File RP1-1 For further resources click M1Maths.com

Introducing Probability
Teaching Notes and Student Resources

Contents

If you are looking at an electronic copy, you can jump to the section of this module you want by clicking on the link below or you can scroll down and work through the module sequentially.

Core learning outcomes, core content and elaborations addressed

This module provides students with opportunities to demonstrate the core learning outcomes, core content and elaborations that are shown in black in the table below.

		Students know:	**Students can:**
Students identify all possible outcomes of familiar situations and actions and describe them as, more likely, less likely, or equally likely.	**Concept of chance** • everyday language of chance • connections with everyday experiences • multiple outcomes **Making judgments** • judgments about the likelihood of an outcome	many situations and actions have more than one possible outcome and these can have the same or different likelihoods	• Identify all of the possible outcomes linked to a familiar situation or action such as when dropping a drawing pin, it can land point up or point down • classify familiar events as impossible, possible and certain and as more, less and equally likely such as when using a spinner with one red sector, two blue sectors, two green sectors and three yellow sectors
		Students know:	**Students can:**
Students make judgments about the likelihood of outcomes, supporting these with experimental	**Concept of chance** • connections with everyday experiences • multiple outcomes **Making judgments**	judgments about the likelihood of events can be supported by data	• make and justify judgments about the likelihood of events based on previous experience • collect experimental and pre-existing data and use it to make judgments about the relative likelihood of events

data or previous experience.	• judgments about the likelihood of an outcome based on experimental and pre-existing data		• identify factors that might affect the fairness of a situation such as starting first in a game of 'noughts and crosses'
		Students know:	**Students can:**
Students determine and interpret numerical probabilities and use them to make informed decisions.	**Concept of chance** • probability **Making judgments** • determining probability using data • determine probability using indifference **Making decisions** • using statements of probability in making decisions	the meaning of probability as a quantitative measure of likelihood between 0 and 1 ways of determining and using probabilities for simple events	• interpret probability as the fraction of times an outcome will occur if the situation or action is repeated a very large number of times • explain that the fraction of times an outcome will occur in a small number of repetitions might be quite different from the probability • design and conduct experiments to determine approximate probabilities • use pre-existing data to determine approximate probabilities • comment on the reliability of a probability estimate, considering the amount of data used, eg after tossing a coin only ten times, how reliable would be the resulting estimate of probability? • use the numerical probability of an outcome to predict the number of times the outcome is likely to occur and to make an informed decision about a future action • determine probabilities using indifference, eg. deciding that the probability of getting a head when tossing a coin is 50% because there is no difference between the two possible outcomes that could make one more likely than the other • decide whether indifference can be used in a particular case, eg that indifference does not apply when tossing a matchbox because the faces are different
		Students know:	**Students can:**
Students calculate probabilities for multi-outcome and multi-stage events and use their knowledge to support the making of informed choices.	**Concept of chance** • probability **Making judgments** • determining probability for multi-outcome and multi-stage events **Making decisions** • using probability in making decisions in social contexts	the probability of compound events can be determined the sum of the probabilities of complementary events equals 1 common social applications of probability	• determine probabilities for multi-outcome events by adding the probabilities of the component events • determine probabilities for multi-stage events by tree diagrams and multiplication • determine probabilities using the idea of complementary events • explain the connection between odds and probabilities • analyse activities involving chance in common social contexts such as medicine and gambling

Notes on the approach to learning in this topic

Students first begin to develop the concept that some events are unpredictable, that some outcomes are more likely than others, that some outcomes are impossible or certain and that outcomes can be equally likely. They have learnt to make judgments about the relative likelihood of two events based on previous experience and data.

Then they learn to quantify likelihood, obtaining a measure of likelihood called probability.

The approach to learning to quantify probability taken in this syllabus and sourcebook module is basically the same as that used in the past. However, the methods used are named and classified somewhat differently. The aim is to make the concepts and the distinction between methods clearer to students.

There are basically four things to learn about quantifying probability:
- the definition of probability
- approximating probabilities using data
- determining probabilities using indifference
- deriving probabilities from other probabilities

The definition of probability

For school mathematics, probability is best defined as the fraction of times the event would occur if tried a very large number of times.

Approximating probabilities using data

The above definition leads to experiment as the obvious way to determine probability. Of course, experiment will only ever give an approximation to the true probability. Students should have plenty of hands-on practice at finding probabilities by experiment, exploring issues like reliability of results from different numbers of trials and fallacies like memory, the dependence of the outcome of tossing a coin on which way up it was before it was tossed etc.

The essence of experimental determination of probability is that data are used. In some cases the data might be generated for the particular purpose of finding the probability. In other cases, the data may already be in existence. Medical probabilities are usually determined from existing data rather than by deliberately infecting healthy people for the sake of an experiment. To be inclusive, it is suggested that the approach be referred to as 'approximating probabilities using data' rather than approximating probabilities by experiment.

Determining probabilities using indifference

In approximating probabilities using data, students will notice that certain probabilities seem to come to simple fractions, while other seem not to. For example the probability of getting a head when tossing a coin seems to come to one half, while the probability of a matchbox landing on its side might come to about 28.7%. Many students will have an explanation for the nice fractions. The explanation lies in the fact that there is no physical difference between the two sides of the coin that will make one side any more likely to land upwards than the other. Thus the coin will give heads and tails an equal number of times in the long run and the probabilities will have to be equal at 50%. Similarly, when rolling a die, there is no such difference between the six faces. When throwing a matchbox, however, there is a material difference between the different faces – some faces are bigger than others – and this will make it more likely that the matchbox will land on some faces than on others.

Because of the lack of difference between the outcomes with the coin, the outcomes are said to be 'indifferent'. This indifference can be used to determine probabilities for coins and dice, but not for the matchbox. The Collins Reference Dictionary of Mathematics defines it as 'the principle that in the absence of any reason to the contrary, each possible outcome of an experiment is to be treated as equiprobable'. Indifference is the only way to determine the probability of simple events other than using data.

Deriving probabilities from other probabilities

This is the calculation of probabilities of compound events from known probabilities of simple events.

The first method is the calculating of the probability of a multi-outcome event such as the probability of a matchbox landing on its side or on its end. First, we need to know the probabilities of these simple events. Let's say they are 0.287 and 0.132 respectively. The probability of landing on its side or its end is then 0.287 + 0.132, ie. 0.419. Another example is the probability of drawing a king or queen from a pack of cards. There are four kings and four queens, each with a probability of 1/52, giving a total of 8/52.

The second method is the calculating of the probability of a multi-stage event such as the probability that two matchboxes will both land on their sides. This is 0.287×0.287, i.e. 0.082.

These techniques are crystallised into the addition and multiplications rules, though other techniques like tree diagrams and tables are often used to lead up to multiplication.

The third method is calculating the probability of an event using the probability that the event will not happen. This is the use of complementarity.

The terms ‘experimental probability’ and ‘theoretical probability’

Traditionally the data method is referred to as determining probability experimentally. The indifference method and determining probabilities from other probabilities are grouped under the heading of determining probability theoretically. This grouping, however, tends to obscure the important distinction between the two methods.

Another problem is the use of the terms 'experimental probability' and 'theoretical probability'. These terms convey the impression that there are two types of probability. There are not; there are just different ways of finding probabilities.

Relation to levels

Level 5 covers the first three things to be learnt:
- the definition of probability
- approximating probabilities using data
- determining probabilities using indifference

Level 6 covers the last one:
- deriving probabilities from other probabilities

Common difficulties and misconceptions

50-50

Many students use the expression 50-50 (or similar expressions) to describe situations where two outcomes are both possible, even though their probabilities might be quite different. During Activity 3, which introduces the concept of indifference, it is worth establishing that students have the concept of ‘equally likely’ (and thus ‘unequally likely’) and that they use the term 50-50 only for situations where two outcomes are equally likely.

n(E)/n(S)

Through traditional approaches to the teaching of probability, many students develop the concept of probability as the number of outcomes in the event we are interested in divided by the number of outcomes in the sample space. This is correct as long as all the outcomes are equally likely (indifferent). Often, however, students are unaware of the requirement for equal likelihood, and indeed, may be unaware that outcomes can be unequally likely.

The approach to probability taken here begins with situations where outcomes are not equally likely. This way it is hoped that students will see equal likelihood as something special. The n(E)/n(S) rule is a short cut to addition of identical fractions. It is suggested that, if n(E)/n(S) is to be introduced at all in Years 1 to 10, then it is introduced only after the derivation of probabilities by addition is mastered at Level 6. It is preferred that the formula is not introduced until senior.

Common fallacies

There are a number of fallacies which are held by significant numbers of people, including adults as well as students. Some of these are:
- that after a run of heads, a coin is more likely to come down tails
- that the way a coin lands depends on the way up it was before it was tossed
- that a six is the hardest number to get on a die
- that 1, 2, 3, 4, 5, 6 is a less likely lotto result than 34, 12, 42, 25, 22, 7, 30
- that knowledge of past lotto results can help in picking the winning numbers

These need not be deliberately debunked by the teacher. In fact they provide opportunities for realistic investigations. The logic can be discussed after the fallacies have been investigated by collecting data. These investigations form some of the activities at Level 5 and Level 6.

Overview of this module

This module contains six activities and an extension investigation. Activities 1 to 4 cover the basics. Activities 5 and 6 are designed to reinforce and broaden students' concepts and to develop their mathematical thinking. The extension investigation is designed to allow the more able and enthusiastic students to explore further ideas in probability.

Activity 1 – Goldfish

In this activity, students play a game that involves chance. The purpose is to introduce the idea that likelihood can be quantified, the meaning of 'probability' and the idea that probability is a good predictor in the long run, but that it may not be so good in the short run.

Activity 2 – Matchboxes and thumb tacks

In this activity students play the game they played in Activity 1, but using the tossing of matchboxes and thumb tacks rather than dice to determine wins and losses. The purpose is to reinforce the concepts of probability developed in Activity 1.

Activity 3 – Rainy days and measles

In this activity students learn that probabilities can be determined using pre-existing data rather than by generating data in an experiment. Sometimes, using pre-existing data is the only possible or only appropriate way to determine a probability.

Activity 4 – Indifference

In this activity students determine the probabilities associated with the tossing of a coin. The results are then used to develop the concepts of and the distinctions between the two ways of determining probabilities – using data and using indifference.

Activity 5 – Coin toss simulation

In this activity students use a pre-made spreadsheet to simulate a large number of tosses of a coin and to construct a graph of the fraction of heads to date after each toss. It is designed:
- to reinforce the idea that the fraction of heads corresponds well with the probability of getting a head for a large number of trials, but that it may not correspond well for a small number of trials
- to reinforce the idea that the probability of heads is 0.5
- to informally introduce students to the idea of simulations (though this concept is not required at Level 5 or Level 6).

Activity 6 – True or not true

In this activity students investigate one of a choice of common fallacies in probability by designing an experiment and collecting appropriate data. This helps students to debunk the fallacies, but more importantly it gives them practice in designing a research investigation and reporting the results.

Extension Investigation

Those students who choose to engage with this investigation will explore the concept of deriving probabilities from other probabilities. This allows students to do some preliminary thinking on the Level 6 Chance outcome.

The module also contains the How Tos for each elaboration for the outcome and a set of Review Questions for each elaboration.

Activity 1 – Goldfish

Overview and purpose

In this activity, students play a game that involves chance. The purpose is to introduce the idea that likelihood can be quantified, the meaning of 'probability' and the idea that probability is a good predictor in the long run, but that it may not be so good in the short run.

This activity might typically take about 1½ hours of class time.

Equipment and resources

- You will need about 20 paper goldfish. These can be made by photocopying Resource Sheet 1a (on orange paper if possible) and cutting out the fish. This is not essential but adds a bit of realism to the activity.
- You will need a copy of Resource Sheets 1b, 1c and 1d enlarged to at least A3 for display.
- You will need, for each student, an A4 sheet with Resource Sheet 1e copied onto both sides (or two separate sheets).
- You will need two dice for each one or two students.

Teacher instructions

Part 1 – Reinforces the idea of more and less likely

Explain that you are at the show. You want to win a goldfish and you are at a stall where you can do just that.

Display a cut-out goldfish and the large copy of Gameboard 1. Ask for a volunteer to play the game. Explain how the game works. The player pays $1, selects either the grey squares or the white squares on the board, then rolls two dice. (You might like to have some plastic money available or just act out the payment.) Check the number rolled to see what colour square it is in on the board. If it is the colour the player chose, they win the fish. If it isn't, they don't.

Play half a dozen games with different players, then discuss their reasons for their choice of colour. Most students will choose white and will probably explain why along the lines that there is more chance of winning with the white because there are more white squares.

Part 2 – Relates 'being more likely to win' to 'winning more often'

Change to Gameboard 2 and repeat the game with about a dozen players. Initially, students will probably go for white because there are 7 white squares and only 5 grey squares. But after a few games, they should start to notice that grey wins more often and should start to choose grey. Again, discuss the reasons for their choice of colour. Make sure that all students are aware of the distinction between there being more white squares and grey winning more often. Students should decide that grey is the best bet because it wins more often. Help students to see that 'being more likely to win' and 'winning more often' are just different ways of saying the same thing.

To check, play a couple of dozen more games, keeping a tally of which colour wins. These games can just simulated by rolling the dice and marking of the result on the tally.

Part 3 – Introduces probability of winning as the fraction of times one wins in the long run

Display Gameboards 2 and 3. Pose the question: "If there were two goldfish stalls beside each other at the fair, one using Gameboard 2 and one using Gameboard 3, which one would you go to . . . and why?" Give them a hint: grey is more likely to win on both Boards.

Bring out of the discussion the fact that we are after the board that will give us a goldfish the greatest fraction of the times we play. Point out that using the fraction of successes gives us a way of putting a number (fraction) on the likelihood and thus of comparing different likelihoods. Point out that we use the term 'probability' for the fraction of successes.

Explain that they have to find the probabilities for:
- rolling a grey square on Gameboard 2
- rolling a grey square on Gameboard 3

Hand out a sheet of paper with Resource Sheet 1e copied onto both sides. Copy the first half a dozen lines of the table and the leftmost few columns of the graph onto the board. Start with Gameboard 2. Roll the dice. Complete the first row of the table and plot the point on the graph. (It will be at 100% or at 0%.) Get students to put the same information on one side of their sheets. Then repeat for the second roll. (The graph point will now be at 100%, 0% or 50%.) Connect the points on the graph. Continue until students are comfortable with what is happening. Use a calculator to work out the difficult percentages and check that everyone knows how to do this.

Then let them continue individually or in pairs to finish their graphs. The Gameboard will need to be kept on display.

When they have finished, discuss the results. Discuss the fact that the graphs start with a lot of variation, but settle down to a fairly steady value. Also, compare the value that different students' graphs seem to be settling down to. Make sure that all realise that the fraction of wins is fairly unpredictable for a small number of rolls, but becomes more predictable for larger numbers of rolls. Ask them to predict what would happen if we did say 100 rolls – or 1000. Explain that we call the long term fraction of wins the probability of winning. Make sure that they realise that the fraction of wins will be close to the probability for a large number of games, but that it can be quite different for a small number of games.

Then remind them that they are trying to find which is the better Gameboard – 2 or 3. Get them to repeat the experiment for Gameboard 3 using the other side of the sheet.

When they have finished this, compare the probabilities for the two Gameboards and decide which would be the best game to play.

Extension

Some students might like to find out if it is possible to make a gameboard which gives a 50% probability of winning with grey and a 50% probability of winning with white. Some may do this by experiment and trial and error; others might choose to try a theoretical approach. Both are of value. If different students take different approaches, they might compare their work afterwards and explain their methods.

Variations

If the activity needs to be repeated, it could be varied in the following ways.

- The arrangement of colours and numbers on the gameboards could be changed.
- A gambling context could be used if this is acceptable with the group of students. For example, the teacher and student start with an few dollars each; if the student's colour comes up the teacher gives the students $1; if it doesn't, the student gives the teacher $1.

Long-term fraction is always the same for a given situation, short-term fraction varies, so use long-term for probability.
This probability is the most likely fraction for the short term.

Student instructions

Work through the following instructions answering the questions on paper as you come to them.

Part 1

Imagine you are at the show. You want to win a goldfish and you are at a stall where you can do just that.

Get a copy of Gameboard 1 (Resource Sheet 1b). You can cut out a few goldfish from Resource Sheet 1a as well if you want to make the game a bit more realistic.

The game works like this. The player pays $1, selects either the grey squares or the white squares on the board, then rolls two dice. (You might like to have some money available or just act out the payment.) Check the number rolled to see what colour square it is in on the board. If it is the colour the player chose, they win the fish. If it isn't, they don't. Play half a dozen games,

Q1. Explain in writing why you chose the colour(s) you did. Was this the best choice?

Part 2

Change to Gameboard 2 and play the game a few dozen times. Decide which colour you will use at the start and stay with that colour throughout.

Q2 Keep a tally of how many times you win and how many times you lose.

Q3 Which colour did you choose? Did it turn out to be the best choice? Why or why not?

Initially, you might have gone for white because there are 7 white squares and only 5 grey squares. But after a few games, you should start to notice that grey wins more often. In fact grey is the best bet because it wins more often. 'Being more likely to win' and 'winning more often' are just different ways of saying the same thing. This gives a more specific meaning to the term 'likelihood'.

Part 3

Get copies of Gameboards 2 and 3. Here is a question which you will now work on: *If there were two goldfish stalls beside each other at the fair, one using Gameboard 2 and one using Gameboard 3, which one would you go to . . . and why?* Here is a hint: grey is more likely to win on both Boards.

What you want is to find out which board gives you a goldfish the greatest fraction of the times you play (assuming you choose grey). Using the fraction of successes gives us a way of putting a number (fraction) on the likelihood and thus of comparing different likelihoods. We use the term 'probability' for the fraction of successes.

You have to find the probabilities for:

- rolling a grey square on Gameboard 2
- rolling a grey square on Gameboard 3

Get two copies of Resource Sheet 1e. Start with Gameboard 2. Roll the dice. Complete the first row of the table and plot the point on the graph. (It will be at 100% or at 0%.) Then repeat for the second roll. (The graph point will now be at 100%, 0% or 50%.) Connect the points on the graph. Continue until you have played 50 games. Use a calculator to work out the difficult percentages.

You will probably notice that he graph starts with a lot of variation, but settles down to a fairly steady value. This is because the fraction of wins is fairly unpredictable for a small number of games, but becomes more predictable for larger numbers of games.

Q4. Predict what the graph would look like if you did 100 rolls and if you did 1000 rolls.

We call the ***long term*** fraction of wins the probability of winning. Make sure you understand that the fraction of wins will be close to the probability for a large number of games, but that it can be quite different for a small number of games.

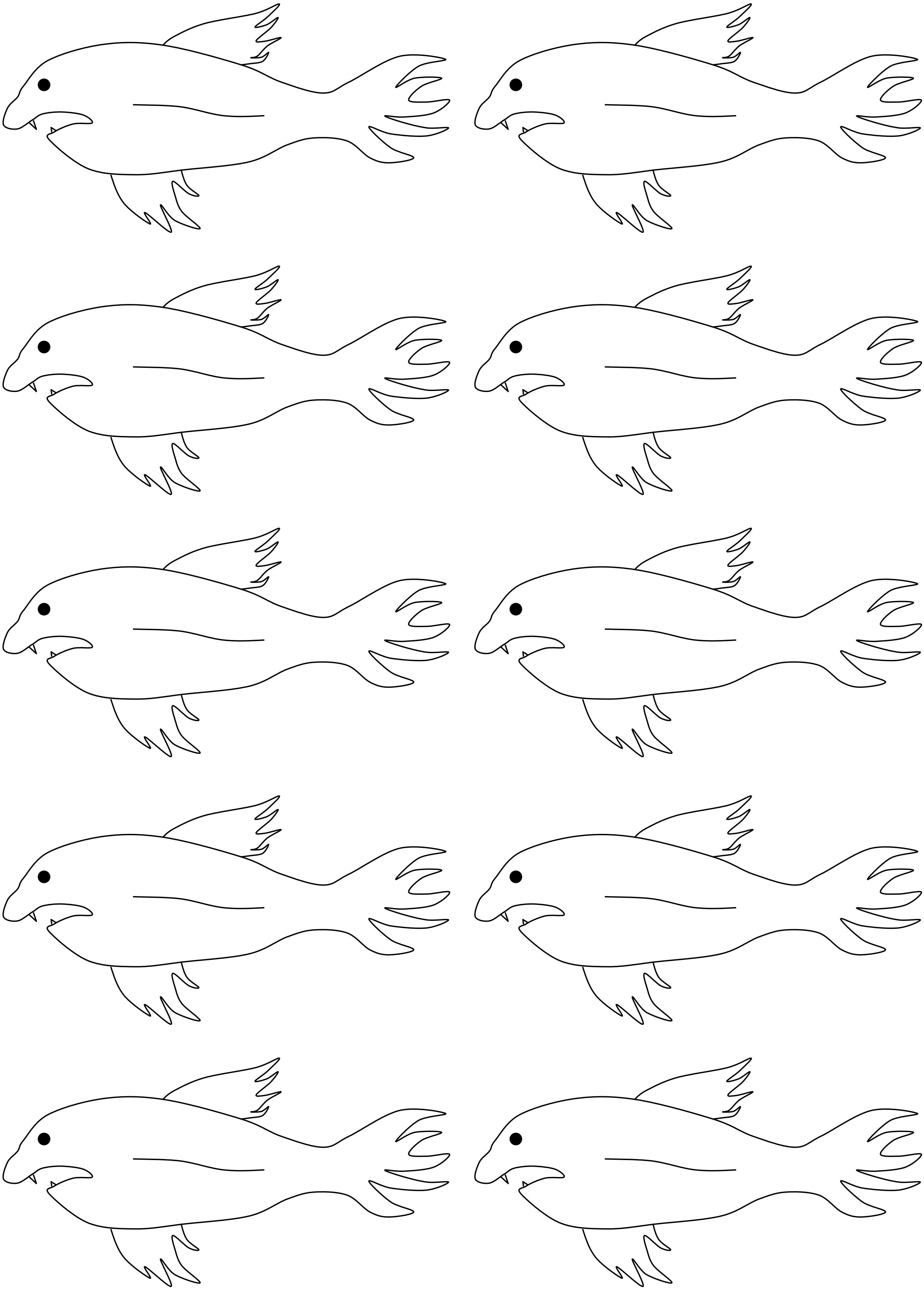
Now you are trying to find which is the better Gameboard – 2 or 3. Repeat the experiment for Gameboard 3 using the other copy of the sheet.

When you have finished this, compare the probabilities for the two Gameboards and decide which would be the best game to play.

Q5. Explain why we use the long-term fraction of wins for the probability and not a short-term fraction.

Extension

You might like to find out if it is possible to make a gameboard which gives a 50% probability of winning with grey and a 50% probability of winning with white. You may do this by experiment and trial and error; or you might choose to try a theoretical approach. Both are of value.



Gameboard 1

Choose the grey boxes or the white boxes

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

Gameboard 2

Choose the grey boxes or the white boxes

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

Gameboard 3

Choose the grey boxes or the white boxes

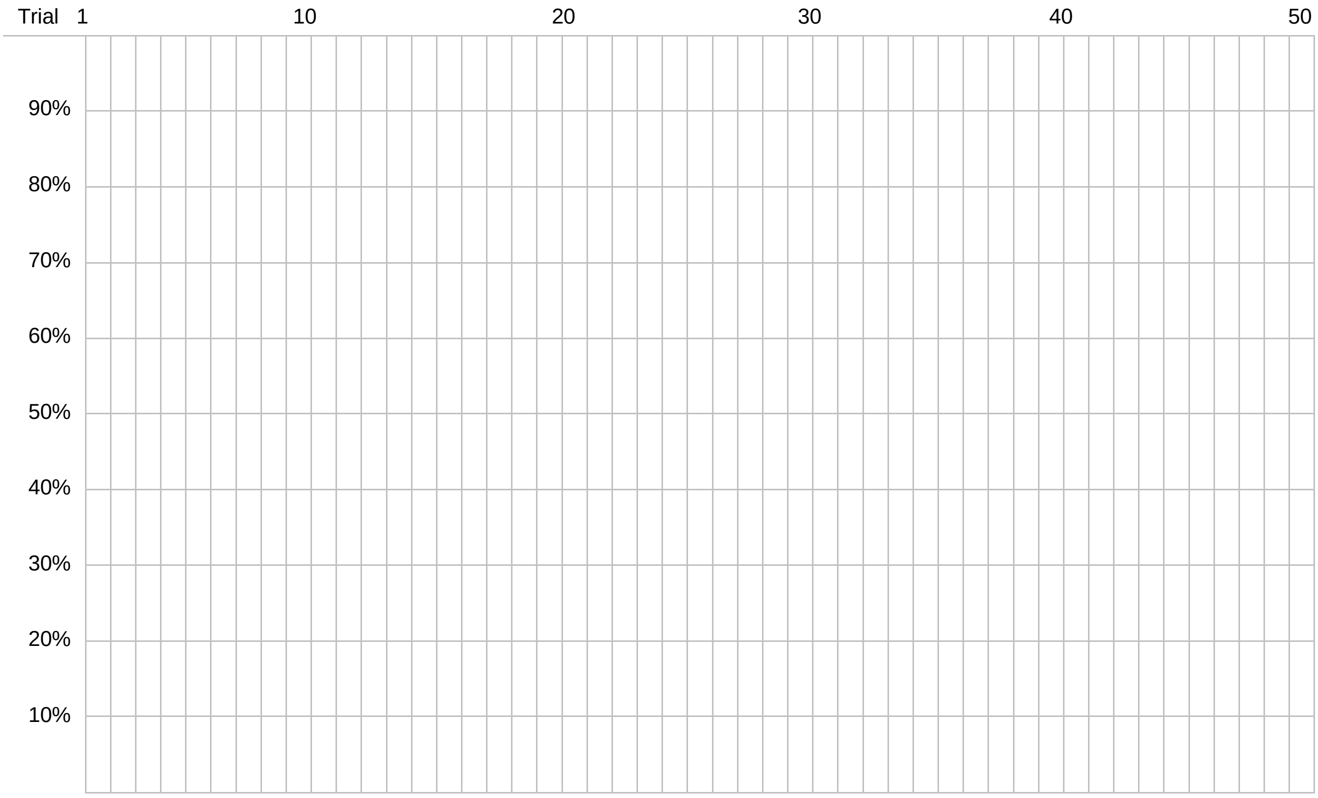
1	5	9
2	6	10
3	7	11
4	8	12

Progressive percentage of wins – table

Throw number	Win ✓ or ×	Wins so far	Percentage wins
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			

Throw number	Win	Wins so far	Percentage wins
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			

Progressive percentage of wins – graph



Activity 2 – Matchboxes and Thumb Tacks

Overview and purpose

In this activity students play the game they played in Activity 1, but using the tossing of matchboxes and thumb tacks rather than dice to determine wins and losses. The purpose is to reinforce the concepts of probability developed in Activity 1.

This activity might typically take about 1 hour of class time.

Equipment and resources

Each group of two or three students will need a matchbox and a thumb tack.
The student worksheet if you decide to use it.

Teacher instructions

The black line master on the next page, entitled *Goldfish, Matchboxes and Thumb Tacks*, details an investigation which the students should pursue as independently as possible. The aim of the investigation is to give students independent practice at using data to find probabilities.

You might hand the black line master to the students or you might present the information on the black line master to them orally.

Get them to work in groups of two or three. Leave them to work out how they will attack the problem. Encourage groups, but give guidance only if necessary. Try to avoid telling them how to do it.

Variations

The matchbox and/or the thumb tack could be replaced by other situations which have two outcomes and where students will not be able to predict the probabilities of the different outcomes. Examples are

- tossing a bent coin
- throwing a construction made of four or five Multilinks
- drawing three cards from a pack and getting three different suits
- etc

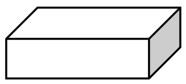
Student instructions

Follow the instructions on Resource Sheet 2.

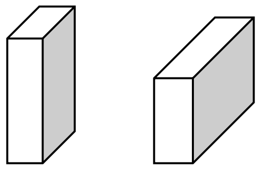
Goldfish, Matchboxes and Thumb Tacks

You will be playing the goldfish game. Instead of using dice and the number boards, however, you will have a choice of using a matchbox or a thumb tack.

If you use the matchbox, you have to chose ‘flat’ or ‘not flat’. The matchbox gets thrown on the floor. ‘Flat’ means that the matchbox lands lying on it largest face. ‘Not flat’ means that it lands lying on its edge or on its end.

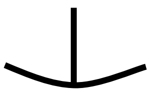


flat

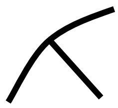


not flat

If you choose the thumb tack, you have to choose ‘point up’ or ‘point down’. The tack gets dropped on the floor.



point up



point down

You have to work out how to give yourself the best chance of winning a fish and what the probability of winning is.

You should present a report explaining:
- what the problem is
- how you went about solving it
- the data you collected
- your conclusion from the data

Activity 3 – Rainy Days and Measles

Overview and purpose

In this activity students learn that probabilities can be determined using pre-existing data rather than by generating data in an experiment. Sometimes, using pre-existing data is the only possible or only appropriate way to determine a probability.

This activity might typically take about 1 hour of class time.

Equipment and resources

Resource Sheets 3a, 3b and 3c

Teacher instructions

Part 1 – Introduces the idea of using pre-existing data where it is impossible to generate it by experiment.

Give students Resource Sheet 3a. Ask them to read it and to produce the required report. Help them as necessary.

Part 2 – Introduces the idea of using pre-existing data where it is unethical to generate it by experiment.

Give students Resource Sheet 3b. Ask them to read it and to answer the questions. Help them as necessary.

Then discuss the use of pre-existing data and ask the students for other situations where pre-existing data might be used instead of performing an experiment. These might include:

- probabilities related to outcomes of other diseases, medical procedures etc
- probabilities related to accidents
- probabilities related to cyclones, earthquakes and other natural disasters
- probabilities related to collapse of buildings or crashes of aircraft because of mechanical problems

These can be grouped into two categories:

- those where it is morally or legally unacceptable to perform an experiment and
- those where it is impossible to collect experimental data or more difficult to collect experimental data than to find pre-existing data.

To calculate probabilities, we need data. It doesn’t matter whether we generate the data ourselves or whether we use pre-existing data.

Make sure that these ideas are clear to the students. Then go on to Part 3.

Part 3 – Asks students to decide how they would collect data in different situations – whether they would conduct an experiment or whether they would search for existing data.

Give students Resource Sheet 3c. Ask them to read it and to answer the questions. Help them as necessary.

Variations
Other situations and data could be used. Medical data is abundant and interesting. Using real data on an issue of concern to the students can be very motivating.

Student instructions

Follow the instructions on the following pages.

Rainy Days

The Groddleston Primary School P and C were discussing the school fete. It is held every year in November, but it had rained on the previous two fete days.

‘Would it be possible to hold it in a different month?’ said Harry, the P and C president.

‘How would that help?’ said Chris.

‘Well, it’s less likely to rain in some months than others’ said Harry, ‘and maybe we should find a month when it is less likely to rain.’

‘I did probability at school.’ said Kate. ‘Maybe we could do an experiment to find the probability of it raining in different months.’

‘How would we do that?’ said Darryl.

‘We could try about 100 November days and see how many of them are wet’ said Kate, ‘then we could do the same with 100 October days and so on.’

‘You mean just wait for the days to come round?’ said Darryl.

‘Er . . .’ said Kate. ‘I suppose that would take us a few years, wouldn’t it’

‘Yes and we’ve got to decide this year.’ said Harry.

‘I know.’ said Kate, ‘We’ll got hold of the rainfall data for the past few years. We can work out the probabilities from that.’

So they agreed that Kate would get hold of the data, work out the probabilities of it raining on a day in each of the months, and bring a recommendation to the next P and C meeting.

Then they discussed the toilets.

The data Kate collected is shown below. Imagine you are Kate. Work out the probabilities the P and C wants (explain how you worked them out), then make a recommendation for the timing of the fete.

Rainfall in millimetres at Groddleston from 1995 to 2000. The first number is the date on which it rained. The second number is the number of millimetres.												
1995												
Jan	3-12	4-7	2-25	8-1	10-3	14-4	15-1	17-4	22-1	28-6		
Feb	7-2	11-6	12-88	13-8	14-42	15-53	16-11	18-6	19-2	25-6	26-5	
Mar	7-14	11-1	17-1	25-12	26-17	27-4	29-26					
Apr	5-1	14-4	15-1	16-2	23-1	26-24	27-4	29-26				
May	4-2	8-1	9-6	10-3	11-2	17-12	18-2	21-9	22-1			
June	1-2	6-1	15-12	16-16	28-7							
July	7-4											
Aug	15-32	16-19										
Sep	3-5	21-1	24-2	26-1								
Oct	1-12	3-2	7-2	8-22	9-1	22-3	26-14					
Nov	1-1	5-55	6-11	16-21	17-27	18-4	20-68	21-1	22-20	23-15		
Dec	1-40	6-5	11-43	12-3	16-4	17-24	18-52	19-7	23-69	24-2	28-11	
1996												
Jan	2-4	3-53	4-42	5-4	9-57	10-36	20-1	23-6	27-3	29-2		
Feb	3-20	7-6	9-9	15-7	16-1	18-18	23-5	24-4	25-1	27-1	28-1	29-4
Mar	1-8	10-2	11-2	12-2	23-1	24-5	26-1					
Apr	23-15	24-7	27-61	30-9								
May	1-69	2-144	3-150	4-27	5-110	6-29	7-3	14-3	17-4	18-2	20-3	
June	1-10	2-6	3-7	6-1	7-2	15-19	16-1					
July	12-1	27-2	28-36									
Aug	6-1	16-2	17-2	18-14	30-42							
Sep	20-8	30-18										
Oct	6-8	7-6	20-4	29-4	30-5							
Nov	5-2	6-10	7-5	8-4	17-12	20-9	21-10	22-1	24-76	25-1		
Dec	4-37	7-20	8-6	9-14	10-3	11-11	12-23	18-29	19-2	21-1	31-1	

1997												
Jan	1-1	4-6	5-2	18-22	19-3	25-34	26-24	27-61	28-32	29-2	30-3	31-1
Feb	2-3	14-14	15-35	16-2	18-4	19-7	27-42					
Mar	6-1	7-2	9-3	26-1	31-14							
Apr	3-13	12-1	20-11	26-1	29-11	30-3						
May	1-16	2-1	3-28	4-15	5-29	7-10	8-5	16-8	17-34	30-8	31-12	
June	16-9	21-7	25-3	28-2								
July	12-2	13-15	24-5	25-1	26-1	27-16	28-1	30-1				
Aug	7-1	21-5	24-1									
Sep	2-1	20-19	21-18	24-30	29-3							
Oct	6-3	7-27	8-26	19-8	20-30	29-4	30-5					
Nov	5-6	6-49	7-2	12-3	15-18	16-8	17-3	18-33	19-2	20-4	30-17	
Dec	7-10	8-2	10-42	15-12	24-12	27-12						
1998												
Jan	9-1	14-4	15-1	28-57	31-51							
Feb	1-1	7-18	9-7	10-18	11-5	14-4	16-3	17-1				
Mar	1-3	4-1	8-2	19-25								
Apr	1-1	2-2	3-16	10-5	11-1	14-11	15-22	16-1	17-3	18-2	21-2	23-14
	24-8	28-3										
May	2-18	5-45	14-18	15-8	16-70	17-1	31-1					
June	1-4	16-4	23-2	29-3	30-2							
July	4-4	8--1	19-10	22-3	25-2	26-7	27-2	28-12				
Aug	5-48	6-4	16-3	20-1	21-10	22-18	23-5	24-1	27-9	28-1		
Sep	9-8	10-12	11-61	13-5	14-18	21-3	25-51					
Oct	13-22	25-3	26-1									
Nov	13-28	14-7	17-1	18-20	22-17	24-15	26-4	30-6				
Dec	1-1	5-6	16-12	19-2	20-1	23-5	24-147					
1999												
Jan	1-21	2-27	3-8	5-2	9-32	13-10	20-1	21-11	22-2	28-8	29-2	31-2
Feb	1-31	2-3	4-3	7-12	8-42	9-72	10-10	11-10	12-1	25-4	28-25	
Mar	1-13	2-5	3-36	10-3	12-9	20-44	23-21					
Apr	4-31	10-53	11-26	12-2	15-8	18-5						
May	6-15	7-7	8-27	10-9	11-6	13-8	19-7	20-3	21-16	22-10	23-7	
June	5-52	6-12	9-32	23-17	24-10	26-2	27-32	28-23	29-30			
July	1-32	9-6	10-9	24-34	27-6							
Aug	3-5	23-2	24-13	25-1	27-1	28-13	29-25	31-14				
Sep	2-2	10-3	11-31	27-19	28-9	29-18	30-5					
Oct	1-20	3-7	4-30	14-30	15-2	18-4	24-20	27-20				
Nov	6-1	7-9	8-18	9-11	12-2	17-6	23-10					
Dec	10-26	11-31	18-45	27-21	28-26							
2000												
Jan	5-3	13-2	15-19	17-11	28-25	30-5						
Feb	7-8	13-13	14-1	26-7	28-7	29-12						
Mar	2-2	4-1	8-31	9-25	10-1	18-4	19-2	21-3	22-4			
Apr	2-1	5-1	11-5	15-3	23-2	27-28	28-9	29-3	30-2			
May	1-10	2-28	4-2	24-14								
June	10-7	11-36										
July	12-9	27-10										
Aug	7-2	8-3										
Sep	2-1											
Oct	10-6	11-1	15-15	24-5	25-1	30-11						
Nov	2-9	3-1	4-2	5-1	13-22	15-1	17-24	20-10				
Dec	8-30	26-4	27-3	28-19	29-20	30-4						

Measles

Using existing data

In the Rainy Days part of this activity, we saw a situation where someone needed to calculate some probabilities, but where it was not possible to collect the required data by experiment – in this case because there wasn't enough time before the next fete. Because of this it was necessary to use existing data to work out the probabilities.

Another situation where we might need to use existing data rather than collect it by experiment would be if we wanted to work out the probability of a person living to the age of 90. We wouldn't pick a bunch of babies and wait 90 years to see how many of them were still alive. Instead we would use data which governments keep on ages at which people die.

These are situations where it is impossible to collect the data in the time available. There are other situations where it is theoretically possible to collect the data by experiment, but where it is not ethical to do so. Here is an example.

Ross River Fever is an illness which people get by being bitten by infected mosquitos. Most people get over it in a few weeks, but in some it develops into Chronic Fatigue Syndrome which can last for years. Suppose we wanted to know the probability that a person who gets Ross River Fever develops Chronic Fatigue Syndrome as a result.

It would be possible to get 100 people, infect them all with Ross River Fever and see how many develop Chronic Fatigue Syndrome. But most people would not consider this ethical – and besides, it would be hard to get 100 volunteers!

Instead, we would use data from existing medical records.

Sharon's baby

Sharon has had a baby boy called Kieran. The nurses have encouraged Sharon to get Kieran immunised against various diseases. Sharon has heard, however, of some babies having adverse reactions, including seizures) to the MMR (measles-mumps-rubella) vaccine. She is worried that her baby might have such a reaction and wonders whether she should get him immunised. On the other hand, she doesn't want him to get measles because she knows measles can have serious complications.

[Measles is an illness which usually occurs in children. It involves inflammation of the air passage and lungs, fever and a rash. Complications include pneumonia, inflammation of the brain and eyes and seizures. The immunisation works by injecting weakened measles viruses. These can cause mild versions of the same symptoms in some children.]

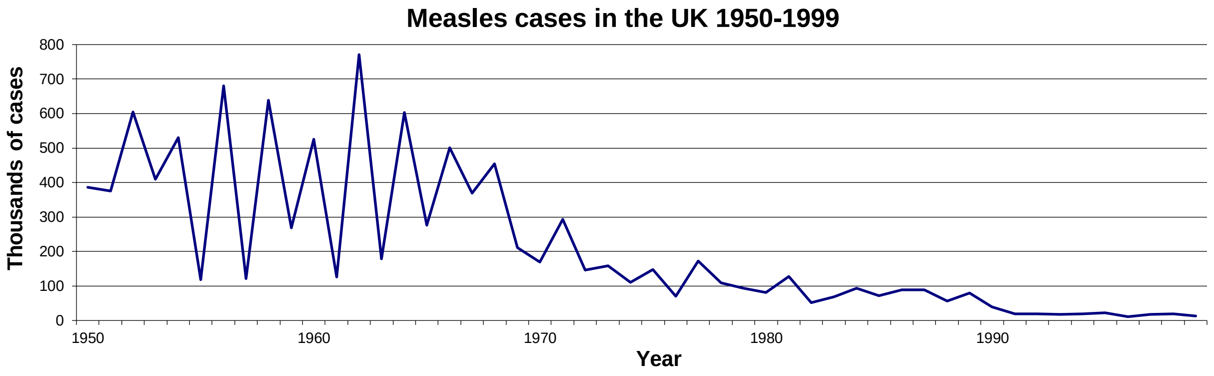
Sharon would like to know things like:

- the probability of getting measles if Kieran is not immunised
- the probability of getting measles if Kieran is immunised
- the probability of getting serious complications if he gets measles
- the probability of getting serious reactions if he is immunised.

Obviously we do not find the probability of getting serious complications of measles by infecting a couple of hundred kids and seeing how many get them. Instead we find existing data from public medical record.

Sharon decided to search the Internet for information. She found various sites and from them put together the following information.

From Site 1



- If people get measles it is usually in childhood. It is rare to get it more than once.
- The first measles vaccine was introduced in the late 1960s. The more effective MMR vaccine was introduced in the late 1980s .
- Between 1950 and 1959 inclusive, 9.1 million people were born in the UK. In the same period, there were 4370 deaths from measles.
- Between 1990 and 1999 inclusive, 8.2 million people were born in the UK. In the same period, there were 530 deaths from measles.

From Site 2

The occurrence of some side effects of measles and Measles-Mumps-Rubella immunisation in Australia in the 1990s

	Measles	MMR Immunisation
Pneumonia	1 in 20 cases	0
Fits/convulsions	1 in 200 cases	1 in 1000 cases
Brain inflammation	1 in 500 cases	1 in 1 million cases
Severe allergic reaction	0	1 in 100 000 cases
Death	1 in 9 000	0

From Site 3

Many countries, particularly third world countries, do not have public measles immunisation programs. In 1999 about 130 million children were born, 43 million got measles and 1 million of those died of it.

Questions

Use the information from Site 1 to answer the following questions.

1. In the United Kingdom in the 1950s, what was the approximate annual number of measles cases?
2. In the United Kingdom in the 1950s, what was the approximate annual number of births?
3. In the United Kingdom, what was the probability that someone born in the 1950s would get measles?
4. In the United Kingdom in the 1950s, what was the probability that someone who got measles would die of it?
5. In the United Kingdom, what was the probability that someone born in the 1950s would die of measles? (This is not the same as the last question.)

6-10. Answer the same questions for the 1990s.

11. Was the situation with measles better, worse or the same after the introduction of immunisation?

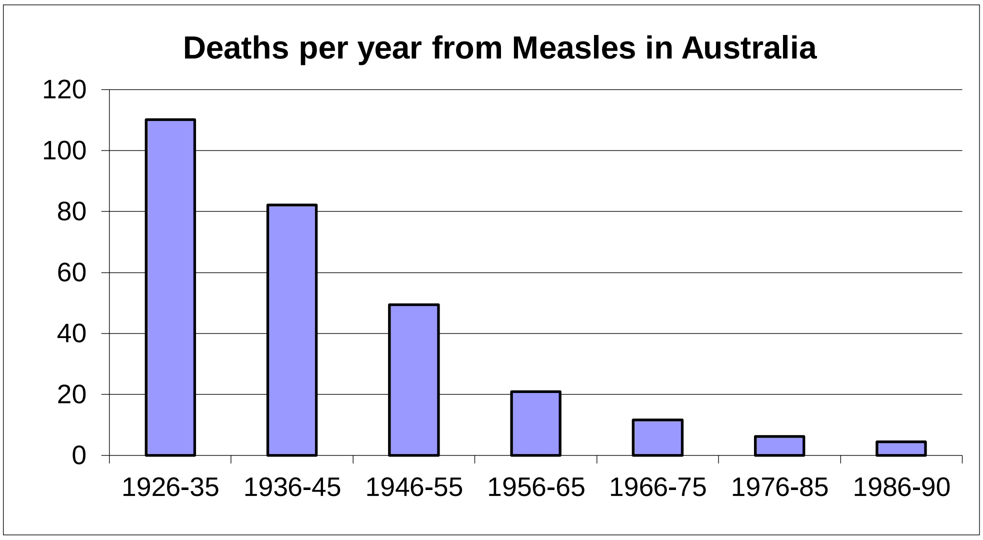
Use the in from Site 2 to answer the following question.

12. Rewrite the table to show probabilities as percentages.

Use the information from Site 3 to answer the following questions.

13. What was the probability that someone born in the world in the 1990s would get measles?
14. What was the probability that someone born in the world in the 1990s would die of measles?
15. Why might the probability of dying from measles be higher in the world now than it was in the UK prior to immunisation?

16. Remember that Sharon is concerned about Kieran getting measles, but that she is also worried about possible side effects of immunisation. Use the data above to write a paragraph or two to Sharon to convince her either to get Kieran immunised or not to get him immunised.



What data to use

Imagine that you needed to find the following probabilities. In some cases you might try to find existing data because performing an experiment would be impossible or immoral. In other cases you might perform an experiment because it is easy to do so or because relevant data may not exist. In some cases you might be able to do it either way.

For each situations, say whether you would conduct an experiment to collect the data you need or whether you would try to find existing data. In each case explain way.

1. You are given a bent coin and want to know the probability that it will land heads up when tossed.
2. You want to know the probability that a cyclone will cross the Queensland coast next summer.
3. You want to know the probability of dying if you get tetanus.
4. You want to know the probability that someone will yawn if you yawn in front of them.
5. You want to know the probability that a two-year-old will cry if you give them an ice block, then take it away again.
6. You want to know the probability that a particular type of brick will break if you drop it onto a concrete floor from a height of 2 metres.
7. You want to know the probability that there will be an accident on a particular stretch of road in a given week.
8. You want to know the probability that a person who is now 60 years old will live to be 90.
9. You want to know the probability that a sample of a given brand of matchbox will contain more than 50 matches.
10. You want to know the probability that a woman will give birth to twins if she gives birth.
11. You want to know the probability that the number 14 will come up in the next Lotto draw.
12. You want to know the probability that you will get peas with your dinner.
13. You want to know the probability that a 20 cent piece picked at random will be more than 10 years old.
14. You want to know the probability that an adult picked at random is a teacher.
15. You want to know the probability that a student picked at random from your class gets the same answer as the teacher on all 15 of these questions.

Activity 4 – Indifference

Overview and purpose

In this activity students determine the probabilities associated with the tossing of a coin. The results are then used to develop the concepts of and the distinctions between the two ways of determining probabilities – using data and using indifference.

This activity might typically take about 1 hour of class time.

Equipment and resources

Resource Sheet
Coins (one per group of two or three)

Teacher instructions

This activity is presented as a student sheet. Able and literate students will be able to read Part 1, conduct the activity and develop the required concepts. Less able, less literate or less motivated students may need to be lead through the activity to some degree.

It is advisable for the class to work through the activity at about the same speed. The class will need to work together at the end of Part 1 where students are asked to pool their data to get a more reliable class approximation of the probabilities. It will be necessary to set aside a time when most students have reached that stage to do this.

It is also recommended that students read the text of Part 2 together, then discuss its meaning as a whole group. It is also suggested that the class tackle each of the ten questions in turn. For each question, students should discuss the question in small groups and come to a conclusion, then groups should compare positions and discuss any differences. This will help consolidate a strong common understanding of the concepts involved.

Variations

Students could extend Part 2 by listing other situations in which indifference can be used and other situations in which it cannot. They could do this in the form of further questions for other students and/or class discussion.

Student Instructions

Follow the instructions on the following Resource Sheet.

In Part 1 you will not be able to put the class's data together. You may like to do a larger number of tosses yourself instead.

Indifference – Part 1

You could have played the goldfish game with a coin instead of a matchbox or thumb tack. When the coin is tossed, it can land heads up or tails up.

You will now find the probability of getting a head and of getting a tail when you toss a coin. A tally chart is provided below.

		Tally	Total	Percent
Coin	heads			
	tails			

You should be able to do a couple of hundred tosses. Working with a friend is a good idea. One of you can toss while the other keeps the tally. As you know, you would probably get a more accurate result if you did a few thousand tosses. That would take you too long and get boring, but what you can do is put your results together with everyone else's in the class. The teacher will arrange this once you have collected enough data.

Indifference – Part 2

You will hopefully have noticed that the probabilities for heads and tails come to be about the same. They should both be 50%, meaning that if you tossed the coin very many times, you would get heads on very close to half the tosses and tails on very close to half the tosses. This probably won't come as a shock. The reason people use a coin to toss before a game is that it is fair. In other words, in the long run there will be as many heads as tails. The probabilities of heads and tails are equal. They are both 50%.

But you probably wouldn't have got a nice round 50% for the matchbox or thumb tack in the last activity.

Why is it that the coin gives you exactly 50%, but the matchbox and thumb tack give probabilities like 63.9%?

Have a think about this before you read on.

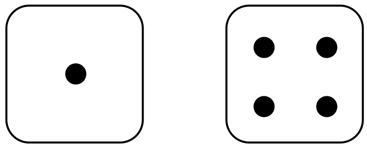
The answer lies in the fact that with the coin there is no physical difference between the two sides which could make one of them more likely to land upwards than the other. So, in the long run, heads and tails must both occur 50% of the times. In other words, both have a probability of 50%.

The mathematical word '*indifference'* is used for this lack of difference. A situation is said to be indifferent if all the possible outcomes are equally likely.

There is indifference in a coin, but not in a matchbox. With a matchbox, there is a physical difference between the various faces that might make the box more likely to land up some ways than others – some faces are bigger than others.

This indifference allows us to find the probability of getting a head or a tail without actually tossing a coin. So we now have two ways of finding probabilities. We can collect data. Or, if the situation involves indifference, we can use this indifference to find the probability.

Note that we can always use the data method, but we can only use the indifference method if there is indifference between all the possible outcomes.

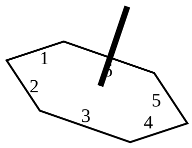


Basically, if we are throwing an object and are interested in the way up it lands, then to be indifferent, the object must look the same when looking at each of the faces. Apart of course from any pictures on the faces which won't affect the way it lands. So a die is indifferent. If you look straight at the 'one' face, the die will look the same as if you look straight at the 'four' face (or any other face). The only difference will be the spots and they don't affect the way a die lands.

So every face of a die has the same probability of landing upwards. In the long run each number will come up one sixth of the times, ie 16.66% of the times. So the probability of getting a four, say, on a die is one sixth or one in six or 16.66%.

For each of the following situations, explain whether the idea of indifference can be used to find probabilities or whether we would have to collect data. If there is indifference, work out the probabilities involved.

1. A bent coin is tossed. We want to know the probabilities of getting a head and of getting a tail
2. A spinner is made by sticking a match through the middle of a regular hexagon made of card. Each edge of the hexagon has a number written on it. The number obtained is the number on the edge that is resting on the table. We want to know the probabilities of getting the different numbers.



3. A bag has three marbles in it. They are the same size, weight and feel, but different colours (one is blue, one is green and one is red). We shake the bag then put in a hand without looking and pick out a marble. We want to know the probability of getting the blue one.
4. A bag has three marbles in it. They are the same colour, but different sizes. We shake the bag then put in a hand without looking and pick out a marble. We want to know the probability of getting the biggest one.
5. A student throws her maths book at the teacher from the other side of the classroom. The book might hit the teacher or it might miss. We want to know the probability that it will hit.
6. A pack of 52 playing cards is shuffled, then someone picks a card without looking. We want to know the probability that they will get the Ace of Hearts.
7. Two old ladies, Edith and Martha, catch the plague. We want to know the probability that Martha will die of it.
8. A rat is offered two pieces of food, an apple and a piece of dried zucchini. We want to know the probability that it will bite the zucchini first.
9. A rat is dropped into a square box. There are two exits. They are at opposite ends of the box. Both are the same size and shape and nothing can be seen through either except darkness. We are interested in the probability of it taking each exit.
10. A girl selects a letter as follows. She opens a book at a random page, and sticks a pin into the page without looking. She then looks to see what letter is closest to the pin hole. We are interested in which letter of the alphabet she selects.

Activity 5 – Coin toss simulation

Nominal time

About 30 minutes.

Overview and purpose

In this activity students use a pre-made spreadsheet to simulate a large number of tosses of a coin and to construct a graph of the fraction of heads to date after each toss. It is designed:

- to reinforce the idea that the fraction of heads corresponds well with the probability of getting a head for a large number of trials, but that it may not correspond well for a small number of trials
- to reinforce the idea that the probability of heads is 0.5
- to informally introduce students to the idea of simulations (though this concept is not required at Level 5 or Level 6).

This activity might typically take about ½ hour of class time.

Equipment and resources

Spreadsheet file (coin-toss-simulation.xls)

Teacher instructions

Explain to students that we want to produce a graph like the one we produced in Activity 1 (Resource Sheet 1e), but for the percentage of heads when tossing a coin. Explain that we will take a short cut by using a spreadsheet.

Ask them to predict what the graph will look like by sketching it before running the spreadsheet. By this stage, they should have a reasonable idea that the graph will start with considerable fluctuations, but then settle down to something around 0.5. Check whether they predict this correctly.

Then run the spreadsheet for them or get them to run it individually (depending on how many computers you have access to) a few times for 50 tosses and then a few times for 1000 tosses.

Afterwards, discuss

- the fact that the long-term percentage is always about 50%, but that even after 1000 tosses, it can be a few percent either side
- the fact that the short-term percentage is much less predictable.

Extension

- Students can work out percentages of heads for 10 000 or more tosses by running the sheet 10 or more times and combining the results.
- They can work out how the spreadsheet works and modify for other numbers of tosses and for other probabilities.
- They should realise that the fraction of heads generally gets closer and closer to 0.5 as the number of tosses increases. They can investigate whether the number of heads tends to get closer to or further from half the number of tosses as the number of tosses increases.

Student instructions

In Activity 1 you produced graphs of progressive percentages of wins using Gameboards 2 and 3. You might like to look at the graphs you produced.

Now imagine you were using a coin instead of dice and gameboard to decide whether you won the fish.. Imagine a head is a win and a tail is a loss. Sketch what the graph of progressive percentage of wins might look like in this case.

If you don't have access to a computer, that is the end of this activity. If you do, read on.

You will now use a pre-make spreadsheet to simulate 50 tosses of a coin and to produce a graph of progressive percentage of wins.

Load the file Cc5 Coins.xls onto you computer and open it.

It should show a graph of the progressive percentage of wins for 50 tosses of a coin.

The spreadsheet uses a random number generator to produce a random number between 0 and 2. It counts it as a win if the number is greater than 1, a loss if it is less than 1. If you are interested, have a look at the formulae used and try to understand how the spreadsheet works.

How does the graph compare to the graphs you produced in Activity 1?

You can get the spreadsheet to do another 50 tosses by placing the cursor on a blank cell and pressing the Delete key. Do this several times to see what variation you get.

How close to 50% does the total usually end up after 50 tosses?

There is a second sheet to the spreadsheet called '1000 tosses'. This does the same experiment but with 1000 tosses instead of 50. Get the spreadsheet to perform the 1000-toss experiment a few times.

How close to 50% does the total usually end up after 1000 tosses?

Extension

You can do the following if you are interested.

Work out the percentage of heads after 10 000 tosses by running the 1000-toss sheet 10 times. The exact number of wins from 1000 tosses will be shown in cell C1001. How close is this to 50%. Repeat the exercise a few times if you like.

You should by now be convinced that as the number of tosses increases, the percentage of wins tends to get closer to 50%. Now look at the actual number of wins above or below the half expected and see if that tends to increase or decrease with more tosses.

Activity 6 – True or not true?

Overview and purpose

In this activity students investigate one of a choice of common fallacies in probability by designing an experiment and collecting appropriate data. This helps students to debunk the fallacies, but more importantly it gives them practice in designing a research investigation and reporting the results.

This activity might typically take about 1 hour of class time.

Equipment and resources

Resource Sheet 6 if you choose to use it.
Dice and coins.

Teacher instructions

You might photocopy the student worksheet for the students or you might write the activities on the board or show them on an overhead projector.

You may wish to let the students choose the investigation they wish to pursue or you might prefer to assign them. For the presentation of the results, it can be worth ensuring that some students work on each of the three investigations.

Get the students, as far as possible, to work out their own approach to the investigation, and to collect, organise, interpret and present their data. Encourage them to attempt to work through problems themselves and to ask you for help only as a last resort.

It is strongly advised that the students work in small groups for this activity.

Variations and extensions

If, after the presentations, students still hold different opinions about the truth of any of the statements, this can be used as a basis for a class debate. The debate could be restricted to logical arguments rather than invoking experimental data.

The class might also discuss the wisdom of various strategies for picking lottery numbers. Some of these strategies are:

- choosing the numbers that have come up most frequently in the past
- choosing the numbers that have come up least frequently in the past
- choosing the numbers 1, 2, 3, 4, 5, 6, 7 and 8
- choosing the same numbers each week rather than changing them each week.

Student instructions

Follow the instructions on the following resource sheet.

True or not true?

Despite what you learnt in the last activity, there are some things which are fairly well known, but which may or may not be true. Three of these are:

1. When you roll a die, a six is the hardest number to get.
2. When you toss a coin, the probability of getting a head will depend on whether the coin was head up before you tossed it or head down.
3. If you toss a coin three times in a row and get all heads, you are more likely to get a tail than a head on the next toss.

Your job is, as a small group, to investigate one of these by collecting data and then to write a report explaining how you investigated it, what data you obtained and how you came to your conclusion. Your teacher may then ask some groups to present their findings to the class.

It is your job to work out how to go about the investigation. Your teacher may give you some help if you are totally stuck, but your request for help will be taken down and may be used in evidence against you!

Extension Investigation

Prerequisite activities

Students will need to have completed at least Activities 1 and 2 in this module before engaging with this investigation.

Overview and purpose

This activity is designed to extend students who wish to go beyond Activities 1 to 6. It is an open investigation. The stimulus is provided on Resource Sheet E.

No program time needs to be assigned. The activity is just for those students who wish to engage with it, either in their own time or while other students are completing Activities 1 to 6.

Equipment and resources

Resource Sheet E if you choose to use it.
Students may ask for coins to simulate the probabilities involved.

Teacher instructions

The stimulus can be given individually to those students who wish to engage in this activity. It can be given as a sheet, orally or via the board. Students should be able to proceed without further teacher input, though some students may benefit from suggestions of things to think about if they get stuck. Working in small groups would be beneficial.

Student instructions

Follow the instructions on the following resource sheet.

A wet weekend

You might have heard expressions like ‘Chance of a shower’ on the weather forecast. In the US, weather forecasters often attempt to be more precise by making statements like

‘20% chance of a shower’

or

‘90% chance of snow’.

A news reader on an American news show once read the weather forecast as

‘A 50% chance of rain on Saturday and a 50% chance of rain on Sunday’.

Off the cuff, he then added

‘That means 100% chance of rain on the weekend’

Was he right? Try to develop some mathematical ideas that can be used to work out such probabilities correctly.

To help you on your way, you might think about the following question: If you know the probability of one thing happening and the probability of another thing happening, how do you find the probability that

- either the first thing or the second thing will happen
- both the first thing and the second thing will happen
- the first thing will not happen
- neither the first thing nor the second thing will happen?

Meanings of New Words

The meanings of words which are new to this module are explained below. The meanings of other mathematical words can be found in the **Glossary**

Outcome An outcome is something that can happen as the result of a situation or action. Some situations or actions can result in a number of different possible outcomes. For example, catching the plague can result in death or recovery, tossing a coin can result in a head or a tail, rolling a die can result in a one, a two, a three, a four, a five or a six. Different people might be interested in different sets of outcomes from the same situation or action. For instance, if a die is rolled, someone might be interested in the outcomes 'one', 'two', etc. Someone else might be interested in the outcomes 'die stays on the table' and 'die falls on the floor'.

Probability Given a situation or action, the probability of an outcome occurring is the fraction of times that it will occur if the situation or action is repeated a very large number of times. For example, if you drop a matchbox a million times and it lands on its end 92 783 times, then the probability of it landing on its end is about 93 000 out of 1 000 000 or 0.093 or 9.3%. If dropped only a small number of times, the fraction of times it land on its end might be quite different from the probability. For example, it is quite possible that, out of 10 drops, it could land on its end twice or 20% of the times it was tried.

Data One of the two methods of determining probabilities is by collecting data and using it to calculate the fraction of times the outcome occurred. The data can be produced by performing an experiment. For example, to find the probability that a matchbox will land on its end, we might drop one a thousand times and count the times it lands on its end. Or we might use pre-existing data. For example, if we want to know the probability that taking a certain pain-killing drug during pregnancy will cause birth defects, we wouldn't give the drug to a thousand pregnant women to find out. But if there are a thousand women who have used the drug while pregnant and if records were kept, we can use this data to find the probability.

Experiment An experiment is when we cause a certain situation or action to happen a large number of times so that we can find out what fraction of the times the different outcomes occur. Experiments are used to find the probabilities of the outcomes.

Trial In an experiment, each time we cause the situation or action to occur is called a trial. The larger the number of trials, the more reliable will be the probability obtained.

Indifference One of the two methods of determining probabilities is indifference. This can be used if there is no difference between the outcomes that could make any of them any more or less likely to occur than any other. For example, if a normal coin is tossed, it can land heads up or tails up and there is no difference between the two sides of the coin that will make one side more likely to land upwards than the other. So it will land heads up 50% of the time and the probability of getting a head is thus 50%. If a bent coin is tossed, on the other hand, there is a difference between the two sides that could make it more likely to land up one way than up the other. To find the probabilities in this case, we have to collect data.

Note that in the case of a coin, there is a difference between the two sides in that they have different designs on them – one has a queen's head, the other has something else. But this difference is not likely to affect the way the coin lands. Some dice have small indentations for the holes. This would make the 'six' face slightly lighter than the 'one' face and it might be expected that this would make it more likely for the 'six' face to end up upwards. This would mean that the probability of a 'six' would be higher than the probability of a 'one'. The difference, however, would be so minute, that it would probably take a billion rolls to detect it. In practice, if there is no physical difference that would make a significant difference to the likelihoods of the different outcomes, then we ignore the difference.

How To

- **interpret probability as the fraction of times an outcome will occur if the situation or action is repeated a very large number of times,**
- **explain that the fraction of times an outcome will occur in a small number of repetitions might be quite different from the probability**

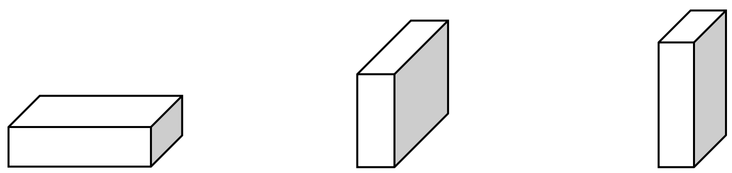
Required prior knowledge

You need an intuitive concept of likelihood. You also need to understand percentages and common and decimal fractions.

Method

Consider the action of dropping a matchbox. Various things can happen: it might land on the table or it might land on the floor; it might land on a white floor tile, it might land on a black one or it might land partly on each; it might land flat, land on its side or land on its end. These different things that can happen as a result of dropping it are called outcomes of dropping it.

We probably would not be concerned about all these possible outcomes. Let's say we are interested in whether it lands flat, on its side or on its end.



flat | on its side | on its end

Experience and intuition should tell us that the matchbox is more likely to land flat than on its end.
What do we mean when we say it is more likely? What we mean is that it lands flat more often. So we can say that one outcome (landing flat) is more likely than another outcome (landing on its end) because it happens more often.

The fraction of times that it lands flat will give us a measure of the likelihood. If we find it lands flat 62% of the times we throw it, then we can say that the likelihood of it landing flat is 62%. Up to now, we have just talked about whether one outcome is more or less likely than another. Now we can put a figure on the likelihood.

A likelihood expressed like this as a fraction is called a probability. So probability means likelihood expressed as a number or fraction. The fraction can be a common fraction, a decimal or a percent. Obviously the probability must be between 0% and 100% (inclusive), that is between 0 and 1.

But if we throw the matchbox twice, we might well find that it lands flat both times. 2 out of 2 is 100%, so we might say that the probability of it landing flat is 100%. But we would be wrong. 100% means that it will land flat every time we throw it. So it would have to land flat the next two times. It might well land on its end the next time.

If, on the other hand, we throw it 10 times, it is most unlikely that it will land flat every time. We might find that it lands flat 7 times. This is 70%. This is a better indicator of how it will land the next 10 times, but it may land flat 5 times out of the next 10 rather than 7.

If we throw it 1000 times, it is extremely unlikely that it will land flat every single time. We might find that it lands flat 634 times. This is about 63%. This means that if we threw it another 100 times we should find it lands flat about 63% of the times. It may be 61% or 66%, but it is unlikely to be 50%.

In a million throws it might land flat 619 136 times. This is about 62%. The chance are that the next million throws would also give 62%. It is unlikely to be 64%.

For a billion throws it might land flat 621 409 210 times. This is also 62%. We could then be fairly sure that the next billion throws would also give 62% (except that the box will probably be worn out from all the throwing!). In other words, in the long run, the box will land flat 62% of the time.

This is the crucial point – probability is the fraction of times something will happen **in the long run**. The fraction of times in the short run might be quite different. The fewer trials, the more different it is likely to be.

We can use the following definition of probability.

The probability of an outcome is the fraction of times it will happen if it is tried a very large number of times.

It is not only actions which have outcomes. Situations can have outcomes too and these outcomes have probabilities. An example of a situation is that it is a day in March. We might be interested in the probability that it will rain, so we would consider the two outcomes ‘it rains’ and ‘it doesn’t rain’.

If it rains on 24% of March days, then the probability of it raining on a day in March is 24%. As with actions, to determine this we have to have data on a very large number of March days. For a small number of days, the fraction of wet days might be quite different from 24%.

How To

- **design and conduct experiments to determine approximate probabilities**
- **use pre-existing data to determine approximate probabilities**
- **comment on the reliability of a probability estimate, considering the amount of data used, eg after tossing a coin only ten times, how reliable would be the resulting estimate of probability?**

The main way to find probabilities is to collect data on a large number of times when the situation or action occurred. We find out what fraction of those occasions the outcome we are interested in occurred. This is an approximation of the probability of that outcome.

If we use a very large number of occasions, then the value we get for the probability will be fairly accurate. It we use a small number, our value might be fairly inaccurate.

The data can be obtained in a couple of ways. It might already be available or we might have to produce it ourselves in an experiment.

For example, we might be interested in the probability that someone who gets lung cancer will die from it. We might collect data on a large number of people who have contracted lung cancer and, in each case, whether they died or whether, with the help of radiotherapy etc., they got better. Suppose we find data on 2578 people who got lung cancer and find that 1893 of them died and 685 got better, then we can say that the probability of dying is $\frac{1893}{2578}$ or 0.74 or 74%. Likewise, the probability of getting better is $\frac{685}{2578}$ or 0.26 or 26%.

In this case, we probably wouldn't do an experiment by getting a couple of thousand people to smoke 200 cigarettes a day till they got lung cancer, then seeing how many died. But in other situations an experiment might be quite OK.

A crockery manufacturer might want to know the probability of a plate breaking if dropped from one metre onto a wooden floor. Data on the plates probably wouldn't already exist, so they might produce some by dropping 100 plates and seeing how many break.

How To

- **use the numerical probability of an event to predict the number of times the outcome is likely to occur and to make an informed decision about a future action**

If we know the probability of a certain outcome, we might be able to use this information in a number of ways.

Firstly we can use it to predict the number of times the outcome is likely to occur if a situation or action is repeated a number of times. For example, you are running are stall at the fete where people pay 50c, roll three dice and win a prize if they roll less than a total of 6. You want to know how much to spend on the prizes, assuming the stall wants to make a profit, not a loss.

You might assume that 600 games will be played through the day. This will make $300. You might plan, therefore to spend $200 on prizes. But how many and at what price each? What you need to know is how many prizes you expect to be won. For this you need to know the probability of winning. It is 0.046 or 4.6%. [You will find out at Level 6 how to calculate these sorts of probabilities. You may be cluey enough to be able to do it now, but otherwise, just accept that it is about 4.6%.]

This means that the people will win about 4.6% of the games. For 600 games this is about 28 wins (4.6% of 600). So you should buy about 28 prizes at about $7 each.

Secondly we can use knowledge of the probability of an outcome to pick the best outcome. For example, when playing the goldfish game with Board 2, the probability of grey winning is $^{2}/_{3}$; the probability of white winning is $^{1}/_{3}$. If you know this, it will help you to make the best choice. [Again, you will learn at Level 6 how to work out these probabilities.]

Or suppose you have to travel 450 km between two remote towns in central Africa. You may know from statistics that the probability of having an accident if you drive is 0.06% and the probability of having an accident if you fly is 0.11%. This would be a reason to consider driving.

Thirdly, and probably least often, we might use the probabilities of the outcomes of more than one action or situation to choose between the actions or situations. An example of this was in choosing which board to use in the goldfish game. Another example might be.

How To

- **determine probabilities using indifference, eg. deciding that the probability of getting a head when tossing a coin is 50% because there is no difference between the two possible outcomes that could make one more likely than the other**
- **decide whether indifference can be used in a particular case, eg that indifference does not apply when tossing a matchbox because the faces are different**

If you drop a matchbox, it can land flat on its front, land on its side or land on its end. The front of the box is bigger than the end and this might make it more likely for the matchbox to land on its front than on its end.

If you toss a coin, it can land heads up or tails up. Both faces are the same size and shape and the coin looks the same shape from both directions. In this case there is nothing about the two faces that could make it more likely that it will land up one way than up the other. We say that the two outcomes, heads and tails, are indifferent.

It might be said that the two faces have different pictures on them – one has the Queen's head; the other has something else. But the pictures aren't likely to have much effect on the way the coin lands. So we can assume that the outcomes 'heads' and 'tails' are indifferent.

Because the two outcomes are indifferent, we can assume that in the long run, the coin will land heads up and tails up an equal number of times. Both outcomes will occur 50% of the times the coin is tossed. In other words the probability of heads is 50% and the probability of tails is 50%.

Thus we can work out the probabilities without having to collect any data. In any situation where there is indifference between the possible outcomes in a situation, then it is possible to work out the probabilities of the outcomes without having to collect any data. By indifference we mean that there is nothing about the outcomes that could make any of them more or less likely than any others.

The probability of getting a four on a die can be worked out using indifference because all the faces of a dice are the same (apart from the spots which won't affect the way the die lands) and thus all faces will end up upwards one sixth of the times the die is rolled. So the probability of getting a four is one sixth or about 17%.

But the probability of a matchbox landing on its end cannot be worked out using indifference, because there is a difference between the different faces. The ends are smaller than the front and back, so the matchbox might be expected to land on its end less often than on its front or back. With the matchbox, the only way to find the probabilities is by collecting data.

There are other situations where indifference can be used. For example if you put equal sized but different coloured marbles in a bag and pull one out without looking, the possible outcomes might be that you get the blue one, the green one, the red one or the white one. There is nothing about any of these outcomes that can make one more likely than the others, so we can say that all outcomes have probabilities of 25%. If, however the marbles are different sizes, it may be that you would be more likely to find the biggest one than the smallest one. So the outcomes are not indifferent and indifference cannot be used. We have to use data and we might find that the probability of picking the biggest one is 28.9%.

Suppose a toilet has eight identical cubicles in a row. We want to know the probability that a woman walking into the toilet will use the cubicle closest to the door. Do you think we can use indifference here? To some extent this might be a matter of opinion, but it is quite possible than more people would use the first cubicle they came to rather than walk down to the third last one. So indifference shouldn't really be used.

If, however, a rat is dropped into the middle of a square box and there is an identical door in the middle of each of the four walls, then there would not be a difference between the four doors that would make the rat more likely to use one than any of the others. So we could say that the probability of it using each door is 25%.

Review Questions

- **interpret probability as the fraction of times an outcome will occur if the situation or action is repeated a very large number of times,**
- **explain that the fraction of times an outcome will occur in a small number of repetitions might be quite different from the probability**
- **design and conduct experiments to determine approximate probabilities**
- **use pre-existing data to determine approximate probabilities**
- **comment on the reliability of a probability estimate, considering the amount of data used, eg after tossing a coin only ten times, how reliable would be the resulting estimate of probability?**

1 If you toss two 20c coins, they can land

- both heads
- both tails
- one of each

(a) Estimate the probability of the outcome 'both heads'.

(b) Toss two 20c coins 12 times and use the data to find the probability of 'both heads'. Express it as a common fraction, as a decimal fraction and as a percentage.

(c) How confident are you that the result from (b) is correct?

(d) Repeat the experiment with 120 tosses.

(e) How confident are you about the result from (d) ?

(f) Does it seem that your prediction in (a) was correct?

2. Up to 1996, about 4000 people had attempted to climb Mount Everest. Of these 660 succeeded and 142 died. Find:

(a) the probability that a person who attempts to climb Everest will make it

(b) the probability that a person who attempts to climb Everest will die in the attempt.

3. Give an example of an outcome of a situation or action with each of the following probabilities:

(a) 1

(b) about 0.999

(c) about 0.9

(d) about 0.5

(e) about 0.2

(f) about 0.01

(g) about 0.000 001

(h) 0

4. Damien and Nigel bought two second hand computers from a flea market for $50 each. When they got them home to Damien's place and tried them, they found one was a lot better than the other. They had to decide who was going to have the good one. Damien suggested that they open a book at a random page and, without looking, stick a pin in it, and that they keep doing this until the pin hits a D or an N. If it hits a D, Damien gets the good one. If it hits an N, Nigel gets the good one. Nigel agreed. If this method were fair, both would have a 50% chance of winning it. But the method wasn't. Find a page of a book and use it to find the probability of each winning.

Review Questions

- **use the numerical probability of an event to predict the number of times the outcome is likely to occur and to make an informed decision about a future action**

1. The probability of it raining on a December day in Brisbane is 38%. How many wet December days would you expect in a typical year?

2. At Greasy Louie's, the probability of getting food poisoning from the burgers is 28%; the probability of getting food poisoning from the hot dogs is 41%. Which is the safest option?

3. In a piece of English text containing 179 419 letters, there were found to be the following numbers of E's, T's, C's and X's

   E 22 619   T 15 862   C 4 618   X 287

   (a) Find the probability that a letter picked at random is an E

   (b) Find the prob ability that a letter picked at random is an X

   (c) Estimate the numbers of Es in a similar piece of text with 61 509 letters

   (d) Estimate the number of Cs in a piece of similar text with 7 letters

4. The probability of an average motorist having an accident in 1000 km of driving is 0.37%. There are 11.2 million motorists in Australia, driving an average of 11 900 km per year.

   (a) Estimate the number of motoring accidents in Australia per year.

   (b) The average repair bill after an accident is $3 300. This is generally paid for by an insurance company. In addition to paying claims, an insurance company needs to spend about 60% as much on administration (wages, buildings etc.). What should the annual insurance premium be? (Assume all motorists pay the same premium.)

5. At a stall at a fair you have a choice of two games. With the two-dice game, you pay $1 and roll two dice. If you score more than 10, you win a stuffed newt. With the three-dice game, you pay $1 and roll three dice. If you score more than 15, you win the same stuffed newt. Assuming you want a stuffed newt, which game should you play to give yourself the best chance of winning?

   (The probability of getting >10 with 2 dice is 1 in 12 or $^{1}/_{12}$; the probability of getting >15 with 3 dice is 10 in 216 or $^{10}/_{216}$.)

Review Questions

- **determine probabilities using indifference, eg. deciding that the probability of getting a head when tossing a coin is 50% because there is no difference between the two possible outcomes that could make one more likely than the other**
- **decide whether indifference can be used in a particular case, eg that indifference does not apply when tossing a matchbox because the faces are different**

For each of the following situations, explain whether the idea of indifference can be used to find probabilities or whether we would have to collect data. If there is indifference, work out the probabilities involved.

1. A regular dodecahedron has twelve identical pentagonal faces. These are numbered 1 to 12 and the shape is used as a die. We want to know the probability of rolling a 5.

2. A three-legged stool is thrown onto the floor. It can land on its feet, on its side, or upside down. We want to know the probability that it will land upside down.

3. A three legged stool is thrown onto the floor. We want to know the probability that it will break.

4. A man and two women compete against one another on a television quiz show. There will be one winner. We want to know the probability that it is the woman.

5. A Siamese cat and a Chihuahua dog are sitting outside the front door. The owner throws them a bone. We want to know the probability that the cat will end up with the bone.

6. A tin of dog meat is used instead of tossing a coin at the start of a cricket game. The tin is tossed. It is agreed that if it lands up the right way, team A wins the toss; if it lands upside down, team B wins the toss; if it lands on its side, it will be tossed again. We want to know the probability that team A wins the toss.

7. Mrs Bright buys four 60 watt light globes, all of the same brand, from the supermarket. She puts them in the four identical light fittings out on the veranda of her house. She turns them all on together. We want to know the probability that the one at the left end lasts longest.

8. A six-year-old boy walks across the thin ice on a frozen lake. He might fall through the ice and he might not. We want to know the probability that he does.

9. We want to know the probability that Wednesday will be the warmest day next week.

10. We want to know the probability that Wednesday next week will be warmer than both Thursday and Friday.

11. We toss two coins. We want to know the probability that both come down heads.

12. We toss 5 coins. We want to know the probability that at least 3 come down heads.

13. We toss 6 coins. We want to know the probability that they all come down heads.

14. We roll 2 dice. We want to know the probability that the blue one shows a higher number than the red one.

15. Australia plays England at cricket. We want to know the probability that England wins.