

AS and A Level Biology



Teacher Maths Guide: 2019

Contents

Introduction	3
Basic Maths for Biologists	4
SI Units and Prefixes	4
Ordinary and Standard Form	5
Significant Figures	6
Ratios	7
Formulae	8
Magnification	8
Using an Eyepiece Graticule	10
Mitotic Index	11
Counting Cells	12
Water Potential	14
Percentage Change	15
Temperature coefficient (Q_{10})	16
Surface Area and Volume	17
Gradients and Rates	18
Equation for a straight line ($y = mx + c$)	20
Logarithms	21
Exponential Growth	22
The Hardy-Weinberg Equation	23
Cardiac Output and Minute Ventilation	24
Respirometers and Potometers	25
Calculations using Quadrats	26
The Lincoln Index	27
Simpson's Diversity Index	28
Productivity and Trophic Efficiency	29
Calculations without Formulae	30
Summary Statistics	32
Calculating the Central Value of a Sample	32
Calculating the Dispersion of a Sample	34
Statistical Tests	37
The Unpaired t -test	40
The Paired t -test	42
The Spearman Correlation Coefficient	44
The Chi squared (χ^2) Test	46
Practice Questions	48
Appendices	65
Appendix 1 – Using a Calculator for Stats Calculations	65
Appendix 2 – Help from the Field Studies Council	66
Appendix 3 – Ofqual's Mathematical Requirements	67

Introduction

Biology is a quantitative experimental science, and mathematical skills are an essential part of AS and A level Biology. The exam regulator Ofqual requires that at least 10% of the marks in A level Biology exams should assess mathematical skills at level 2 (i.e. GCSE Mathematics Higher Tier). The specific mathematical requirements detailed by Ofqual for all examining bodies are given in Appendix 3 at the end of this guide.

Many students of A-level Biology are put off by maths questions in exams and often give up on them too easily. In fact the maths questions should be amongst the easiest marks to gain in a biology exam since the maths is just GCSE standard. This guide shows the tips and tricks that help to make these questions achievable. By reading this guide and practising the examples students should gain confidence in the style and techniques of maths questions and so should approach the exam with a more positive attitude to these all-important marks.

Many of the maths skills are best covered during practical work, especially in the core practicals. The core practical worksheets provided by Edexcel give suggested applications and the Edexcel "Practical Guide" that complements this guide includes examples of maths in practical contexts.

This guide covers both of the Edexcel A-level Biology specifications. Since the maths content is largely dictated by Ofqual, the requirements of the two specifications are very similar, but there are a few formulae that are specific to one or the other:

- Specification A includes Q_{10} , minute ventilation and the heterozygosity index.
- Specification B includes water potential and the exponential growth rate constant.

The student guide is identical to the teachers' guide, but without the answers to the practice questions.

Formulae in exams

It is helpful to know which formulae will be provided to students in exams.

- Formulae will be provided for Simpson's Diversity Index (since there are different versions) and for the statistical tests (s , t , r_s , χ^2), but not for mean or degrees of freedom. Formulae have so far always been provided for volume and surface area of solid objects and for exponential growth and water potential.
- Formulae won't be provided for simple relationships (e.g. area of a circle, percent change) or where the formula is integral to the understanding of the biology (e.g. magnification, Hardy-Weinberg, productivity).
- One-off formulae that are not in the specification may also be given in exams, and students should expect to encounter formulae they have not seen before.

Two Top Tips

If you learn nothing else from this guide, learn these two top tips:

- Always show your workings with labels for each step, e.g. "surface area ="
- Always write the units at each stage. They show you what to do next.

Use the worked examples here as a guide to good practice.

Basic Maths for Biologists

SI Units and Prefixes

Biological measurements are always made using international standard "SI" units. The main units used in biology are:

Quantity	Unit	Symbol	
length	metre	m	
mass	kilogram	kg	
amount of substance	mole	mol	
time	second	s	note 1
temperature	degree celcius	°C	note 2
force	newton	N	
pressure	pascal	Pa	
energy	joule	J	
volume	cubic decimetre	dm ³	note 3

Note 1 The symbol for second is s not sec. The units min, h and y are also allowed units for time.

Note 2 The kelvin (K) is the SI unit of temperature, but is rarely appropriate in biology, where °C is much more common. Never write °c or use the term centigrade.

Note 3 Litre (L) is also commonly used in biology for volumes and is often easier. It's useful to remember that 1 dm³ = 1 L, 1 cm³ = 1 mL and 1 mm³ = 1 µL.

Names of units are always spelt with a small letter, even if they're named after scientists (e.g. joule). Symbols do not need a full stop (like an abbreviation) or an s (like a plural) e.g. 3 min not 3 mins. There should be a space between the value and its symbol (e.g. 6 g not 6g). Use the index ⁻¹ for division in units, not a slash (e.g. ms⁻¹ not m/s).

Unit Prefixes

All SI units can take prefixes in front of them to make them smaller or larger. The common prefixes are shown in this chart applied to grams (g) but obviously they can be applied to any unit e.g. km, MJ, ms, µmol. There is no space between a prefix and a symbol (e.g. 7 mN not 7 m N)

The prefixes follow the "thousands rule" i.e. each prefix increases or decreases the unit by a factor of a thousand. So by choosing the right prefix, all values can be in the range 1–999.

e.g. 10 mm instead of 0.01 m
2.56 MPa instead of 2 560 000 Pa

mega	Mg	↗ x 1000
kilo	kg	↗ x 1000
unit	g	↗ x 1000
milli	mg	↗ x 1000
micro	µg	↗ x 1000
nano	ng	↗ x 1000
pico	pg	↗ x 1000

Top Tip. Measure lengths in mm, not cm. This forces you to measure accurately, but it also makes converting units easier, because you can use the thousands rule (which doesn't work for cm). Using a clean transparent ruler also helps to measure accurately.

Ordinary and Standard Form

Values in biology can often be very large or very small, so to make them easier to write we can use standard form (also called scientific notation or normalised form), e.g. 0.0028 mm in ordinary form becomes 2.8×10^{-3} mm in standard form. Standard form has two parts: the significand (or mantissa) and the exponent:

$$\begin{array}{ccc} \text{ordinary form} & & \text{standard form} \\ 8\ 460\ 000 & = & 8.46 \times 10^6 \\ & & \underbrace{\hspace{1cm}} \quad \underbrace{\hspace{1cm}} \\ & & \text{significand} \quad \text{exponent} \end{array}$$

In standard form there is always just one digit to the left of the decimal point and the index indicates how many positions the decimal point has moved. A positive index means a value larger than one (move the decimal point to the right), while a negative index means a value smaller than one (move the decimal point to the left). You should be able to convert quickly and easily between these forms, e.g.:

- 574 281 000 g = $5.742\ 81 \times 10^8$ g = 574.281 Mg (megagrams)
- 0.005 438 m = 5.438×10^{-3} m = 5.438 mm

To make long numbers easier to read a space is used to separate groups of three digits, not a comma (e.g. 72 000 not 72,000).

Top Tip. Practise converting between ordinary and standard form, and using prefixes, until it becomes second nature.

Calculations using standard form

1. To **add** or **subtract** numbers in standard form, they must have the same exponent. Add (or subtract) the significands and keep the exponent the same. It's usually easiest to convert the smaller exponents to the biggest exponent. Note that adding or subtracting a much smaller number will have very little effect on a larger number.
2. To **multiply** numbers in standard form multiply the significands and add the exponents.
3. To **divide** numbers in standard form divide the significands and subtract the exponents.

Worked Examples

1. $(5.7 \times 10^8) + (8.3 \times 10^7) = (5.7 \times 10^8) + (0.83 \times 10^8) = 6.53 \times 10^8$
2. $(5.7 \times 10^8) - (3.3 \times 10^7) = (5.7 \times 10^8) - (0.33 \times 10^8) = 5.37 \times 10^8$
3. $(5.82 \times 10^4) \times (2.81 \times 10^3) = 16.35 \times 10^7 = 1.64 \times 10^8$
4. $(5.82 \times 10^4) \div (2.81 \times 10^3) = 2.07 \times 10^1 = 20.7$

Significant Figures

The number of significant figures tells us the precision of a measurement. For example 3.72 m is more precise than 3.7 m – it has been measured to more significant figures. Each digit in a measurement counts as a significant figure, except zero when used as a placeholder (usually leading and trailing zeros). So:

806	has 3 significant figures.
8.06×10^3	has 3 significant figures. The exponent doesn't add any precision.
806 001	has 6 significant figures.
806 000	has 3 significant figures. The trailing zeros are assumed to be just placeholders, and the number could be written as 8.06×10^3 . It is however possible that that this value was measured precisely and the last three digits just happen to be all zeros, so it has 6 significant figures, like the previous example.
806.0	has 4 significant figures. You wouldn't write that trailing zero unless it was measured.
0.00806	has 3 significant figures. The leading zeros are just placeholders and the number could have been written as 8.06×10^{-3} .
80.600	has 5 significant figures. You wouldn't write the trailing zeroes unless they were measured.

It's important to realise that 2 and 2.0 are not the same.

- 2 has one significant figure and means the value is between 1.5 and 2.5.
- 2.0 has two significant figures and means the value is between 1.95 and 2.05.

Top Tip. When writing the results of a calculation, always use the same number of significant figures as the least significant piece of raw data.

Worked Example

Calculate the mean of 2.43, 3.71, 3.55 and 2.9. Give your answer to an appropriate number of significant figures.

Answer

The mean is $\frac{(2.43+3.71+3.55+2.9)}{4} = 3.1475$, but this result has 5 significant figures. The least significant piece of raw data (2.9) has only two significant figures, so the answer should be rounded to 3.2.

Ratios

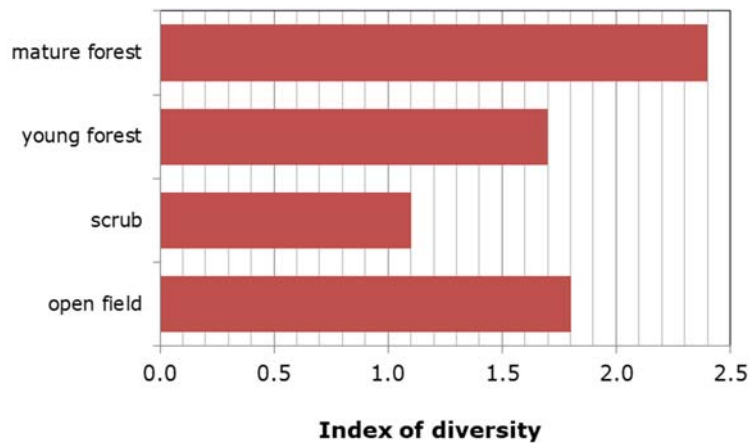
Ratios should always be written in their simplest form, usually $x : 1$, or sometimes $1 : x$. In this form ratios can easily be compared.

For example if two lengths are measured as 600 mm and 250 mm then write their ratio as $\frac{600}{250} = 2.4 : 1$, rather than $6 : 2.5$.

Top Tip. Make sure you get the ratio the right way round. If you are asked for the ratio of a to b , the calculation is $\frac{a}{b} : 1$ or, if required, $1 : \frac{b}{a}$.

Worked Example

This chart shows the diversity index for four different ecosystems.



Calculate the ratio of the diversity indices of mature forest to young forest.

Answer

$$\text{ratio} = \frac{2.4}{1.7} = 1.4:1$$

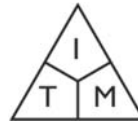
If the question had been the ratio of the diversity indices of young forest to mature forest, then answer would have been $0.71:1$.

Formulae

Magnification

Microscope drawings and photographs (micrographs) are usually magnified, and you have to be able to calculate the actual size, or true size, of the object from the image. Magnification is simply how much bigger the image is than the true object. It has no units and is calculated using the formula:

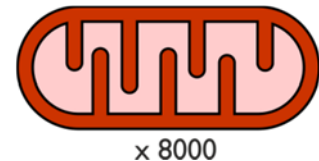
$$\text{magnification} = \frac{\text{image length}}{\text{true length}}$$



For this formula to work, the image length and true length must be in the same units, usually μm for cells and organelles. Normally you have to calculate the true size from an image, but you can be asked to calculate any of the three values, given the other two. "TIM's triangle" above can help to rearrange the equation as needed. Note that magnifications can be less than one (e.g. $\times 0.1$), which means that the drawing is smaller than the actual object.

Worked Example 1

Calculate the true length of this mitochondrion. Give your answer in μm .



Answer

This image has a magnification factor so we can use the formula directly.

1. First we have to measure the length of the image on the paper, which in this case is 40 mm.
2. We want the true length in μm , so we must convert the image length to μm before it goes in the formula. There are 1000 μm in a mm (remember the thousands rule), so

$$40 \text{ mm} = 40\,000 \mu\text{m}.$$

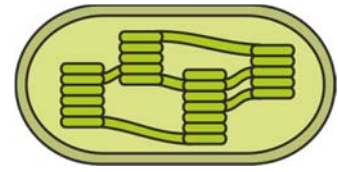
3. Now we use the equation:

$$T = \frac{I}{M} = \frac{40\,000 \mu\text{m}}{8000} = 5 \mu\text{m}.$$

Top Tip. Remember, always measure in mm, never cm. This makes converting units easier because you can use the thousands rule.

Worked Example 2

If the true length of this chloroplast is 25µm, what is the magnification of the drawing?

**Answer**

1. Measure the image length, which in this case is 42 mm.
2. The image and true lengths must be in the same units, so we convert to 42 000 µm.
3. We rearrange the formula to make M the subject and put the numbers in:

$$M = \frac{I}{T} = \frac{42\,000\ \mu\text{m}}{25\ \mu\text{m}} = \times 1680.$$

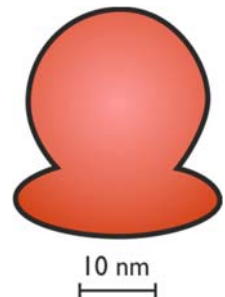
Top Tip. Magnification doesn't have units, but often has an × (for "times") before it.

Worked Example 3

Calculate the true height of this ribosome.

Answer

This ribosome picture doesn't have a magnification, but has a scale bar instead. So we first use the scale bar to calculate the magnification, and then apply this magnification to the drawing.



1. Measure the length of the scale bar in the image, which in this case is 10 mm.
2. Since the true length is given in nm, we convert the image length to nm:
10 mm = 10 000 000 nm.
3. We know the true length of this scale bar is 10 nm, so the magnification is:

$$M = \frac{I}{T} = \frac{10\,000\,000\ \text{nm}}{10\ \text{nm}} = \times 1\,000\,000.$$

4. Now we apply this magnification to the drawing of the ribosome. The image height is 30 mm, which we convert to nm (30 000 000 nm).
5. Finally we put the values into the formula:

$$T = \frac{I}{M} = \frac{30\,000\,000\ \text{nm}}{1\,000\,000} = 30\ \text{nm}.$$

Top Tip: With any magnification problem, it's always worth checking that your answer is sensible, to avoid silly mistakes. It's very easy to be out by a factor of 1000, but common sense should tell you that a mitochondrion is not 5 cm long! Scale bars make this much easier, since you can judge roughly how many scale bars fit into the object, and so roughly how big it is. For example, in the last question, we can fit about three scale bars into the object, so our answer of 30 nm is sensible.

Using an Eyepiece Graticule

The size of objects under the microscope can be measured accurately using an eyepiece graticule – a fine scale built into the eyepiece lens. The graticule scale is 10 mm long with 100 small divisions, so each division is 0.1 mm (100 μm) long.

The length represented by the graticule scale depends on the magnification of the objective lens: the higher the objective magnification, the smaller the length covered by the graticule. So the eyepiece graticule needs to be calibrated for each objective lens.

Rough Calibration

A rough calibration can be made by simply dividing the real length of each division (100 μm) by the objective lens magnification:

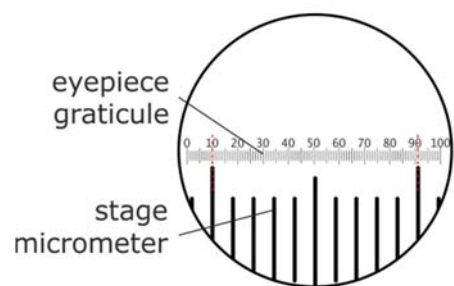
Objective lens magnification	$\times 4$	$\times 10$	$\times 40$	$\times 100$
length of each smallest division	25 μm	10 μm	2.5 μm	1 μm

Top Tip. Do not include the eyepiece magnification in this calculation.

Accurate Calibration

The rough calibration isn't always accurate due to slight differences in microscopes. But an eyepiece graticule can be calibrated accurately using a stage micrometer. This is a slide with a scale on it exactly like the eyepiece graticule – 10 mm long with 100 small divisions – so each division is 100 μm long.

1. To calibrate a microscope, place the stage micrometer on the microscope stage.
2. Turn to the objective you want and focus on the stage micrometer. Move the slide and rotate the eyepiece graticule until the two scales are aligned. The stage micrometer has thick lines because it is magnified.
3. Choose a whole number of stage divisions that align with the eyepiece graticule. In this example 10 stage micrometer divisions match up with 81 eyepiece graticule divisions (from 10 to 91 on the graticule, marked with red lines).



$$\therefore 81 \text{ divisions} = 10 \times 100 \mu\text{m} = 1000 \mu\text{m}$$

$$\therefore 1 \text{ division} = 1000 / 81 = 12.4 \mu\text{m}$$

This is the calibration for this objective lens.

4. Remove the stage graticule slide and replace with a real slide. To measure the size of an object in μm , measure it in eyepiece divisions and multiply by the calibration factor.

e.g. if a cell is 7 eyepiece divisions long, its true length = $7 \times 12.4 = 86.8 \mu\text{m}$.

Mitotic Index

The mitotic index is the percentage of cells in a tissue undergoing mitosis. It is found by observing under a microscope a sample of the tissue, suitably treated to stop mitosis and stain the chromosomes. A large number of cells are counted and the number of cells with visible chromosomes (i.e. in mitosis) is counted. The mitotic index is calculated using the formula:

$$\text{mitotic index} = \frac{\text{number of cells in mitosis}}{\text{total number of cells counted}} \times 100$$

The mitotic index shows how actively the cells are dividing, as the higher the mitotic index, the greater the proportion of dividing cells in the sample at the moment in time when the tissue was fixed. And the higher the proportion of dividing cells at any moment, the faster the cells must be dividing. It is therefore used as a diagnostic tool in cancer treatment – cancer cells divide unusually quickly and the effect of a cancer therapy can be measured using the mitotic index.

Worked Example

This table shows the number of cells in different stages of mitosis in a tissue.

stage	number of cells
Interphase	554
Prophase	28
Metaphase	29
Anaphase	5
Telophase	17

Calculate the mitotic index.

Answer

First we add up the total number of cells counted (633) and the number in the different phases of mitosis (79), then we put these numbers into the formula:

$$\text{mitotic index} = \frac{79}{633} \times 100 = 12.5 \%$$

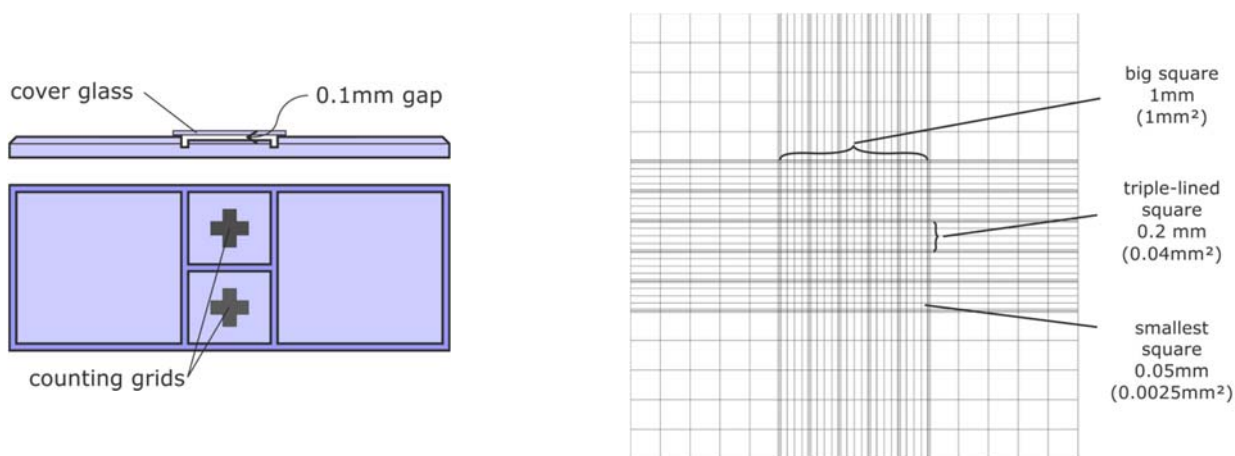
Top Tip. Common mistakes are not including all the stages of mitosis in the mitotic cells, and not including the mitotic cells in the total.

Counting Cells

In microbiology work we often want to know the population of cells in a broth culture, for example to obtain a growth curve. This is done by counting the number of cells in a known volume, to get a cell density in cells cm^{-3} . There are two methods for counting cells: the cell counter and dilution plating.

Cell Counter (Haemocytometer)

The cell counter (or haemocytometer) is a large microscope slide with one or two very accurate grids drawn in the centre. The grids mark out squares with 1 mm, 0.2 mm and 0.05 mm sides. There is an accurate gap of 0.1 mm between the grid and the thick coverslip, so the volume of liquid above the grid is known. All the cells in that volume fall onto the grid and can be counted under the microscope.



Worked Example

A culture of yeast cells is diluted 1000 $\times$ and a small drop placed on a haemocytometer. Five triple-lined squares are examined, with counts of 74, 82, 76, 88 and 80 cells. What is the cell density (in cells cm^{-3}) of the undiluted culture?

Answer

1. First we work out the mean cell count: $(74+82+76+88+80) / 5 = 80.0$ cells
2. Now we calculate the volume above a triple-lined square (0.2 mm sides)

$$\text{volume} = 0.2 \times 0.2 \times 0.1 = 0.004 \text{ mm}^3$$
3. So we have 80 cells in 0.004 mm^3 so:

$$\text{density} = \frac{80}{0.004} = 20\,000 = 2 \times 10^4 \text{ cells mm}^{-3}$$
4. But this was a 1000 $\times$ dilution, so the undiluted suspension had

$$2 \times 10^4 \times 10^3 = 2 \times 10^7 \text{ cells mm}^{-3}$$
5. We need the answer in cells cm^{-3} , and there are 1000 mm^3 in a cm^3 (because one $\text{cm}^3 = 10 \text{ mm} \times 10 \text{ mm} \times 10 \text{ mm} = 1000 \text{ mm}^3$), so

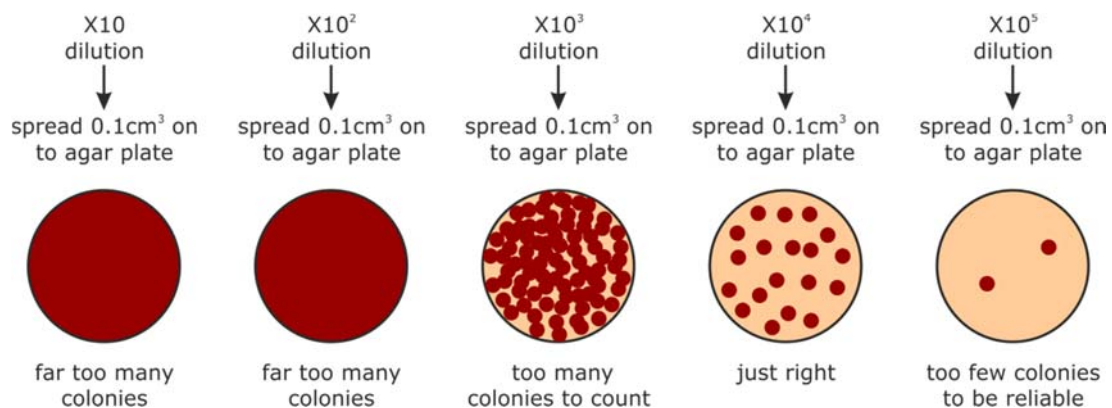
$$2 \times 10^7 \times 10^3 = 2 \times 10^{10} \text{ cells cm}^{-3}$$

Top Tip. Since the numbers are so large, it makes the calculations easier to use standard notation throughout. Multiplying becomes a simple matter of adding exponents.

Dilution Plating

This is another technique for measuring cell density in a culture, but this one only counts viable (living) cells, whereas the haemocytometer counts living and dead cells.

1. A broth culture is diluted serially to make a range of different dilutions.
2. A known small volume (typically 0.1 cm^3) is taken from each dilution and spread evenly onto a sterile agar plate.
3. The agar plates are incubated at 30°C for 24h. Each viable cell in the sample will multiply to form a visible circular colony.
4. Choose a dilution plate with a suitable number of visible colonies (not too many, not too few) and count the colonies. Since each colony arose from a single living cell, this is the number of viable cells in the initial small volume.



Worked Example

83 colonies were counted in the $\times 10^4$ dilution agar plate. What was the density of viable cells in the original culture? Give the answer in cells cm^{-3} .

Answer

1. 83 colonies means there were 83 viable cells in the 0.1 cm^3 sample, so

$$\text{density} = \frac{83}{0.1} = 830 \text{ cells cm}^{-3}$$

2. But this was diluted $\times 10^4$ from the original culture, so

$$\text{density} = 830 \times 10^4 = 8.3 \times 10^6 \text{ cells cm}^{-3} \text{ in the original culture.}$$

Water Potential

Water potential is a way of quantifying osmosis, so we can calculate the forces acting on water and determine which way it will move. Water potential is measured in units of pressure (Pa, or usually kPa), and is abbreviated by the symbol Ψ (Greek letter psi). Water potential is made up of two components: turgor pressure (P) and osmotic potential (π), related by this equation:

$$\text{water potential } (\Psi) = \text{turgor pressure } (P) + \text{osmotic potential } (\pi)$$

- **Water potential (Ψ)** is the overall pressure acting on water. Ψ can be positive or negative. Water always “falls” from a high to a low water potential. In a fully turgid plant cell the inflow of water is exactly balanced by the outflow of water, so there is no net movement of water and the overall water potential is zero ($\Psi = 0$).
- **Turgor pressure (P)** is the hydrostatic pressure due to an external force. P can be positive, which means a compression force (e.g. due to a cell wall) or negative, which means a tension force (e.g. in xylem vessels). In a plasmolysed plant cell the cell wall is not under any pressure, so the turgor pressure is zero ($P = 0$). Also, in any open system (like a beaker) there is no external pressure, so $P = 0$. The term turgor pressure has been replaced by the term pressure potential (Ψ_p) in modern books, but Edexcel still uses the old term.
- **Osmotic potential (π)** is the pressure due to dissolved solutes. The higher the solute concentration, the lower the osmotic potential. π is always negative since pure water has $\pi = 0$, and any solute decreases π . The term osmotic potential has been replaced by the term solute potential (Ψ_s) in modern books, but Edexcel still uses the old term.

Worked Examples

1. A plasmolysed plant cell has an osmotic potential of -750 kPa. Calculate its water potential.

Answer

Since the cell is plasmolysed, the turgor pressure is zero, so the equation is:

$$\Psi = P + \pi = 0 + -750 = -750 \text{ kPa}$$

2. A turgid plant cell has an osmotic potential of -600 kPa. Calculate its turgor pressure.

Answer

First we make P (the turgor pressure) the subject of the equation:

$$P = \Psi - \pi$$

Since the cell is turgid, the water potential is zero, so the equation is:

$$P = \Psi - \pi = 0 - -600 = +600 \text{ kPa}$$

Top Tip. With so many negative numbers, you have to be careful that you add and subtract them correctly.

Percentage Change

Experimental results can be shown as percentage change to make it easier to compare different results. There are two versions of the formula:

$$\% \text{ change} = \frac{(\text{final value} - \text{starting value})}{\text{starting value}} \times 100 \quad \text{or} \quad \% \text{ change} = \left(\frac{\text{final value}}{\text{starting value}} - 1 \right) \times 100$$

Percentage changes can be positive (an increase) or negative (a decrease), so it can be helpful to always write a + or a - in front of the value. It is also possible to have changes greater than 100%.

Worked Examples

1. Calculate the percentage change if the skin temperature of an athlete changes from 37.0 °C to 37.5 °C.

$$\% \text{ change} = \left(\frac{37.5}{37.0} - 1 \right) \times 100 = +1.35 \% \text{ (using second equation)}$$

2. Calculate the percentage change if the area of a decomposing leaf changes from 150 mm² to 80 mm².

$$\% \text{ change} = \frac{(80 - 150)}{150} \times 100 = -46.7 \% \text{ (using first equation)}$$

3. Calculate the percentage change if the leaf area of a growing tree changes from 0.6 m² to 3.7 m².

$$\% \text{ change} = \left(\frac{3.7}{0.6} - 1 \right) \times 100 = +516.7 \% \text{ (using second equation)}$$

Top tip. A common mistake is to put the final value as the denominator. Using the second equation might make this mistake less likely.

Temperature coefficient (Q_{10})

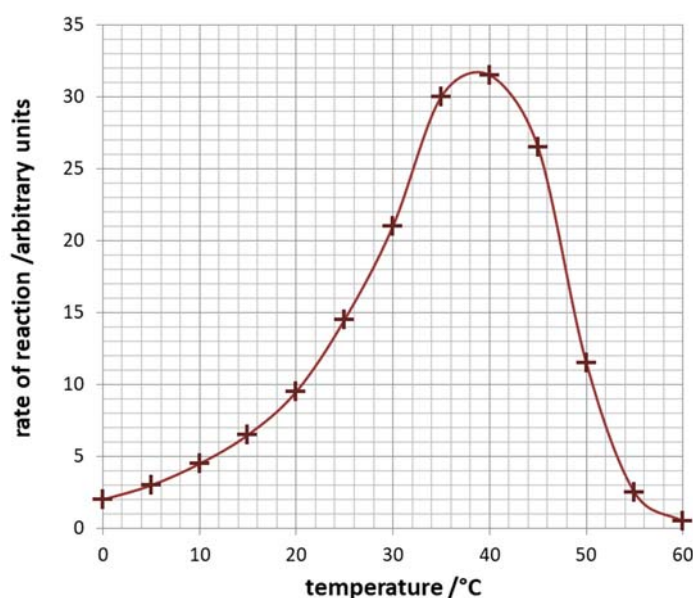
Q_{10} is a way of quantifying the effect of temperature on rate of reaction. It is the factor by which the rate increases for a 10°C rise in temperature, and is calculated using the formula:

$$Q_{10} = \frac{R_{t+10^{\circ}}}{R_{t^{\circ}}} \quad (\text{where } R_{t^{\circ}} \text{ is the rate at } t^{\circ}\text{C})$$

You have to be told which temperatures to use for a Q_{10} . For enzyme reactions, Q_{10} is usually about 2, which indicates that the rate doubles for every 10°C increase in temperature.

Worked Example

This graph shows the effect of temperature on the rate of an enzyme reaction.



- a) Using the graph, calculate Q_{10} as the temperature increases from 20° to 30°C .

Answer

Rate of reaction at 20°C is 9.5 and at 30°C it is 21.

$$\text{So } Q_{10} = \frac{R_{30^{\circ}}}{R_{20^{\circ}}} = \frac{21}{9.5} = 2.2$$

- b) What is Q_{10} if the temperature increases from 10°C to 20°C ?

A 2.0 **B** 2.1 **C** 2.2 **D** 2.4

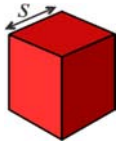
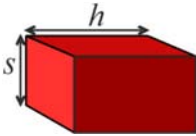
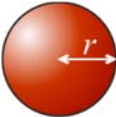
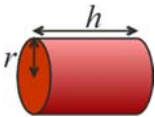
Answer

$$\text{The answer is B, because } Q_{10} = \frac{R_{20^{\circ}}}{R_{10^{\circ}}} = \frac{9.5}{4.5} = 2.1$$

Top Tip. A common mistake is to simply divide 20°C by 10°C and produce an answer of 2.0, rather than finding the rate at each of these two temperatures.

Surface Area and Volume

You could be asked to calculate the surface areas and volumes of common three-dimensional objects, often so you can calculate the surface area : volume ratio. You will be given these formulae in an exam, but it helps if you are familiar with them:

3-D Shape		Surface area	Volume
cube		$6s^2$	s^3
cuboid (square prism)		$2s^2 + 4sh$	s^2h
sphere		$4\pi r^2$	$\frac{4}{3}\pi r^3$
cylinder (circular prism)		$2\pi r^2 + 2\pi rh$	πr^2h

Don't forget to give the units, and make sure that all the measurements are in the same units.

Surface Area : Volume Ratio

With these formulae you can calculate the surface area : volume ratio of an object by dividing the surface area by the volume. The surface area and volume must have the same base units (e.g. mm^2 and mm^3 , or dm^2 and dm^3 , or m^2 and m^3). The units of surface area : volume ratio are then mm^{-1} , (or dm^{-1} or m^{-1}), but these units are not usually required.

Worked Example

Calculate the surface area : volume ratio of an earthworm, assuming its shape is a cylinder 5 mm in diameter and 12 mm long:

$$\text{surface area} = 2\pi r^2 + 2\pi rh = (2 \times \pi \times 2.5^2) + (2 \times \pi \times 2.5 \times 12) = 228 \text{ mm}^2$$

$$\text{volume} = \pi r^2 h = \pi \times 2.5^2 \times 12 = 236 \text{ mm}^3$$

$$\text{surface area : volume ratio} = \frac{228}{236} = 0.97 : 1$$

Top Tip. You should know how to calculate the area and perimeter of simple two-dimensional shapes, like rectangles and circles.

Gradients and Rates

You may have to calculate the gradient (or slope) of a line on a graph. A common example is when the horizontal axis represents time, in which case the gradient is rate (e.g. growth rate, velocity, rate of reaction, etc.).

The gradient is calculated by $\frac{\Delta y}{\Delta x}$, where Δ (delta) means “change in”. We’ll look at four different examples of gradient calculations.

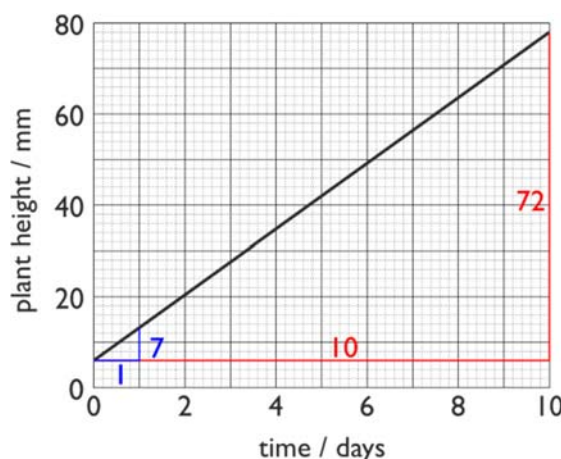
Example 1 – Gradient of a Straight Line

The graph on the right shows the growth of a plant over 10 days. Calculate the rate of growth.

Answer

Draw a right-angled triangle from the graph line and find the x and y lengths: $x = 10$ and $y = 72$. So:

$$\text{rate of growth} = \frac{\Delta y}{\Delta x} = \frac{72 \text{ mm}}{10 \text{ days}} = 7.2 \text{ mm day}^{-1}$$

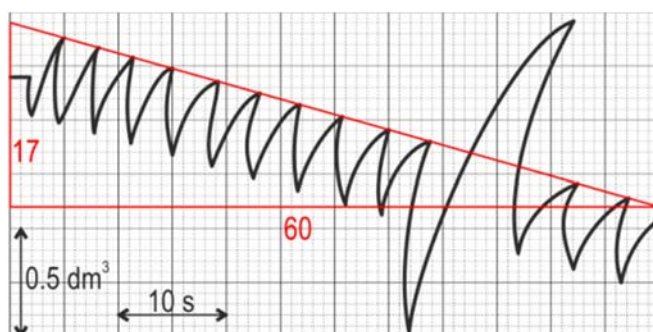


Don't forget to use the correct units from the graph axes.

Top Tip. Always choose the longest straight line possible to minimise errors. For example if you just measured the growth from the graph over one day you would calculate the growth rate as $7/1 = 7 \text{ mm day}^{-1}$, which is less accurate.

Example 2 – Gradient of a Noisy Line

If the graph is “noisy” you will have to draw a straight trend line (line of best fit) through the data first, and then calculate the gradient of this trend line. A common example of this is a spirometer trace, which shows the volume changes each breath but also falls steadily due to oxygen uptake. In this case it is often easier to draw the straight trend line through the peaks or troughs.



What is the mean rate of oxygen uptake from this spirometer trace?

Answer

Draw a straight trend line through the peaks, measure its height and width in graph squares, then use the scales to convert to real units.

Vertical: 17 squares. 10 squares = 0.5 dm^3 , so 17 squares = $17 \times 0.5 / 10 = 0.85 \text{ dm}^3$

Horizontal: 60 squares. 10 squares = 10 s, so 60 squares = 60 s = 1 min

$$\text{rate of oxygen uptake} = \frac{\Delta y}{\Delta x} = \frac{0.85}{1} = 0.85 \text{ dm}^3 \text{ min}^{-1}$$

Example 3 – Gradient of a Curved line

Calculating the gradient is more complicated if the graph is a curve. In this case the gradient changes at different points on the curve and we need to be told where to take the gradient. Place a ruler on the curve at that point and turn it until it is symmetrical with the curve at that point. Draw this straight line, which is now a tangent to the curve at that point, and use this straight line to calculate the gradient.

The graph shows the volume of gas evolved during an enzyme reaction. Calculate

- the initial rate of reaction,
- the rate of reaction at 1 min.

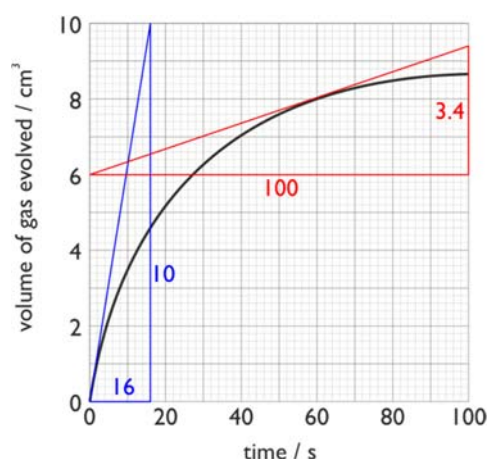
Answer

- We draw a tangent through the origin for the initial rate of reaction. The straight line should match the curve for the first two seconds (blue line). The gradient of this line is

$$\frac{10}{16} = 0.625 \text{ cm}^3 \text{ s}^{-1}$$

- We draw another tangent through the 60s point, making the tangent symmetrical at that point (red line). The gradient of this line is:

$$\frac{3.4}{100} = 0.034 \text{ cm}^3 \text{ s}^{-1}$$

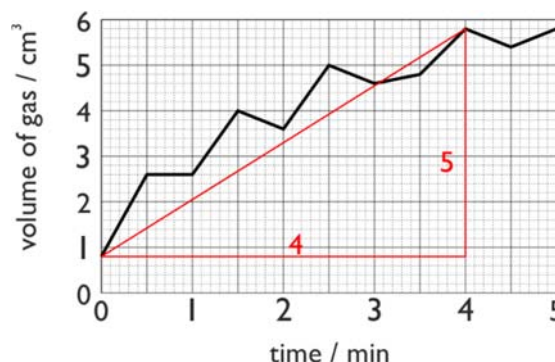
**Example 4 – Mean Gradient of a Curved Line**

Sometimes you might want to know the overall mean gradient of a curved line, rather than the gradient at one point. In this case simply join the start and end points with a straight line and calculate its gradient.

In this example the volume of gas given off by an enzyme reaction was measured every 30 s. What is the mean rate for the first 4 minutes?

Answer

$$\text{mean rate} = \frac{\Delta y}{\Delta x} = \frac{5 \text{ cm}^3}{4 \text{ min}} = 1.25 \text{ cm}^3 \text{ min}^{-1}$$



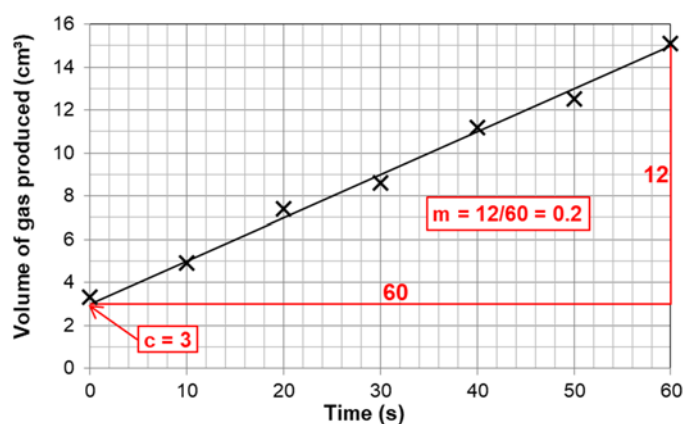
Equation for a straight line ($y = mx + c$)

If a straight trend line is drawn on a graph then you can measure the line's gradient and intercept (m and c in the equation $y = mx + c$). You can then use these values to calculate y for any given value of x (or x for any given value of y).

Worked Example

This graph shows the evolution of gas by an enzyme reaction. Use the equation $y = mx + c$:

- to predict the volume of gas produced at 100 s
- to predict the time taken to produce 20 cm³ of gas.



Answer

First we find c and m . The intercept on the y axis (c) can be read off as 3, and the gradient of the line (m) can be calculated as $\frac{12}{60} = 0.2$.

- To predict the volume of gas at 100 s we simply put the values into the equation:

$$y = mx + c = 0.2 \times 100 + 3 = 23 \text{ cm}^3$$

- To predict the time it would take to produce 20 cm³ of gas, we first rearrange the equation to make x the subject, then we fill in the numbers:

$$\begin{aligned}
 y = mx + c &\Rightarrow y - c = mx \Rightarrow \frac{(y - c)}{m} = x \\
 x = \frac{(y - c)}{m} &= \frac{(20 - 3)}{0.2} = 85 \text{ s}
 \end{aligned}$$

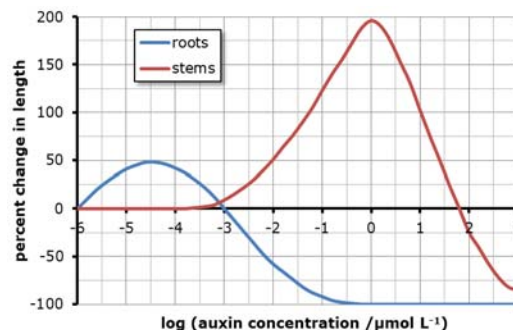
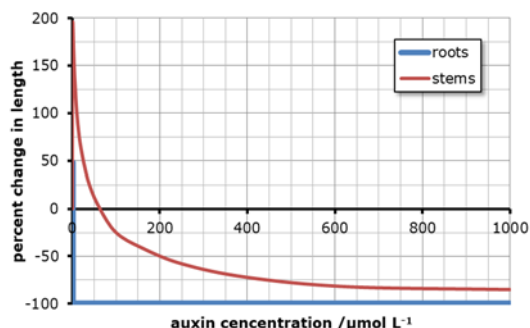
Logarithms

A logarithm (or log) is a mathematical method for scaling a number. It is actually the exponent of the value when it is written in standard form. For example:

- the log of 1000 is 3, because $1000 = 1 \times 10^3$
- the log of 0.01 is -2, because $0.01 = 1 \times 10^{-2}$
- the log of 50 is 1.699, because $50 = 1 \times 10^{1.699}$

Just like exponents, logs can be positive or negative – numbers greater than one have positive logs while numbers less than one have negative logs. Logs can only be found for positive numbers. There are different kinds of logs, with different bases, but we shall always use base 10 logs ($\log_{10}$, or just log). Logs can be found using the **log** button on a calculator, and the antilog is found using the **10^x** button (**SHIFT** **log**). Logs don't have units. You have already come across logs without realising it: pH, absorbance and the Richter scale are all examples of log scales.

Logs are useful in biology because they compress wide-ranging values to a much smaller range, so log scales can be used on graphs when plotting values with a large range. For example, these two graphs show the effect of different concentrations of auxin on the growth of plant roots and shoots. When auxin concentration (on the x-axis) is plotted on a linear scale the effects of low concentrations are obscure, but become much clearer when plotted on a logarithmic scale.



Worked Example

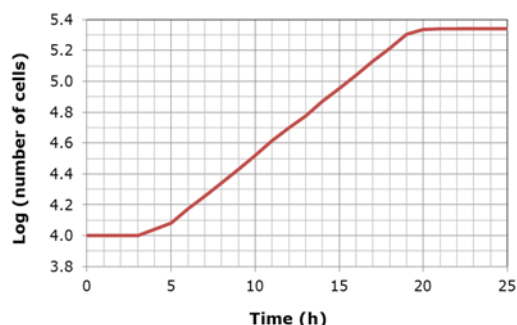
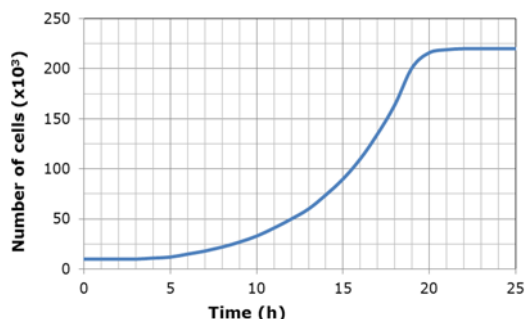
How many times more concentrated does auxin have to be in order to produce maximum growth of stems compared to roots?

Answer

- For roots, the concentration that gives maximum growth is -4.5 on the log scale, which is $10^{-4.5} \mu\text{mol L}^{-1}$.
- For stems, the concentration that gives maximum growth is 0 on the log scale, which is 10^0 (or 1) $\mu\text{mol L}^{-1}$.
- To find how many times more this is, we divide the stem concentration by the root concentration: $10^0 \div 10^{-4.5} = 10^{(0 - (-4.5))} = 10^{4.5}$ (remember from p 5 that to divide numbers in standard form we subtract the exponents).
- To get the answer in ordinary form we use the **10^x** button on our calculator:
 $10^{4.5} = 31\,622$ times.

Exponential Growth

The growth of microbes over time also benefits from being plotted on a log scale, this time on the y-axis.



In this case the advantage is that the exponential growth phase appears as a straight line. This is useful because the exponential growth rate constant (k) can be calculated from the gradient of this straight line.

$$k = \frac{\log_{10} N_t - \log_{10} N_0}{0.301 \times t}$$

The value of 0.301 in the formula is $\log_{10}(2)$ and is needed to give the correct value of k . k is also called the generation rate constant because it is the reciprocal of the generation or doubling time T , i.e. $k = 1/T$. The generation time is the time taken for one generation, or for the population to double in size. A feature of exponential growth is that it has a constant doubling time.

Worked Example 1

What is the exponential growth rate constant from the log graph between 5 and 15 h?

Answer

We first read off the log values at 5 h (4.1) and 15 h (4.95), and then we put these into the equation:

$$k = \frac{4.95 - 4.1}{0.301 \times 10} = 0.282 \text{ h}^{-1}$$

Top Tip. We don't need to take the logs of 4.1 and 4.95, because they are already logs.

Worked Example 2

Sometimes we are given actual cell counts, so we do need to take logs. For example what is the exponential growth rate constant and generation time (in minutes) if a cell count increases from 2.5×10^3 to 4.0×10^6 over the space of 8 hours?

Answer

Using our calculators, $\log_{10} (2.5 \times 10^3) = 3.40$ and $\log_{10} (4.0 \times 10^6) = 6.60$. So we put these log values into the formula:

Exponential growth rate constant: $k = \frac{6.6 - 3.4}{0.301 \times 8} = 1.329 \text{ h}^{-1}$

Generation time: $T = \frac{1}{k} = \frac{1}{1.329} = 0.75 \text{ h} = 45 \text{ min}$

The Hardy-Weinberg Equation

The Hardy-Weinberg equation calculates allele frequencies in a gene pool. It applies to the case of a single gene with full dominance. We define p as the frequency of the dominant allele and q is the frequency of the recessive allele in a population. There are two equations relating p and q :

$$p + q = 1 \quad \text{and} \quad p^2 + 2pq + q^2 = 1$$

We can use the Hardy-Weinberg equations to calculate frequencies of alleles, genotypes and phenotypes, with these formulae:

Allele frequencies	recessive allele (b)	q	} sum = 1
	dominant allele (B)	p	
Genotype frequencies	homozygous recessive (bb)	q^2	} sum = 1
	homozygous dominant (BB)	p^2	
	heterozygous (Bb)	$2pq$	
Phenotype frequencies	recessive phenotype	q^2	} sum = 1
	dominant phenotype	$p^2 + 2pq$	

Usually you start with observed phenotype frequencies. From the recessive phenotype frequency, which is q^2 , you can calculate q and then p , and then anything else.

Top Tip. A question can ask you to calculate any of these frequencies, so read the question carefully to make sure you work out the correct frequency.

Worked Example

Coat colour in cats is controlled by a single gene with two alleles – black (B) and white (b). The black allele is dominant over the white allele. In a population of cats 840 were black and 160 white. How many cats are heterozygous for the coat colour gene?

Answer

1. The total number of cats is 1000 (840 + 160).
2. First we calculate the phenotype frequencies. The phenotype frequency for black is 0.84 (840 / 1000) and for white is 0.16 (160 / 1000).
3. Then we can calculate the frequency of the recessive allele q . We know that white is the recessive allele so the white cats are homozygous recessive (bb). The frequency of this genotype is q^2 , so $q^2 = 0.16$, so

$$q = \sqrt{q^2} = \sqrt{0.16} = 0.4$$

4. Now we can calculate the frequency of the dominant allele, p .

$$p + q = 1, \text{ so } p = 1 - q = 1 - 0.4 = 0.6$$

5. We want the frequency of the heterozygous cats, which is $2pq$:

$$2pq = 2 \times 0.6 \times 0.4 = 0.48$$

6. Finally we convert this frequency to an actual number. The total population is 1000, so:

$$\text{number of heterozygous cats} = 0.48 \times 1000 = 480.$$

Cardiac Output and Minute Ventilation

These are two common quantities in physiological investigations.

Heart: Cardiac Output

Cardiac Output is the volume of blood pumped out of each ventricle in one minute:

$$\text{cardiac output} = \frac{\text{stroke volume}}{(\text{dm}^3 \text{ min}^{-1})} \times \frac{\text{heart rate}}{(\text{beats min}^{-1})}$$

The heart rate is often obtained from a chart of the cardiac cycle by measuring the time taken for one cycle and using the formula:

$$\text{heart rate (beats min}^{-1}\text{)} = \frac{60}{\text{cycle time (s)}}$$

Lungs: Minute Ventilation

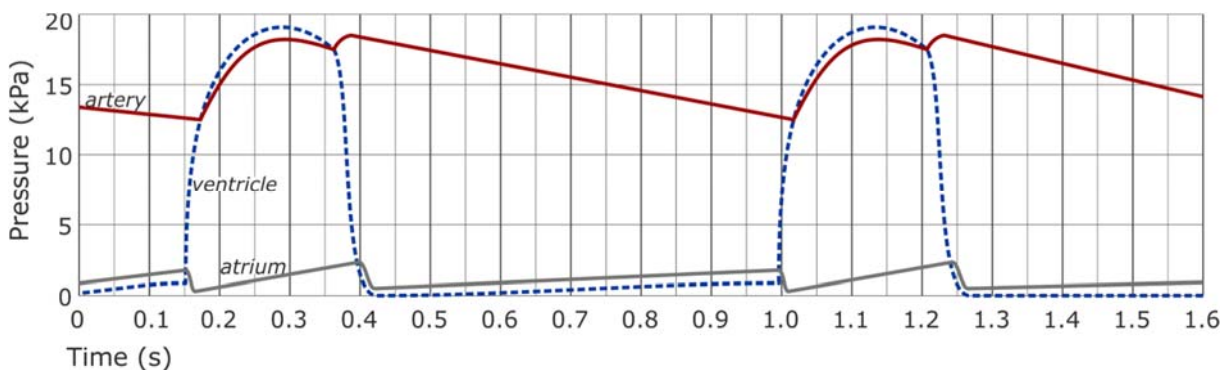
Minute ventilation is the volume of air breathed in or out of the lungs each minute:

$$\text{minute ventilation} = \frac{\text{tidal volume}}{(\text{dm}^3 \text{ breath}^{-1})} \times \frac{\text{ventilation rate}}{(\text{breaths min}^{-1})}$$

The ventilation rate is often obtained from a pressure graph by measuring the time taken for one ventilation cycle and using the formula:

$$\text{ventilation rate (breaths min}^{-1}\text{)} = \frac{60}{\text{cycle time (s)}}$$

Worked Example



This chart shows the pressure changes in one side of a heart. If the stroke volume is 75 cm^3 , calculate the cardiac output in $\text{dm}^3 \text{ min}^{-1}$.

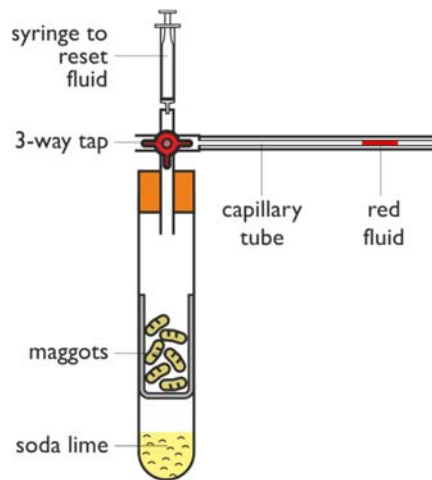
Answer

$$\text{cycle time} = 0.85 \text{ s, so heart rate} = \frac{60}{0.85} = 70.6 \text{ bpm}$$

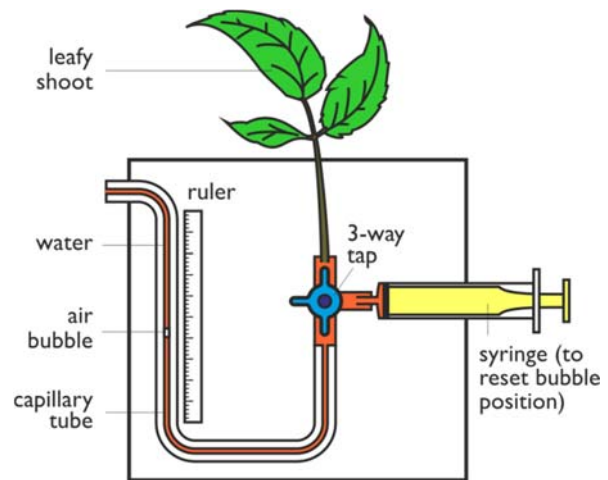
$$\text{cardiac output} = 70.6 \text{ bpm} \times 75 \text{ cm}^3 = 5295 \text{ cm}^3 = 5.30 \text{ dm}^3 \text{ min}^{-1}$$

Top Tip. It's easy to mis-read the cycle time. Don't assume the whole chart is exactly one cycle (or a whole number of cycles). Choose an obvious feature that is repeated each cycle – for example where the ventricle line crosses above the atrium line in this case. This happens at 0.15 s and again at 1.0 s, so the cycle time is 0.85 s.

Respirometers and Potometers



A **respirometer** measures the rate of respiration in small organisms. Soda lime absorbs carbon dioxide and the organisms take in oxygen as they respire, so the pressure falls and the fluid is pulled in.



A **potometer** measures the rate of transpiration of plant shoots. As the leaves transpire, water is sucked up the stem and the capillary tube, so the air bubble is moved towards the leafy shoot.

In both cases we need to calculate the volume taken up in a certain time. For respirometers this is the volume of oxygen used and for potometers this is the volume of water transpired. The capillary tube is a cylinder, so its volume (in mm^3) can be calculated by multiplying the distance (in mm) by the cross-sectional area of the capillary tube (in mm^2). Capillary tubes usually have an internal diameter of 1 mm, so:

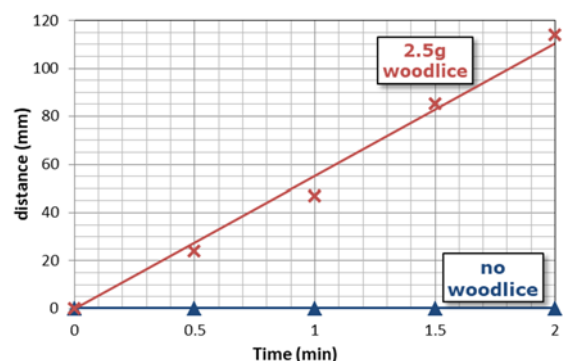
$$r = 0.5 \text{ mm, so cross-sectional area} = \pi r^2 = \pi \times 0.5^2 = 0.79 \text{ mm}^2.$$

This volume is divided by the time to calculate a rate (in $\text{mm}^3 \text{ min}^{-1}$) and may further be standardised by dividing by the mass of organisms or area of leaves.

Worked Example

The chart shows the movement of the fluid in a respirometer with woodlice in it.

- Calculate the rate of oxygen uptake by the woodlice in $\text{mm}^3 \text{ min}^{-1} \text{ g}^{-1}$ of woodlice. The capillary tube has a diameter of 1 mm.
- Why are these units ($\text{mm}^3 \text{ min}^{-1} \text{ g}^{-1}$) a useful way to present the rate?

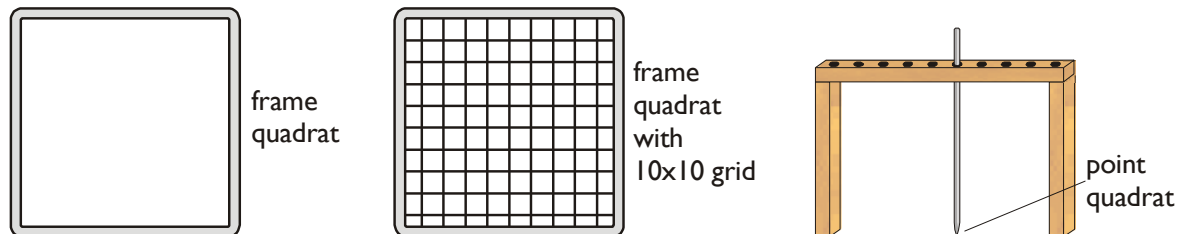


Answer

- distance moved in 2 min = 110 mm (from trend line)
 volume in 2 min = $110 \times \pi \times 0.5^2 = 86.4 \text{ mm}^3$
 rate = $86.4 \text{ mm}^3 / 2 \text{ min} = 43.2 \text{ mm}^3 \text{ min}^{-1}$
 mass of woodlice = 2.5 g, so rate = $43.2 / 2.5 = 17.3 \text{ mm}^3 \text{ min}^{-1} \text{ g}^{-1}$
- rate in $\text{mm}^3 \text{ min}^{-1} \text{ g}^{-1}$ allows us to compare organisms of different size.

Calculations using Quadrats

Quadrats are commonly used in ecology investigations to sample plants and some animals. Quadrats come in a variety of sizes, commonly 0.1 m, 0.5 m or 1 m a side, and may be subdivided into 25 or 100 smaller squares. The smallest quadrat is the point quadrat, which is a needle, with the point of the needle being the actual tiny quadrat.



Quadrats allow us to make quantitative measurements of the abundance of plants. There are different ways to do this, involving different calculations:

- **Density.** This is the number of individual organisms, divided by the area of the quadrat. For example, if there is an average of 12 daisies in a 0.5 m quadrat (area 0.25 m²), the density is $12/0.25 = 48$ daisies m⁻². The advantage of density is that it can be scaled up to estimate the population of a whole habitat by simply multiplying by the total area. Density isn't appropriate when individuals are difficult to identify.
- **Species Frequency.** This is the fraction of quadrats in which a species was found (its frequency). For example if a species was found in 12 quadrats out of a total of 40, then the frequency is $12/40 = 30\%$. Alternatively the fraction of small squares in a grid quadrat in which the species is found is recorded.
- **Percent Cover.** This is appropriate when it is difficult to identify individual plants (such as grasses). The percentage area of the quadrat covered by that particular species is estimated. Since plants can be layered the total percentage cover can be more than 100%.
- **Point Quadrats** provide an alternative way to measure percent cover. A needle is randomly placed many times and the fraction of times it hits the species gives the percent cover. For example if a species was hit 66 times out of a total of 200 needles then the percent cover is $66/200 = 33\%$. This is less subjective than using a frame quadrat.
- **Abundance Scale.** This is a qualitative way to assess abundance. A common scale is the five-point "ACFOR" scale where A = Abundant; C = Common; F = Frequent; O = Occasional; R = Rare. With practice, this is a very quick way to collect data, but it is subjective and not quantitative. The scale can be made semi-quantitative by making the points correspond to ranges of percent cover, and the categories can be ranked (A = 5, C = 4, F = 3 etc.), allowing statistical analysis to be used.

The Lincoln Index

The Lincoln Index is used to estimate the size of a population of animals using the capture-mark-recapture technique:

1. Capture a sample of animals using a suitable trapping technique.
2. Count all the animals in this sample (S_1) and mark them in a way that doesn't harm them (e.g. a spot of paint, leg-rings, a shaved patch of hair).
3. Release all the animals where they were caught and give them time to mix with the rest of the population.
4. Capture a second sample of animals in the same place using the same trapping technique. Count the animals in the second sample (S_2), and the number of marked (i.e. recaptured) animals in the second sample (R).

We estimate the size of the population (N , the Lincoln index) using the formula:

$$N = \frac{S_1 S_2}{R}$$

Worked Example

A student wished to estimate the population size of one species of aquatic snail in a pond. She collected 110 snails on day 1 and marked the shells of them all with a small spot of waterproof paint. On day 2 she collected another sample of snails. She found 35 with paint on them and 230 without.

Calculate the population size in the pond for this species.

Answer

$S_1 = 110$, $S_2 = 265$ (35+230) and $R = 35$

So $N = \frac{110 \times 265}{35} = 833$ snails

Top Tip. Read the question carefully. In this question, the number recaptured on day 2 is not only the 230 unmarked individuals but also the 35 marked ones.

Simpson's Diversity Index

Species diversity means the variety of species in a community. Species diversity is useful because it tells us about the complexity, quality and stability of an ecosystem. Species diversity can be quantified using the Simpson's Diversity Index (D), which takes into account the species richness (the number of species) and the abundance of each species population (the number of individuals of each species). The higher the index, the higher the species diversity. The Simpson diversity index is given by the formula:

$$D = \frac{N(N-1)}{\sum n(n-1)}$$

where N = total number of individuals (total abundance)
 n = number of individuals in each species
 $\sum$ = sum

D has no units. To calculate the D you usually start with a table of species and their abundances (measured by a suitable sampling technique).

Worked Example

This table shows the number of plants growing on a school playing field. Calculate the index of diversity.

species	abundance (number of individuals in 100m ² field)	$n(n-1)$
daisy	349	121 452
dandelion	283	79 806
buttercup	313	97 656
total	945	298 914



Answer

1. Add an extra column to this table to calculate $n(n-1)$ for each species. For the daisy this is $349 \times 348 = 121\,452$.
2. Then calculate the total of this column to give $\sum n(n-1)$, and the total of the abundance column to give N .
3. We can now calculate D: $D = \frac{(945 \times 944)}{298914} = 2.98$

Top Tip. Use the memory function of your calculator (**M+**) then (**RCL**) (**M+**) to sum the $n(n-1)$ values as you go. This saves typing the values in again.

Teacher comment. Simpson's original index (also called D) was an index of similarity, with higher values meaning lower diversity. A diversity index is the opposite of the original index, but there are different ways of calculating this opposite. Some ecologists use the reciprocal ($1 / D$), used by Edexcel, while others use the complement ($1 - D$), used by other awarding bodies. Unfortunately they're all called D! This explains why there are different formulae in different exams, and why Edexcel will always give the formula to the students.

Productivity and Trophic Efficiency

Productivity

The productivity of an ecosystem is the amount of biomass produced by that ecosystem each year, measured in units of biomass (e.g. $\text{kg m}^{-2} \text{y}^{-1}$) or energy (e.g. $\text{kJ m}^{-2} \text{y}^{-1}$). The amount of energy fixed by producers in photosynthesis is called the gross primary productivity (GPP). But some of this gross productivity is lost as heat via respiration (R), and the amount of energy actually accumulated in producer biomass is called the net primary productivity (NPP). So the two terms are related by the equation:

$$\text{NPP} = \text{GPP} - \text{R}$$

Trophic Efficiency

The trophic efficiency (or ecological efficiency) shows how much of the energy received by a trophic level (as food or light energy) is passed on to the next trophic level. It can be calculated in terms of energy flow or productivity:

$$\text{trophic efficiency (\%)} = \frac{\text{energy passed on to next level}}{\text{energy received from previous level}} \times 100 = \frac{\text{Net productivity of this level}}{\text{Net productivity of previous level}} \times 100$$

As a very rough guide, trophic efficiencies are around 10%, though they vary a lot.

Worked Example

Scientists measured the productivity in a woodland ecosystem.

Trophic Level	Gross Productivity $/\text{g m}^{-2} \text{y}^{-1}$	Net Productivity $/\text{g m}^{-2} \text{y}^{-1}$
Producers	2200	1200
Primary consumers	38	18
Secondary consumers	12	3.2

1. Calculate the overall respiration rate of the producers.

Answer

$$\text{R} = \text{GPP} - \text{NPP} = 2200 - 1200 = 1000 \text{ g m}^{-2} \text{y}^{-1}$$

2. Calculate the trophic efficiency for the two consumers.

Answer

$$\text{primary consumers: efficiency} = \frac{\text{NP out}}{\text{NP in}} \times 100 = \frac{18}{1200} \times 100 = 1.5\%$$

$$\text{secondary consumers: efficiency} = \frac{\text{NP out}}{\text{NP in}} \times 100 = \frac{3.2}{18} \times 100 = 17.8\%$$

Calculations without Formulae

In addition to all the formulae mentioned in this guide, exams often have maths questions that don't require a known formula, like magnification or diversity index. These are problems where you have to work out for yourself what to do with the numbers to arrive at the answer, without any guidance from a known formula. You have to use common sense to work out what to do, and it can help to look at the units of the answer (if given). For example if the units of the answer are kg h^{-1} , then you know you'll need to divide some mass in kg by time in hours.

Here are two examples, modified from real exam questions.

Worked Example 1

The table shows the effect of air temperature on the metabolic rate in the arctic ground squirrel.

Calculate the change in metabolic rate for an arctic ground squirrel, with a body mass of 950g, as the air temperature increases from -16°C to 4°C . Give your answer in $\text{dm}^3 \text{ oxygen day}^{-1}$.

Air temperature / $^{\circ}\text{C}$	Metabolic rate / $\text{cm}^3 \text{ oxygen g}^{-1} \text{ h}^{-1}$
-16	0.25
-8	0.11
-4	0.06
0	0.03
4	0.03
8	0.03
12	0.03

Answer

- We're asked to calculate the change in metabolic rate as the air temperature increases from -16°C to 4°C , so the first thing is to read these two values from the table and calculate the change.
rate at -16°C = 0.25; rate at 4°C = 0.03; change = $0.22 \text{ cm}^3 \text{ g}^{-1} \text{ h}^{-1}$
- The units of the answer don't have g^{-1} in them, but we're given the mass of the squirrel, so if we multiply by grams, we'll get rid of g^{-1} .
rate = $0.22 \text{ cm}^3 \text{ g}^{-1} \text{ h}^{-1} \times 950 \text{ g} = 209 \text{ cm}^3 \text{ h}^{-1}$
- The answer has units of dm^3 , not cm^3 , so we have to convert the units. There are 1000 cm^3 in a dm^3 (this is a useful thing to know), so we divide by 1000.
rate = $209 \text{ cm}^3 \text{ h}^{-1} / 1000 = 0.209 \text{ dm}^3 \text{ h}^{-1}$
- The answer also has units of day^{-1} , not h^{-1} , so we have to convert the units. There are 24 hours in a day, so we multiply by 24.
rate = $0.209 \text{ dm}^3 \text{ h}^{-1} \times 24 = 5.02 \text{ dm}^3 \text{ day}^{-1}$

And this is our answer in the correct units.

Top Tip. You have to use common sense when converting units. In step 3 we divided by 1000 because the volume in dm^3 is going to be a smaller number than the same volume in cm^3 . In step 4 we multiplied by 24 because more oxygen will be used in one day than in one hour.

Worked Example 2

This table shows data for four diving mammals.

Calculate the total mass of haemoglobin in a Northern elephant seal.

Species	Mass of animal /kg	Volume of oxygen in body /cm ³ kg ⁻¹	Concentration of haemoglobin in blood /g dm ⁻³	Total blood volume /cm ³ kg ⁻¹
bottlenose dolphin	200	36	14	71
harbour seal	24	57	21	132
Weddell seal	400	87	210	173
Northern elephant seal	400	97	216	207

Answer

There's a lot of data here, and the difficulty is deciding which bits we need. Obviously we're looking at the Northern elephant seal row, but which columns do we need? The third column is concentration of haemoglobin in blood, but it's in units of g dm⁻³ of blood, so we need the blood volume. That's in the fourth column, but it's in units of cm³ kg⁻¹ of body mass, so we need to convert the volume to dm³, and we need the body mass. That's in the first column, and it's in the right units! Now we know what to do.

1. The blood volume is 207 cm³ kg⁻¹ and the body mass is 400 kg, so we multiply those to get the blood volume.
Blood volume = 207 cm³ kg⁻¹ × 400 kg = 82 800 cm³
2. We need to convert that blood volume to dm³. There are 1000 cm³ in a dm³, so we divide by 1000.
Blood volume = 82 800 cm³ / 1000 = 82.8 dm³
3. Now we can multiply by the haemoglobin concentration of 216 g dm⁻³.
Haemoglobin mass = 216 g dm⁻³ × 82.8 dm³ = 17 884 g
4. This number is a bit big, so we should convert to kg and round to 3 significant figures.
Haemoglobin mass = 17 884 g / 1000 = 17.9 kg

Top Tip. If you have to calculate some value in a table, check to see if you understand how the existing entries were calculated, then just repeat the procedure for your value.

Summary Statistics

Calculating the Central Value of a Sample

Any good investigation will have many replicates, and we can summarise the sample of replicates by finding their central value. There are three different ways of measuring the central value: the mean, the median and the mode.

We'll demonstrate how to calculate each for this sample showing the resting heart rate (in beats per min) of 16 men:

74	74	69	76	74	72	70	77	75	71	72	73	78	73	74	70
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Mean

The mean (or arithmetic mean) is the best choice if the data are normally-distributed (and they usually are). The mean is represented by the symbol $\bar{x}$ ("x bar") and calculated using the formula: $\bar{x} = \frac{\sum x}{n}$.

So to calculate the mean heart rate we add up all the heart rates and divide by 16.

$$\text{mean} = \frac{74+74+69+76+74+72+70+77+75+71+72+73+78+73+74+70}{16} = \frac{1172}{16} = 73.25 \text{ bpm}$$

You can also calculate the mean using the statistical function on a calculator (see p65).

Top tip. Don't call the mean the average. The term average is too vague for A-level.

Median

The median is the middle value of a sample, i.e. the one with equal numbers of replicates greater than and less than it. The median is the best measure of the central value if the data are not normally-distributed (e.g. ranked data). There is no formula to calculate the median, so we have to find the median by ranking all the replicates from smallest to largest and then finding the middle value. If there is an even number of values, then take the mean of the middle two. So, for the heart rate data:

1. First we re-write the values but in rank order (smallest to largest). It's a good idea to cross off each value in the original table as you go, to avoid missing any values:

69 70 70 71 72 72 73 73 74 74 74 74 75 76 77 78 

2. Count in from the two ends to find the middle value. If there is an odd number of values there will be a single middle value, which is the median. But if there is an even number of values (as here) there will be two middle values:

69 70 70 71 72 72 73 73 74 74 74 74 75 76 77 78

3. The median in this case is the arithmetic mean of the two middle values:

$$\text{median} = \frac{(73 + 74)}{2} = 73.5 \text{ bpm}$$

The Casio classwiz calculator (but not the GT plus) can calculate the median, which is a much easier way to do it (see p65).

Mode

The mode is the most common, or frequent, replicate value. It is most useful when there are a very large number of replicates. You often make a frequency table or plot a histogram to explore the distribution of the replicates. A mode can apply to quantitative data (in which case it's called a modal value) or qualitative data (in which case it's called a modal class). There can be more than one mode (if more than one value is equally frequent, or there are two widely-separated peaks), or no mode (if all values are equally frequent).

If there are a fairly small number of replicates, the mode can be found by re-writing the values in rank order, and seeing which value is the most common.

69 70 70 71 72 72 73 73 74 74 74 74 75 76 77 78

For this heart rate data the mode is 74 bpm.

If there are a large number of replicates, then it is helpful to draw a frequency table, like this one.

With a larger sample of men we can see that the modal value is 73 bpm, because it has the highest frequency.

Top Tip. A common mistake is to read the frequency value from a frequency table and say that the mode is 14.

heart rate (bpm)	tally	frequency
68	### /	6
69	### ///	8
70	### ### //	12
71	### ### /	11
72	### ////	9
73	//// //// ////	14
74	### ////	9
75	### ///	8
76	### ###	10
77	### ### //	12
78	### ###	10

A mode can also be found for categorical data.

This table shows the frequencies of six plant species in a field measured using a point quadrat.

The mode (or modal class) is buttercup.

plant	tally	frequency
buttercup	### ### ### ### /	21
daisy	### ### ### ///	18
dandelion	### ### //	12
dock	### ///	8
plantain	### ### /	11
thistle	### ###	10

Top Tip. You may be expected to remove any anomalies (outliers) from a set or replicates before calculating a mean. But only do this if the anomaly is very different from all the other results (say by more than a factor of 5). And never remove anomalies when calculating median or mode – these values are not affected by anomalies.

Calculating the Dispersion of a Sample

As well as calculating the central value of a sample, we also need to know the dispersion of the sample, or how variable it is, for that will tell us how reliable the central value is. There are two ways to measure the dispersion of a sample that you need to know about.

Range

The range is simply the difference between the maximum and minimum values of a sample, and so is very easy to calculate. However it is only based on two values (which may be outliers) rather than the whole sample, so is not a good indicator of dispersion and is not widely used.

Standard Deviation

The standard deviation (s) measures the dispersion of the raw data around the mean. The standard deviation is calculated from all the replicate data, so is more representative than the range. The larger the standard deviation the larger the spread of the data about the mean, so the less reliable the sample mean.

- 68% of the sample replicates will be found in the range $\text{mean} \pm s$
- 95% of the sample replicates will be found in the range $\text{mean} \pm 1.96 s$.

Whenever a mean is calculated a standard deviation should also be calculated to indicate the quality of the sample mean. The summary can then be written as $\bar{x} \pm s$, e.g. mean height = 1.62 ± 0.08 m.

The standard deviation is calculated by the formula on the right. This shows that it is calculated by summing the squared differences between each value (x) and the mean ($\bar{x}$). Obviously the bigger the differences, the bigger the standard deviation.

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

This shows another version of the formula. This looks more complicated, but may actually be easier to use, since you don't have to calculate the mean first.

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$$

But by far the easiest way of calculating the mean and standard deviation is to use the statistical function of a scientific calculator (see p65).

Worked Example

The table shows the lengths of a sample of 12 snail shells.

Calculate the mean and standard deviation for this sample.

We'll demonstrate the use of both equations.

snail	shell length / mm (x)	new column for first equation ▼ $(x - \bar{x})^2$	new column for second equation ▼ x^2
1	32	134.6	1024
2	51	54.8	2601
3	26	309.8	676
4	46	5.8	2116
5	38	31.4	1444
6	65	458.0	4225
7	50	41.0	2500
8	36	57.8	1296
9	41	6.8	4761
10	48	19.4	2304
11	46	5.8	2116
12	44	0.2	1936
Total	523	1124.9	23919

First Equation

1. Calculate the mean by adding up all the lengths (523) and dividing by n (12).

$$\bar{x} = \frac{\sum x}{n} = \frac{523}{12} = 43.6$$

2. Make a new column called $(x - \bar{x})^2$ and, for each row, subtract the mean from the length value and square the answer. So for the first row do $(32 - 43.6)^2 = 134.6$.
3. Add up this column to give $\sum(x - \bar{x})^2$ (1124.9).
4. Fill in the values in the formula:

$$s = \sqrt{\frac{1124.9}{11}} = \sqrt{102.3} = 10.1$$

Second Equation

1. Calculate the mean by adding up all the lengths (523) and dividing by n (12).

$$\bar{x} = \frac{\sum x}{n} = \frac{523}{12} = 43.6$$

2. Make a new column called x^2 and, for each row, write the square of the length value. So for the first row do $32^2 = 1024$.
3. Add up this column to give $\sum x^2$ (23 919)
4. Fill in the values into the formula.

$$s = \sqrt{\frac{23919 - \frac{523^2}{12}}{11}} = \sqrt{\frac{1124.9}{11}} = 10.1$$

Calculator Method

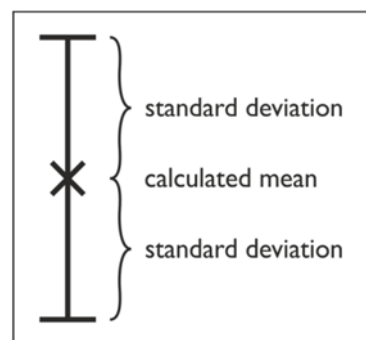
In stats mode, simply type in the 12 length values, and then the calculator will give you the mean and standard deviation without any calculations (see p65).

Whichever method we use, we can now write that the mean shell length is 43.6 ± 10.1 mm.

Error Bars

Standard deviations can be shown on charts as error bars, to give a visual indication of the spread of the data about the mean and therefore the uncertainty of the mean (in fact they would be better called uncertainty bars).

Using error bars emphasises the point that a calculated mean value is not perfect and is not the true value – it's just our best provisional estimate of the true value from our data so far. The true value could easily be higher or lower than our mean. Error bars allow us to put limits on that range: the true value is likely to be somewhere within the error bars, but is very unlikely to be outside that range.

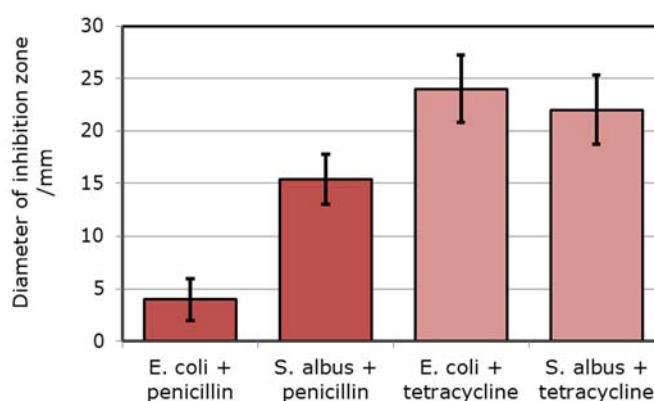


In exams, whenever error values are given (either as numbers in a table or as error bars in a chart), you should look for an overlap. If the error bars overlap, then you can conclude that there is no significant difference between the mean values. Conversely, if the error bars don't overlap, then you can conclude that the mean values are significantly different. (Actually this isn't strictly true for standard deviation error bars, but it's the right general idea, and it is what is expected in A level exams.)

Top Tip. In the next section we'll learn a better way to test for significant differences.

Worked Example

The effectiveness of two antibiotics (penicillin and tetracycline) on the growth of two species of bacteria (*S. albus* and *E. coli*) was investigated using a disc diffusion assay. There were 24 replicates ($n = 24$) and the bar chart shows mean diameters of the inhibition zones around each antibiotic disk. The error bars show ± 1 standard deviation. What conclusions can be drawn from these results?



Answer

- The error bars for the two penicillin results don't overlap, so penicillin is significantly more effective against *S. albus* than against *E. coli*.
- The error bars for the two tetracycline results do overlap, so there is no significant difference in the effect of tetracycline on *S. albus* and *E. coli*.
- The error bars for tetracycline don't overlap with those for penicillin, so tetracycline is significantly more effective than penicillin against both species of bacteria.

Statistical Tests

In A level biology you have to be able to carry out statistical tests. Statistical tests tell us whether an observed effect is real or just due to chance. Suppose we find that mean light intensity is lower in a deciduous wood than in a coniferous wood. Is this difference real, or did we by chance choose darker sites in the deciduous wood?

The Null Hypothesis

All statistical tests are built on a mathematical model called the null hypothesis, which states that there is no effect (e.g. no difference between groups, or no correlation between variables). The null hypothesis is fixed for a given test, and we must state the null hypothesis when we've chosen our test so that we are clear about what exactly we're testing for.

The null hypothesis has nothing to do with (and can be quite different from) any scientific hypothesis we may be making about the result of the experiment. You may also hear about the alternate hypothesis, which is the opposite of the null hypothesis, but the alternate hypothesis is not really needed for a conclusion, so we shall do without it. Note that the word "significant" cannot be used in the null hypothesis as it has no meaning here, though it is used later in the conclusion, when we have a *P*-value.

P-values

The statistical test returns a probability (or *P*-value), on a scale of 0–1. This *P*-value is the probability that our results would be obtained if the null hypothesis were true, and that any observed effect is just due to random chance.

So we're hoping for a low *P*-value: it means that the effect in the results is probably not due to chance, so we say that the effect is significant. In biology we usually take a probability of 0.05 (5% or 1 in 20) as the cut-off. This cut-off is called the significance level of the test.

So we have two possible conclusions to choose from:

- If $P < 0.05$ this means that the probability is < 0.05 that the observed effect is due to chance. We reject the null hypothesis and conclude that there is a significant effect.
- If $P \geq 0.05$ this means that the probability is ≥ 0.05 that the observed effect is due to chance. We accept the null hypothesis and conclude that there is no significant effect.

Top Tip. Learn all the parts of these alternative conclusions and use them when you have to write a conclusion in an exam. They're all important and could gain marks.

Effect Size

If the effect is significant, then we should also report the effect size, e.g. the size of a difference (expressed either as an absolute value or as a percentage change) or the strength of a correlation (i.e. the correlation coefficient).

Critical Value Tables

Unfortunately we can't calculate P -values directly. Instead we calculate a test statistic, which we compare with a critical value to determine whether the P -value is greater or less than 0.05. For this we need a set of critical value tables, and these will be provided in any exam question that needs them. Below are tables giving critical values of t , χ^2 and r_s . The tables in exams may look a bit different, but will have basically the same features.

To use these tables you need two pieces of information:

- Significance level (α). This is the level of significance or certainty we want in our conclusion. As already mentioned, the usual level in biology is 0.05, but you can choose other levels of significance, and they are found in tables.
- Degrees of freedom (dof) or number of pairs (n) for the Spearman correlation. Degrees of freedom is generally one less than the number of results you have ($n - 1$), but you have to learn the exact dof formula for each test.

Find the column for your significance level (usually 0.05) and the row for your dof and read the value where they meet. This is the critical value for your test. For example, for a t -test with $\alpha = 0.05$ and dof = 7 the critical value is 2.36.

Then you compare this critical value with your calculated value to find the P -value for this test.

- If the critical value at $\alpha=0.05 >$ the calculated value, then $P > 0.05$
- If the critical value at $\alpha=0.05 <$ the calculated value, then $P < 0.05$

Critical Values of t			
dof	significance level (α)		
	0.1	0.05	0.01
1	6.31	12.71	63.66
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.02	2.57	4.03
6	1.94	2.45	3.71
7	1.89	2.36	3.50
8	1.86	2.31	3.36
9	1.83	2.26	3.25
10	1.81	2.23	3.17
11	1.80	2.20	3.11
12	1.78	2.18	3.05
13	1.77	2.16	3.01
14	1.76	2.14	2.98
15	1.75	2.13	2.95
16	1.75	2.12	2.92
17	1.74	2.11	2.90
18	1.73	2.10	2.88
19	1.73	2.09	2.86
20	1.72	2.09	2.85

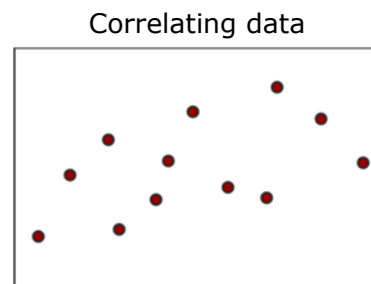
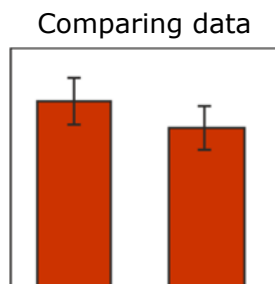
Critical Values of χ^2			
dof	significance level (α)		
	0.1	0.05	0.01
1	2.71	3.84	6.63
2	4.61	5.99	9.21
3	6.25	7.81	11.34
4	7.78	9.49	13.28
5	9.24	11.07	15.09
6	10.64	12.59	16.81
7	12.02	14.07	18.48
8	13.36	15.51	20.09
9	14.68	16.92	21.67
10	15.99	18.31	23.21
11	17.28	19.68	24.72
12	18.55	21.03	26.22
13	19.81	22.36	27.69
14	21.06	23.68	29.14
15	22.31	25.00	30.58
16	23.54	26.30	32.00
17	24.77	27.59	33.41
18	25.99	28.87	34.81
19	27.20	30.14	36.19
20	28.41	31.41	37.57

Critical Values of r_s			
n	significance level (α)		
	0.1	0.05	0.01
5	0.90	1.00	1.00
6	0.83	0.89	1.00
7	0.71	0.79	0.93
8	0.64	0.74	0.88
9	0.60	0.70	0.83
10	0.56	0.65	0.79
11	0.54	0.62	0.82
12	0.50	0.59	0.78
13	0.48	0.57	0.75
14	0.46	0.55	0.72
15	0.44	0.53	0.69
16	0.43	0.51	0.67
17	0.41	0.49	0.65
18	0.40	0.48	0.63
19	0.39	0.46	0.61
20	0.38	0.45	0.59
21	0.37	0.44	0.58
22	0.36	0.43	0.56
23	0.35	0.42	0.55
24	0.34	0.41	0.54

Choosing the Correct Test

There are many different statistical tests available for different purposes, but for A-level Biology you only have to know about four: the t -test (also called the Student's t -test), the paired t -test, the Spearman correlation coefficient and the chi squared (χ^2) test.

To choose the correct test, you first need to decide what kind of analysis you want to do. There are basically two kinds of experiment, characterised by different charts:



You also need to decide what kind of data you will be collecting. Again there are basically two kinds: quantitative (numeric) data or qualitative (categorical) data. Categorical data are pretty rare and mostly occur in genetic crosses.

Then you can use this table to choose the best test:

What do I want to do?	Quantitative (numeric) data	Qualitative (categorical) data
Summarise data	Mean and standard deviation	Tally frequencies
Compare two sets of data (bar chart)	Unpaired t -test (independent samples) Paired t -test (matched samples)	Chi-squared test
Correlate two sets of data (scatter graph)	Spearman correlation coefficient	<i>not needed in exams*</i>

*Note: There is another chi-squared test that tests categorical data for an association. You may come across it in books or old exam papers, but it's not needed in current A-level biology exams.

Example: suppose you want to know if grazing affects the abundance of certain grassland plants. You might compare two areas (a grazed field and an un-grazed field) and you would use quadrats to collect numeric abundance data (such as plant density). You could then plot a bar chart with two bars, so the correct test is the unpaired t -test.

Top Tip: If you have to choose a statistical test for an investigation, it helps to sketch a chart with labelled axes of what your results might look like. Would the results look best as a bar chart or a scatter graph?

Whatever kind of experiment you are doing, all the statistics tests work better with a large sample size (n). The examples in exams and in this guide often have sample sizes of 10 or less just for convenience, but for a real investigation plan for at least 30 and preferably 100 repeats. Three repeats may be acceptable for a school lesson but it's quite unacceptable for a serious investigation!

The Unpaired t -test

The unpaired t -test (or just t -test, and also known as the Student's t -test) tests for differences between two sets of independent data. The unpaired t -test tests the null hypothesis that there is no difference between the means of the two sets.

The formula for unpaired t is:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where:

$|x|$ = absolute value (ignore sign)

$\bar{x}_1, \bar{x}_2$ = means

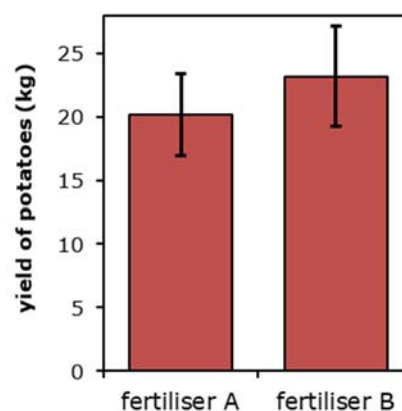
s_1, s_2 = standard deviations

n_1, n_2 = number of repeats of the two sets

Worked Example

A farmer compared the effect of two different fertilisers on the yield of potatoes. He grew potato plants in many plots, and treated each plot with equal masses of the two fertilisers A or B. After a set time he dug up the potatoes and recorded the yield from each plot, with the results shown in the table below. One of the plots was damaged by animals and yielded no potatoes. Which fertiliser is better?

	yield of potatoes /kg	
	fertiliser A	fertiliser B
	27	28
	20	19
	16	18
	18	21
	22	24
	19	20
	23	25
	21	27
	17	29
	19	
$\bar{x}$	20.20	23.44
s	3.22	4.10
n	10	9



Answer

1. First we choose the test. We are comparing two unmatched groups, and would plot a bar chart, so we use the t test.

2. State the null hypothesis that we are testing:

There is no difference in mean yield between the two fertilisers

3. Calculate the mean of each set ($\bar{x}$) and put it in a new row as shown above.

4. Calculate the standard deviation (s) of each set, and put it in a new row. It saves time if you use the stats function of a calculator for these two steps.

5. Count the number of measurements in each set (n), and put it in a new row.

6. Put all the values into the formula to calculate unpaired t :

$$t = \frac{|20.20 - 23.44|}{\sqrt{\frac{3.22^2}{10} + \frac{4.10^2}{9}}} = \frac{3.24}{\sqrt{1.04 + 1.87}} = \frac{3.24}{1.71} = 1.89$$

7. Calculate the degrees of freedom: $\text{dof} = (n_1 - 1) + (n_2 - 1) = 9 + 8 = 17$

Note that the two samples don't have to be the same size.

8. Look up the critical value of t for $\alpha = 0.05$ and $\text{dof} = 17$ ($= 2.11$).

9. Compare this critical value of t (2.11) with our calculated value of t (1.89).

$$t_{\text{crit}} (2.11) > t_{\text{calc}} (1.89) \text{ so } P > 0.05$$

10. Now that we have our P -value, we can state the conclusion:

The probability is greater than 0.05 that the observed difference in the means is just due to chance, so we accept the null hypothesis. There is no significant difference between the two fertilisers.

Top Tip. In an investigation it always helps to plot a chart of your results as well as carrying out a statistical test. This makes the patterns easier to understand. In this case the statistical conclusion agrees with the observation that the error bars on the chart overlap.

The Paired t -test

In some experimental designs the data are matched, or paired. This can happen when we look at the same locations at different times, or in a “before and after” experiment. In this case we can use a more powerful version of the t -test called the paired t -test, which looks at the differences between each pair. The paired t -test tests the null hypothesis that the mean of the paired differences is zero, in other words that there is no difference between the two groups. Note that this is a different test from the unpaired t -test, so it has a slightly different null hypothesis.

The formula for paired t is:

$$t = \frac{|\bar{d}|\sqrt{n}}{s_d}$$

where

$|\bar{d}|$ = absolute mean of the differences (ignore sign)

n = number of pairs

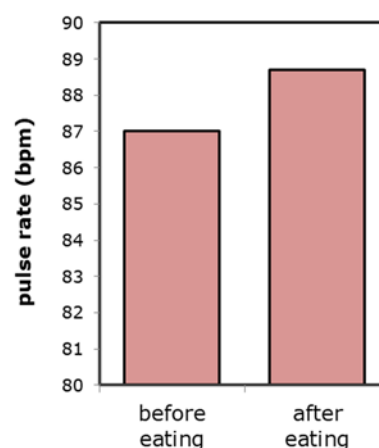
s_d = standard deviation of the differences

Top Tip. Only use the paired t -test if you’re sure that each row in the table is a linked pair. If the experiment still makes sense when you re-order one of the columns then it’s not a paired design. The paired t -test is less common than the unpaired test, so if you’re not sure, just choose the unpaired t -test.

Worked Example

The effect of eating a meal on pulse rate was investigated. Ten subjects had their pulse rates measured before and immediately after eating a meal, with these results:

pulse rate (bpm)			
subject	before eating	after eating	difference
1	105	108	+3
2	83	83	0
3	86	89	+3
4	101	100	-1
5	90	94	+4
6	75	75	0
7	93	96	+3
8	78	77	-1
9	71	76	+5
10	88	89	+1
$\bar{d}$			+1.7
s_d			2.16
n			10



Answer

1. First we choose the test. This is a paired design, so we use the paired t -test.
2. State the null hypothesis we are testing:

The mean of the paired differences is zero
(or there is no difference between the two conditions)

3. Calculate the differences between each pair, and put them in a new column. It doesn't matter what way round you calculate the differences, so long as you do the same for each pair, and you mustn't forget the signs.
4. Calculate the mean of the differences ($\bar{d}$) and put this in a new row at the bottom of the table.
5. Calculate the standard deviation of the differences (s_d) and put this in a new row. It saves time to use the stats function of a calculator for these two steps.
6. Count the number of paired measurements (n), and put that in a new row.
7. Put all the values into the formula to calculate paired t :

$$t = \frac{1.70 \times \sqrt{10}}{2.16} = \frac{5.38}{2.16} = 2.49$$

8. Calculate the degrees of freedom. Since this is a paired test, n is the number of pairs, so $\text{dof} = n - 1 = 10 - 1 = 9$
9. Look up the critical value of t for $\alpha = 0.05$ and $\text{dof} = 9$ (2.26).
10. Compare this critical value of t (2.26) with our calculated value of t (2.49).

$$t_{\text{crit}} (2.26) < t_{\text{calc}} (2.49) \text{ so } P < 0.05$$

11. Now that we have our P -value, we can state the conclusion:

The probability is less than 0.05 that the observed difference between the paired difference and zero is due to chance, so we reject the null hypothesis. There is a significant difference between the pulse rates before and after a meal. The pulse rate after the meal was significantly higher, by 1.7 bpm on average.

Top Tip. If there is a significant difference in a test, don't forget to give the effect size and say which way the difference lies, i.e. which group is larger.

The Spearman Correlation Coefficient

Sometimes we are looking for a correlation between two sets of numeric data (e.g. if this goes up does that go up?). The Spearman rank correlation coefficient (r_s) gives the strength of a correlation, on a scale of -1 to $+1$.

- -1 is a perfect negative correlation (as x goes up y goes down)
- $+1$ is a perfect positive correlation (as x goes up y goes up)
- 0 is no correlation at all

The Spearman rank correlation coefficient tests the null hypothesis that there is no correlation between the two sets (i.e. $r_s = 0$).

The formula for r_s is:

$$r_s = 1 - \frac{6 \times \sum D^2}{n^3 - n}$$

where:

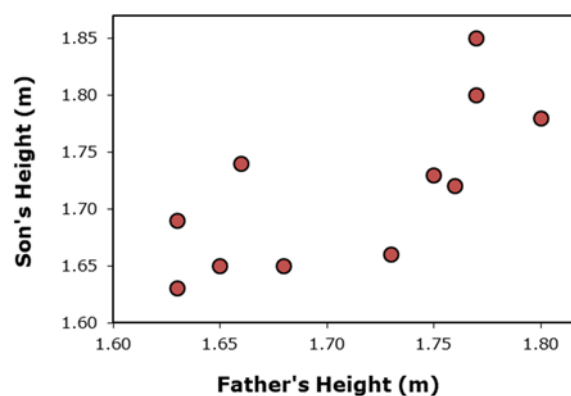
D = difference in ranks

n = number of pairs

Worked Example

A scientist wanted to know if there was a correlation between the heights of men and their sons. The heights from 11 father-son pairs are shown in the first two columns of the table and plotted on a scatter graph.

pair no.	father's height /m	son's height /m	rank of father's height	rank of son's height	(difference in ranks) ²
1	1.76	1.72	4	6	4
2	1.66	1.74	8	4	16
3	1.68	1.65	7	9.5	6.25
4	1.75	1.73	5	5	0
5	1.63	1.63	10.5	11	0.25
6	1.65	1.65	9	9.5	0.25
7	1.77	1.80	2.5	2	0.25
8	1.73	1.66	6	8	4
9	1.80	1.78	1	3	4
10	1.63	1.69	10.5	7	12.25
11	1.77	1.85	2.5	1	2.25
sum					49.5



Answer

1. First we choose the test. We're looking for a correlation between two sets of numeric data and we would plot a scatter graph, so we use the correlation coefficient.
2. State the null hypothesis that we are going to test:

There is no correlation between fathers' and sons' heights.

3. We rank the data in the first column (fathers' heights), writing the ranks in a new column. We start with the biggest value (so 1.80m is ranked 1) and then the second biggest and so on. There is a tie for the next tallest height (1.77m), so we give them both the mean of their ranks (2.5 instead of 2 and 3). The next rank is 4. There is a tie for the shortest height too.
4. We rank the data in the second column (sons' heights) in exactly the same way, writing the ranks in a new column. We can rank the data in either direction (highest to lowest or lowest to highest), so long as we do the same for both columns.
5. For each row we then calculate the difference in each of the paired ranks and square it, writing that value in a new column. For the first row we have $(4 - 6)^2 = 4$.
6. We add up the squared differences (ΣD^2) in this column and write the sum at the bottom of the column.
7. We can now put all the values into the formula to calculate r_s :

$$r_s = 1 - \frac{6 \times 49.5}{1331 - 11} = 1 - \frac{297}{1320} = 0.775$$

8. Look up the critical value of r_s for $\alpha = 0.05$ and $n = 11$ pairs ($= 0.62$).
9. Compare this critical value of r_s (0.62) with the absolute value of our calculated value of r_s (0.775). Although r_s can be positive or negative, we ignore the sign here.

$$r_{\text{crit}} (0.62) < r_{\text{calc}} (0.775) \text{ so } P < 0.05$$

10. Now that we have our P -value, we can state the conclusion:

The probability is less than 0.05 that the observed correlation is just due to chance, so we reject the null hypothesis. There is a significant positive correlation between fathers' and sons' heights with a correlation coefficient of +0.78. Tall fathers generally have tall sons.

Top Tip. If the correlation is significant, say whether it positive or negative and give the strength of the correlation by quoting the correlation coefficient, including the sign.

Top Tip. Remember that a correlation does not necessarily mean that there is a causal relationship between the factors, though in this case there are good genetic reasons to believe that there is.

The Chi squared (χ^2) Test

Sometimes we collect categoric data, i.e. observations using words rather than numbers (e.g. colours, shapes, species). Collecting categoric data is not common in biology as we normally endeavour to make quantitative measurements, which can be analysed in more depth. The most common example of categoric data is from a genetic cross, where offspring are categorised by their phenotype (like "red flower", "wingless" or "female"), but other examples of categoric data include choice chamber experiments (where the categories might be "dark zone" or "light zone" and prey selection (where the categories might be "banded shell" or "pink shell").

If a large number of these categoric observations are made then the number of observations of each category can be counted to give frequencies. The Chi squared (χ^2) test compares the frequencies observed from an experiment with the frequencies expected from a theory. You need to be able to calculate these expected frequencies: either assuming an equal distribution in each category or based on a theory such as Mendel's laws of genetics.

The chi squared test tests the null hypothesis that there is no difference between the observed and expected frequencies. But it's helpful to include what the expected frequencies are based on, so we might state a null hypothesis as: there is no difference between the observed frequencies and an expected ratio of 1:1:1:1.

The formula for χ^2 is:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where:

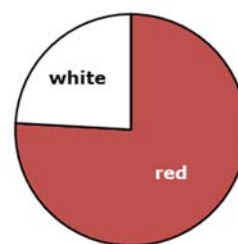
O = observed frequencies

E = expected frequencies

Worked Example

In one of Mendel's experiments with pea plants, he observed the flower colours of 929 offspring from a cross between two heterozygous red-flowered plants and observed 705 red flowers and 224 white flowers. From his theory he expected a ratio of 3 red : 1 white. Do his results agree with this theory?

flower colour	observed frequency (O)	expected frequency (E)	$\frac{(O - E)^2}{E}$
red	705	696.75	0.10
white	224	232.25	0.29
sum	929	929	0.39



Answer

1. First we choose the test. We have observed and expected frequencies of categorical data, so we use the chi squared test.

2. State the null hypothesis that we are going to test:

There is no difference between the observed and expected frequencies,
or there is no difference between the observed frequencies and a ratio of 3:1

3. Add a new column is for "expected frequencies". We want a 3:1 ratio here, so $\frac{3}{4}$ of the 929 plants (696.75) are expected to be red, and $\frac{1}{4}$ are expected to be white (232.25). Note that you can have fractions in the expected frequencies, though not in the observed frequencies. It's worth checking that the expected frequencies add up to the correct total (929 in this case).

4. Add another new column and calculate $(O - E)^2 / E$ for each row.

5. Add up these values at the bottom to give the sum. This is the χ^2 value (0.39).

6. Calculate the degrees of freedom: $\text{dof} = \text{number of categories} - 1 = 2 - 1 = 1$.

7. Look up the critical value of χ^2 for $\alpha = 0.05$ and $\text{dof} = 1$ (= 3.84).

8. Compare this critical value of χ^2 (3.84) with our calculated value of χ^2 (0.39).

$$\chi^2_{\text{crit}} (3.84) > \chi^2_{\text{calc}} (0.39) \text{ so } P > 0.05$$

9. Now that we have our P -value, we can state the conclusion:

The probability is greater than 0.05 that the observed difference between the observed frequencies and a ratio of 3:1 is just due to chance, so we accept the null hypothesis. There is no significant difference between the observed frequencies and a ratio of 3:1, i.e. the observed frequencies of flower colours are consistent with the 3:1 ratio predicted by Mendel's law.

With genetic crosses, agreeing with the predicted ratios (i.e. agreeing with the null hypothesis) is often the desired result (as here), so we may be quite happy with $P > 0.05$. This is different from most experiments where we're hoping for interesting results that disagree with the null hypothesis ($P < 0.05$). But be careful: a high P -value (>0.05) does not provide any support for the null hypothesis, far less prove it; we just haven't disproved the null hypothesis.

Summary of degrees of freedom formulae

Test	t -test	paired t -test	χ^2 test	Spearman
Dof	$(n_1 - 1) + (n_2 - 1)$	$n_{\text{pairs}} - 1$	$n_{\text{categories}} - 1$	tables use n_{pairs}

Top Tip. In an exam, you will always be given the formulae for the test, but not for the degrees of freedom, so you'll have to remember these. And you need to know how to find the correct critical value in a table and how to compare it correctly with your calculated value.

Practice questions (with answers)

SI Units and Prefixes

- Using standard form, how many:
 - mg in an kg? 10^6
 - nmol in a mol? 10^9
 - μm in a Mm? 10^{12}
 - mm^3 in a dm^3 ? 10^6
- Convert 0.2 mm into μm $200 \mu\text{m}$
 - Convert 6000 μm into mm 6 mm
 - Which is bigger, 0.005 mm or 50 μm ? $50 \mu\text{m}$ (0.005 mm is 5 μm)
 - Which is bigger, 0.5 dm^3 or 50 cm^3 ? 0.5 dm^3 (it is 500 cm^3)
- Convert these values to units with more suitable prefixes i.e. with values in the range 1–999. For example $0.0035 \text{ kg} = 3.5 \text{ g}$.
 - 0.002 56 mm $2.56 \mu\text{m}$
 - 5 840 000 mm 5.84 km
 - $2.62 \times 10^{-5} \text{ g}$ $26.2 \mu\text{g}$
 - $1.98 \times 10^7 \text{ ms}$ 330 min or 5.5 h

Ordinary and Standard Form

- Convert these values to metres and write them in standard form.
 - 1 mm $1 \times 10^{-3} \text{ m}$
 - 47 nm $4.7 \times 10^{-8} \text{ m}$
 - 7693 mm $7.693 (\times 10^0) \text{ m}$
 - 181 795 μm $1.817 95 \times 10^{-1} \text{ m}$
- Write the following as decimals, choosing a suitable prefix.
 - $1 \times 10^{-4} \text{ m}$ $100 \mu\text{m}$ or 0.1 mm
 - $3.9 \times 10^5 \text{ m}$ 390 km
 - $2.698 \times 10^{-6} \text{ m}$ $2.698 \mu\text{m}$
 - $4.55 \times 10^2 \text{ m}$ 455 m
- Calculate
 - $(4.8 \times 10^4) \times (9.3 \times 10^6)$ 4.46×10^{11}
 - $(5.6 \times 10^8) \div (2.8 \times 10^2)$ 2.0×10^6
 - $(5.6 \times 10^4) + (1.5 \times 10^5)$ 2.1×10^5
 - $(9.8 \times 10^3) - (1.7 \times 10^4)$ -7.2×10^3

Significant Figures

7. State the number of significant figures in each of the following numbers.
- | | |
|--------------|--------------------------------|
| (a) 1302 | 4 |
| (a) 600 | 1 (or 3 if zeros are measured) |
| (a) 0.005601 | 4 |
| (a) 0.04500 | 4 |
8. Round the following numbers to the specified number of significant figures.
- | | |
|--|----------|
| (a) 1865 to two significant figures | 1900 |
| (b) 0.358 to one significant figure | 0.4 |
| (c) 0.09076 to three significant figures | 0.0908 |
| (d) 0.000 49648 to two significant figures | 0.000 50 |

Magnification

9. (a) If this plant cell is 32 μm long, what is the magnification of the drawing?

$$M = I/T = 56\,000\,\mu\text{m} / 32\,\mu\text{m} = \times 1750$$

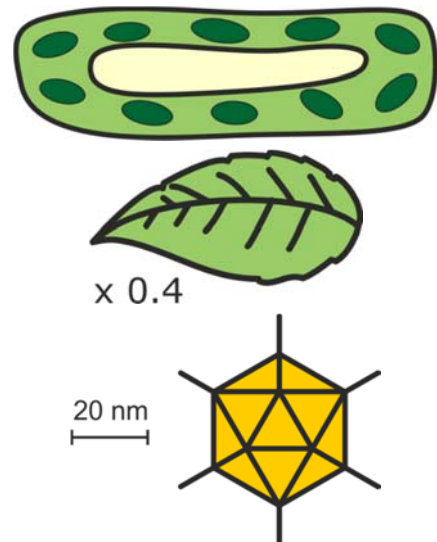
- (b) How long is this leaf in real life?

$$T = I/M = 36\text{mm} / 0.4 = 90\text{mm}$$

- (c) Calculate the true length of this virus.

$$M = \frac{I}{T} = \frac{10\text{ mm}}{20\text{ nm}} = \frac{10\,000\,000\text{ nm}}{20\text{ nm}} = \times 500\,000$$

$$T = \frac{I}{M} = \frac{40\text{ mm}}{500\,000} = \frac{40\,000\,000\text{ nm}}{500\,000} = 80\text{ nm}$$



Using an Eyepiece Graticule

10. A student calibrated a microscope and found that one division of the stage micrometer (0.1 mm) matched up with 21 eyepiece graticule divisions. The student then viewed some plant tissue using the same objective lens and measured one cell as 35 eyepiece divisions long.

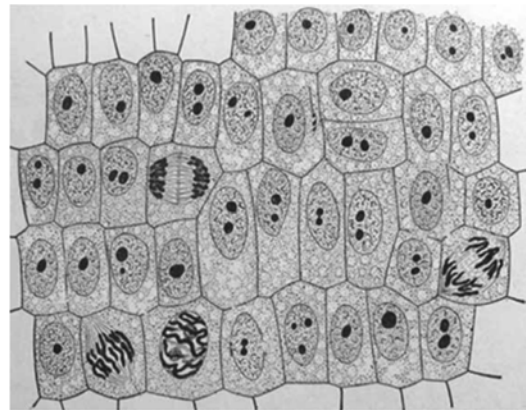
How long is the cell? Give your answer in μm .

calibration: $21\text{ divisions} = 100\,\mu\text{m}$, so $1\text{ division} = \frac{100}{21} = 4.76\,\mu\text{m}$

Cell length: $35\text{ divisions} \times 4.76\,\mu\text{m} = 167\,\mu\text{m}$

Mitotic Index

11. A student made a drawing of tissue from a plant root tip. Some of these cells are dividing.



Calculate the mitotic index from this drawing.

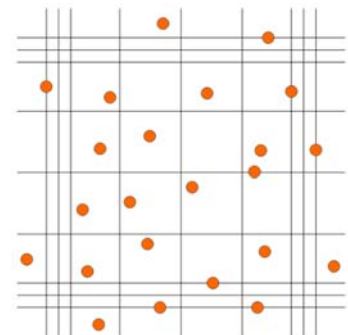
total number of cells = 40

number of cells in mitosis = 4

mitotic index = $4 / 40 \times 100 = 10\%$

Counting Cells

12. The image on the right shows a haemocytometer grid with a sample of yeast cells on it. The culture of yeast cells was diluted 10^5 times before it was placed on the haemocytometer. The middle line of the triple-lined square shown has sides 0.2 mm long and the coverslip is 0.1 mm above the grid.



From this information, calculate the cell density (in cells cm^{-3}) of the undiluted yeast culture.

Cell count = 14 cells

volume = $0.2 \times 0.2 \times 0.1 = 0.004 \text{ mm}^3$

density = $\frac{14}{0.004} = 3.5 \times 10^3 \text{ cells mm}^{-3}$

undiluted suspension = $(3.5 \times 10^3) \times 10^5 = 3.5 \times 10^8 \text{ cells mm}^{-3}$

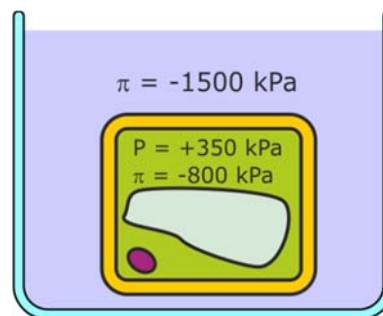
convert to $\text{cm}^3 = (3.5 \times 10^8) \times 10^3 = 3.5 \times 10^{11} \text{ cells cm}^{-3}$

Water Potential

13. A plant cell was immersed in pure water until it was fully turgid. It was then placed in a concentrated salt solution until it was plasmolysed. The table shows some of the values of the potentials in the cell at equilibrium under these two different conditions. Fill in the missing values in the table.

	Ψ (kPa)	P (kPa)	π (kPa)
Fully turgid	0	+300	-300
plasmolysed	-500	0	-500

14. This diagram shows a cell of plant tissue, immersed in a sucrose solution. Some potentials are marked on the diagram.



- (a) Calculate the water potential of the cell.
 (a) $\Psi = 350 - 800 = -450 \text{ kPa}$
- (b) Calculate the water potential of the solution.
 (b) $\Psi = -1500 + 0 = -1500 \text{ kPa}$
- (c) State whether water will enter or leave the cell and explain your answer.
 (c) leave, as solution has lower Ψ .
- (d) Calculate the water potential of this cell when it is plasmolysed.
 (d) $\Psi = 0 + -800 = -800 \text{ kPa}$

Percentage Change

15. Calculate the percentage change if:
- (a) a carrot chip changes from 5 mm long to 4.8 mm long in salt solution.

$$\frac{(4.8-5)}{5} \times 100 = -4\%$$
- (b) a reaction produces 22 mL of gas per min at 20°C and 48 mL of gas per min at 30°C.

$$\frac{(48-22)}{22} \times 100 = +118\%$$
- (c) the number of bacterial cells in a colony changes from 5×10^4 to 2×10^5 after an hour.

$$\frac{(20\,000-5\,000)}{5\,000} \times 100 = +300\%$$
- (d) an athlete's pulse rate changes from 165 to 75 beats per min after a race.

$$\frac{(75-165)}{165} \times 100 = -54.5\%$$

Temperature coefficient (Q_{10})

16. A student investigated the effect of temperature on an enzyme using a colorimeter to measure the rate of a colour change. From her results she calculated the Q_{10} for this enzyme as 2.17.

If the rate at 35°C was 24 absorbance units s^{-1} , what was the rate at 5°C?

$$\text{rate at } 25^\circ\text{C} = 24 / 2.17 = 11.06$$

$$\text{rate at } 15^\circ\text{C} = 11.06 / 2.17 = 5.10$$

$$\text{rate at } 5^\circ\text{C} = 5.10 / 2.17 = 2.3 \text{ absorbance units s}^{-1}$$

Surface Area and Volume

17. A scientist investigated the effect of size on surface area : volume ratio using a mathematical model. Fill in the missing values in this table.

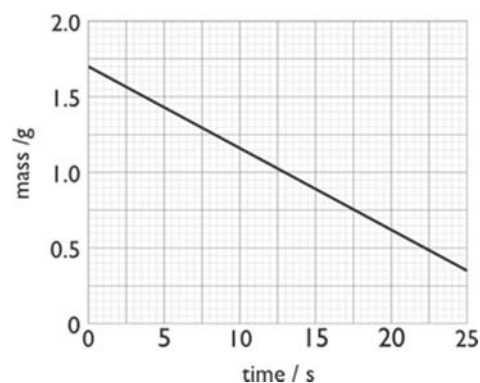
	number of cubes	length of side /mm	total surface area /cm ²	total volume /cm ³	surface area : volume ratio
large cube	1	100	600	1000	0.6 : 1
small cube	10 ⁶	1	60 000	1000	60 : 1

Gradients and Rates

18. The graph shows the change in mass when gas is evolved during a chemical reaction.

Calculate the rate of reaction. Don't forget the units.

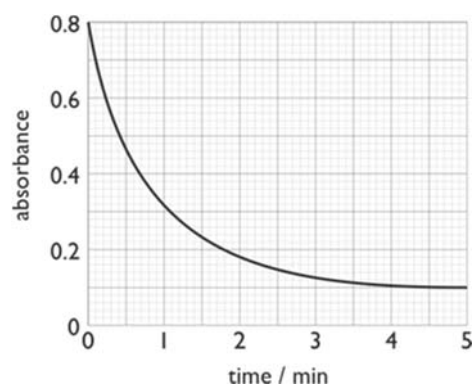
$$\text{rate} = (1.7 - 0.35) / 25 = 0.054 \text{ g s}^{-1}$$



19. The graph shows the change in absorbance when an enzyme is mixed with a coloured substrate.

(a) Calculate the initial rate of reaction over the first 0.2 min.

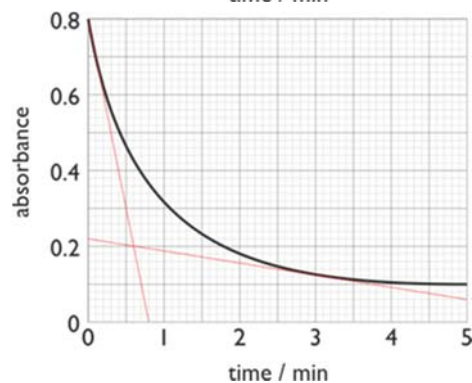
(b) Calculate the rate of reaction at 3 min.



(a) $0.8 / 0.8 = 1 \text{ units min}^{-1}$

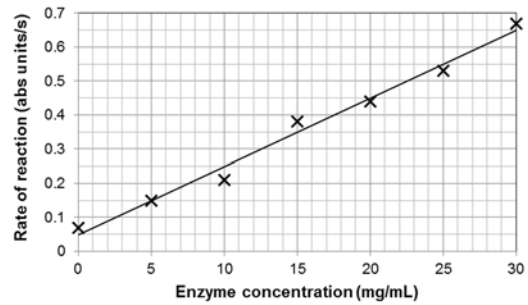
(b) $(0.22 - 0.06) / 5 = 0.032 \text{ units min}^{-1}$

Allow similar values



Equation for a straight line ($y = mx + c$)

20. The graph shows the results of an investigation into the effect of enzyme concentration on the rate of reaction, measured using a colorimeter. A line of best fit has been drawn.



- (a) Using the equation $y = mx + c$, calculate m and c .

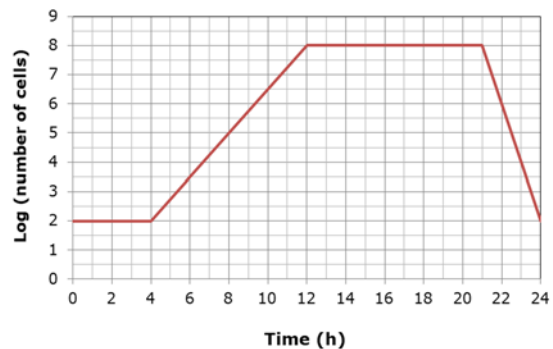
$$c = 0.05 \quad m = (0.65 - 0.05) / 30 = 0.02$$

- (b) Use these values to determine the rate of reaction at an enzyme concentration of 50 mg ml^{-1} .

$$\text{rate} = 0.02 \times 50 + 0.05 = 1.05 \text{ units s}^{-1}$$

Exponential Growth

21. The growth of a bacterial cell culture was monitored every two hours by dilution plating, with the results shown in the graph.



- (a) Use the equation $k = \frac{\log_{10} N_t - \log_{10} N_0}{0.301 \times t}$ to calculate the exponential growth rate constant for this culture.

- (b) What is the generation time in minutes for these bacteria?

(a) At 4h $\log(N) = 2$ and at 12h $\log(N) = 8$, so $k = \frac{8-2}{0.301 \times 8} = 2.49 \text{ h}^{-1}$

(b) $T = \frac{1}{k} = \frac{1}{2.49} = 0.4016 \text{ h} \times 60 = 24 \text{ min}$

The Hardy-Weinberg Equation

22. In fruit flies, wing length is controlled by a single gene. Fruit flies with the homozygous recessive genotype have short wings. In a sample of 350 fruit flies, 73 were found to have short wings. Use the Hardy-Weinberg equation to calculate the expected number of fruit flies in this sample with the homozygous dominant genotype.

$$\text{Phenotype frequency of homozygous recessive} = 73/350 = 0.2086 = q^2$$

$$q = \sqrt{0.2086} = 0.4567$$

$$p = 1 - 0.4567 = 0.5433$$

$$\text{homozygous dominant frequency} = p^2 = 0.5433^2 = 0.2951$$

$$\text{number of flies} = 350 \times 0.2951 = 103.3 = 103 \text{ (rounded down to nearest whole fly)}$$

23. The incidence of the single-gene recessive disorder cystic fibrosis in humans is one in 2500. Use the Hardy-Weinberg equation to show that one in 25 people are heterozygous carriers of the disease allele.

Phenotype frequency of homozygous recessive = $1/2500 = 0.0004 = q^2$

$q = \sqrt{0.0004} = 0.02$

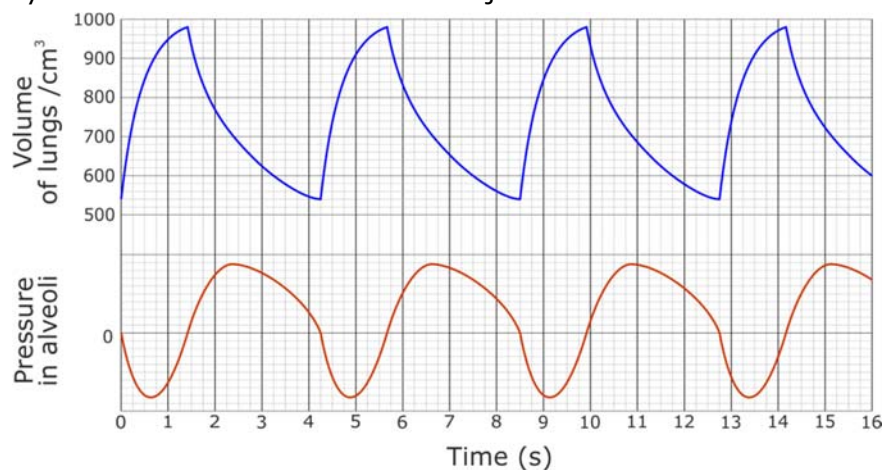
$p = 1 - 0.02 = 0.98$

heterozygous frequency = $2pq = 2 \times 0.98 \times 0.02 = 0.0392$

$0.0392 = 1 / 25.5$, so about 1 in 25 are carriers.

Cardiac Output and Minute Ventilation

24. This spirometer trace shows the lung volume and pressure changes during normal breathing by an adult male. Calculate the subject's minute ventilation in $\text{dm}^3 \text{ min}^{-1}$.

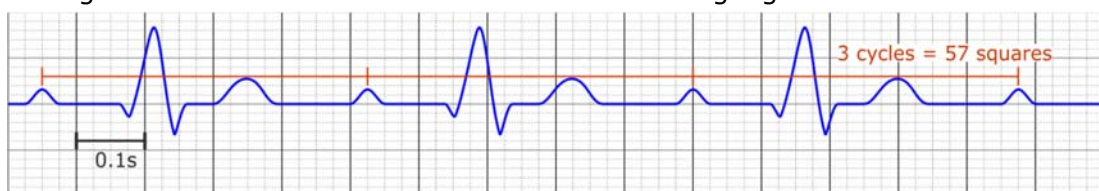


Tidal volume = $980 - 540 = 440 \text{ cm}^3 = 0.44 \text{ dm}^3 \text{ breath}^{-1}$

Cycle time = 4.25 s, so breathing rate = $60 / 4.25 = 14.1 \text{ breaths min}^{-1}$

minute ventilation = $0.44 \times 14.1 = 6.20 \text{ dm}^3 \text{ min}^{-1}$

25. The diagram shows an ECG trace from a student during vigorous exercise.



- (a) Calculate the heart rate of this student during exercise.

(a) 4 small squares = 0.1 s, so 1 small square = 0.025 s

3 cycles in 57 small squares = $57 \times 0.025 = 1.425 \text{ s}$

so one cycle time = $1.425 / 3 = 0.475 \text{ s}$

heart rate = $60 \text{ s} / 0.475 \text{ s} = 126.3 \text{ bpm}$

- (b) If the student's resting heart rate was 80 bpm, what is their exercise heart rate as a percentage of their resting heart rate?

(b) $126.3 / 80 \times 100 = 158\%$

Respirometers and Potometers

26. A potometer was used to investigate the effect of covering different surfaces of leaves with waterproof petroleum jelly on the rate of transpiration. The internal diameter of the capillary tube in the potometer was 0.8 mm.

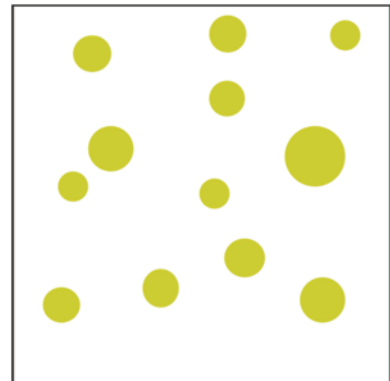
Surface covered with petroleum jelly	leaf area /cm ²	distance moved by bubble in 2 min /mm
upper	28	65
lower	31	8
both	37	2
neither	24	83

- (a) Calculate the rate of transpiration of the leaf with no petroleum jelly. Give the answer in mm³ min⁻¹ cm⁻² of leaf.
- (a) volume in 2 min = $83 \times \pi \times 0.4^2 = 41.7 \text{ mm}^3$
 rate = $41.7 \text{ mm}^3 / 2 \text{ min} = 20.9 \text{ mm}^3 \text{ min}^{-1}$
 rate per leaf area = $20.9 / 24 \text{ cm}^2 = 0.87 \text{ mm}^3 \text{ min}^{-1} \text{ cm}^{-2}$
- (b) Calculate the ratio of the transpiration rate of the leaf with no petroleum jelly to the leaf with the lower surface covered.
- (b) rate of lower surface leaf = $8 \times \pi \times 0.4^2 / (2 \times 31) = 0.065 \text{ mm}^3 \text{ min}^{-1} \text{ cm}^{-2}$
 ratio = $0.87 / 0.065 = 13.4 : 1$

Calculations using Quadrats

27. A student used a 50 cm × 50 cm quadrat to compare the abundance of white clover (*Trifolium repens*) in a trampled and an untrampled area of grassland. Each area measured 90 m × 45 m.

The diagram shows the distribution of white clover plants in one quadrat from the area of trampled grassland. Each circle represents one clover plant.



Use the results from this quadrat to calculate the total number of white clover plants present in the area of trampled grassland.

$$\begin{aligned} \text{Number of plants} &= 12 \\ \text{Density} &= 12 / (0.5 \times 0.5) = 48 \text{ plants m}^{-2} \\ \text{Area} &= 90 \times 45 = 4050 \text{ m}^2 \\ \text{Total in area} &= 48 \times 4050 = 194\,400 \text{ plants} \end{aligned}$$

The Lincoln Index

28. Scientists investigated whether the population malarial mosquitoes changed between the wet and dry seasons in Africa. They used the capture-mark-recapture technique to estimate the populations in the same area in the two seasons of the same year. The results are shown in the table.

Season	Number of mosquitoes trapped and marked on first day	Number of mosquitoes trapped on second day	Number of mosquitoes trapped on second day who were marked
Wet season	175	227	4
Dry season	88	89	4

Using the data provided, state whether the population of mosquitoes was larger in the wet or dry season.

$$\text{population in wet season} = (175 \times 227) \div 4 = 9931$$

$$\text{population in dry season} = (88 \times 89) \div 4 = 1958$$

so population greater in wet season

Simpson's Diversity Index

29. The species diversity of ground vegetation in a woodland was investigated by randomly placing 20 quadrats in the woodland and counting the number of plants found in each quadrat. The results are shown in this table.

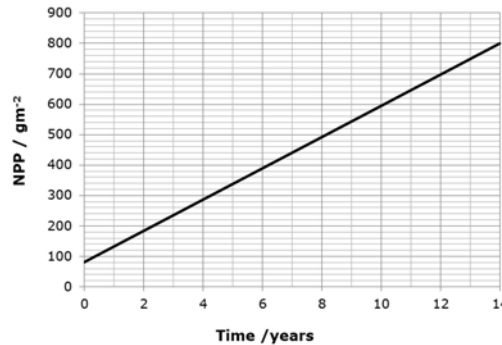
Species	Total number for 20 quadrats	$n(n-1)$
Woodrush	194	37 442
Holly	236	55 460
Bramble	88	7 656
Yorkshire Fog	4	12
sum	522	100 570

Using the formula $D = \frac{N(N-1)}{\sum n(n-1)}$, calculate Simpson's diversity index for this woodland.

$$D = \frac{N(N-1)}{\sum n(n-1)} = \frac{(522 \times 521)}{100\,570} = \frac{271\,962}{100\,570} = 2.70$$

Productivity and Trophic Efficiency

30. When a playing field was fenced off and left untouched for many years, bushes and small trees began to grow. Each year the net primary productivity (NPP) was estimated by measuring the dry biomass of small samples of the vegetation. The change in NPP is shown in this chart.



The average solar radiation on this field was measured as $3000 \text{ MJ m}^{-2} \text{ y}^{-1}$. On average, the energy content of the vegetation was 22 kJ g^{-1} . Calculate the trophic efficiency of the vegetation at converting solar radiation into plant biomass.

$$\text{NPP (as mass)} = (800 - 80) / 14 = 51.43 \text{ g m}^{-2} \text{ y}^{-1}$$

$$\text{NPP (as energy)} = 51.43 \times 22 = 1131.4 \text{ kJ m}^{-2} \text{ y}^{-1} = 1.131 \text{ MJ m}^{-2} \text{ y}^{-1}$$

$$\text{efficiency} = \frac{1.131}{3000} \times 100 = 0.0377\%$$

31. The ratio of NPP : GPP can be used by scientists to study productivity in different habitats. Use the data below to calculate the NPP : GPP ratio for these two habitats.

Grassland: $\text{GPP} = 400 \text{ g m}^{-2} \text{ y}^{-1}$ $\text{R} = 140 \text{ g m}^{-2} \text{ y}^{-1}$

Scrubland: $\text{NPP} = 220 \text{ g m}^{-2} \text{ y}^{-1}$ $\text{R} = 130 \text{ g m}^{-2} \text{ y}^{-1}$

grassland: $\text{NPP} = 400 - 140 = 260 \text{ g m}^{-2} \text{ y}^{-1}$. $\text{NPP:GPP ratio} = \frac{260}{400} = 0.65:1$

scrubland: $\text{GPP} = 220 + 130 = 350 \text{ g m}^{-2} \text{ y}^{-1}$. $\text{NPP:GPP ratio} = \frac{220}{350} = 0.63:1$

Calculations without Formulae

32. (a) The oxygen consumption of a mouse is $200 \text{ cm}^3 \text{ kg}^{-1} \text{ h}^{-1}$. How much oxygen is used up by a 100 g mouse in 24h?

(a) $200 \text{ cm}^3 \text{ kg}^{-1} \text{ h}^{-1} \times 0.1 \text{ kg} \times 24 \text{ h} = 480 \text{ cm}^3$

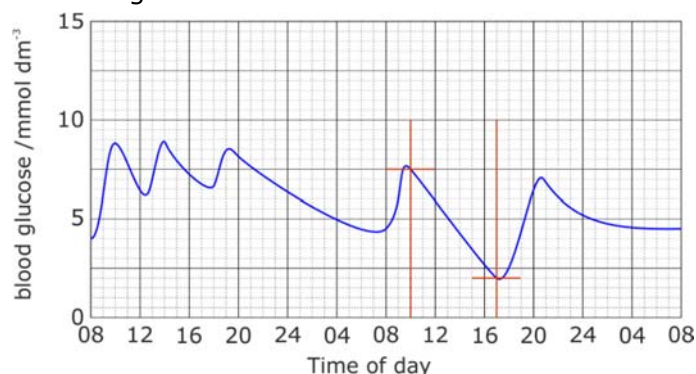
- (b) If air contains 180 cm^3 of oxygen per dm^3 of air, humans can absorb 25% of this oxygen, and 1 dm^3 of air is breathed in each breath, what volume of oxygen will be absorbed in 20 breaths?

(b) $1 \text{ dm}^3 \text{ breath}^{-1} \times 180 \text{ cm}^3 \text{ O}_2 \text{ dm}^{-3} \times 25\% \times 20 \text{ breaths} = 900 \text{ cm}^3$

- (c) DNA is replicated at a rate of $1000 \text{ base pairs s}^{-1}$. How many days will it take to replicate 3×10^9 base pairs?

(c) $\frac{3 \times 10^9}{1 \times 10^3} = 3 \times 10^6 \text{ s} = 833.3 \text{ h} = 34.7 \text{ days}$

33. People with type 1 diabetes have to inject insulin to control their blood glucose levels. Blood glucose levels are normally maintained between 4 and 8 mmol dm⁻³. The graph shows the changes in blood glucose over the course of two days for a person with diabetes being treated with insulin.



If insulin is injected and a meal is missed, blood glucose levels may fall too low. Calculate the change in mass of glucose in 1 dm³ of blood between 10:00 h and 17:00 h on Day 2. A blood glucose level of 4 mmol dm⁻³ is equivalent to 72 mg 100 cm⁻³.

concentration at 10:00 h = 7.5; concentration at 17:00 h = 2.0

so change in concentration = 7.5 – 2.0 = 5.5 mmol dm⁻³

conversion: 100 cm³ = 0.1 dm³, so 72 mg 100 cm⁻³ = 720 mg dm⁻³

so conversion is 4 mmol dm⁻³ = 720 mg dm⁻³

so change in mass = 5.5 mmol dm⁻³ × $\frac{720 \text{ mg dm}^{-3}}{4 \text{ mmol dm}^{-3}}$ = 990 mg dm⁻³

Summary Statistics

34. In a test of two drugs 10 patients were given one drug and 10 patients another drug. The number of hours of relief from symptoms was measured with the results shown.

Calculate the mean and standard deviation of each. You can use this formula or the statistical functions of your calculator.

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Drug A	Drug B
3.2	3.8
2.6	4.0
4.7	7.4
2.8	3.6
4.5	5.0
2.2	3.5
6.1	7.3
3.7	4.9
3.9	7.8
2.9	4.8

Drug A: $\bar{x} = 3.7$ $s = 1.18$

Drug B: $\bar{x} = 5.2$ $s = 1.67$

35. The flow rate of a river was measured at ten stations on two different days.

Calculate the median and range for each day.

Day 1 Median = 265.5 Range = 44
Day 2 Median = 240 Range = 34

Station	Day 1	Day 2
1	268	236
2	260	231
3	243	239
4	280	265
5	284	252
6	270	253
7	263	241
8	254	234
9	240	236
10	269	249

36. Find the mode for each of the following data sets:

(a) Seed masses (g) 4, 3, 3, 3, 3, 2, 2, 4, 3, 3, 2, 4, 3, 3, 2, 3, 2, 4, 3, 3, 2, 3, 2, 4, 3, 5

3 g

(b) Blood groups A, B, AB, AB, O, A, B, AB, O

AB

(c) DNA fragment lengths (kb) 53, 43, 39, 76, 11, 93

There isn't one

37. An investigation into plant heights obtained the results shown.

Condition	mean plant height $\pm$ standard deviation (m)
shade	0.36 $\pm$ 0.08
semi-shade	0.47 $\pm$ 0.10
open	0.77 $\pm$ 0.14

What conclusions can be made from these results?

- There is no significant difference between the shade and semi-shade heights, because the standard deviations overlap.
- The open plants are significantly taller than the others because the standard deviations don't overlap.

38. An investigation was carried out to compare the size of eggs laid by one species of lizard. The length of 80 eggs was measured with the results shown in this tally chart.

Egg length / mm	Frequency
26	///
28	### ## ///
30	### ## ## ##
32	### ## ## ///
34	### ## ///
36	### ///
38	///

Using the information provided, calculate the difference between the mode and the mean. Show your working.

$$\text{mode} = 30$$

$$\text{mean} = \frac{(26 \times 3) + (28 \times 14) + (30 \times 20) + (32 \times 19) + (34 \times 13) + (36 \times 8) + (38 \times 3)}{3 + 14 + 20 + 19 + 13 + 8 + 3} = \frac{2522}{80} = 31.5$$

$$\text{difference} = 31.5 - 30 = 1.5$$

Statistical Tests

39. A statistical test on a computer gives a result of $P = 0.002$.

- (a) P is a probability value. What is it the probability of?

P is the probability that our results are consistent with the null hypothesis and that any observed effect is just due to chance.

- (b) What can you conclude from this value of 0.002?

The probability is 0.002 that the observed difference or correlation is due to chance. Since this is < 0.05 , we reject the null hypothesis and conclude that there is a significant effect.

40. You are going to investigate the effects of two different drugs on blood pressure.

- (a) What kind of chart would you plot of the results?

Bar chart (with error bars)

- (b) Which statistical test would you use to analyse the results, and why?

The t-test, because we are comparing two sets of numeric data.

- (c) State the null hypothesis for this investigation.

There is no difference between the effects of two different drugs on mean blood pressure.

- (d) The critical value of t at $\alpha = 0.05$ is 1.99. If the calculated value of t was 2.31, what is your conclusion?

$$t_{\text{crit}} (1.99) < t_{\text{calc}} (2.31) \text{ so } P < 0.05$$

The probability is < 0.05 that the observed difference is due to chance. We reject the null hypothesis and conclude that there is a significant difference in the effect of the two drugs.

41. An investigation was carried out into the relationship between length and mass in a species of tuna caught in the Pacific Ocean. Twelve specimens were caught and the length and mass of each fish was recorded. The table below shows the results of this investigation.

Fish number	1	2	3	4	5	6	7	8	9	10	11	12
Length / cm	100	112	130	143	103	98	93	118	140	138	122	82
Mass / kg	56	45	62	50	27	16	32	54	80	60	38	20
rank length	9	7	4	1	8	10	11	6	2	3	5	12
rank mass	4	7	2	6	10	12	9	5	1	3	8	11
difference ²	25	0	4	25	4	4	4	1	1	0	9	1
sum of diff ²												78

- (a) What kind of chart would be best for displaying these results?
scatter graph
- (b) Which statistical test would be best for analysing these results, and why?
Spearman correlation coefficient, because we are investigating a correlation between two sets of numeric data.
- (c) State the null hypothesis for this investigation.
There is no correlation between the length and mass of the tuna fish.
- (d) Calculate the test statistic.
 $\Sigma D^2 = 78$
 $r_s = 1 - \frac{6 \times 78}{1728 - 12} = 1 - \frac{468}{1716} = 0.727$
- (e) Find the critical value in the tables for $\alpha = 0.05$ on p38.
 $n = 12$ $r_{crit} = 0.59$
- (f) Use your values to write a conclusion.
 $r_{crit} (0.59) < r_{calc} (0.727)$ so $P < 0.05$
 The probability is < 0.05 that the observed correlation is due to chance. We reject the null hypothesis and conclude that there is a significant positive correlation between the length and mass of the tuna fish, with a correlation coefficient of +0.73.

42. Dog's mercury is a plant that grows in woodland and open clearings. An investigation was carried out to determine whether the leaves from dog's mercury plants were different sizes in shaded areas and open clearings of a wood. 12 leaves of dog's mercury were collected at random from plants growing in a shaded area and 12 from plants growing in an open clearing, and the surface area of each leaf was measured. The table shows the results of this investigation.

	Surface area of leaves / cm ²	
	Shaded area	Open clearing
1	21	15
2	14	17
3	16	18
4	18	17
5	19	17
6	21	19
7	19	13
8	22	14
9	18	21
10	16	13
11	13	16
Mean	17.91	16.36
<i>s</i>	2.91	2.5
<i>n</i>	11	11

- (a) Calculate the mean, standard deviation and *n* for the open clearing.
see table
- (b) Which statistical test would be best for analysing these results, and why?
t-test, because we are comparing two sets of numeric data.
- (c) State the null hypothesis for this investigation.
There is no difference between the mean leaf areas in the shaded and open woodland.

- (d) Calculate the test statistic.

$$t = \frac{|17.91 - 16.36|}{\sqrt{\frac{2.91^2}{11} + \frac{2.50^2}{11}}} = \frac{1.55}{\sqrt{0.77 + 0.57}} = \frac{1.55}{1.16} = 1.34$$

- (e) Calculate the degrees of freedom and find the critical value in the tables on p38.
dof = 10 + 10 = 20 $t_{crit} = 2.09$
- (f) Use your values to write a conclusion.

$$t_{crit} (2.09) > t_{calc} (1.34) \text{ so } P > 0.05$$

The probability is > 0.05 that the observed difference is due to chance. We accept the null hypothesis and conclude that there is no significant difference in leaf area in the two woodlands.

43. A test cross was carried out between two maize plants, one with yellow, smooth cobs and the other with white, wrinkled cobs. The table shows the frequencies of each phenotype in the offspring of this cross.

Type of maize cob	smooth yellow	wrinkled yellow	smooth white	wrinkled white	total
Number of cobs	286	333	301	349	1269
Expected frequency	317.25	317.25	317.25	317.25	1269
$(O - E)^2 / E$	3.08	0.78	0.83	3.18	7.87

Mendel's law predicted that there should have been a 1 : 1 : 1 : 1 ratio of the four phenotypes.

- (a) Which statistical test would be best for analysing these results, and why?

Chi squared test, because these are frequencies of categorical data.

- (b) State the null hypothesis for this investigation.

There is no difference between the observed and expected frequencies, or the observed frequencies agree with an expected ratio of 1 : 1 : 1 : 1.

- (c) Calculate the test statistic.

$$\chi^2 = 7.87$$

- (d) Calculate the degrees of freedom and find the critical value in the tables on p38.

$$\text{dof} = 3$$

$$\chi^2_{\text{crit}} = 7.81$$

- (e) Use your values to write a conclusion.

$$\chi^2_{\text{crit}} (7.81) < \chi^2_{\text{calc}} (7.87) \text{ so } P < 0.05$$

The probability is < 0.05 that the difference between the observed and expected frequencies is due to chance. We reject the null hypothesis and conclude that there is a significant difference between the observed frequencies and a ratio of 1 : 1 : 1 : 1. So these results do not agree with Mendel's law.

44. A group of nine athletes (A to I) wanted to see if training for two weeks at a mountain camp 2000 m above sea level had an effect on the number of red blood cells in their blood. Samples of blood were taken from each of the athletes at their normal training camp at sea level. Blood samples were taken again after two weeks of training at the mountain camp. The following table shows the results.

	Number of red blood cells $\times 10^{12}$ per dm^3 of blood		
Athlete	Before training	After training	difference
A	5.0	4.9	0.1
B	5.1	5.3	-0.2
C	4.9	5.7	-0.8
D	5.3	5.5	-0.2
E	5.4	5.6	-0.2
F	5.0	5.4	-0.4
G	4.8	5.3	-0.5
H	5.1	5.6	-0.5
I	5.5	5.1	0.4
mean	5.1	5.4	-0.26
<i>s</i>			0.35
<i>n</i>			9

- (a) Calculate the percentage increase in the number of red blood cells after mountain training.
 $(5.4 - 5.1) / 5.1 \times 100 = 5.9\%$
- (b) Which statistical test would be best for analysing these results, and why?
 Paired t-test, because we are comparing two sets of numeric data and this is a paired design with the same subjects before and after training.
- (c) State the null hypothesis for this investigation.
 The mean of the paired differences is zero,
 or there is no difference in red blood cell count between the two groups.
- (d) Calculate the test statistic.

$$t = \frac{0.26 \times \sqrt{9}}{0.35} = \frac{0.78}{0.35} = 2.23$$
- (e) Calculate the degrees of freedom and find the critical value in the tables on p38.
 $\text{dof} = 9 - 1 = 8$ $t_{\text{crit}} = 2.31$
- (f) Use your values to write a conclusion.
 $t_{\text{crit}} (2.31) > t_{\text{calc}} (2.23)$ so $P > 0.05$
 The probability is > 0.05 that the observed difference is due to chance. We accept the null hypothesis and conclude that there is no significant difference in red blood cell count before and after training at altitude.

Appendices

Appendix 1 – Using a Calculator in Exams

Many scientific calculators have a stats mode, where some of the statistics calculations can be done more easily. These calculators are allowed in exams, so students can make their lives easier by learning how to use the stats functions of their own calculators.

These instructions show how to calculate mean and standard deviation on the Casio GT Plus and EX (classwiz) calculators, two of the most popular school calculator models.

	Casio GT PLUS / ES PLUS	Casio EX (Classwiz)
To enter stats mode	MODE 2	MENU 6
To choose single-variable mode	1	1
To enter data	Type in each value followed by = Press AC when finished	
To display stats menu	SHIFT 1	OPTN
To display statistical results	4 for variance menu, then 2 = for mean ($\bar{x}$) 4 = for stand. dev. (s_x)	2 for 1-variable calc. Shows all results on scrolling screen

Note that in biology we always want the sample standard deviation (s_x), not the population standard deviation (σ_x), because we are always dealing with samples and never the whole population.

The Classwiz also calculates median and other useful statistics. Both calculators can calculate a correlation coefficient in regression mode. But it is a different correlation coefficient (Pearson's rather than Spearman's), so you can't use it in biology exams, which require you to calculate Spearman's correlation coefficient.

Top tip. Take time to learn how to use all functions of your calculator before your exams.

Appendix 2 – Help from the Field Studies Council

The Field Studies Council (FSC) is an environmental education charity providing field trips and ecology resources for schools. Their [Biology fieldwork website](#) has good information and resources for fieldwork at both GCSE and A level. In this site there is a [maths and statistics page](#), which contains advice on statistical tests and other maths relevant to ecology.

Maths Skills for Biologists Fold-Out Guide

The FSC has also produced a “Maths Skills for Biologists” fold-out guide. This guide contains lots of useful information on maths and statistics for ecology, presented in a compact and colourful way. The guide can be downloaded for free from [FSC](#) or from [Edexcel](#).

Printed copies are also available to [buy from the FSC](#), priced at £4 each, or £60 for a set of 20 (June 2019 prices).

Appendix 3 – Ofqual's Mathematical Requirements

These are the *mathematical requirements and exemplifications*, from Appendix 6 of Ofqual's *GCE Subject Level Conditions and Requirements for Biology* document (April 2014). They outline all the maths requirements for A-Level Biology for all the exam awarding bodies. The list here has been re-arranged into topics to make it easier to follow. Items in bold are for A level only; the rest are for A level and AS level.

Basic Maths	
Units	Recognise and make use of appropriate units in calculations. Convert between units, e.g. mm^3 to cm^3 as part of volumetric calculations.
Algebra	Understand and use the symbols: =, <, <<, >>, >, $\propto$, ~. Change the subject of an equation. Substitute numerical values into algebraic equations using appropriate units for physical quantities. Solve algebraic equations.
Standard form	Recognise and use expressions in decimal and standard form. Use an appropriate number of decimal places in calculations, e.g. for a mean. Carry out calculations using numbers in standard and ordinary form, e.g. use of magnification. Understand standard form when applied to areas such as size of organelles. Convert between numbers in standard and ordinary form.
Significant figures	Use an appropriate number of significant figures. Report calculations to an appropriate number of significant figures given raw data quoted to varying numbers of significant figures. Understand that calculated results can only be reported to the limits of the least accurate measurement. Understand that significant figures need retaining when making conversions between standard and ordinary form, e.g. $0.0050 \text{ mol dm}^{-3}$ is equivalent to $5.0 \times 10^{-3} \text{ mol dm}^{-3}$.
Sampling	Understand the principles of sampling as applied to scientific data. Analyse random data collected by an appropriate means.
Estimating	Estimate results to sense check that the calculated values are appropriate.
Formulae	
Percent	Use percentages and calculate percentage yields.
Ratios	Use ratios and fractions.
SA : Vol	Calculate the circumferences, surface areas and volumes of regular shapes e.g. the circumference and area of a circle; the surface area and volume of rectangular prisms, of cylindrical prisms and of spheres; the surface area or volume of a cell. Calculate surface area to volume ratio.
Magnification	Make order of magnitude calculations e.g. use and manipulate the magnification formula.
Rate	Calculate rate of change from a graph showing a linear relationship e.g. rate of transpiration. Draw and use the slope of a tangent to a curve as a measure of rate of change e.g. amount of product formed plotted against time when the concentration of enzyme is fixed. Work out the unit for a rate e.g. breathing rate.
Cardiac output	cardiac output = stroke volume $\times$ heart rate.
Diversity	Use Simpson's index of diversity to calculate the biodiversity of a habitat. $[D = 1 - (\Sigma(n/N)^2)]$.

Biology Teacher Maths Guide

Genetics Represent phenotypic ratios (monohybrid and dihybrid crosses). Understand the probability associated with genetic inheritance.

Logs Use logarithms in relation to quantities that range over several orders of magnitude e.g. growth rate of a microorganism such as yeast. Use calculators to find and use power, exponential and logarithmic functions e.g. estimate the number of bacteria grown over a certain length of time.

Tables and Charts

Tables Represent a range of data in a table with clear headings, units and consistent decimal places. Interpret data from a variety of tables, e.g. data relating to organ function.

Understand charts Translate information between graphical, numerical and algebraic forms understand that data may be presented in a number of formats and be able to use these data, e.g. dissociation curves.

Selecting charts Plot two variables from experimental or other data. Select an appropriate format for presenting data, bar charts, histograms, graphs and scattergrams.

Histograms Construct and interpret frequency tables and diagrams. Construct and interpret histograms.

Bar charts Construct and interpret bar charts.

Line graphs Plot a range of data in an appropriate format, e.g. enzyme activity over time represented on a graph. Interpret data for a variety of graphs, e.g. explain electrocardiogram traces.

Straight line equation Understand that $y = mx + c$ represents a linear relationship. Predict/sketch the shape of a graph with a linear relationship, e.g. the effect of substrate concentration on the rate of an enzyme-controlled reaction with excess enzyme.

Intercept Determine the intercept of a graph e.g. compensation point in plants.

Correlation Use a scatter diagram to identify a correlation between two variables e.g. the effect of life style factors on health.

Summary Statistics

Mean Understand the terms mean, median and mode e.g. height/mass/size of a group of organisms. Find arithmetic means e.g. the mean number of stomata in the leaves of a plant.

Dispersion Understand measures of dispersion, including standard deviation and range, including calculate the standard deviation and understand why standard deviation might be a more useful measure of dispersion for a given set of data e.g. where there is an outlying result.

Uncertainty Identify uncertainties in measurements and use simple techniques to determine uncertainty when data are combined e.g. calculate percentage error where there are uncertainties in measurement.

Statistical Tests

Probability Understand simple probability. Use the terms probability and chance appropriately.

Stats Select and use a statistical test e.g. the chi squared test to test the significance of the difference between observed and expected results; the Student's t-test; the correlation coefficient.

