Mathematics Level 7

Chapter 7 — Probability

Aboriginal and Torres Strait Islander content has been excluded as accuracy cannot be guaranteed, and it is better to be respectful than cover the entire curriculum inaccurately.

Significant effort has been made to ensure this content is legal. Please send any areas of concern to admin@vwx.au with appropriate document/legal references so errors can be verified and changes be made.

Contents

Sample spaces and probabilities for single-stage experiments

Content description VC2M7P01: identify the sample space for single-stage experiments; assign probabilities to the possible outcomes and predict relative frequencies for related experiments

Achievement standard (extract): Students list sample spaces for single-step experiments, assign probabilities to outcomes of events and predict relative frequencies for related events.

Theory

The 0–1 scale, brought back from Level 6. At Level 6 you placed probabilities on a scale running from 0 to 1: a probability of 0 means the event can never happen (**impossible**) and a probability of 1 means it always happens (**certain**). Figure 1 brings that scale back. An event with probability $\frac{1}{2}$ has an **even chance**, a probability close to 0 is **unlikely**, and a probability close to 1 is **likely**. The bigger the probability, the more likely the event is to happen.

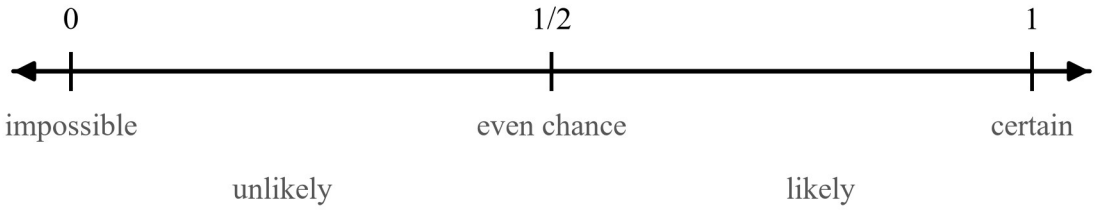

Figure 1. A probability scale from 0 to 1, with 0 labelled impossible, 1/2 labelled even chance and 1 labelled certain, and the words unlikely and likely placed between them.

The words of this subtopic. A chance set-up, such as throwing a dice or spinning a spinner, is called an **experiment**. One go of the experiment is a **trial**, and the result of one trial is an **outcome**. The list of *all* the outcomes that could happen is the **sample space**. An **event** is any outcome, or group of outcomes, that we care about, and the outcomes that make the event happen are its **favourable outcomes**. A **probability** is a number between 0 and 1 that says how likely an event is.

Term	Meaning here	Related words
experiment	a chance set-up, like throwing a dice	game, chance set-up
trial	one go of the experiment	one throw, one spin, one toss
outcome	the result of one trial	result
sample space	the list of all possible outcomes	all results
event	an outcome, or group of outcomes, we care about	"rolling an even number"
favourable outcome	an outcome that makes the event happen	a winning result
probability	a number from 0 to 1 saying how likely an event is	chance, likelihood

Listing sample spaces. Every experiment in this subtopic is **single-stage**: it is one action with one result — one throw of a dice, one toss of a coin, one spin of a spinner, or one pick from a lucky dip. Listing the sample space means writing down every outcome that could happen, once each.

- Throwing a dice: the sample space is $\{1, 2, 3, 4, 5, 6\}$ (Figure 2).
- Tossing a coin: the sample space is $\{H, T\}$ for heads and tails (Figure 3).
- Spinning the spinner in Figure 4: the sample space is $\{R, B, G, Y\}$.
- Picking one ticket from a lucky dip holding red and blue tickets: the sample space is $\{\text{red, blue}\}$.

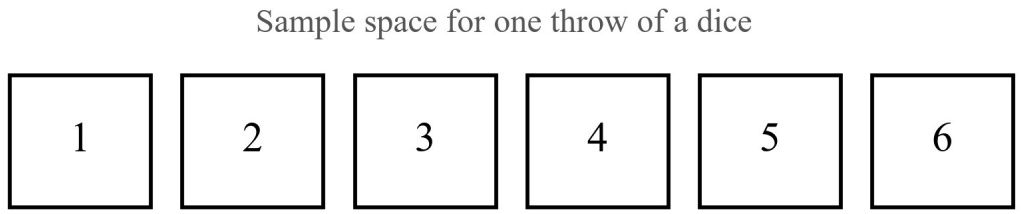

Figure 2. Six boxes in a row labelled 1 to 6: the sample space for one throw of a dice.

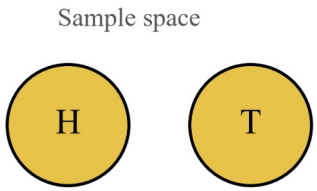

Figure 3. Two coins side by side, one showing H and one showing T: the sample space for one toss of a coin.

Equally likely outcomes. On a fair dice, no face is easier to throw than another: the six outcomes are **equally likely**. The spinner in Figure 4 has four equal sections, so its four outcomes are equally likely too, and each has probability $\frac{1}{4}$. When the outcomes of an experiment are equally likely, every single outcome has the same probability, found by dividing 1 by the number of outcomes.

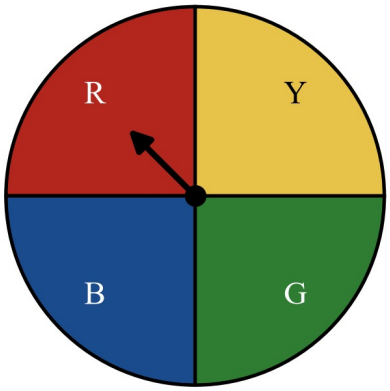

Figure 4. A spinner cut into four equal sections coloured red, blue, green and yellow, with a pointer: four equally likely outcomes.

Assigning probabilities. When the outcomes are equally likely, the probability of an event is the number of favourable outcomes divided by the number of outcomes in the sample space: $P(\text{event}) = \frac{\text{number of favourable outcomes}}{\text{number of outcomes}}$. Here P stands for "probability of". Figure 5 shows some of these fractions sitting at their true places on the 0–1 scale: $\frac{1}{6}$ sits close to 0, $\frac{5}{6}$ sits close to 1, and $\frac{1}{2}$ sits in the middle. Two facts are always true: every probability sits between 0 and 1, and the probabilities of *all* the outcomes in a sample space add up to 1.

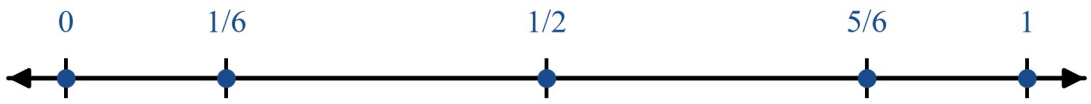

Figure 5. A probability scale from 0 to 1 with the fractions 1/6, 1/2 and 5/6 marked at their true positions between 0 and 1.

Predicting counts for related experiments. A probability also predicts what happens over many trials. If $P(6)=\frac{1}{6}$ on a fair dice, then in 60 throws we predict about $\frac{1}{6}\times 60=10$ sixes, and in 600 throws about $\frac{1}{6}\times 600=100$. The rule is:

$$\text{predicted number}=\text{probability}\times\text{number of trials}$$

The share of trials we predict for the outcome, 10 out of 60, is its predicted **relative frequency**, and it matches the probability: $\frac{10}{60}=\frac{1}{6}$. A prediction is a best guess for the long run, not a promise: a real run of 60 throws will usually give a number *near* 10, not exactly 10.

Worked examples

Example 1 — scaffolded

A fair dice is thrown once. (a) List the sample space. (b) Assign $P(6)$. (c) Assign $P(\text{even})$.

Working	Comment
(a) Sample space $=\{1,2,3,4,5,6\}$: six outcomes.	List every outcome once.
(b) $P(6)=\frac{1}{6}$.	One favourable outcome (a 6) out of six.
(c) The even outcomes are $2,4,6$ (shaded in Figure 6): three favourable outcomes, so $P(\text{even})=\frac{3}{6}=\frac{1}{2}$.	Count the favourable outcomes, then divide by six.

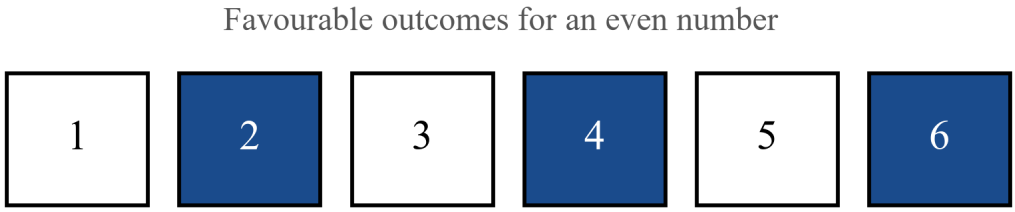

Figure 6. Six boxes labelled 1 to 6 with the boxes 2, 4 and 6 shaded: the favourable outcomes for throwing an even number.

Example 2 — scaffolded

The spinner in Figure 7 has eight equal sections: 3 red, 2 blue, 2 green and 1 yellow. It is spun once. Assign $P(\text{red})$, $P(\text{blue})$ and $P(\text{yellow})$, and check that all four probabilities add to 1.

Working	Comment
The sample space has 8 equally likely sections.	One trial is one spin.
$P(\text{red})=\frac{3}{8}$	3 red sections out of 8.
$P(\text{blue})=\frac{2}{8}=\frac{1}{4}$, $P(\text{green})=\frac{2}{8}=\frac{1}{4}$	2 sections each.
$P(\text{yellow})=\frac{1}{8}$	1 yellow section out of 8.
Check: $\frac{3}{8}+\frac{2}{8}+\frac{2}{8}+\frac{1}{8}=\frac{8}{8}=1$. ✓	All outcomes must add to 1.

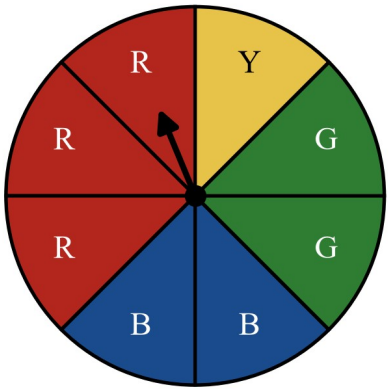

Figure 7. A spinner cut into eight equal sections: three red, two blue, two green and one yellow, with a pointer.

Example 3 — unscaffolded

A bag holds 3 red, 2 blue and 5 green marbles. One marble is drawn from the bag without looking. Write $P(\text{red})$ and $P(\text{green})$ as fractions in simplest form, and $P(\text{blue})$ as a decimal.

The sample space has $3+2+5=10$ equally likely marbles. Then $P(\text{red})=\frac{3}{10}$ and $P(\text{green})=\frac{5}{10}=\frac{1}{2}$, and $P(\text{blue})=\frac{2}{10}=0.2$.

Example 4 — unscaffolded

Assign the probability of throwing a 6 on a fair dice, and use it to predict the number of sixes in 60 throws and in 30 throws.

On a fair dice, $P(6)=\frac{1}{6}$, because one of the six equally likely outcomes is a 6. The predicted number of sixes is the probability times the number of trials. In 60 throws: $\frac{1}{6}\times 60=10$, so we predict about 10 sixes (Figure 8 shades those 10 throws). In 30 throws: $\frac{1}{6}\times 30=5$, so we predict about 5 sixes. Both predictions are best guesses for the long run, not exact promises.

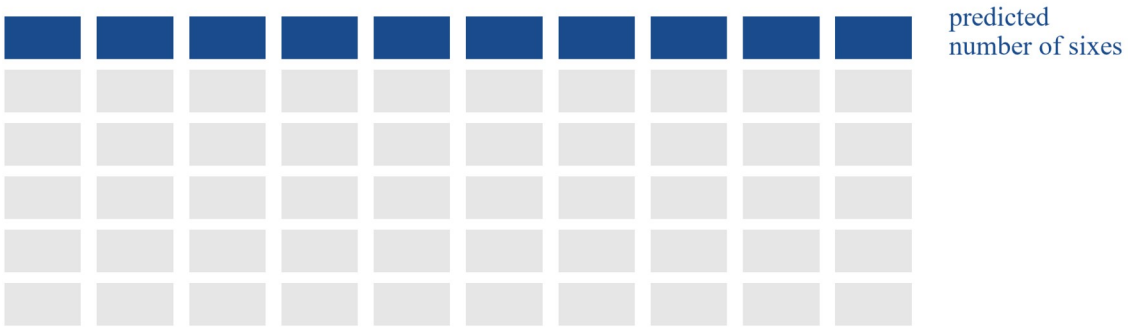

Figure 8. Sixty squares drawn as six rows of ten; the top row of ten is shaded to show the predicted number of sixes in 60 throws of a dice.

Practice questions

Scaffolded

Q1. List the sample space for each single-stage experiment. (a) tossing a coin (b) throwing a dice (c) spinning the spinner in Figure 4 (d) picking one ticket from a lucky dip holding red and blue tickets

Q2. A fair dice is thrown once (Figure 6 helps with part (c)). (a) How many outcomes are in the sample space? (b) Assign $P(3)$. (c) Assign $P(\text{even})$. (d) Assign $P(\text{number less than } 7)$.

Q3. The spinner in Figure 7 is spun once. Assign each probability as a fraction in simplest form. (a) $P(\text{red})$ (b) $P(\text{blue})$ (c) $P(\text{green})$ (d) $P(\text{yellow})$

Q4. The arrows A, B and C on the probability scale in Figure 9 each mark the probability of an event. Match each arrow to the word that best describes the event: *unlikely*, *even chance*, *likely*.

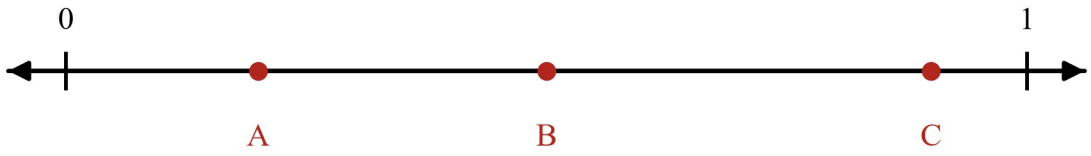

Figure 9. A probability scale from 0 to 1 with three arrows marked A at one fifth of the way, B at the middle and C near 1.

Q5. A bag holds 3 red, 2 blue and 5 green marbles. One marble is drawn without looking. Assign each probability. (a) $P(\text{red})$ (b) $P(\text{blue})$ as a fraction in simplest form (c) $P(\text{green})$

Q6. A lucky dip holds 20 tickets, and 4 of them win a prize. One ticket is picked without looking. (a) Assign $P(\text{prize})$ as a fraction in simplest form. (b) Predict the number of prizes in 100 picks.

Unscaffolded

Q7. A coin is tossed once. Assign $P(\text{heads})$.

Q8. A fair dice is thrown once. Assign each probability in simplest form. (a) $P(\text{a number greater than } 4)$ (b) $P(\text{odd})$

Q9. The spinner in Figure 10 has six equal sections numbered 1 to 6. It is spun once. Assign each probability in simplest form. (a) $P(\text{even})$ (b) $P(\text{a number less than } 3)$

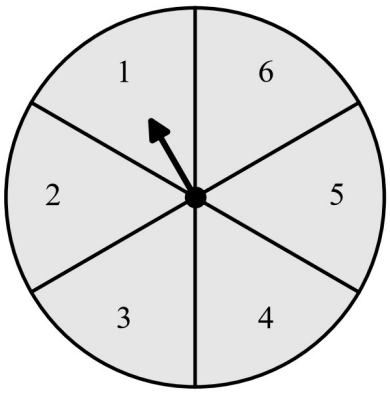

Figure 10. A spinner cut into six equal sections numbered 1 to 6, with a pointer.

Q10. A bag holds 12 marbles, some red and some blue. $P(\text{red}) = \frac{1}{4}$. How many red marbles are in the bag?

Q11. Match each probability to the word that best describes the event: *impossible*, *even chance*, *certain*. (a) 0 (b) $\frac{1}{2}$ (c) 1

Q12. A fair dice is thrown once. Say whether each event is impossible or certain, and give its probability. (a) the dice shows a number less than 7 (b) the dice shows an 8

Q13. A fair dice is thrown 60 times. Predict the number of sixes (Figure 8 shows one way to picture it).

Q14. A coin is tossed 50 times. Predict the number of heads.

Q15. The spinner in Figure 7 is spun 40 times. Predict the number of times it lands on red.

Q16. The bag from Q5 is used 30 times: a marble is drawn, its colour is recorded, and it is put back before the next draw. Predict the number of green marbles drawn.

Q17. Say whether each statement is true or false. Correct every false statement. (a) A probability is always between 0 and 1. (b) The probabilities of all the outcomes in a sample space add up to 1. (c) $P(\text{even}) = 1$ for one throw of a fair dice.

Q18. $P(6) = \frac{1}{6}$ on a fair dice. Predict the number of sixes in 600 throws, and explain in one sentence why the prediction for 600 throws is ten times the prediction for 60 throws.

Answers

Q1. (a) $\{H, T\}$; (b) $\{1, 2, 3, 4, 5, 6\}$; (c) $\{R, B, G, Y\}$; (d) {red, blue}.

Q2. (a) 6; (b) $\frac{1}{6}$; (c) $\frac{1}{2}$; (d) 1.

Q3. (a) $\frac{3}{8}$; (b) $\frac{1}{4}$; (c) $\frac{1}{4}$; (d) $\frac{1}{8}$.

Q4. A: unlikely; B: even chance; C: likely.

Q5. (a) $\frac{3}{10}$; (b) $\frac{1}{5}$; (c) $\frac{1}{2}$.

Q6. (a) $\frac{1}{5}$; (b) 20.

Q7. $\frac{1}{2}$.

Q8. (a) $\frac{1}{3}$; (b) $\frac{1}{2}$.

Q9. (a) $\frac{1}{2}$; (b) $\frac{1}{3}$.

Q10. 3.

Q11. (a) impossible; (b) even chance; (c) certain.

Q12. (a) certain, $P = 1$; (b) impossible, $P = 0$.

Q13. about 10.

Q14. about 25.

Q15. about 15.

Q16. about 15.

Q17. (a) true; (b) true; (c) false — $P(\text{even}) = \frac{1}{2}$.

Q18. about 100.

Fully-worked solutions

Q1. Each list holds every outcome once. (a) A coin lands heads or tails: $\{H, T\}$. (b) A dice shows 1 to 6: $\{1, 2, 3, 4, 5, 6\}$. (c) The spinner lands on one of its four colours: $\{R, B, G, Y\}$. (d) The ticket is red or blue: {red, blue}.

Q2. (a) The sample space $\{1, 2, 3, 4, 5, 6\}$ has 6 outcomes. (b) One outcome is a 3, so $P(3) = \frac{1}{6}$. (c) The even outcomes $2, 4, 6$ are 3 of the 6, so $P(\text{even}) = \frac{3}{6} = \frac{1}{2}$. (d) Every outcome is less than 7, so the event is certain: $P = 1$.

Q3. There are 8 equal sections. (a) 3 are red: $P(\text{red})=\frac{3}{8}$. (b) 2 are blue: $P(\text{blue})=\frac{2}{8}=\frac{1}{4}$. (c) 2 are green: $P(\text{green})=\frac{2}{8}=\frac{1}{4}$. (d) 1 is yellow: $P(\text{yellow})=\frac{1}{8}$.

Q4. A sits at $\frac{1}{5}$, close to 0: unlikely. B sits at $\frac{1}{2}$: even chance. C sits at $\frac{9}{10}$, close to 1: likely.

Q5. The bag holds 10 marbles. (a) $P(\text{red})=\frac{3}{10}$. (b) $P(\text{blue})=\frac{2}{10}=\frac{1}{5}$. (c) $P(\text{green})=\frac{5}{10}=\frac{1}{2}$.

Q6. (a) 4 of the 20 tickets win: $P(\text{prize})=\frac{4}{20}=\frac{1}{5}$. (b) Predicted number $=\frac{1}{5}\times 100=20$ prizes.

Q7. The sample space $\{H, T\}$ has 2 equally likely outcomes, one of them heads: $P(\text{heads})=\frac{1}{2}$.

Q8. (a) The numbers greater than 4 are 5 and 6: 2 favourable outcomes, so $P=\frac{2}{6}=\frac{1}{3}$. (b) The odd numbers $1, 3, 5$ are 3 favourable outcomes, so $P=\frac{3}{6}=\frac{1}{2}$.

Q9. (a) The even numbers $2, 4, 6$ are 3 of the 6 sections: $P(\text{even})=\frac{3}{6}=\frac{1}{2}$. (b) The numbers less than 3 are 1 and 2: $P=\frac{2}{6}=\frac{1}{3}$.

Q10. $P(\text{red})=\frac{1}{4}$ means $\frac{1}{4}$ of the 12 marbles are red: $\frac{1}{4}\times 12=3$ red marbles.

Q11. (a) 0 means the event can never happen: impossible. (b) $\frac{1}{2}$ is the middle of the scale: even chance. (c) 1 means the event always happens: certain.

Q12. (a) Every outcome is less than 7, so the event is certain and $P=1$. (b) A dice cannot show 8, so the event is impossible and $P=0$.

Q13. $P(6)=\frac{1}{6}$, so the predicted number of sixes is $\frac{1}{6}\times 60=10$, about 10 sixes.

Q14. $P(\text{heads})=\frac{1}{2}$, so the predicted number of heads is $\frac{1}{2}\times 50=25$, about 25 heads.

Q15. $P(\text{red})=\frac{3}{8}$, so the predicted number of reds is $\frac{3}{8}\times 40=15$, about 15 reds.

Q16. $P(\text{green})=\frac{1}{2}$, so the predicted number of greens is $\frac{1}{2}\times 30=15$, about 15 greens.

Q17. (a) **True**: every probability sits between 0 and 1. (b) **True**: all the outcomes must add to 1. (c) **False**: the even outcomes are 3 of the 6, so $P(\text{even})=\frac{3}{6}=\frac{1}{2}$, not 1.

Q18. The predicted number of sixes is $\frac{1}{6}\times 600=100$, about 100 sixes. 600 is ten times 60, and the predicted number is the probability times the number of trials, so ten times the trials gives ten times the prediction.

Teacher note

This subtopic is single-stage throughout (D4.16): every experiment is one throw, one toss, one spin or one pick, and no question combines two stages or uses the complement of an event — both are Level 8 work owned by VC2M8P01. The 0–1 scale is recalled from Level 6 (VC2M6P01), not re-taught. No digital tool is used here: VC2M7P01 does not name one, and the simulation work with a large number of trials belongs to 7B (VC2M7P02), where the tool is compulsory. VCAA writes *a dice* for the singular, and this book follows that usage (D12).

Repeated experiments and simulations; the effect of sample size

Content description VC2M7P02: conduct repeated chance experiments and run simulations with a large number of trials using digital tools; compare predicted with observed results, explaining the differences and the effect of sample size on the outcomes

Achievement standard (extract): They conduct repeated single-step chance experiments and run simulations using digital tools, giving reasons for differences between predicted and observed results.

Theory

What we bring from 7A. In subtopic 7A we met the words this subtopic uses every line: a chance **experiment** is a set-up whose result we cannot know in advance; one go of it is a **trial**; the result of a trial is an **outcome**; the list of all possible outcomes is the **sample space**; an **event** is one or more outcomes we are watching for; and an outcome that makes the event happen is a **favourable outcome**. When all outcomes are equally likely, the probability of an event is the number of favourable outcomes divided by the number of outcomes in the sample space. That probability is a *prediction*: for a fair six-sided dice, $Pr(6)=\frac{1}{6}$, and 7A used this to predict about 10 sixes in 60 rolls. Subtopic 7B asks the next question: when the experiment is actually run, do the results match the prediction — and what changes as we run more trials?

Predicted results and observed results. The **predicted result** is what probability says should happen; the **observed result** is what actually happens when the trials are run. Flip a coin 10 times and count the heads: the predicted result is 5 heads, but one run might give 6, another 4, another 7. To compare runs of different sizes we use the **relative frequency** of an event:

$$\text{relative frequency} = \frac{\text{number of trials in which the event happened}}{\text{number of trials run}}$$

Twelve heads in 20 flips is a relative frequency of $\frac{12}{20}=0.6$. The predicted relative frequency is simply the probability: for a fair coin it is $\frac{1}{2}=0.5$ whatever the number of flips.

A chance experiment to repeat. Figure 1 shows a fair spinner with four equal sectors. Each spin is one trial, the four colours are the equally likely outcomes, and "red" is an event with probability $\frac{1}{4}$.

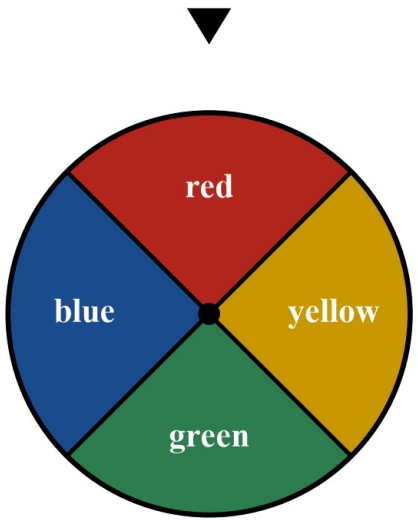

Figure 1. A spinner divided into four equal sectors labelled red, blue, green and yellow, with a pointer at the top.

Do it yourself. With a partner, make this spinner from cardboard and a paper clip, or use a digital spinner tool. Each of you spins it 20 times and records the colour of every spin. Almost certainly your two tables will disagree with each other and with the predicted count of 5 per colour. That disagreement is not a mistake. Every trial is governed by chance, and with only 20 trials the counts can wander quite far from the prediction. In probability language, the **sample size** — the number of trials run — is small, and small samples vary a lot.

Small samples vary a lot. Figure 2 shows the same idea with a coin. Eight people each flip the same fair coin 10 times and record the heads as filled squares.

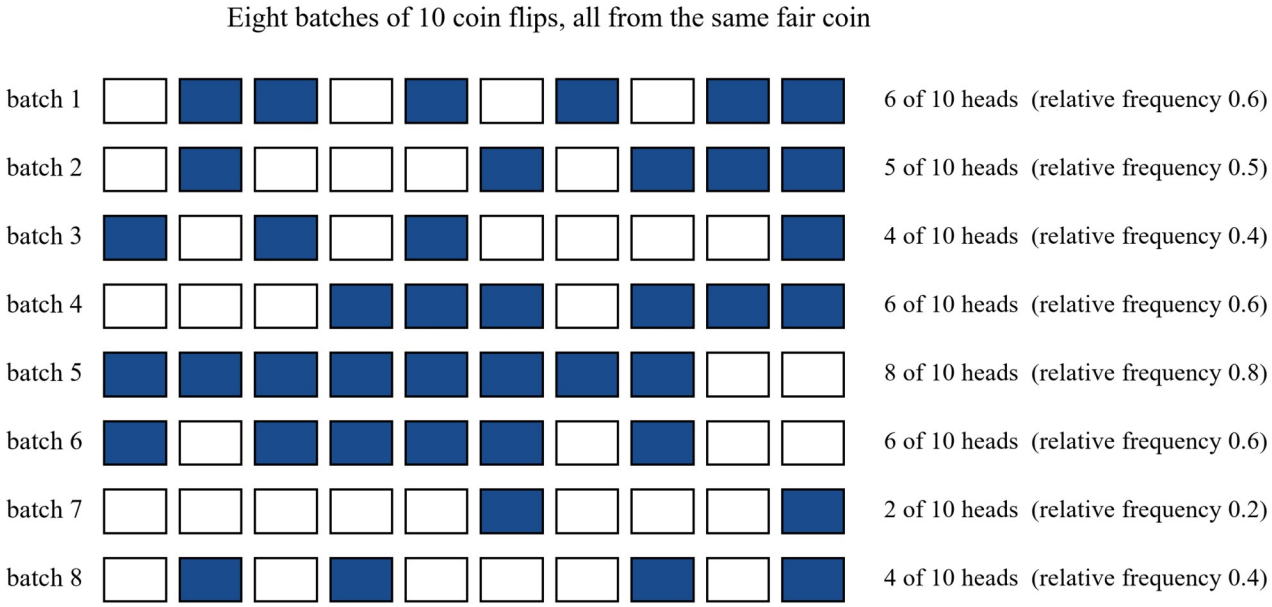

Figure 2. Eight rows of ten squares, one row for each batch of ten flips of the same fair coin; filled squares are heads. The eight batches show 6, 5, 4, 6, 8, 6, 2 and 4 heads out of ten.

Every batch lands somewhere different: from 2 heads (relative frequency 0.2) to 8 heads (relative frequency 0.8), against the predicted 0.5. Nobody's coin was unfair — 10 trials are simply too few for the results to show the $\frac{1}{2}$ clearly.

The effect of sample size. Figure 3 rolls the same fair dice a small number of times and a large number of times, and plots the relative frequency of each face.

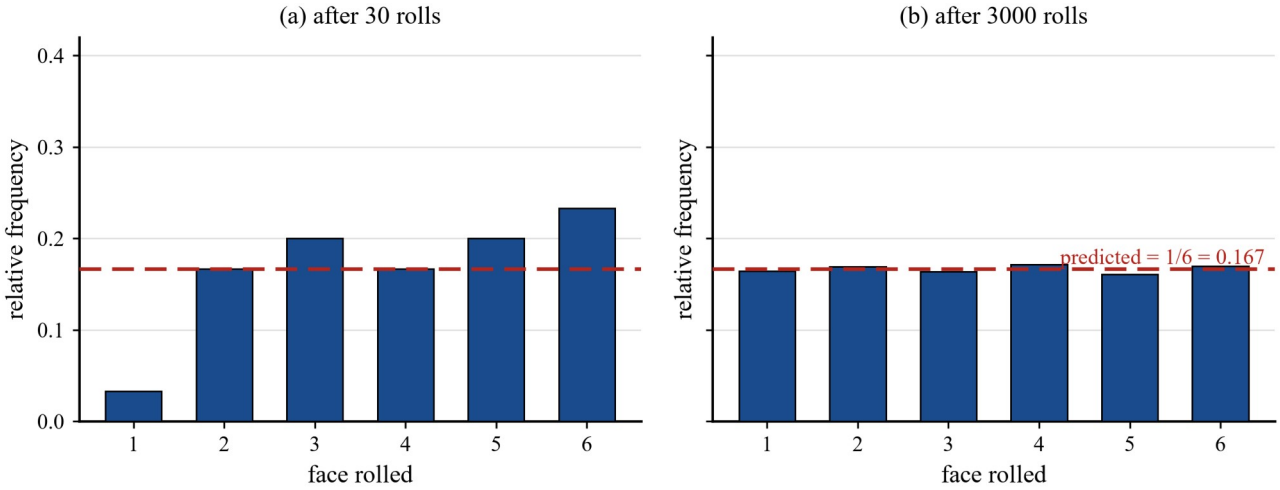

Figure 3. Two bar charts of the relative frequency of each face of the same fair dice: after 30 rolls the bars are very uneven (face 1 shows only 0.033), while after 3000 rolls every bar sits close to the predicted value 0.167.

After 30 rolls, face 1 has appeared only once, a relative frequency of about 0.03 against the predicted $\frac{1}{6} \approx 0.167$ — it looks as if the dice never shows a 1. After 3000 rolls, every face sits between 0.161 and 0.172, all close to 0.167. Bigger samples give observed results that sit closer to the prediction. Figure 4 shows why this is trusted: three separate runs of 500 coin flips take completely different paths, yet each one ends up near the predicted 0.5.

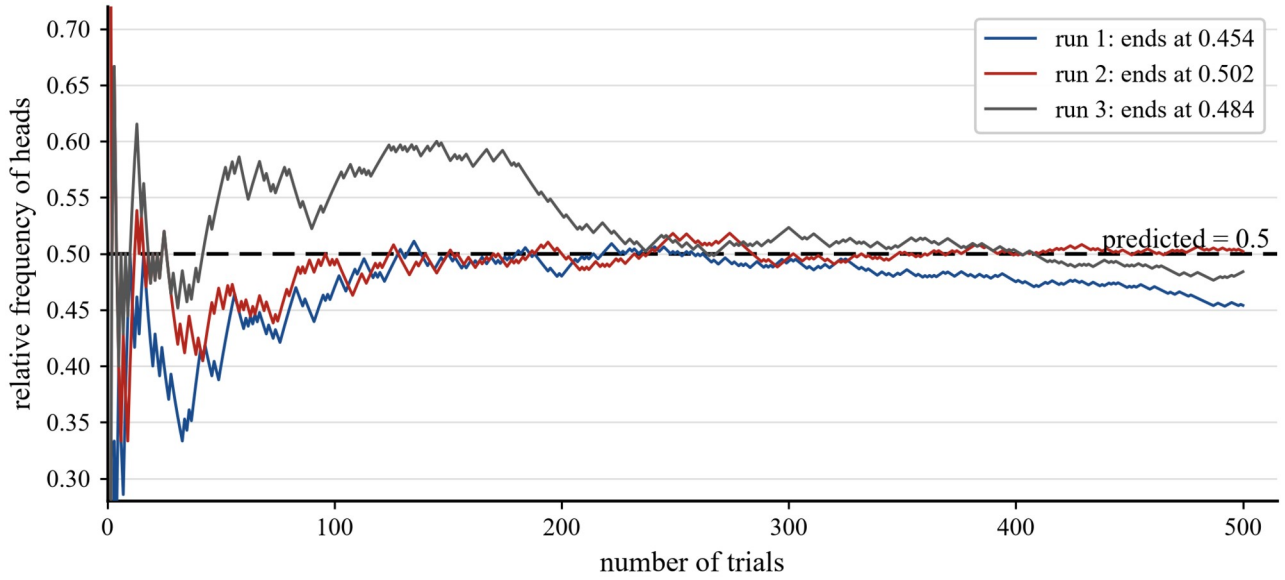

Figure 4. Three line graphs, one for each run of 500 simulated coin flips: the relative frequency of heads starts far from 0.5 in every run and settles near the predicted line at 0.5, ending at 0.454, 0.502 and 0.484.

This pattern — observed relative frequencies settling near the predicted probability as the sample size grows — is so reliable that it has a name, the **law of large numbers**. We use the name here to describe the pattern you can see in the figures; it is studied formally in later years. It is the reason a casino, a bank or a scientist trusts an average built from thousands of trials but not one built from ten.

Simulations with digital tools. Flipping a coin 5000 times by hand takes days; a digital tool does it in a moment. A **simulation** is a run of a chance experiment carried out by a digital tool — a spreadsheet or simulation software that produces a random result for each trial in the same way the real experiment would. This content description requires the tool: the large runs in this subtopic are all done with one, and each states the number of trials it ran. A hundred trials is not a large number for a simulation; the runs here use 500 to 5000 trials. A simulation does not replace the real experiment — it repeats it quickly enough for the law of large numbers to show.

Comparing predicted with observed, and explaining the difference. Three ideas summarise the whole subtopic:

- An observed result is *expected* to differ from the predicted result in any single run. Chance makes each run wander, and a difference is not evidence that anything is wrong.
- The larger the sample size, the smaller that difference usually is. Observed relative frequencies from small samples bounce about; from large samples they sit near the predicted probability.
- A difference that stays large across a large number of trials is a different matter. Then the sensible explanation is that the experiment is not fair — the outcomes are not equally likely — because chance alone would not keep a large sample that far from the prediction.

The words of this subtopic, with synonyms you may meet:

Term	Meaning here	Synonyms and related words
trial	one go of a chance experiment	one spin, one roll, one flip
observed result	what actually happens when the trials are run	actual result, data
predicted result	what probability says should happen	expected result
relative frequency	trials in which the event happened ÷	observed proportion

Term	Meaning here	Synonyms and related words
	trials run	
sample size	the number of trials run	number of goes, number of repeats
simulation	a digital tool repeating a chance experiment	computer model of the experiment
law of large numbers	larger samples give observed results closer to the prediction	the settling pattern

Worked examples

Example 1 — scaffolded

Use the spinner of Figure 1. (a) Write the predicted probability of red on one spin. (b) Predict the number of reds in 40 spins. (c) Predict the number of reds in 100 spins.

Working	Comment
(a) $Pr(\text{red}) = \frac{1}{4}$.	Four equal sectors, one of them red.
(b) $40 \times \frac{1}{4} = 10$ reds.	Multiply the number of trials by the probability.
(c) $100 \times \frac{1}{4} = 25$ reds.	The same prediction rule for any number of spins.

Example 2 — scaffolded

You spin the spinner of Figure 1 twenty times. The observed counts are red 8, blue 5, green 3, yellow 4 (Figure 5). (a) Find the observed relative frequency of red. (b) State the predicted count for red and compare it with the observed count. (c) Is the difference a reason to doubt the spinner? Explain.

Working	Comment
(a) relative frequency of red $= \frac{8}{20} = 0.4$.	Event count ÷ trials run.
(b) Predicted count $= 20 \times \frac{1}{4} = 5$; observed 8, a difference of 3.	Predicted relative frequency is 0.25.
(c) No. 20 trials is a small sample, and small samples often sit this far from the prediction by chance alone.	A difference is expected, not a fault.

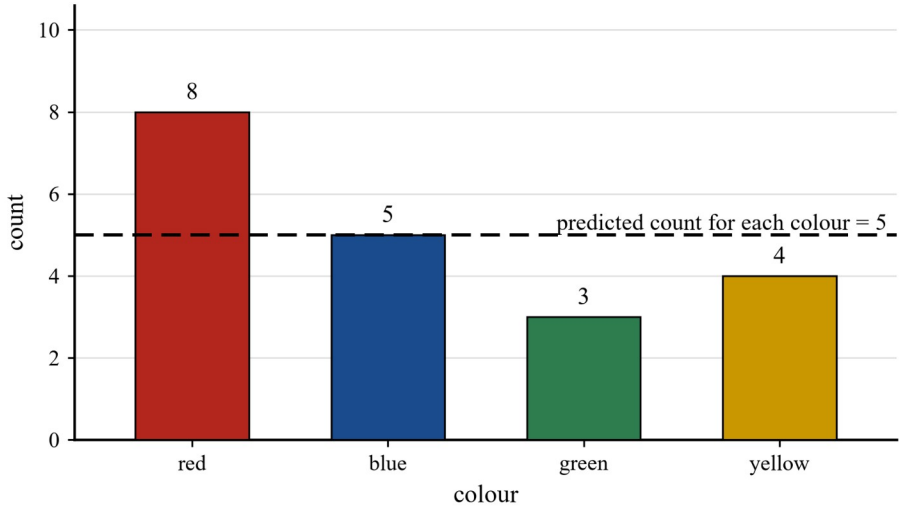

Figure 5. Bar chart of 20 spins of the four-colour spinner: red 8, blue 5, green 3, yellow 4, with the predicted count of 5 for each colour shown as a dashed line.

Example 3 — unscaffolded

A simulation rolls a fair dice 1000 times and records the relative frequency of the event "roll a 6" after every trial (Figure 6). Read the marked values, compare each with the predicted probability, and describe the effect of the sample size.

After 100 trials the relative frequency is 0.210 (21 sixes); after 500 trials it is 0.174 (87 sixes); after 1000 trials it is 0.165 (165 sixes). The predicted probability is $\frac{1}{6} \approx 0.167$, which for 1000 trials predicts about 167 sixes. The gaps from the prediction are $0.210 - 0.167 = 0.043$ at 100 trials, 0.007 at 500 trials and 0.002 at 1000 trials: as the sample size grows, the observed relative frequency settles closer and closer to the predicted probability. That is the law of large numbers at work.

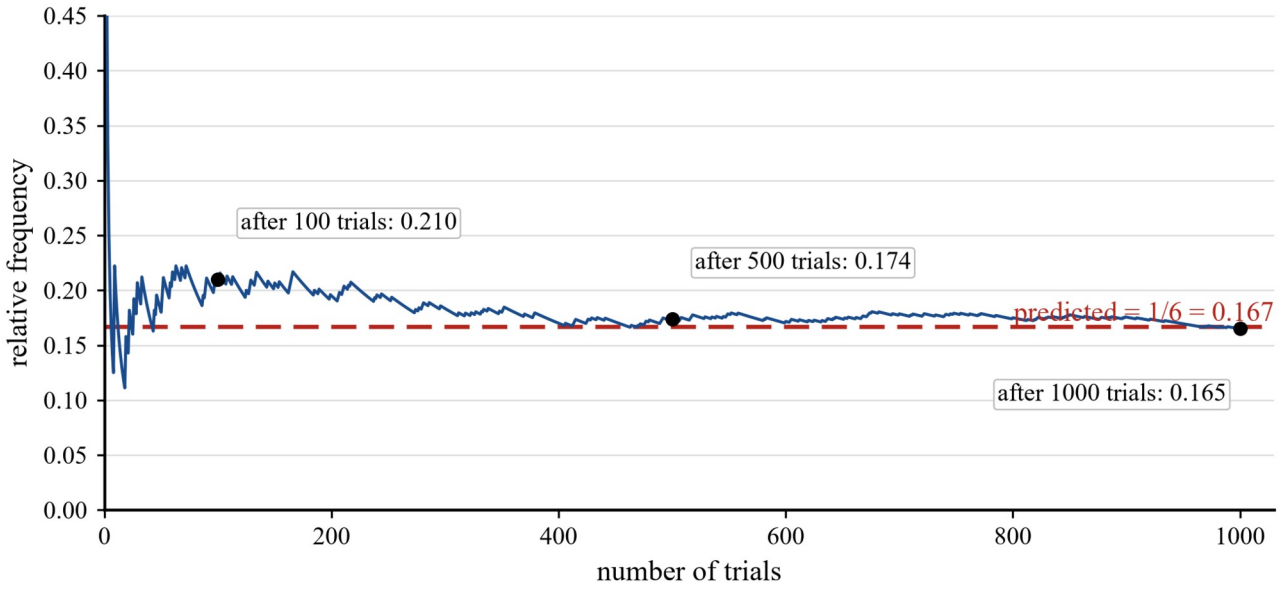

Figure 6. Line graph of the running relative frequency of rolling a 6 across 1000 simulated dice rolls, marked after 100 trials (0.210), after 500 trials (0.174) and after 1000 trials (0.165), with the predicted value 0.167 as a dashed line; the large swings early on rise above the top of the chart.

Example 4 — unscaffolded

A simulation rolls two different dice 600 times each. Dice A gives the counts $95, 103, 113, 88, 102, 99$ for faces 1–6; dice B gives $29, 27, 30, 463, 23, 28$ (Figure 7). For a fair dice the predicted count for each face is $600 \times \frac{1}{6} = 100$. Decide which dice appears fair and which does not, giving reasons.

Dice A: every count sits between 88 and 113, close to the predicted 100. Differences of at most 13 in 600 rolls are what chance usually produces, so dice A appears fair. Dice B: face 4 appeared 463 times, more than four times its predicted count, while every other face is far below 100. Chance would not keep a sample of 600 rolls that far from the prediction, so dice B does not appear fair — its outcomes are not equally likely.

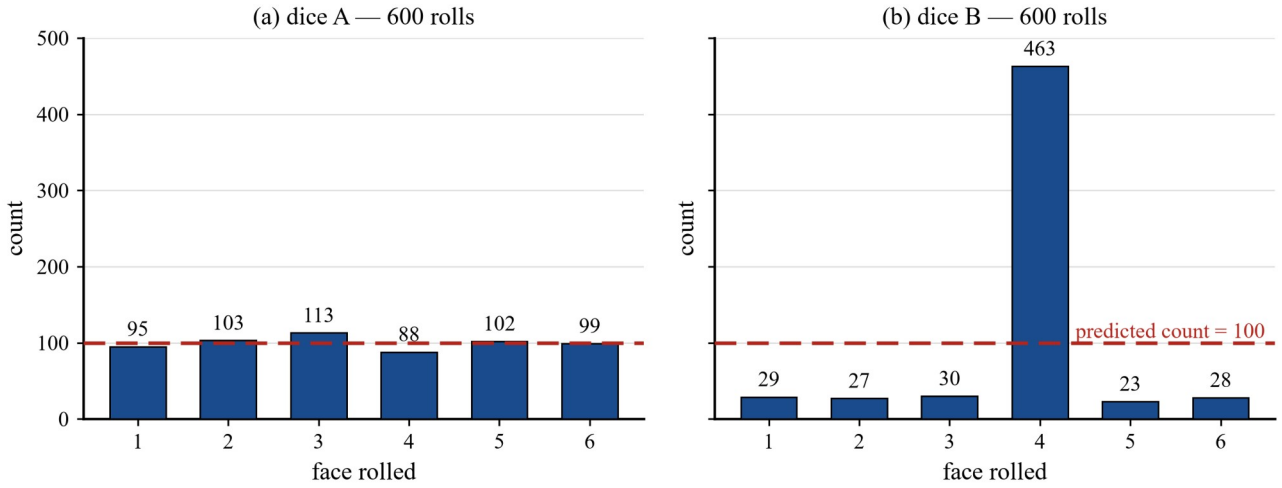

Figure 7. Two bar charts of 600 simulated rolls each. Dice A shows counts 95, 103, 113, 88, 102 and 99, all close to the predicted count of 100 shown as a dashed line. Dice B shows counts 29, 27, 30, 463, 23 and 28, with face 4 far above the predicted count.

Practice questions

Scaffolded

Q1. Replace each blank with the correct word: *trial, simulation, sample size, relative frequency, predicted*. (a) One spin of a spinner is one _____. (b) A digital tool repeating a chance experiment many times is a _____. (c) The number of trials run is the _____. (d) The event happened 14 times in 50 trials, so its observed _____ is 0.28. (e) For a fair coin, the _____ relative frequency of heads is 0.5 for any number of flips.

Q2. A fair coin is flipped. The predicted probability of a head is $\frac{1}{2}$. Predict the number of heads in: (a) 30 flips (b) 100 flips (c) 250 flips

Q3. A dice is rolled 30 times. The counts for faces 1–6 are $8, 5, 4, 4, 4, 5$. (a) Find the observed relative frequency of the event "odd number". (b) State the predicted probability of "odd number". (c) Compare your answers. Is the difference a surprise for 30 trials? Explain in one sentence.

Q4. Figure 8 shows the running relative frequency of heads in a simulation of 500 coin flips.

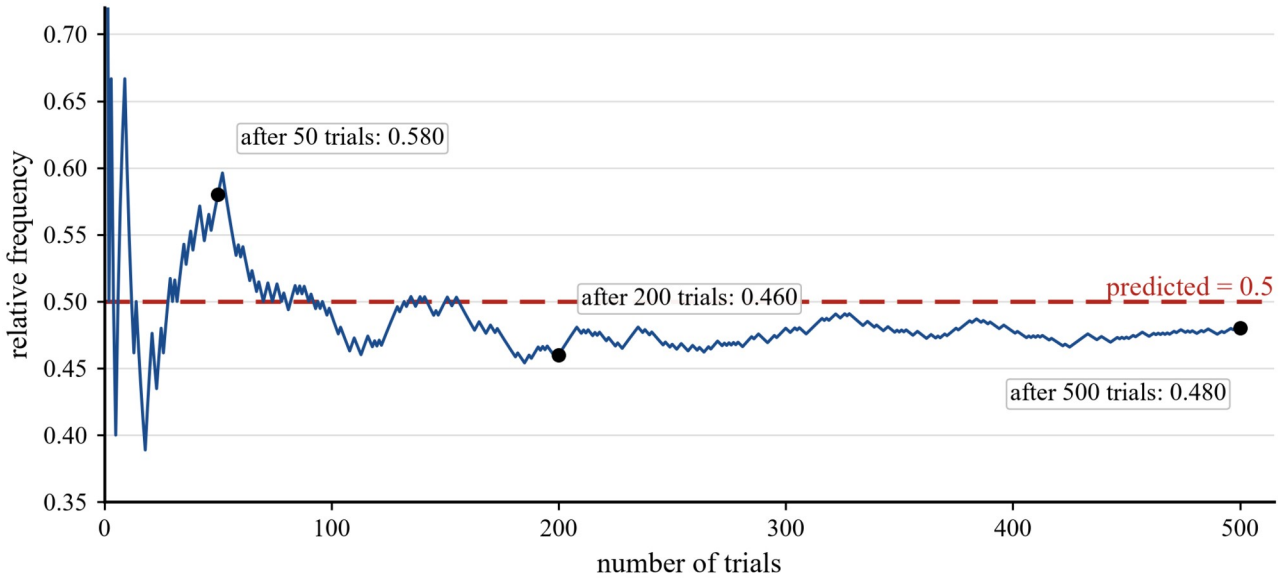

Figure 8. Line graph of the running relative frequency of heads across 500 simulated coin flips, marked after 50 trials (0.580), after 200 trials (0.460) and after 500 trials (0.480), with the predicted value 0.5 as a dashed line; the large swings early on rise above the top of the chart.

(a) Read the observed relative frequency of heads after 50, 200 and 500 trials.

(b) State the predicted relative frequency of heads.

(c) At which of the three sample sizes is the observed value closest to the prediction? What does this show about sample size?

Q5. State whether each statement is true or false. Correct every false statement. (a) In 10 trials, the observed relative frequency can be quite far from the predicted probability. (b) In 1000 trials, the observed relative frequency is usually closer to the predicted probability than in 10 trials. (c) If a fair coin shows 6 heads in 10 flips, the next 10 flips must show 4 heads to bring the results back to 0.5. (d) A simulation of 20 trials gives just as good a check of a prediction as a simulation of 2000 trials. (e) In one run of 100 flips of a fair coin, the relative frequency of heads will always be exactly 0.5.

Q6. In a simulation, an event has observed relative frequency 0.28 after 250 trials. How many trials did the event happen in?

Unscaffolded

Q7. Figure 9 shows a spinner with five equal sectors numbered 1–5.

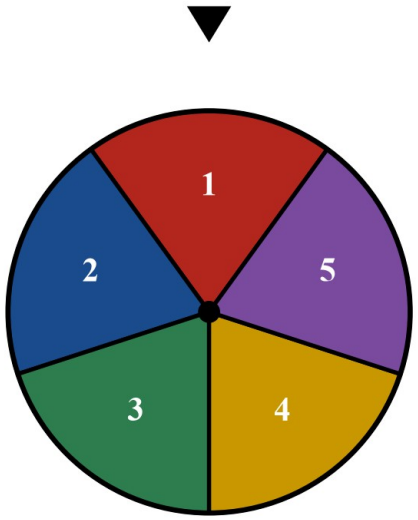

Figure 9. A spinner divided into five equal sectors numbered 1 to 5, with a pointer at the top.

(a) Write the predicted probability of spinning a 5.

(b) Predict the number of 5s in 300 spins.

(c) A simulation of 300 spins gives 54 fives. Is this difference from the prediction a reason to doubt the spinner? Explain.

Q8. A coin is flipped 80 times and shows 47 heads. (a) Find the observed relative frequency of heads. (b) State the predicted relative frequency of heads. (c) Should the 80 flips have shown exactly 40 heads? Explain in one sentence.

Q9. A simulation rolls a dice 1000 times. The counts for faces 1–6 are 170, 163, 169, 172, 159, 167. The predicted count for each face of a fair dice is about 167. Does the dice appear fair? Give a reason.

Q10. Two simulations test the same event, whose predicted probability is 0.5. The first runs 10 trials and gets relative frequency 0.7; the second runs 400 trials and gets relative frequency 0.52. Which result would you trust more as a check of the prediction? Explain.

Q11. A simulation spins a fair three-sector spinner (sectors A, B, C) 900 times. The counts are A 274, B 317, C 309. (a) Find the observed relative frequency of each sector, to three decimal places. (b) State the predicted probability of each sector. (c) Explain the differences between the observed and predicted values.

Q12. A dice is rolled 600 times in a simulation. The counts for faces 1–6 are $46, 52, 49, 341, 55, 57$. Does the dice appear fair? Give a reason.

Q13. A student runs 30 trials of a simulation and says: "My observed relative frequency is not equal to the predicted probability, so the tool must be broken." Explain what is wrong with this reasoning.

Q14. A game says that 1 play in 4 wins a prize. A student plays 8 times, wins 1 prize, and concludes the game is not keeping its word. Explain why 8 plays cannot test the claim.

Q15. A bag holds 10 counters, 3 of them red. A simulation draws a counter, records its colour, and returns it, repeating this 5000 times. The table shows the number of reds drawn at four sample sizes. Complete the relative frequency column and answer (b) and (c).

Trials run	Reds drawn	Observed relative frequency of red
100	32	
500	147	
1000	308	
5000	1498	

(a) Complete the table.

(b) Which sample size gives the relative frequency closest to the predicted 0.3?

(c) Describe the pattern in the completed column.

Q16. Complete the sentence with the words *closer, larger, predicted*: "As the sample size gets _____, the observed relative frequency usually gets _____ to the _____ probability."

Q17. You suspect a four-colour spinner like the one in Figure 1 may be unfair. Describe how you would use a digital tool to check, including the number of trials you would choose and how you would compare the observed results with the prediction.

Q18. In a simulation of 600 rolls of a dice, the observed relative frequency of "roll a 5" is 0.16. (a) How many 5s were rolled? (b) State the predicted number of 5s in 600 rolls of a fair dice. (c) What does the size of the difference suggest about the dice?

Answers

Q1. (a) trial; (b) simulation; (c) sample size; (d) relative frequency; (e) predicted.

Q2. (a) 15; (b) 50; (c) 125.

Q3. (a) $\frac{16}{30} \approx 0.53$; (b) $\frac{1}{2} = 0.5$; (c) the observed value is 0.03 above the prediction — not a surprise, because 30 trials is a small sample.

Q4. (a) 0.580, 0.460, 0.480; (b) 0.5; (c) after 500 trials (0.480 is 0.020 from 0.5, closer than 0.080 at 50 and 0.040 at 200); larger samples sit closer to the prediction.

Q5. (a) true; (b) true; (c) false — the coin owes nothing to earlier flips; any count is possible in the next 10; (d) false — 2000 trials give a much better check than 20; (e) false — in one run of 100 flips the relative frequency is usually near 0.5 but rarely exactly 0.5.

Q6. $250 \times 0.28 = 70$ trials.

Q7. (a) $\frac{1}{5}$; (b) $300 \times \frac{1}{5} = 60$; (c) no — a difference of 6 in 300 spins is well within what chance produces.

Q8. (a) $\frac{47}{80} = 0.5875$; (b) 0.5; (c) no — 40 heads is the prediction on average, and any single run of 80 flips usually differs from it a little.

Q9. Yes, it appears fair: the predicted count is about 167 per face and every count sits within 8 of it, which is what chance usually gives in 1000 rolls.

Q10. The 400-trial result: 0.52 is close to 0.5, while 0.7 from only 10 trials is the kind of large swing small samples often show.

Q11. (a) A 0.304, B 0.352, C 0.343; (b) $\frac{1}{3} \approx 0.333$ each; (c) chance makes every run differ a little from the prediction, and at 900 trials the differences are small (at most 0.02).

Q12. No: the predicted count is 100 per face, but face 4 appeared 341 times; chance would not keep 600 rolls that far from the prediction, so the outcomes are not equally likely.

Q13. A difference after 30 trials is expected — small samples wander. The tool should be judged on a large run (say 1000 trials), where the observed relative frequency should sit near the prediction.

Q14. The prediction for 8 plays is $8 \times \frac{1}{4} = 2$ prizes; winning 1 is only one prize short, and 8 plays is far too small a sample to tell chance from a false claim.

Q15. (a) 0.320, 0.294, 0.308, 0.2996; (b) 5000 trials (0.2996 is 0.0004 from 0.3); (c) as the sample size grows, the observed relative frequency settles closer to the predicted 0.3.

Q16. larger; closer; predicted.

Q17. Run the spinner in a simulation tool for a large number of trials (say 1000); count the spins of each colour; the predicted count for a fair spinner is $1000 \times \frac{1}{4} = 250$ each; if all four counts sit close to 250 the spinner appears fair, while a colour far from 250 after 1000 trials suggests it is unfair.

Q18. (a) $600 \times 0.16 = 96$; (b) $600 \times \frac{1}{6} = 100$; (c) a difference of 4 in 600 rolls is small — the dice behaves like a fair one.

Fully-worked solutions

Q1. (a) One go of an experiment is a **trial**. (b) A digital tool repeating the experiment is a **simulation**. (c) The number of trials run is the **sample size**. (d) $14 \div 50 = 0.28$ is the observed **relative frequency**. (e) The **predicted** relative frequency of heads is the probability, 0.5, for any number of flips.

Q2. Multiply the number of flips by $\frac{1}{2}$: (a) $30 \times \frac{1}{2} = 15$; (b) $100 \times \frac{1}{2} = 50$; (c) $250 \times \frac{1}{2} = 125$.

Q3. (a) The odd faces $1, 3, 5$ appeared $8 + 4 + 4 = 16$ times, so the relative frequency is $\frac{16}{30} \approx 0.53$. (b) Three of the six faces are odd: $\frac{3}{6} = 0.5$. (c) The observed value is 0.03 above the prediction. With only 30 trials a gap this size is common, so it is not a surprise.

Q4. (a) The marked points read 0.580 after 50 trials, 0.460 after 200 trials and 0.480 after 500 trials. (b) The predicted relative frequency of heads is 0.5. (c) The gaps from 0.5 are 0.080, 0.040 and 0.020: the largest sample, 500 trials, is closest. This is the effect of sample size — larger samples usually sit closer to the prediction.

Q5. (a) **True** — small samples vary a lot. (b) **True** — this is the law of large numbers. (c) **False** — each flip is a fresh trial; the coin does not owe the earlier results anything, so the next 10 flips can show any count. (d) **False** — the 2000-trial simulation gives a far better check, because larger samples sit closer to the prediction. (e) **False** — one run of 100 flips usually gives a relative frequency near 0.5, but exactly 0.5 is not guaranteed and in fact is unusual.

Q6. Relative frequency = event count ÷ trials, so event count $= 250 \times 0.28 = 70$.

Q7. (a) One of five equal sectors: $Pr(5) = \frac{1}{5}$. (b) $300 \times \frac{1}{5} = 60$. (c) The observed count is 54, a difference of 6 from the predicted 60. In 300 spins a difference of 6 is well within the wander chance produces, so there is no reason to doubt the spinner.

Q8. (a) $\frac{47}{80} = 0.5875$. (b) $\frac{1}{2} = 0.5$. (c) No. The prediction says what happens on average over many runs; a single run of 80 flips is expected to differ from 40 heads by a small amount.

Q9. For a fair dice the predicted count is $1000 \times \frac{1}{6} \approx 167$ per face. The observed counts $170, 163, 169, 172, 159, 167$ all sit within 8 of 167. Over 1000 rolls, chance usually keeps the counts this close, so the dice appears fair.

Q10. Trust the 400-trial result. With a predicted probability of 0.5, a relative frequency of 0.7 after 10 trials is the kind of large swing a small sample often shows, while 0.52 after 400 trials sits close to the prediction, as a large sample usually gives.

Q11. (a) A: $\frac{274}{900} \approx 0.304$; B: $\frac{317}{900} \approx 0.352$; C: $\frac{309}{900} \approx 0.343$. (b) $\frac{1}{3} \approx 0.333$ for each sector. (c) Every run differs a little from the prediction by chance; because 900 trials is a large sample, the differences are small — at most $0.352 - 0.333 = 0.019$ from the prediction.

Q12. For a fair dice the predicted count is $600 \times \frac{1}{6} = 100$ per face. Face 4 appeared 341 times and every other face far fewer than 100. A chance difference would not stay this large across 600 rolls, so the dice does not appear fair — its outcomes are not equally likely.

Q13. The reasoning treats one small run as if it must match the prediction exactly. With 30 trials the observed relative frequency usually differs from the predicted probability — that is what small samples do. A fair check needs a large

run: after, say, 1000 trials the observed relative frequency should sit near the prediction, and only a difference that stays large then would point to a fault.

Q14. The predicted number of prizes in 8 plays is $8 \times \frac{1}{4} = 2$, so the student expected 2 and won 1. A gap of one prize over 8 plays is nothing more than the usual small-sample wander; hundreds of plays would be needed before the observed relative frequency of prizes could really test the claim of $\frac{1}{4}$.

Q15. (a) $32 \div 100 = 0.320$; $147 \div 500 = 0.294$; $308 \div 1000 = 0.308$; $1498 \div 5000 = 0.2996$. (b) The gaps from 0.3 are 0.020, 0.006, 0.008 and 0.0004: the 5000-trial sample is closest. (c) As the sample size grows, the observed relative frequency settles closer to the predicted 0.3 — the law of large numbers.

Q16. "As the sample size gets **larger**, the observed relative frequency usually gets **closer** to the **predicted** probability."

Q17. A good answer names each step: use a simulation tool (a spreadsheet or spinner software) set to the four equal colours; run a large number of trials, say 1000; count the spins of each colour; the predicted count for a fair spinner is $1000 \times \frac{1}{4} = 250$ per colour; compare each observed count with 250; if all four are close to 250 the spinner appears fair, and if one colour is far from 250 after 1000 trials the spinner is probably unfair, because chance alone would not do that in a large sample.

Q18. (a) $600 \times 0.16 = 96$ fives. (b) $600 \times \frac{1}{6} = 100$ fives. (c) The observed count is 4 below the prediction — a small difference for 600 rolls, so the dice behaves like a fair one.

Teacher notes

- **First Nations exclusion (CONTENT_POLICY.md Rule 1, recorded 2026-08-15).** `VC2M7P02.E03` invites exploring the instructive game *Koara* of the Jawi and Bardi Peoples of Sunday Island, Western Australia. Aboriginal and Torres Strait Islander content is excluded from this book as accuracy cannot be guaranteed; this subtopic therefore teaches the content description on other contexts, and the elaboration is recorded as an accepted exclusion rather than a gap. The front matter carries the book-wide notice.
- **Digital tools are compulsory (D9).** `VC2M7P02` says "using digital tools" without qualification; the large runs here (500, 600, 900, 1000, 3000 and 5000 trials) are presented as simulation output, and Q17 asks students to plan a simulation themselves. Tools are named generically (spreadsheet or simulation software), never by product. A hundred trials is *not* treated as large.
- **7A vocabulary is recalled, not re-taught.** Trial, outcome, sample space, event and favourable outcome belong to `VC2M7P01` (subtopic 7A); the opening paragraph recalls them in one place so this section survives reordering (D6).
- **Level boundaries (D4.16).** All experiments are single-stage; complementary events, two-event outcomes and compound simulations are Level 8 and do not appear. "Sample size" is used in the sense the content description uses it — the number of trials — and is defined at first use; populations and samples in the statistical sense are Level 8 (`VC2M8ST01`) and are not introduced.
- **Reproducibility of the simulation data.** Every figure and quoted simulation result in this subtopic is the output of a fixed-seed run documented in `src/gen_7b.py` (numpy `default_rng`, seeds 7101–7107, 7109–7111, 7114 and 7299), so the numbers in the prose and the figures cannot drift apart.

Test

Chapter test T7 — Probability

Content descriptions assessed (for diagnostic marking against the Level 7 achievement standard): `VC2M7P01` (7A Sample spaces and probabilities for single-stage experiments) · `VC2M7P02` (7B Repeated experiments and simulations; the effect of sample size)

> This is a school-designed paper. It is not a VCAA instrument. Achievement in the Victorian Curriculum is reported against the Level 7 achievement standard, not against a mark on this paper.

Time	15 minutes — 5 minutes reading time, then 10 minutes working time (about 1 minute per mark)
Total	**10 marks**
Questions	5 — answer every question
Equipment	Access to a digital simulation tool (a spreadsheet or simulation software) for Section B — `VC2M7P02` says *using digital tools*, so the tool is **required** for the Section B items. No calculator is needed anywhere on this paper; `VC2M7P01` does not name a digital tool, so Section A is worked without one.

Subtopic	Content description	Marks	Questions
7A Sample spaces and probabilities for single-stage experiments	`VC2M7P01`	5	1–3
7B Repeated experiments and simulations; the effect of sample size	`VC2M7P02`	5	4–5

Show your working wherever a question asks for it — marks are given for the working, not just the final answer.

Section A — Sample spaces and probabilities (7A, `VC2M7P01`)

Every experiment in this section is single-stage: one action, one result. All outcomes on this paper are equally likely — no outcome is easier to get than another.

Q1. A bag holds 4 red counters, 1 blue counter and 3 yellow counters. One counter is picked without looking. *(2 marks)*

(a) List the **sample space** — the list of all possible outcomes — for this experiment. *(1 mark)*

(b) Assign $P(\text{blue})$, the probability that the counter picked is blue. *(1 mark)*

Q2. A fair spinner has 8 equal sectors numbered 1 to 8. It is spun once. *(2 marks)*

(a) Assign $P(\text{odd})$ as a fraction in simplest form. *(1 mark)*

(b) The spinner is now spun 40 times. Predict the number of odd results. *(1 mark)*

Q3. In a game, the probability of winning a prize on one play is 0.15. Which word best describes the chance of winning on one play? *(1 mark)*

A. impossible B. unlikely C. even chance D. certain

Section B — Repeated experiments and simulations (7B, VC2M7P02)

*The results in this section come from a simulation — a digital tool running a chance experiment many times. One run of the experiment is one **trial**. The **observed relative frequency** of an event is the number of trials it happened in, divided by the number of trials run. The **sample size** is the number of trials run.*

Q4. A simulation tool spins a fair spinner 400 times. The spinner has four equal sectors coloured red, blue, green and yellow. *(2 marks)*

(a) Predict the number of times the spinner lands on red. *(1 mark)*

(b) The run gives 108 reds. Find the observed relative frequency of red. *(1 mark)*

Q5. A simulation tool rolls a fair dice 40 times, then 400 times, then 4000 times, and counts the even numbers in each run. The counts are shown below. *(3 marks)*

Trials run	Even rolls	Observed relative frequency
40	23	
400	196	
4000	1988	

(a) Complete the observed relative frequency column. *(1 mark)*

(b) For a fair dice, the predicted relative frequency of an even number is 0.5. Which sample size gives the observed relative frequency closest to 0.5? *(1 mark)*

(c) In one sentence, explain why observed relative frequencies usually sit closer to the predicted probability when the sample size is larger. *(1 mark)*

End of test — 10 marks.

Solutions

Fully worked solutions for every question, with the mark breakdown, followed by a compact answer key for self-checking.

Worked solutions

Q1.

(a) The sample space lists every outcome that could happen, once each. The counter picked must be red, blue or yellow:

$$\text{sample space} = \{\text{red, blue, yellow}\}$$

1 mark for the three outcomes, each listed once.

(b) Count the counters first: $4 + 1 + 3 = 8$ counters, all equally likely. Exactly 1 of them is blue, so there is 1 favourable outcome — an outcome that makes the event happen — out of 8:

$$P(\text{blue}) = \frac{1}{8}$$

1 mark for $\frac{1}{8}$.

Q2.

(a) The favourable outcomes for "odd" are $1, 3, 5$ and 7: 4 outcomes out of the 8 equal sectors. So

$$P(\text{odd}) = \frac{4}{8} = \frac{1}{2}$$

(divide top and bottom by 4 for simplest form). *1 mark for* $\frac{1}{2}$; $\frac{4}{8}$ *also earns the mark.*

(b) The predicted number is the probability times the number of trials:

$$\frac{1}{2} \times 40 = 20$$

We predict about 20 **odd results** in 40 spins. A prediction is a best guess for the long run, not a promise. *1 mark for* 20.

Q3. B — unlikely. On the 0–1 scale, 0.15 sits close to 0 but is not 0: events close to 0 are unlikely. It is not impossible, because impossible means probability exactly 0; it is not an even chance ($\frac{1}{2}$) and not certain (1). *1 mark for B.*

Q4.

(a) Four equal sectors, one of them red, so $P(\text{red}) = \frac{1}{4}$. The predicted number of reds is the probability times the number of trials:

$$400 \times \frac{1}{4} = 100$$

We predict about 100 **reds**. *1 mark for* 100.

(b) Observed relative frequency = number of reds ÷ number of trials:

$$108 \div 400 = 0.27$$

The observed relative frequency of red is 0.27. (Check: the predicted relative frequency is $\frac{1}{4} = 0.25$, and 0.27 sits a little above it — an ordinary chance difference over 400 trials.) *1 mark for* 0.27.

Q5.

(a) Divide each count by its number of trials:

$$23 \div 40 = 0.575,\ 196 \div 400 = 0.49,\ 1988 \div 4000 = 0.497$$

Trials run	Even rolls	Observed relative frequency
40	23	0.575
400	196	0.49
4000	1988	0.497

1 mark for all three relative frequencies correct.

(b) Compare each observed value with the predicted 0.5:

$$0.575 - 0.5 = 0.075,\ 0.5 - 0.49 = 0.01,\ 0.5 - 0.497 = 0.003$$

The 4000-**trial** sample is closest to the prediction — its gap of 0.003 is the smallest. *1 mark for* 4000 *trials.*

(c) Sample answer: as the sample size grows, the ups and downs of chance even out, so the observed relative frequency settles closer to the predicted probability. This is the effect of sample size — the law of large numbers, described informally. *1 mark for any sentence that says larger samples vary less around the prediction, or that the observed results settle towards the prediction as more trials are run.*

Answer key

Q	Answer	Marks
1	(a) $\{$red, blue, yellow$\}$; (b) $\frac{1}{8}$	2
2	(a) $\frac{1}{2}$; (b) about 20	2
3	B — unlikely	1
4	(a) about 100; (b) 0.27	2
5	(a) 0.575, 0.49, 0.497; (b) 4000 trials; (c) larger samples settle closer to the prediction	3
	Total	**10**

Total: 10 marks — Section A (7A, `VC2M7P01`): 5 · Section B (7B, `VC2M7P02`): 5.