Before you start

You should know how to:

1 Find simple probabilities

A number is chosen at random from the set {1, 2, 3, …, 100}.

Find the probability of choosing a cube number.

Cube numbers in the set are: 1, 8, 27, 64.

Probability of a cube number $= \frac{4}{100} = 0.04$

2 Find the mean from a grouped frequency table without technology

eg

x_i	4	7	8	10
f_i	2	1	5	7

$$\text{Mean} = \frac{\sum f_i x_i}{\sum f_i} = \frac{125}{15} \approx 8.33$$

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 a number is:

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

2 A student investigates whether there is a relationship between gender and prevalence of smoking.

She 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.

3

x_i	1	2	3	6	11
f_i	9	7	3	2	1

Find the mean value of x.

7.1 Theoretical and experimental probability

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

Subjective probability

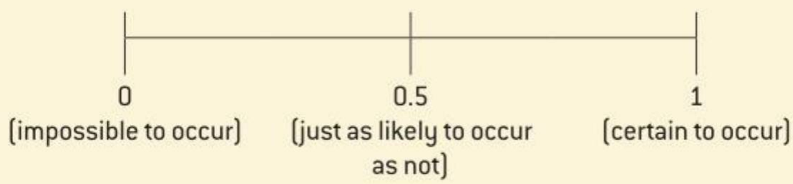
You may judge that you are more likely to get to school on time if you take a particular route, based on your experience with traffic. Subjective probabilities are based on past experiences and opinions rather than formal calculations.

Investigation 1

1 **a** Discuss the likelihood of these outcomes.

A: The team winning the FIFA World Cup in 2030 will be from the Americas.	B: It will rain tomorrow.	C: Humans will reach Mars by 2050.
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 thousand digits of pi.

b Display your answers by plotting them on this probability scale:



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

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

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

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.

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, and 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: $\{2, 3, 5, 7\}$.

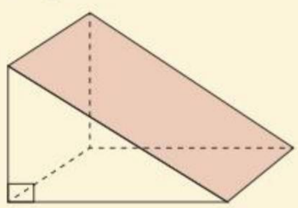
Erin can either write **P(throw a prime)** to represent the probability of her event occurring or **P(A)** if A denotes the set $\{2, 3, 5, 7\}$.

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 as equally likely as any other.

Investigation 2

1 Use card to build a triangular prism where the constant cross-sectional area is a right-angled triangle.



2 In your class, decide how to label the faces with A, B, C, D and E.

- Which face(s) do you predict are least likely to land face down when your prism is thrown?
- Which are most likely?
- **Discuss** in a group and try to rank the five outcomes in order of likelihood.

3 Throw your prism for 50 trials and record the frequency of each outcome. Store your data so that it can be used for further investigation in Section 7.3.

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:

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.

4 Find the experimental probability of each outcome by completing a table:

Outcome	Frequency	Relative frequency (exact)	Relative frequency to 4 sf
A			
B			
C			
D			
E			
Totals	50 trials		

5 **Factual** What do the experimental probabilities add up to?

6 **Factual** What is the range of values that an experimental probability can have?

7 **Factual** What do you notice about how the smaller and larger experimental probabilities relate to each other in the context of the die?

8 **Collaborate** by adding your data to that of a friend to make a set of 100 trials. Do any patterns emerge? Then add up all the data in your class. Compare and contrast your experimental probabilities with your subjective probabilities from **2**.

9 **Conceptual** How can you have a more accurate estimate of the experimental probability?

Theoretical probability

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

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.

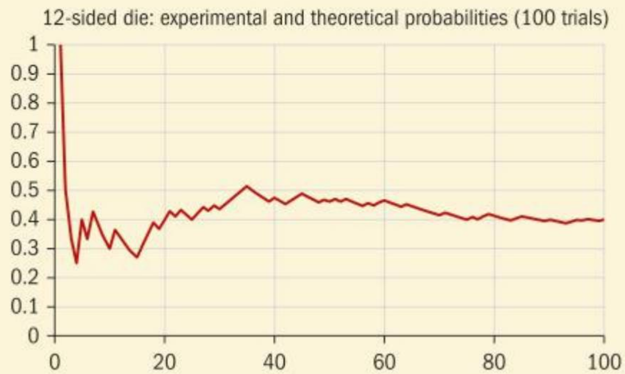
Whenever $\mathrm{P}(A)$ represents a subjective, experimental or theoretical probability, then $0 \le \mathrm{P}(A) \le 1$.

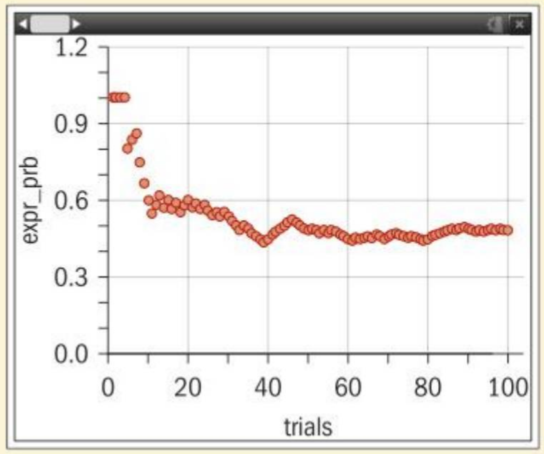
Investigation 3

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

2 Follow the instructions on p.307 to build a spreadsheet or GDC document that shows the sequence of experimental probabilities of A after 1, 2, 3, ..., 100 throws. You should be able to create an image like one of these:

Continued on next page





Number of trials (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
13	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. Carry out another 100 trials using your spreadsheet or GDC.

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

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

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

9 **Conceptual** Does random behaviour involve predictability in the short term or unpredictability?

10 **Conceptual** Does random behaviour involve predictability in the long term 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 decide 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 six-sided die.

I: throw a multiple of 5 on a 20-sided die.

All the dice are fair.

$n(T) = n(\{1, 2, 3, 4\}) = 4 = n(U)$ so $P(T) = 1$	Every element of $\{1, 2, 3, 4\}$ is a factor of 24. $P(T) = \frac{n(T)}{n(U)} = \frac{4}{4} = 1$, so T is certain to happen.
$n(O) = n(\{2, 3, 5, 7\}) = 4$, $n(U) = 8$ so $P(O) = \frac{4}{8} = 0.5$	"at least 11" means "11 or more".
$n(D) = n(\{11, 12\}) = 2$, $n(U) = 12$ so $P(D) = \frac{2}{12} = \frac{1}{6} = 0.1\dot{6}$	
$n(C) = n(\{1, 2, 3\}) = 3$, $n(U) = 6$ so $P(C) = \frac{3}{6} = 0.5$	"at most 3" means "3 or fewer".
$n(I) = n(\{5, 10, 15, 20\}) = 4$, $n(I) = 20$ so $P(I) = \frac{4}{20} = 0.25$ Hence D is the least likely event.	

Exercise 7A

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

2 A dartboard has 20 sectors 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 Ann and Ruth are designing a game for a CAS project. The numbers 1, 2, 3, …, 11 are written on identical tickets and one ticket is drawn at random from an envelope. Find the probability that the number on the ticket drawn is:

a odd
b square
c prime
d square and odd
e square and prime
f prime and odd
g a prime and even.

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 is:

a equal to 0000
b less than 8000 and more than 7900
c divisible by 10
d at least 13.

Investigation 4

On 18 August 1913, at the casino in Monte Carlo, gamblers betting on whether the roulette ball would fall into a black or red slot witnessed a very unusual event. (For simplicity, assume that the roulette wheel only has red or black slots for the ball to fall into, and these are equal in number.)

The ball fell into the black slot 26 times in succession, a record. After 15 successive blacks, gamblers excitedly rushed to bet heavily on red, hugely increasing their bets with every time the ball ended up in black once more. Consequently, at the end of the sequence of 26 blacks, the casino made a huge profit because of the choices made by the gamblers.

Discuss in a group: Why were the gamblers panicking? What would the most rational response be to the 15 consecutive blacks? Have you ever experienced an unusual set of outcomes like the one in this story? Can we predict how long "unusual runs" are and when to expect them?

Ideas for further exploration: What is the maturity doctrine? How is it related to the representativeness fallacy? What about cognitive illusions and biases?

Discuss these questions in a group:

1 Discuss the validity of these three statements that could have been made after 15 blacks:

P: "It has to be black next—black is more likely today!"

Q: "It has to be red next—it's red's turn now!"

R: "There is clear evidence of cheating on the part of the casino."

T: "P and Q are just as poor examples of reasoning as each other."

2 How would you classify how we quantify probabilities? Assign each word from the top row to one or more of the three types and complete the table.

Rational, Emotional, Empirical, Subjective, Objective, Abstract, Concrete, Intuitive, Random, Predictive, Descriptive, Interpretive, Exact, Approximate, Data, Trials, Sample space, Equally likely outcomes, Personal knowledge, Shared knowledge		
Subjective probability	**Experimental probability**	**Theoretical probability**

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.

TOK

Do ethics play a role in the use of mathematics?

Investigation 5

You've established that P(throw a prime number on a fair 12-sided die) $= \frac{5}{12}$.

1 If we threw such a die 2512 times, how many times do you **expect** that a prime number will be thrown?

2 **Think** about and then **write down** your own intuitive answer to this question. **Discuss** your answer with a friend then **share** your ideas with your class. **Test** your prediction with a spreadsheet by following these steps:

- Enter "=RANDBETWEEN(1, 12)" in cell A1 of a new spreadsheet document.
- Copy this down to cell A2512 to create 2512 trials.
- Enter "=IF(OR(A1 = 2, A1 = 3, A1 = 5, A1 = 7, A1 = 11), 1, 0)" in cell B1 and copy down to cell B2512.
- Enter "=SUM(B1:B2512)" in cell C1. This is the frequency of primes occurring in your 2512 trials.
- Press F9 to refresh the trials and observe how the frequency compares to what you expected.
- Adapt your spreadsheet for dice with different numbers of sides and a different number of trials.

3 **Predict** other numbers of occurrences using a spreadsheet.

4 **Factual** For 128 trials, what is the expected number of occurrences of the event "throw a square number on a fair 12-sided die"?

5 Show that the formula

Expected number of occurrences of $\mathrm{A} = n\mathrm{P}(A)$ follows from the formula

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

and the fact that you model relative frequency with theoretical probability in the long term.

6 **Conceptual** Why is the formula for an **expected** value and not just a value?

7 **Conceptual** How may we interpret and apply the number quantified by the formula for the expected number of occurrences?

Statistics and probability

Example 2

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

b Statistical data built up over five 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 you expect to be absent on any given day and interpret your answer.

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

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

Example 3

A coastal town carries out a survey of tourists in July 2018. The accommodation type chosen and the age of the tourist are recorded. The information is given in the table below:

		Accommodation type			
		Hotel	Campervan	Tent	Apartment
Age	18–30	67	81	125	32
	31–50	107	230	73	119
	51–70	87	76	34	89
	>70	109	32	15	54

a A tourist is chosen at random from the survey to receive a promotional offer. Determine what age group the tourist is most likely to belong to.

b Find the probability that the tourist chosen belongs to this age group, expressing your answer as a decimal.

c In July 2019, 16 000 tourists are predicted to visit the town. If the town's hotels have capacity for 5000 predict if the town has enough hotel capacity to meet the demand.

a The table shows that the total number of tourists in the 31–50 age group is $107 + 230 + 73 + 119 = 529$, and this is the most frequent age group.

Apply technology to manage tables of data efficiently. Cell E1 has the formula shown:

E1 | fx | =SUM(A1:D1)

	A	B	C	D	E
1	67	81	125	32	305
2	107	230	73	119	529
3	87	76	34	89	286
4	109	32	15	54	210
5	370	419	247	294	1330

b Finding the totals of all the other columns and adding them together gives 1330. Hence the probability required is $\frac{529}{1330} \approx 0.398$	Apply the formula for theoretical probability.
c The table shows that the total number of tourists choosing a hotel for their accommodation is $67 + 107 + 87 + 109 = 370$. Hence the expected number of hotel places is $\frac{370}{1330} \times 16\,000 \approx 4451$	Apply the formula for the expected value. Expressing the value to the nearest whole number is appropriate.
There is sufficient capacity.	State the conclusion.

Exercise 7B

1 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 was randomly selected.

a Find the probability that this shopper travelled by car.

One male shopper was randomly selected.

b Find the probability that this male shopper travelled on foot.

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

2 In an experiment, a number a is chosen at random from the set $\{2, 3, 4, 5\}$ and a number b is chosen at random from the set $\{3, 4, 5, 6\}$.

a Find the probability that $\frac{b}{a}$ is a natural number.

b The experiment is repeated 320 times. Find the expected number of times that $a - b$ will be positive.

3 A fair dodecahedral die has faces numbered 1, 2, 3, …, 12. The die is thrown 154 times. Find the expected number of times that the die will show:

a a factor of 12 **b** a prime number

c a prime factor of 12.

4 The probabilities of each outcome of a biased die are modelled with the following theoretical probabilities:

Outcome	1	2	3	4	5	6
Probability	0.15	0.2	0.25	0.2	0.12	0.18

The die is thrown 207 times. Find the expected number of odd outcomes.

5 A large jar contains 347 marbles, 125 of which are red. A marble is chosen at random and replaced.

Find the expected number of times a non-red marble is chosen in 531 trials.

6 A letter is chosen at random from the word "ICOSAHEDRAL". Find the expected number of times a vowel is chosen in 79 trials.

7 Each day in June, Maged records types of cars passing a point on a road popular with tourists. The percentages of each type of car are given in the table below:

Type of car	Classic	Luxury	Compact	Family saloon	Estate	SUV	Other
Percentage	0.5%	1.2%	23.1%	30.9%	15.4%	19.8%	9.1%

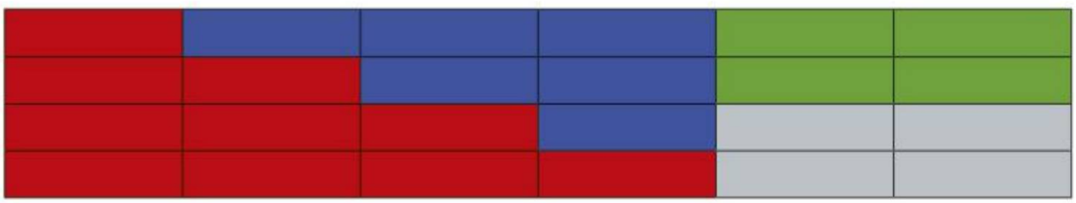
a Use this information to predict the number of classic or luxury cars Maged would expect to observe in July given that he observes 573 cars in July.

b State the assumption made in your answer.

8 Quality control is being carried out in a clothing factory. 1.37% of garments produced by machine A have defects and 0.41% of garments produced by machine B have defects.

A quality control manager inspects 67 garments from machine A and 313 garments from machine B. Find the expected number of defects.

9 An artist chooses an area of her neighbourhood to photograph by throwing a dart each morning at a large map pinned to her studio wall.



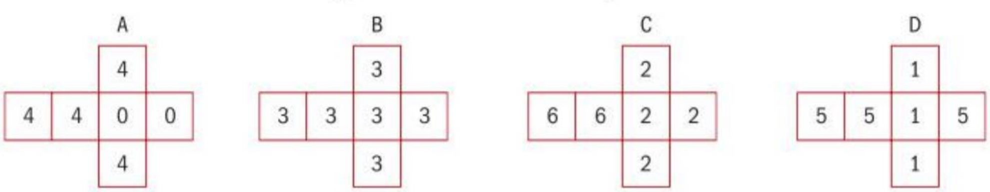
Key:

Red = central business district Blue = government buildings

Green = park Grey = housing

Find the expected number of times the artist will not photograph government buildings given that she takes photographs each day during August.

10 These dice compete in the "Dice world cup". The semifinals are A vs B and C vs D. The winners of each semifinal go in to the world cup final.



Write down your subjective judgment of which die you feel will be most likely to win the "Dice world cup".

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?