affect in any way** the occurrence of the other. Hence:

- Knowing that $A$ has occurred does not affect the probability that $B$ does, and knowing that $B$ occurs does not affect that probability that $A$ does.
- $P(A|B) = P(A)$ and $P(B|A) = P(B)$
- $P(A \cap B) = P(A) \times P(B)$
- If $A$ and $B$ are not independent, they are **dependent**.

If events $A$ and $A'$ are **complementary** then $A'$ is the set of outcomes that are not in the event $A$:

- $A$ and $A'$ are mutually exclusive.
- $P(A) + P(A') = 1$.

**Reflect** Are complementary events independent events?

Are complementary events mutually exclusive?

Can mutually exclusive events be independent?

Can non-mutually exclusive events be independent?





You can use the laws of probability to justify other statements.

## Example 7

Marcelo is playing in a cricket match and a game of hockey at the weekend.

The probability that his team will win the cricket match is 0.2, and the probability of winning the hockey match is 0.6. Assume that the results in the matches are independent.

**a** Find the probability that Marcelo's team wins both matches.

**b** Find the probability that Marcelo's team wins the cricket match or the hockey match.

**c** Determine if winning the cricket match and winning the hockey match are mutually exclusive. Justify your answer.

| | |
|---|---|
| **a** $P(C \cap H) = 0.2 \times 0.6 = 0.12$ | Let $C$ be the event "wins the cricket match" and $H$ be "wins the hockey match". Since you are given that $C$ and $H$ are independent, you can use $P(C \cap H) = P(C) \times P(H)$. |
| **b** $P(C \cup H) = 0.2 + 0.6 - 0.12 = 0.68$ | Apply the formula. |
| **c** $C$ and $H$ are not mutually exclusive because $P(C \cap H) \neq 0$. | Write a complete and clear reason. |

## Example 8

Let $H$ and $G$ be events such that $P(H) = \frac{1}{3}$, $P(G) = \frac{3}{7}$ and $P(H \cup G) = \frac{13}{21}$.

Find $P(H \cap G)$ and $P(H|G)$ and hence, determine if $H$ and $G$ are independent.

| | |
|---|---|
| $P(H \cup G) = P(H) + P(G) - P(H \cap G)$<br>So $\frac{13}{21} = \frac{1}{3} + \frac{3}{7} - P(H \cap G)$,<br>hence $P(H \cap G) = \frac{1}{7}$ | Write down the appropriate formula. |
| $P(H\|G) = \dfrac{P(H \cap G)}{P(G)} = \dfrac{\frac{1}{7}}{\frac{3}{7}} = \dfrac{1}{3}$ | Write down the appropriate formula. |
| Since $P(H\|G) = P(H)$, $H$ and $G$ are independent. | Write a complete and clear reason. |

Statistics and probability

## Exercise 5C

1 For these pairs of events, state if they are mutually exclusive, independent or neither.

a $A$ = throw a head on a fair coin
$B$ = throw a prime number on a fair die numbered 1, 2, 3, 4, 5, 6.

b $C$ = it rains tomorrow, $D$ = it rains today

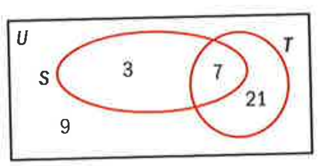
c $D$ = throw a prime number on a fair die numbered 1, 2, 3, 4, 5, 6, $E$ = throw an even number on the same die.

d $F$ = throw a prime number on a fair die numbered 1, 2, 3, 4, 5, 6, $G$ = throw an even number on another die.

e $G$ = choose a number at random from {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} that is at most 6, $H$ = choose a number from the same set that is at least 7.

f $M$ = choose a number at random from {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} that is no more than 5, $H$ = choose a number from the same set that is 4 or more.

g $S$ = choose a Spanish speaker at random from a set of students represented by this set, $T$ = choose a Turkish speaker at random from this set.




2 Use a Venn diagram to confirm that

a $A = (A \cap B) \cup (A \cap B')$ and that $(A \cap B)$ and $(A \cap B')$ are mutually exclusive events.

b $P(A) = P((A \cap B) \cup (A \cap B'))$

c $P(A) = P(A)\,P(B) + P(A \cap B')$

d Hence, $P(A \cap B') = P(A)(1 - P(B))$ so $A$ and $B'$ are independent.

3 Daniel throws a fair dice numbered with {1, 2, 3, 4, 5, 6} five times.

a Write down the probability that Daniel throws the sequence

i 4, 1, 3, 5, 2  ii 1, 1, 1, 1, 1

iii 6, 5, 4, 3, 2

b Find the probability that he throws at least one 3.

c Daniel simulates this experiment with a spreadsheet. Predict how many times in 10 000 trials he would expect to throw the sequence 1, 1, 1, 1.

d Find the probability that he throws a Yahtzee (all five numbers are equal).

e After how many throws would he expect to have thrown 3 Yahtzees?

4 Events $G$ and $H$ are such that $P(G) = 0.3$ and $P(H) = 0.6$.

a Find $P(G \cup H)$ when $G$ and $H$ are mutually exclusive;

b Find $P(G \cup H)$ when $G$ and $H$ are independent.

c Given that $P(G \cup H) = 0.63$, find $P(H \mid G)$.

5 Achille and Barbora throw a fair octahedral die until one of them throws an eight.

a Find the probability Barbora wins on one of her first two throws, if she throws first.

b Investigate if throwing first gives Barbora an advantage in this game.

6 $P(A) = 0.35$, $P(A \cup B) = 0.75$ and $P(A|B) = 0.35$. Find $P(B)$.

## Developing inquiry skills

Which of the events in the first opening scenario are independent? Which are mutually exclusive?





# 5.4 Complete, concise and consistent representations

You can use diagrams as a rich source of information when solving problems. Choosing the correct way to represent a problem is a skill worth developing. For example, consider the following problem:

In a class of 15 students, 3 study art and 6 study biology of which 1 studies art. A student is chosen at random. How many simple probabilities can you find? How many combined probabilities can you find?

Let $A$ represent the event "An art student is chosen at random from this group" and $B$ "A biology student is chosen at random from this group".

**HINT**

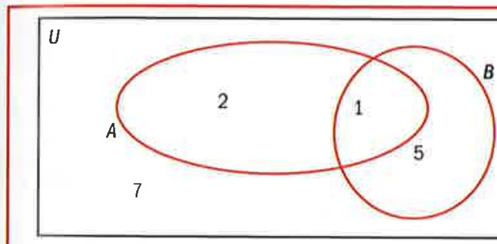
If you represent the problem only as text, the simple probabilities $P(A) = \frac{1}{5}$ and $P(B) = \frac{2}{5}$ can be found easily, but calculating these do not show you the whole picture of how the sets relate to each other.

**TOK**

In TOK it can be useful to draw a distinction between shared knowledge and personal knowledge. The IB use a Venn diagram to represent these two types of knowledge.

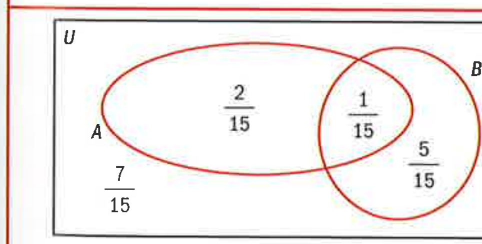
If you are to think about mathematics (or any subject, in fact), what could go in the three regions illustrated in the diagram?


Represent this information as follows in a **Venn diagram** to see more detail:




The rectangle represents the sample space $U$ for which $P(U) = 1$, the total probability.

The diagram allows us to find $P(B|A) = \frac{1}{3}$, $P(B'|A) = \frac{2}{3}$ etc easily.



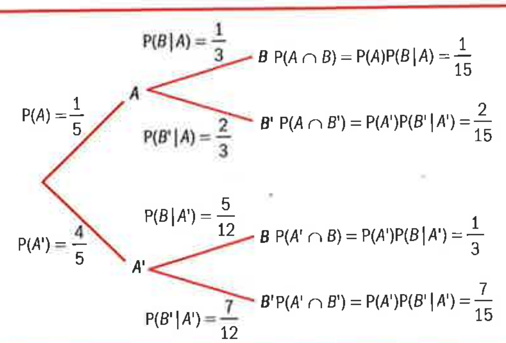

The Venn diagram can be adapted to show the distribution of the total probability in **four** regions that represent mutually exclusive events: $P(A \cap B) = \frac{1}{15}$, $P(A' \cap B) = \frac{5}{15}$, $P(A \cap B') = \frac{2}{15}$ and $P(A' \cap B') = \frac{7}{15}$.

Hence, the probability that a randomly chosen student studies neither biology nor art is $P(A' \cap B') = \frac{7}{15}$. The simple probability $P(B) = P(A \cap B) + P(A' \cap B) = \frac{1}{15} + \frac{1}{3} = \frac{2}{5}$ is represented as a union of two mutually exclusive events in the Venn diagram.

The Venn diagram can be used to find all the simple, combined and conditional probabilities.

Statistics and probability

A tree diagram representation of this problem is as follows:




The total probability of 1 is distributed along the branches of the tree using the multiplication law of probability.

The probability at the end of a branch is found by applying the multiplication law of probability. For example,

$P(A \cap B) = P(A)P(B|A) = \frac{1}{5} \times \frac{1}{3} = \frac{1}{15}$.

Notice that the simple probability $P(B) = P(A \cap B) + P(A' \cap B) = \frac{1}{15} + \frac{1}{3} = \frac{2}{5}$ can be found from summing the probabilities at the end of two branches of the tree diagram.

A tree diagram is another way to represent all the possible outcomes of an event. The end of each branch represents a combined event.

Choosing an appropriate diagram to represent a probability problem is an important stage in the problem-solving process.

## Investigation 5

Maria considers an experiment in which a bag contains $n_1$ dice of colour $c_1$ and $n_2$ dice of colour $c_2$. Two dice are chosen with replacement.

Maria works out the probability of choosing two dice of different colour in two ways after drawing a tree diagram.

**Method one:** $P(\text{two dice have different colours}) = \frac{n_1}{n_1+n_2} \times \frac{n_2}{n_1+n_2} + \frac{n_2}{n_1+n_2} \times \frac{n_1}{n_1+n_2}$

**Method two:** $P(\text{two dice have different colours}) = 1 - P(\text{two dice have same colour}) = 1 - \left(\frac{n_1}{n_1+n_2} \times \frac{n_1}{n_1+n_2} + \frac{n_2}{n_1+n_2} \times \frac{n_2}{n_1+n_2}\right)$

Maria reflects on the methods and decides that there is not a significant difference in terms of efficiency because in the first method she finds and adds two probabilities whereas in the second she finds two probabilities and subtracts their sum from 1 using complementary events.

Maria explores the same situation and methods but for $n_1$, $n_2$ and $n_3$ dice with distinct colours $c_1$, $c_2$ and $c_3$, respectively.

1 How many probabilities would be found and added using method one? A tree diagram may be useful to count the number of ways.

2 How many probabilities would be found, added and their total subtracted from 1 using method two?





Repeat for $n_1$, $n_2$, $n_3$ and $n_4$ dice with distinct colours $c_1$, $c_2$, $c_3$ and $c_4$, respectively.

3 **Factual** Hence, fill in the table:

| Number of dice | How many probabilities would be found and added from the tree diagram representation | How many probabilities would be found, added and their total subtracted from 1 using a complementary events representation. |
|---|---|---|
| 2 | 2 | 2 |
| 3 | | |
| 4 | | |
| 5 | | |
| $k$ | | |

4 **Factual** Can you generalize for $n_1, n_2, \ldots, n_k$ dice with distinct colours $c_1, c_2, \ldots, c_k$?

5 **Conceptual** Why is it useful to calculate probabilities for the complementary event in some situations?

The complementary probability law $P(A) = 1 - P(A')$ may be more efficient in some situations.

### International-mindedness

In 1933, Russian mathematician Andrey Kolmogorov built up probability theory from fundamental axioms in a way comparable with Euclid's treatment of geometry. His approach forms the basis for the modern theory of probability. Kolmogorov's work is available in an English translation titled *Foundations of the Theory of Probability*.

## Exercise 5D

1 A bag contains 12 green socks, 8 yellow socks and 7 red socks. Two socks are drawn at random without replacement. Find the probability that the two socks have the same colour.

2 A jewellery box contains 13 gold earrings, 10 silver earrings and 12 titanium earrings. Two earrings are drawn at random with replacement. Find the probability that they are made of different metals.

3 A supermarket uses two suppliers of strawberries, C and D. Supplier C provides 70% of the supermarket's strawberries. The strawberries are examined in a quality control inspection (QCI). It is found that 90% of the strawberries supplied by C pass QCI and 95% of the strawberries from D pass QCI.

A strawberry is selected at random.

**a** Find the probability that the strawberry passes QCI.

**b** Given that a strawberry passes QCI, find the probability that it came from supplier D.

**c** In a sample of 2000 strawberries, find the expected number of strawberries that would fail QCI.

**d** The supermarket wants the probability that a strawberry passes QCI to be 0.93. Find the percentage of strawberries that should be supplied by D to achieve this.

4 Chevy plays a game in which she throws a pair of fair cubical dice numbered {1, 2, 3, 4, 5, 6} 24 times. Find the probability that she throws at least one double six.

Statistics and probability

**5** A factory produces a large number of electric cars. A car is chosen at random from the production line as a prize in a competition. The probability that the car is blue is 0.5. The probability that the car has five doors is 0.3. The probability that the car is blue or has five doors is 0.6. Find the probability that the car chosen is not a blue car with five doors.

Pietro solves this problem with a Venn diagram but Maria solves it with a tree diagram. They both get the correct answer. Solve the problem both ways, discuss and then state who used the most efficient method.

**6 a** In a five-a-side football match, there are 10 players on the pitch plus the referee. Find the probability that they all have different birthdays. (Assume all birthdays are equally likely.)

**b** Find the smallest number of people needed on the football pitch so that the probability of at least two of them sharing a birthday is greater than 0.5.

**TOK**

During the mid-1600s, mathematicians Blaise Pascal, Pierre de Fermat and Antoine Gombaud puzzled over this simple gambling problem:

Which is more likely: rolling at least one six on four throws of one die or rolling at least one double six on 24 throws with two dice?

## Chapter summary

Three perspectives on probability:

- Subjective probability is derived from an individual's personal judgment about whether a specific outcome is likely to occur.
- For experimental and theoretical probability, you use the following terminology:
  - Experiment: A process by which you obtain an observation
  - Trials: Repeating an experiment a number of times
  - Outcome: A possible result of an experiment
  - Event: An outcome or set of outcomes
  - Sample space: The set of all possible outcomes of an experiment, always denoted by $U$
- We write $\mathrm{P}(A)$ to represent the probability of event $A$ occurring.
- Experimental probability is given by

$$\text{Relative frequency of } A = \frac{\text{Frequency of occurrence of event } A \text{ in } n \text{ trials}}{n}$$

- The theoretical probability of an event $A$ is $\mathrm{P}(A) = \dfrac{n(A)}{n(U)}$ where $n(A)$ is the number of outcomes that make $A$ happen and $n(U)$ is the number of outcomes in the sample space.





- Whenever $\mathrm{P}(A)$ represents a subjective, experimental or theoretical probability, then $0 \le \mathrm{P}(A) \le 1$.
- The expected number of occurrences of $A = n\mathrm{P}(A)$.

Three probability diagrams:

- A Venn diagram represents the sample space with a rectangle. Within the rectangle, each event is represented by a set of outcomes in a circle or an oval shape and is labelled accordingly.
- A sample space diagram is a useful way to represent the whole sample space and often takes the form of a table. It is especially useful when the situation involves combining two sets in some way to form a sample space.
- A tree diagram is a useful way to represent two or more combined events, often involving choices and conditional probabilities.
- Notation used to represent probabilities of combined events:

| Name | Symbol applied to events | Meaning |
|---|---|---|
| Intersection | $A \cap B$ | Events $A$ and $B$ both occur |
| Union | $A \cup B$ | Events $A$ or $B$ or both occur |
| Complement | $A'$ | Event $A$ does not occur |
| Conditional | $A \mid B$ | Event $A$ given that event $B$ has occurred |

Probability laws which are always true:

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

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

$\mathrm{P}(A) + \mathrm{P}(A') = 1$

Types of combined events and the special cases of probability laws that follow:

| Type of event | Consequences |
|---|---|
| Events $A$ and $B$ are mutually exclusive if they cannot both occur. | $\mathrm{P}(A \cap B) = \mathrm{P}(B \cap A) = 0$<br>$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$ |
| Events $A$ and $B$ are independent if the occurrence of each event does not affect in any way the occurrence of the other. | $\mathrm{P}(A \mid B) = \mathrm{P}(A)$<br>$\mathrm{P}(B \mid A) = \mathrm{P}(B)$<br>$\mathrm{P}(A \cap B) = \mathrm{P}(B \cap A) = \mathrm{P}(A)\mathrm{P}(B)$ |
| The events $A$ and $A'$ are complementary. Hence, $A$ and $A'$ are mutually exclusive. | $\mathrm{P}(A) + \mathrm{P}(A') = 1$ |

- The complementary probability law $\mathrm{P}(A) = 1 - \mathrm{P}(A')$ can give you a quick way to solve problems.

Statistics and probability

## Developing inquiry skills

Apply what you have learned in this section to represent the first opening problem with a tree diagram.

Hence, find the probability that a cab is **identified** as yellow.

Apply the formula for conditional probability to find the probability that the cab was yellow **given that** it was identified as yellow.

How does your answer compare to your original subjective judgement?

## Chapter review

Click here for a mixed review exercise

**1** $B$ and $C$ are independent events. $P(B \cap C) = 0.1$ and $P(B \cap C') = 0.4$. Find $P(B' \cup C)$.

**2** $P(X) = \frac{2}{3}$, $P(X|Y') = \frac{7}{12}$, $P(X|Y) = 0.8$

- **a** Find $P(Y)$.
- **b** Determine if $X$ and $Y$ are independent events.

**3** **a** Find the probability of the outcome "throw at least one six" when a fair cubical die is thrown 1, 2, 3, 4, 5, ..., $n$ times.

- **b** Hence, find the least value of $n$ for which P(throw at least one six in $n$ throws) is 99.5%

**4** A packet of seeds contains 65% green and 35% red seeds. The probability that a green seed grows is 0.85 and that a red seed grows is 0.74. A seed is chosen at random from the packet.

- **a** Calculate the probability that the seed grows.
- **b** Calculate the probability that the seed is green and grows.
- **c** Calculate the probability that the seed is red or it grows.

**5** Each **odd** number from 1 to $5n$ where $n$ is odd is written on a piece of paper and placed in a box.

- **a** Calculate how many pieces of paper there are in the box.
- **b** Find the probability in terms of $n$ that a paper selected at random from the box shows a number that is divisible by 5.

**6** A company fleet has six blue and $n$ white cars. Two cars are chosen without replacement. The probability that two blue cars are chosen is $\frac{1}{7}$. Find the value of $n$.

**7** The genes in human chromosome pairs determine if a child is male or female. Males have an X and Y chromosome pair and females an X and X pair. Inheriting the XY combination causes male characteristics to develop and XX causes female characteristics to develop. Sperm contain an X or Y with equal probability and an egg always contains only an X. An infant inherits one chromosome determining gender from each parent.

- **a** Complete the following *Punnett square* to show the possible outcomes of pairs of chromosomes that can be inherited:

| | | Chromosome inherited from mother | |
|---|---|---|---|
| | | X | X |
| Chromosome inherited from father | X | XX | |
| | Y | XY | |

- **b** Hence, show that the probability that a child is born female is 0.5.

## Exam-style questions

**8** **P1:** A box contains 16 chocolates, of which 3 are known to contain nuts.

Two chocolates are selected at random.

Find the probability that

- **a** exactly one chocolate contains nuts (3 marks)
- **b** at least one chocolate contains nuts. (3 marks)

**9** **P2:** Hamid must drive through three sets of traffic lights in order to reach his place of work. The probability that the first set of lights is green is 0.7.

The probability that the second set of lights is green is 0.4.

The probability that the third set of lights is green is 0.8.

It may be assumed that the probability of any set of lights being green is independent of the others.

- **a** Find the probability that all three sets of lights are green. (2 marks)
- **b** Find the probability that only one set of lights is green. (3 marks)
- **c** Given that the first set of lights is red (ie not green), find the probability that the following two pairs of lights will be green. (2 marks)
- **d** Find the probability that at least one set of lights will be green (3 marks)

**10** **P1:** Jake and Elisa are given the same mathematics problem.

The probability that Jake can solve it is 0.35.

If Jake has solved it, the probability that Elisa can solve it is 0.6, otherwise it is 0.45.

- **a** Draw a tree diagram to illustrate the above situation, showing clearly the probabilities on each branch. (3 marks)
- **b** Find the probability that at least one of the students can solve the problem. (2 marks)
- **c** Find the probability that Jake solves the problem, given than Elisa has. (4 marks)

**11** **P1:** $A$ and $B$ are events such that $P(A) = 0.3$, $P(B) = 0.65$ and $P(A \cup B) = 0.7$

By drawing a Venn diagram to illustrate these probabilities, find:

- **a** $P(A' \cap B)$ (2 marks)
- **b** $P(A \cup B')$ (2 marks)
- **c** $P(A \cap B)'$ (2 marks)

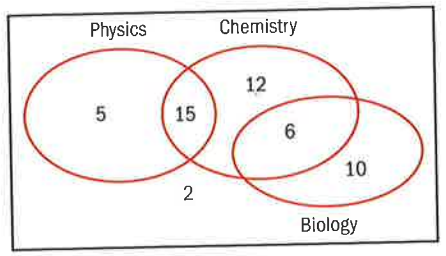
**12** **P1:** In a survey, 48 people were asked about their holidays over the past year. It was found that 32 people had taken a holiday in Europe, and 25 people had taken a holiday in the USA.

Everyone surveyed had been to at least Europe or the USA.

- **a** Determine how many people had taken a holiday in both Europe and the USA. (2 marks)
- **b** Find the probability that a randomly selected person had been to Europe, but not the USA. (3 marks)
- **c** Explain why the events "taking a holiday in Europe" and "taking a holiday in the USA" are not independent events. (3 marks)

Statistics and probability

**13 P2:** The Venn diagram illustrates the number of students taking each of the three sciences: Physics, Chemistry and Biology.




A student is randomly chosen from the group.

Find the probability that

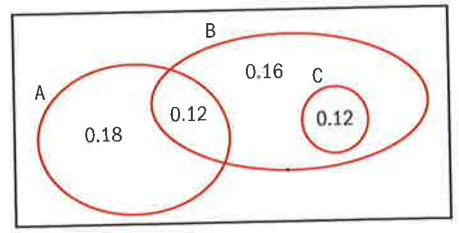
- **a** the student studies Chemistry or Biology (2 marks)
- **b** the student studies neither Physics nor Biology (2 marks)
- **c** the student studies Physics, given that they study Chemistry (2 marks)
- **d** the student studies Biology, given that they study Physics (2 marks)
- **e** the student studies Physics, given that they do not study Biology. (2 marks)

**14 P1:** $A$ and $B$ are independent events, such that $P(A) = 0.3$ and $P(B) = 0.5$.

Find the following probabilities.

- **a** $P(A \cap B)$ (2 marks)
- **b** $P(A \cup B)$ (2 marks)
- **c** $P(B' \cap A)$ (2 marks)
- **d** $P(B|A')$ (3 marks)

**15 P2:** The Venn diagram below shows the probabilities for three events $A$, $B$ and $C$.




- **a** Justify that events $B$ and $C$ are not independent. (2 marks)
- **b** Explain why events $A$ and $C$ are mutually exclusive. (2 marks)
- **c** Determine whether events $A$ and $B$ are independent. (4 marks)
- **d** Determine whether events $A'$ and $C'$ are mutually exclusive. (2 marks)
- **e** Find $P(A \cap C')$. (2 marks)

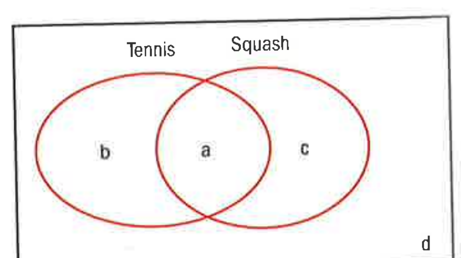
**16 P2:** At a local sports centre, members can either play tennis or squash.

The probability that a member plays tennis is 0.8.

Given that a member plays squash, the probability that they play tennis is 0.8.

The probability that a member does not play squash is 0.1.

The information is illustrated by the following Venn diagram, where $a, b, c$ and $d$ are probabilities:








- **a** Find the value of $a$ and the value of $c$. (5 marks)
- **b** Find the value of $b$ and the value of $d$. (3 marks)
- **c** Find the probability that a randomly chosen member plays only one sport. (2 marks)
- **d** Find the probability that a randomly chosen member plays tennis given that they do not play squash. (2 marks)
- **e** Find the probability that a randomly chosen member does not play squash, given that they do not play tennis. (2 marks)
- **f** Let $S$ be the event that a member plays squash, and let $T$ be the event that a member plays tennis. Determine whether:
  - **i** events $S$ and $T$ are mutually exclusive (1 mark)
  - **ii** events $S$ and $T$ are independent. (2 marks)

# Random walking!

**Approaches to learning:** Critical thinking

**Exploration criteria:** Mathematical communication (B), Personal engagement (C), Use of mathematics (E)

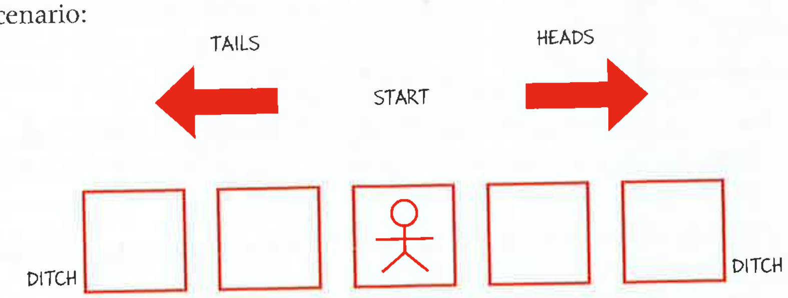
**IB topic:** Probability, Discrete distributions

## The problem

A man walks down a long, straight road. With each step he either moves left or right with equal probability. He starts in the middle of the road. If he moves three steps to the left or three steps to the right, he will fall into a ditch on either side of the road. The aim is to find probabilities related to the man falling into the ditch, and in particular to **find the average number of steps he takes before inevitably falling into the ditch.**

## Explore the problem

Use a counter to represent the man and a "board" to represent the scenario:










Toss a coin.

Let a tail ($T$) represent a left step and a head ($H$) represent a right step.

Write down the number of tosses/steps it takes for the man to fall into the ditch.

Do this a total of 10 times.

Calculate the average number of steps taken.

Construct a spreadsheet with the results from the whole class.

Calculate the average number of steps taken from these results.

How has this changed the result?

Do you know the actual average number of steps required?

How could you be certain what the average is?

## Calculate probabilities

Construct a tree diagram that illustrates the probabilities of falling into the ditch within five steps.

Use your tree diagram to answer these questions:

What is the probability associated with each sequence in which the man falls into the ditch after a total of exactly five steps?

What is the probability that the man falls into the ditch after a total of exactly five steps?

What is the minimum number of steps to fall into the ditch?

What is the maximum number?

What is the probability that the man falls into the ditch after a total of exactly three steps?

Explain why all the paths have an odd number of steps.

Let $x$ be the number of the steps taken to fall into the ditch.

Copy and complete this table of probabilities:

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | .... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X=x)$ | | | | | | | | | | | | | |

Look at the numbers in your table.

Can you see a pattern?

Could you predict the next few entries?

## Simulation

Since there is an infinite number of values of $x$, calculating the expected number of steps to fall into the ditch would be very complicated.

An alternative approach is to run a computer simulation to generate more results, and to calculate an average from these results.

You can write a code in any computer language available that will run this simulation as many times as needed.

This will allow you to improve on the average calculated individually and as a class.

Although this would not be a proof, it is convincing if enough simulations are recorded.

### Extension

Once you have a code written you could easily vary the problem.

What variations of the problem can you think of?

You may also be able to devise your own probability question which you could answer using simulation.