

AS and A LEVEL

Mathematical Skills Handbook

BIOLOGY A BIOLOGY B (ADVANCING BIOLOGY)

H020, H420, H421, H421

For first teaching in 2015

This Mathematical Skills Handbook is designed to accompany the OCR Advanced Subsidiary GCE and Advanced GCE specifications in Biology A and Biology B (Advancing Biology) for teaching from September 2015.

Version 1.3

Contents

OCR will update this document on a regular basis. Please check the OCR website (www.ocr.org.uk) at the start of the academic year to ensure that you are using the latest version.

Version 1.3 – February 2018

Version 1.3

Changes of note made between Version 1.2 and Version 1.3:

1. Change in M4.1 to the geometrical formulae that need to be recalled. Formulae for calculating surface area and volume of spheres and cylinders will now be provided in assessments where needed.
2. Change to Appendix A to provide up to date list of formulae that must be recalled.
3. Change to Appendix B to provide up to date list of formulae that will be provided where needed in assessments.
4. Removal of unnecessary text in M1.7 to make it clearer and simpler.
5. Removal of unnecessary text in M1.10 to make it clearer and simpler.

Version 1.2 - June 2016

Version 1.2

Changes of note made between Version 1.1 and Version 1.2:

1. Alternative method for calculating percentage change added to M0.3.
2. Clarification about what is, and is not, required of students added to the statistics section M1.9
3. Spearman's rank correlation example corrected to refer to n (number of paired data items) rather than degrees of freedom in section M1.9
4. Correction (from 66% to 68%) in the statement about the part of a normal distribution lying within one standard deviation of the mean in section M1.10
5. Species evenness formula removed from the formula appendix
6. Percentage change formula corrected to match the alternative method described in section M0.3

Version 1.1 - March 2016

Version 1.1

Changes of note made between Version 1.0 and Version 1.2:

1. The chi squared example revised to use equal expected frequencies for the five species in section M1.9
2. Rate of change from a linear graph has been revised to remove the 'rise over run' terminology

1 Introduction	6
Definition of Level 2 mathematics	6
M0 – Arithmetic and numerical computation	7
M0.1 Recognise and make use of appropriate units in calculations	7
Mathematical concepts	7
M0.2 Recognise and use expressions in decimal and standard form	9
Mathematical concepts	9
Contexts in biology	11
M0.3 Use ratios, fractions and percentages	12
Mathematical concepts	12
Contexts in biology	13
M0.4 Estimate Results	14
Mathematical concepts	14
Contexts in biology	14
M0.5 Use calculators to find and use power, exponential and logarithm functions A Level only	16
Mathematical concepts	16
Contexts in biology	16
M1 – Handling data	17
M1.1 Use an appropriate number of significant figures	17
Mathematical concepts	17
Contexts in Biology	17
M1.2 Find arithmetic means	18
Mathematical concepts	18
Contexts in Biology	19
M1.3 Construct and interpret frequency tables and diagrams, bar charts and histograms	20
Mathematical concepts	20
Contexts in Biology	20
M1.4 Understand simple probability	22
Mathematical concepts	22
Contexts in Biology	22
M1.5 Understand the principles of sampling as applied to scientific data	24
Mathematical concepts	24
Contexts in Biology	24
M1.6 Understand the terms mean, median and mode	25
Mathematical concepts	25
Contexts in Biology	25
M1.7 Use a scatter diagram to identify a correlation between two variables	26
Mathematical concepts	26
Contexts in Biology	27
M1.8 Make order of magnitude calculations	28
Mathematical concepts	28
Contexts in Biology	28
M1.9 Select and use a statistical test	29
Mathematical concepts	29
Contexts in Biology	29

M1.10 Understand measures of dispersion, including standard deviation and range	35
Mathematical concepts	35
Contexts in Biology	35
M1.11 Identify uncertainties in measurements and use simple techniques to determine uncertainty when data are combined	37
Mathematical concepts	37
Contexts in Biology	37
M2 – Algebra	38
M2.1 Understand and use the symbols: =, <, <<, >>, >, ∞, ~	38
Mathematical concepts	38
Contexts in biology	38
M2.2 Change the subject of an equation	39
Mathematical concepts	39
This mathematical principle should be familiar to the learners.	39
Contexts in Biology	39
M2.3 Substitute numerical values into algebraic equations using appropriate units for physical quantities	40
Mathematical concepts	40
Contexts in Biology	41
M2.4 Solve algebraic equations	42
Mathematical concepts	42
Contexts in Biology	42
M2.5 Use logarithms in relation to quantities that range over several orders of magnitude A Level only	43
Mathematical concepts	43
Contexts in Biology	43
M3 – Graphs	45
M3.1 Translate information between graphical, numerical and algebraic forms	45
Mathematical concepts	45
Contexts in Biology	47
M3.2 Plot two variables from experimental or other data	48
Mathematical concepts	48
Please see the Practical Skills Handbook Appendix 6 for more on the topic of graphs including the use of error bars and range bars.	48
Contexts in Biology	48
M3.3 Understand that $y = mx + c$ represents a linear relationship	51
Mathematical concepts	51
Contexts in Biology	51
M3.4 Determine the intercept of a graph A Level only	53
Mathematical concepts	53
Contexts in Biology	53
M3.5 Calculate rate of change from a graph showing a linear relationship	54
Mathematical concepts	54
Contexts in Biology	54
M3.6 Draw and use the slope of a tangent to a curve as a measure of rate of change	56
Mathematical concepts	56
Contexts in Biology	56

M4 – Geometry and trigonometry	58
M4.1 Calculate the circumferences, surface areas and volumes of regular shapes	58
Mathematical concepts	58
Appendix A – Formulae learners need to be able to recall	60
GCSE (9–1) Mathematical formulae to recall	60
Biological formulae to recall	60
Appendix B – Formulae that will be provided in the assessments where needed	62
Mathematical formulae that will need to be used but not recalled (provided in the assessment where needed).	62
Biological formulae that will need to be used but not recalled (provided in the assessment where needed).	63
Appendix C – Key power laws	64
Appendix D – Statistical Tables	65

1 Introduction

In order to be able to develop their skills, knowledge and understanding in AS and A Level Biology, learners need to have been taught, and to have acquired competence in, the appropriate areas of mathematics relevant to the subject as indicated in Appendix 5e of the specifications:

- **H020/H420 - OCR Biology A**
- **H022/H422 - OCR Biology B (Advancing Biology)**

The assessment of all AS and A Level Biology qualifications will now include at least 10% Level 2 (or above) mathematical skills as agreed by Ofqual (see below for a definition of 'Level 2' mathematics). These skills will be applied in the context of the relevant biology.

This Handbook is intended as a resource for teachers, to clarify the nature of the mathematical skills required by the specifications, and indicate how each skill is relevant to the subject content of the specifications.

The content of this Handbook follows the structure of the Mathematical Requirements table in Appendix 5d of the specifications, with each mathematical skill, *M0.1 – M4.1*, discussed in turn. The discussion of each skill begins with a description and explanation of the mathematical concepts, followed by a demonstration of the key areas of the biological content in which the skill may be applied. Notes on common difficulties and misconceptions, as well as suggestions for teaching, may be included in each section as appropriate.

As this Handbook shows, all required mathematical skills for biology can be covered along with the subject content in an integrated fashion. However, as assessment of the mathematical skills makes up at least 10% of the overall assessment, OCR recommends that teachers aim to specifically assess learners' understanding and application of the mathematical concepts as a matter of course, in order to discover and address any difficulties that they may have. This is particularly relevant for learners who are not taking an AS or A Level Mathematics qualification alongside AS or A Level Biology.

Definition of Level 2 mathematics

Within AS or A Level Biology, 10% of the marks available within the written examinations will be for assessment of mathematics (in the context of biology) at a Level 2 standard, or higher. Lower level mathematical skills will still be assessed within examination papers, but will not count within the 10% weighting for biology.

The following will be counted as Level 2 (or higher) mathematics:

- application and understanding requiring choice of data or equation to be used
- problem solving involving use of mathematics from different areas of maths and decisions about direction to proceed
- questions involving use of A Level mathematical content (as of 2012) e.g. use of logarithmic equations.

The following will not be counted as Level 2 mathematics:

- simple substitution with little choice of equation or data and/or structured question formats using GCSE mathematics (based on 2012 GCSE mathematics content).

As lower level mathematical skills are assessed in addition to the 10% weighting for Level 2 and higher, the overall assessment of mathematical skills will form greater than 10% of the assessment.

M0 – Arithmetic and numerical computation

M0.1 Recognise and make use of appropriate units in calculations

Learners may be tested on their ability to:

- convert between units e.g. mm^3 to cm^3 as part of volumetric calculations
- work out the unit for a rate e.g. breathing rate.

Mathematical concepts

Units indicate what a given quantity is measured in. A measured quantity without units is meaningless, although note that there are some derived quantities in biology that do not have units, for example pH.

At GCSE, learners will have used different units of measurement and would be required to recognise appropriate units for common quantities. For example, whilst cm is appropriate for a length or distance, learners should be able to identify that cm^2 is used for area and cm^3 is used for volume. Learners should be aware that in biology pipettes usually measure volume in ml or μl and should be able to readily convert between ml and cm^3 , i.e. $1 \text{ ml} = 1 \text{ cm}^3$. However, learners should be aware that volume will usually be stated in cm^3 in assessments.

Learners will be expected to be able to convert between different metric units, for example $10 \text{ mm} = 1 \text{ cm}$, without conversion 'facts' being given (i.e. $1 \text{ m} = 1000 \text{ mm}$). Converting between different multiples is a matter of multiplying by the appropriate factor. For example, converting 7 nm to mm requires a *division* by 10^6 , *not* a multiplication. It is a common misconception for learners to believe that because millimetres are 'larger' than nanometres (in the sense that 1 mm is larger than 1 nm) that a multiplication is necessary to go to the larger unit. However, a simple check should reveal that 7 000 000 mm is not equal to 7 nm, and so a division is required.

Typical measures that would have been encountered at GCSE and would come up within AS/A Level Biology are distance, area, volume, density and mass.

Unit prefixes indicate particular multiples and fractions of units. A full list of SI unit prefixes is given in Table 1, with the prefixes that are most likely to be used within the AS/A Level Biology course highlighted.

Factor	Name	Symbol	Factor	Name	Symbol
10^{24}	yotta	Y	10^{-1}	deci	d
10^{21}	zeta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	μ
10^{12}	tera	T	10^{-9}	nano	n
10^9	giga	G	10^{-12}	pico	p
10^6	mega	M	10^{-15}	femto	f
10^3	kilo	k	10^{-18}	atto	a
10^2	hecto	h	10^{-21}	zepto	z
10^1	deca	da	10^{-24}	yocto	y

Table 1: SI unit prefixes

Rates of change would have also already been encountered at GCSE, for example speed in m s^{-1} . At AS and A Level, learners should be able to work out the appropriate unit for a rate, e.g.

{quantity of 'stuff' they are measuring}/{Unit of Time}.

A common biological example is to work out the rate of an enzyme-controlled reaction. For example, in an assay using catalase, in which a volume of oxygen (cm^3) is being measured over time (s), the rate of reaction would be in $\text{cm}^3 \text{s}^{-1}$ as the volume evolved is measured against time.

In the AS and A Level Biology assessments, learners will be expected to recognise and use compound units in the form $\text{cm}^3 \text{s}^{-1}$, rather than cm^3/s (see Appendix C).

Other examples of common rates of change in biology include:

- bacterial growth rates (number of bacteria hour^{-1})
- breathing rate (breaths min^{-1})
- heart/pulse rate (beats min^{-1} , bpm)
- a temperature change over time ($^{\circ}\text{C s}^{-1}$)

Within the OCR GCE Biology qualifications, learners will in general be expected to use and recognise standard SI units. For example, dm^3 is used rather than *l* (litre). However, there are exceptions to this, e.g. degree ($^{\circ}$) for angles, which is used in preference to the radian, minutes (min) and hours (h) in addition to seconds (s), *ml* and μl in pipette use (as discussed above). In general, any other conversion to or from non-standard units that may be required in assessment would be provided in the question.

Contexts in biology

When making measurements in an experiment, biology learners will have to decide on the appropriate units of measure to use, and the unit symbol, for example common biological quantities include,

- **distance: m, cm, mm, μm , nm**
- **area: m^2 , cm^2 , mm^2**
- **volume: dm^3 , cm^3 (*ml*), mm^3 (μl)**
- **density: mg cm^{-3} (mg m^{-3})**
- **mass: g, mg**
- **time: s, min, h**
- **temperature: $^{\circ}\text{C}$**
- **light intensity: lux**
- **microbiology: CFU, PFU**
- **heart rate: bpm**

When approximating gradients on graphs to measure rates of change learners will have to decide what the correct units of measurement are for the rate of change i.e. {Quantity of 'stuff' they are measuring}/{Unit of Time}.

Learners will have to be aware that the choice and correct use of units can have drastic changes on the output answer. For example, in measuring a 10 cm by 10 cm quadrat for ecological analysis, if learners decide to use cm then they could calculate the area of the quadrat in cm^2 as $10 \times 10 = 100 \text{ cm}^2$. Issues can arise if a learner then wishes to use 1 cm = 10 mm to convert this area into millimetres. A common misconception is that if 1 cm is 10 mm then 100 cm^2 is 1000 mm^2 . Actually, $10\text{cm} = 100\text{mm}$ so the correct area in millimetres is $100 \times 100 = 10\,000 \text{ mm}^2$. The difference in these answers is a factor of 10 and can lead to massive calculation errors.

M0.2 Recognise and use expressions in decimal and standard form

Learners may be tested on their ability to:

- 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
- 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 $5.0 \times 10^{-3} \text{ mol dm}^{-3}$.

Mathematical concepts

Standard Form

Many numbers in biology will be written in standard form (scientific notation). Learners are expected to be able to express results in standard form and be able to convert to and from decimal form. For example, the size of a cell may be $100 \mu\text{m}$ or 0.0001 m in ordinary (decimal) form. This would be written in standard form as $1.0 \times 10^{-4} \text{ m}$.

The mathematical notation for a number written in standard form is,

$$a \times 10^n$$

Where n is an integer (whole number) and

$$1 \leq (-)a < 10$$

When using standard form in calculations, a key mistake is for learners to concentrate on the value of n , and forget the rule for the value of a . For example, they may write:

$$(2.0 \times 10^4) \times (8.0 \times 10^3) = 16 \times 10^7$$

Here the learner has correctly dealt with the indices and has got the correct numerical answer, but the number is not in standard form. The 16 is incorrect and has to be 'scaled' down by dividing by 10. However if we divide by 10 we have to multiply the 10^7 by 10 to keep the numerical value the same:

$$16 \times 10^7 = 1.6 \times 10^8$$

Other problems occur if learners are required to apply the power rule (see Appendix C), for example in,

$$(2 \times 10^{-3})^2 = 4 \times 10^{-6}$$

Learners will make a number of errors here. Common mistakes are,

- to forget to square the 2: $(2 \times 10^{-3})^2 = 2 \times 10^{-6}$
- to ignore the negative power: $(2 \times 10^{-3})^2 = 4 \times 10^6$
- to add the powers rather than multiplying them: $(2 \times 10^{-3})^2 = 2 \times 10^{-1}$

Whatever form numbers are given in, learners must use the appropriate number of decimal places or significant figures (as appropriate) in calculations. In the context of converting between standard and ordinary form, learners must appreciate that significant figures need to be retained. For example:

$$0.0050 \text{ mol dm}^{-3} = 5.0 \times 10^{-3} \text{ mol dm}^{-3}$$

Here the final zero in the expression on the left is a significant figure, and so must be retained in standard form.

Decimal Places and Significant Figures

Biology learners are usually expected to record raw data to the same number of decimal places (rather than the same number of significant figures). For example, when recording the following volumes 9.0, 9.5, 10.0 and 10.5 mL, the measurements can be recorded to the same number of decimal places (but they do not have the same number of significant figures),

Measurement	Volume (mL)	Volume (mL)
	X	✓
1	9	9.0
2	9.5	9.5
3	10	10.0
4	10.5	10.5
		Mean = 9.75

Processed data can be recorded to up to one decimal place more than the raw data. For example, if the learners were asked to calculate the mean for the above example, the answer could be recorded as 9.8 or 9.75 mL. (See the Practical Skills Handbook for more guidance on tables).

In the examinations, learners may be asked to record their answer(s) to a particular number of decimal places or to a particular number of significant figures.

Decimal Places in Calculations

When adding and subtracting numbers that are quoted to the appropriate number of significant figures, the answer should be given using the lowest number of decimals used in the calculation. This can sometimes give different results than if the answer was given to the lowest number of significant figures (see *M1.1*).

For example:

$25.5 - 8.3 = 17.2$; answer given to the lowest number of decimal places, not lowest number of significant figures

$105.5 - 93.75 = 11.8$; answer given to the lowest number of decimal places, not the lowest number of significant figures.

Calculator Use

Learners with access to a scientific calculator should be able to use it to convert between different decimal/standard form calculations, as well as enter numbers in standard form. Table 2 shows the required functions for common makes of calculator.

Calculator make	Convert standard / decimal	Enter standard form
Sharp	Change	EXP
Casio	S→D	$\times 10^x$

Table 2: Calculator functions for standard form

For other models encourage learners to investigate the appropriate functions for themselves.

It should be noted that calculators will not necessarily retain the correct number of decimal places required for the calculation. For example, 3.0×10^3 is correct to 2 significant figures, but once entered into a calculator the display could be 3×10^3 , which loses 1 significant figure.

Contexts in biology

There are many areas where learners will be required to recognise and use standard form. Any calculations which involve large or small numbers will require the use of standard form.

For example, in biology learners measuring species population, number of cells in an organism/specimen, or lengths of molecular objects (organelles etc.) would all be examples where large and small numbers occur and where the most convenient representation for them is standard form. Sometimes it may be difficult to appreciate how 'small' or 'large' these standard form numbers are.

For example, the typical length of a mitochondrion in a human is $2\ \mu\text{m} = 2 \times 10^{-6}\ \text{m}$. A typical length for a bacterial cell is also $2\ \mu\text{m} = 2 \times 10^{-6}\ \text{m} = 0.000002\ \text{m}$, hence a mitochondrion is approximately the same size as a bacterial cell, despite the mitochondrion being an organelle within a eukaryotic cell.

Encourage learners to convert to decimal form using a calculator or by hand to get a full appreciation of the size. With experience and exposure, learners will eventually be able to get a grasp of the size of a standard form number without having to convert to decimal form.

Measuring quantities by difference

This is a main area where learners need to consider the role of decimal places in addition and subtraction.

The most common quantities measured by difference in practical work are mass, temperature and volume. The measurements made should be recorded to a specific number of decimal places, depending on the resolution of the instrument (see the Practical Skills Handbook for more on this topic). When calculating the difference between the measurements, this number of decimal places should be maintained.

For example, a learner conducting an investigation may record the following measurements:

Initial temperature	22.5 °C
Final temperature	29.5 °C
Temperature difference	7.0 °C

The temperature difference is given to 1 decimal place, to match the resolution of the measured values. The '0' is significant, so must be included.

M0.3 Use ratios, fractions and percentages

Learners may be tested on their ability to:

- calculate percentage yields
- calculate surface area to volume ratio
- use scales for measuring
- represent phenotypic ratios (monohybrid and dihybrid crosses).

Mathematical concepts

Ratios, fractions and percentages are related concepts. Many problems within biology will require learners to have a good understanding of the relationships between these concepts, and to use them in calculations. The individual skills required will have been covered at GCSE, but they are used in new contexts here.

Percentage change

There are many misconceptions when performing these calculations. This is often because learners are 'over-taught' the method on how to find percentage increases and decreases, actually the calculations are relatively easy and shouldn't produce too much anxiety.

One simple way to calculate the percentage change is to calculate the difference between the new quantity and the original and then express this as a percentage of the original value. For example if the new quantity is 90 and the original quantity is 80 the difference is 10, so the percentage change is

$$\frac{90 - 80}{80} \times 100 = 12.5\%$$

This can be expressed in the formula:

$$\text{percentage change} = \frac{\text{new quantity} - \text{original quantity}}{\text{original quantity}} \times 100$$

A different way to calculate percentage change uses the idea of multipliers. The key is to understand that the multiplier 1 represents a change of 0%. A multiplier of 1.43 therefore represents an increase of 43% whilst a multiplier of 0.83 represents a decrease of 17% (note that it is the difference between the multiplier and 1 which is the change – it isn't a percentage decrease of 83%). Quantities and percentages can then be found using a simple formula:

$$\text{Original Quantity} \times \text{Multiplier} = \text{New Quantity}$$

This formula can be stated in a formula triangle and then applied to situations where the percentage change is required. For example, if the initial mass of a model cell in an osmosis investigation is 5.6 g and after a day it is 4.7 g then to work out the percentage decrease we have:

$$5.6 \times \text{Multiplier} = 4.7$$

$$\text{Multiplier} = \frac{4.7}{5.6} = 0.84$$

This represents a percentage decrease of $(1 - 0.84) = 16\%$ decrease.

Or, say we wanted to know the population of rabbits after a percentage increase of 10%, given that the initial population of rabbits was 80:

$$80 \times \text{Multiplier} = \text{New Population}$$

$$80 \times 1.1 = \text{New} = 88 \text{ rabbits}$$

Contexts in biology

Percentage yield

Learners will be using these skills to find percentage change (increases and decreases) as well as percentage yields in experiments. To gain the percentage yield they will need to use the formula,

$$\% \text{ yield} = \frac{\text{Actual Amount}}{\text{Theoretical Amount}} \times 100$$

For example, if the maximum theoretical yield of insulin produced using a batch of modified bacteria was 120 g but at the end only 90 g had been produced then the percentage yield is:

$$\% \text{ yield} = \frac{90}{120} \times 100 = 75\%$$

Surface area to volume ratio

Learners will also need to calculate surface area to volume ratios and understand the implications that these ratios have. Learners will have come across this at GCSE (for example in the context of insulation) and this concept will be covered across AS and A Level biology (multicellular organisms, exchange services, heat regulation etc.).

The formula to calculate the surface area to volume ratio is:

$$\text{Ratio} = \frac{\text{Surface Area}}{\text{Volume}}$$

See *M4.1* for a list of formulae to calculate the surface areas and volumes of regular shapes.

Scales

Learners will need to use scales for measuring. The ability to use the Unitary Method to label diagrams and pictures is an essential skill. This can be denoted by a sentence, (i.e. 1 cm represents 1 μm) or a ratio (1: 0.00001) or a scale bar.

Phenotypic Ratios

Ratios are commonly used in genetics. And learners will need to represent phenotypic ratios (monohybrid and dihybrid crosses). For example, in a monohybrid cross of two heterozygous parents, where a learner is asked what the phenotype of the offspring will be the answer is presented as a ratio e.g. 3:1.

Reporting Ratios

When presenting information as a ratio of one quantity to another the ratio is reported in the form

$$x : 1$$

where x is found by dividing the first quantity by the second.

For example in a fruit fly genetics experiment the offspring were counted:

Phenotype	Number of offspring
Red eyes	78
Brown eyes	20

What is the ratio of red-eyed flies to brown-eyed flies?

$$78 / 20 = 3.9$$

The ratio of red to brown is **3.9 : 1**

What is the ratio of brown-eyed flies to red-eyed flies?

$$20 / 78 = 0.256 = 0.3$$

The ratio of brown to red is **0.3 : 1**

When more than two quantities are all being compared in this way the order in which the ratio is given once again follows the order in which the different quantities are named and the last one will always be given as 1, with the other numbers all relative to that.

Phenotype	Red eyes	Purple eyes	Scarlet eyes	White eyes
Number of offspring	127	50	32	8
Number of offspring / 8	15.9	6.3	4	1

The ratio of red to purple to scarlet to white is **15.9 : 6.3 : 4.0 : 1**

M0.4 Estimate Results

Learners may be tested on their ability to:

- Estimate results to sense check that the calculated values are appropriate.

Mathematical concepts

Being able to make an estimate for the size of a given measure is a notoriously difficult concept. The best advice for making reasonable estimates is to start at a 'known' quantity and then extrapolate from that fact to the object to be estimated. For example, let's say the mass of a tiger is to be estimated. If the mass of an average human is 'known' to be 70 kg then perhaps a sensible first estimate is to say that the tiger is 2-3 times the mass of the human. Hence the mass of the tiger could be estimated as $70 \times 2.5 = 175$ kg.

In making estimations in calculations care should be taken with how the numbers are rounded as it could lead to confusion as to whether there is an under or over estimation.

Estimating is a valuable skill; if you are able to estimate an approximate answer to a calculation, it is easier to spot if you have made a mistake in carrying out the actual calculation. For example the calculation:

$$4.9/1.10$$

could be estimated as

$$5/1 = 5$$

If the answer is then calculated as 4.45, the estimate gives reassurance that this is a reasonable answer. However, if the calculation gives an answer of 0.45, the estimate will help to recognise that a decimal point error has been made.

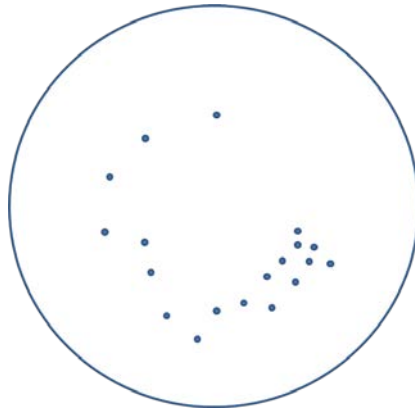
Estimating can become easier if learners are familiar with the types of answer that are typical for a particular situation.

Contexts in biology

Estimating results to sense check actual calculations or to give an idea of what might be expected is a useful tool in biology. For example, in microbiology dilutions are often used to reduce the number of bacteria to a workable number, however it can be useful to know how many bacteria were in the original culture.

So, imagine a 10 ml starter culture of bacteria was left to incubate overnight and then a serial dilution was used to make a 10 ml 1×10^{-3} dilution. A 10 µl sample of this dilution was used to streak an agar plate to isolate individual colonies.

The resulting plate looked like this,



There are 19 colonies on the streak plate. If you want to work out how many bacteria were in the original sample you should multiply by the dilution factor used, however, sometimes learners get this wrong,

19 in 10 μl

$19 \times 1000 = 19,000$ in 10 ml diluted culture

multiplied by the dilution factor $1 \times 10^{-3} = 19$ bacteria in the 10 ml starter culture

If the learner had estimated the final answer they would know this calculated answer cannot be correct. The starter culture must have thousands of bacteria in it as it had been left overnight to incubate.

The dilution factor was 1 in 1000 so the learner must multiply by 1000, so the correct answer is actually,

$19000 \times 1000 = 19,000,000$ bacteria in the 10 ml starter culture

Interestingly, this is in itself an estimate as we have not been able to count all the bacteria in the culture individually but have made a calculation based on a sample.

M0.5 Use calculators to find and use power, exponential and logarithm functions **A Level only**

Learners may be tested on their ability to:

- Estimate the number of bacteria grown over a certain length of time.

Mathematical concepts

The chief difficulty with calculating powers on a calculator is that different models have different ways of entering powers and using them. The learners in your class will potentially own a wide range of calculator models. Different calculator symbols used include 'x^y', '10^x', '^' and 'exp'. It is worth taking the time to become familiar with the different models used in your class, and ensuring that learners understand how to use these functions correctly. Also make sure they understand the different operations for e.g.:

$$3.6^3$$

and 3.6×10^3

The chief difficulty for using logarithms is that in different contexts different bases need to be used. A Level learners will need to use logarithm operations for base 10 and base e. To take the logarithm, the calculator button is usually 'log' for base 10 logarithms, and 'ln' for base e logarithms (natural logarithms).

Contexts in biology

Bacterial cells multiply exponentially. In general if a culture of bacteria of a stated initial number (N_0) increases by a factor of 2 (doubles) in one generation then the formula to find the number of cells after n generations is:

$$N = N_0 \times 2^n$$

For example, one bacterial cell will divide about every 20 minutes under standard conditions. The cells produced will form a sequence of numbers:

$$1, 2, 4, 8, 16, 32, 64, 128 \dots$$

After the 6th division there will be 64 cells. The formula to calculate the n th generation is given by:

$$N_0 \times 2^n$$

where n is the division number (note initially the division number, n is 0).

So how many cells will there be after 4 hours?

First, work out how many divisions will occur in 4 hours:

$$\text{How many divisions in one hour?} \quad 60 \text{ minutes} / 20 \text{ minutes} = 3$$

$$\text{So in 4 hours we have:} \quad 4 \times 3 = 12 \text{ divisions}$$

Then use the formula,

$$N = N_0 \times 2^n$$

$$1 \times 2^{12} = 4096 \text{ bacterial cells}$$

So, there are 4096 bacterial cells after 4 hours.

M1 – Handling data

M1.1 Use an appropriate number of significant figures

Learners may be tested on their ability to:

- 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.

Mathematical concepts

The number of significant figures used to express particular values ultimately derives from the resolution of the measuring apparatus used to determine experimental values. See the Practical Skills Handbook for more on this topic, including the appropriate number of decimal places to use for certain apparatus.

Common rounding errors include:

- Forgetting to include zeroes where they are significant figures, rather than placeholders. For example, 4.99×10^5 rounded to 2 significant figures is 5.0×10^5 , not 5×10^5
- Rounding sequentially; for example rounding 2.4478 first to 2.45 and then to 2.5. This is incorrect; the number should be rounded in a single step, giving 2.4 to 2 significant figures.

Learners must understand that the lowest level of accuracy in the inputs of a calculation will determine the level of accuracy in the answer. If there are 3 inputs to a particular calculation and they are quoted as being correct to 2, 3 and 4 significant figures, respectively then the answer can only be quoted reliably correct to 2 significant figures. Note though that if the calculation only involves addition and subtraction, decimal places should be taken into account rather than significant figures – see Section M0.2.

In these calculations, learners should be aware that certain numbers are ‘exact’, that is, there is no uncertainty in their value. These numbers can be treated as having an infinite number of significant figures; they do not affect the number of significant figures the result of a calculation should be reported to.

Contexts in Biology

Thinking about significant figures is important in any calculation, and learners should be particularly aware of the meaning of significant figures when performing calculations using experimentally determined values.

As discussed in M0.2, biology learners could be asked to report an answer to a certain number of decimal places or to a certain number of significant figures. A representative question is given below:

The diameter of a neutrophil was measured as 12 μm . Calculate the volume of the neutrophil and report your answer to 3 significant figures.

$$\text{Volume of a sphere} = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi 6^3$$

$$\text{Answer} = 904.78 = 905 \mu\text{m}^3 \text{ (to 3 sig. fig.)}$$

M1.2 Find arithmetic means

Learners may be tested on their ability to:

- find the mean of a range of data e.g. the mean number of stomata in the leaves of a plant.

Mathematical concepts

Means and weighted means

The mean (\bar{x}) is calculated using a simple formula:

$$\bar{x} = \frac{\sum x}{n}$$

where $\sum x$ is the sum of the data values and n is the number of data values.

Most learners will be familiar with this from GCSE and it is best taught as a rather ad-hoc message: 'add them all up, divide by how many'. There are few misconceptions with this when dealing with raw, listed data as the calculations involved are quite simple.

Anomalies and selecting data

When treating experimental data, for example in recording the mean value of a number of measurements, learners must be able to identify anomalies (outliers) and exclude them from the calculation. It should be emphasised that there is no hard and fast rule on how to deal with anomalies; they should be treated case by case. The detail of the mathematics involved to identify anomalies goes way beyond the scope of AS/A level Biology.

For experiments a simple checklist is this:

- was the suspected anomaly recorded in error?
- was the suspected anomaly recorded in different conditions to the other values?

If the answer to any of these questions is yes, then the anomaly should be omitted from the data set and the mean should be calculated without this value. If a potential anomaly is spotted at the time of the experiment then learners should question whether the experiment should be repeated (see the Practical Skills Handbook for more on this topic).

Calculator use

Many different scientific calculators have a Statistics mode where the mean can be calculated automatically. Whilst these functions are incredibly useful to gain full summary statistics (standard deviation, sum of squares, correlation coefficients etc.) the computational advantage of doing this for just the mean is nil. Learners can use this mode (if they know how!) but will have to press just as many buttons and therefore have the same risk for error as if they perform the addition and division manually.

Contexts in Biology

Calculating the mean for sets of data in biology can be very useful, for instance when using replicates in experiments or for comparison purposes. For example, learners might be asked to compare the number of stomata on the upper and lower sides of the leaves of a plant. The stomata on five leaves have been recorded:

No. of stomata on underside of leaf	No. of stomata on upper side of leaf
45	6
48	9
47	11
50	7
46	7

To find the mean for each set of data,

$$\frac{(45 + 48 + 47 + 50 + 46)}{5} = 47$$

$$\frac{6 + 9 + 11 + 7 + 7}{5} = 8$$

Then the learners can use the calculated means to compare the numbers of stomata.

As stated in *M0.2*, processed data such as the mean can be recorded to up to one decimal place more than the raw data.

M1.3 Construct and interpret frequency tables and diagrams, bar charts and histograms

Learners may be tested on their ability to:

- 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
- 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.

Mathematical concepts

Learners will be familiar with frequency tables, bar charts and histograms from GCSE. There are a few misconceptions however for the construction of histograms and bar charts as there is a subtle difference.

Bar charts are used when the data is *discrete*. The data can only take specific values. Variables such as eye colour, number of offspring etc. can only take certain values, there are no 'inbetween' values; number of offspring could be 0, 1, 2, etc. but couldn't be 1.5. Eye colour can only be blue, green etc. (unless you specify a wave frequency with each colour and then it becomes continuous!). With bar charts there are gaps between the bars when it is plotted.

Histograms are used when the data is *continuous*. Height, weight, and age are all examples of continuous data where they are measured to a specified accuracy (age is considered continuous because actually it is a measure of time, even though we only state integer years when we say our age). With histograms there are *no* gaps between the bars. For more information on tables and graphs please see the Biology Practical Skills Handbook.

Contexts in Biology

Plotting a histogram

Below is an example of a table of data and the data plotted as a histogram. The heights of different plants of the same species were measured to the nearest cm and displayed in the following table:

Height to the nearest cm	Frequency
$150 \leq x < 160$	3
$160 \leq x < 170$	6
$170 \leq x < 180$	7
$180 \leq x < 190$	6
$190 \leq x < 210$	2

This is an example of continuous data. Because the data is recorded to the nearest cm the groups in the above table are incorrect. $150 \leq x < 160$ should actually be $149.5 \leq x < 159.5$ to account for this rounding.

The table is then written as:

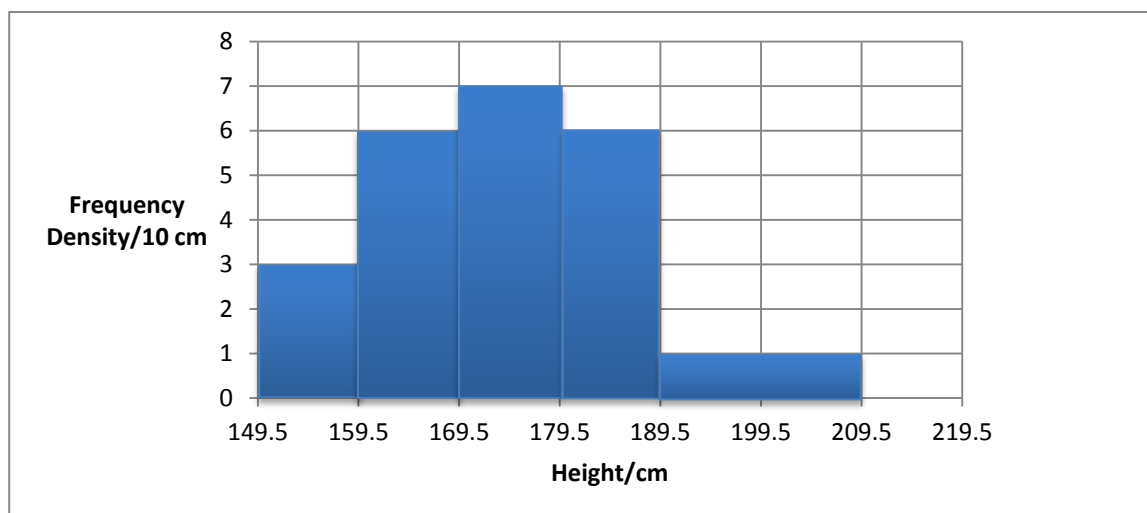
Height to the nearest cm	Frequency
$149.5 \leq x < 159.5$	3
$159.5 \leq x < 169.5$	6
$169.5 \leq x < 179.5$	7
$179.5 \leq x < 189.5$	6
$189.5 \leq x < 209.5$	2

Before this can be plotted it should be noted that the last group is larger than the other groups. The first 4 groups have a class width of 10 whilst the last has a class width of 20. This has to be taken into account otherwise the histogram will appear to show a large proportion in the last group; the *area* of the bars has to be proportional to the frequency.

Frequency density has to be calculated. As the common class width is 10 cm it makes sense to make this Frequency / 10 cm. The frequency densities for the first 4 groups remain the same as they are already 10 cm wide but for the last one, if there are 2 plants in the 20 cm then there must be 1 plant in 10 cm and hence the frequency density / 10 cm for the last group is 1:

Height to the nearest cm	Frequency	Frequency Density / 10 cm
$149.5 \leq x < 159.5$	3	3
$159.5 \leq x < 169.5$	6	6
$169.5 \leq x < 179.5$	7	7
$179.5 \leq x < 189.5$	6	6
$189.5 \leq x < 209.5$	2	1

Finally the histogram can be plotted:



Interpreting data

Learners are often required to interpret data presented in tables or graphs in AS/A Level assessments. These skills are always linked to the biological information these representations contain. For example, it would be expected that learners are familiar with ECG traces and the information these contain.

M1.4 Understand simple probability

Learners may be tested on their ability to:

- use the terms *probability* and *chance* appropriately
- understand the probability associated with genetic inheritance.

Mathematical concepts

The fundamental misconception when dealing with probability is the idea of 'randomness'. There are a number of different definitions for the word 'random' but to keep it simple emphasise to the learners that a random process doesn't produce rare results but it produces results that are impossible to *predict*. Randomness is based on our inability to predict what will happen at a *given* moment of time.

This can be modelled by rolling a die or, if there are just two equally likely outcomes, flipping a coin.

Contexts in Biology

Imagine a particular gene that determines hair colour in a fictional species. Let *b* be the blond hair allele and *B* be the brown hair allele. The *B* allele is dominant. There are 3 possible pairs; *bb*, *bB* and *BB*.

First consider the offspring resulting from sexual reproduction by two individuals. Parent 1 has the homozygous genotype *bb* and so has the phenotype of blond hair. Parent 2 has the heterozygous genotype *Bb* and so has the phenotype of brown hair.

The gametes produced by the first parent will all carry a single *b* allele. Half the gametes produced by the other parent will carry a *B* allele, and half will carry *b*.

The fusion between gametes from one parent and the other happens at random. We do not know whether one particular zygote, formed by such a fertilisation event, will have the genotype *Bb* (and hence the brown hair phenotype) or *bb* (blond hair), but we do know that:

- Each of these outcomes is equally likely
- We can model this by flipping a coin
- If we collect a large sample of these offspring, we expect to find approximately equal numbers of each.

This simple scenario is clearly represented by a Punnett square where the edges represent the gametes produced (in equal numbers) by each parent and the body of the table shows the four equally likely outcomes of random fusion of those gametes to form zygotes.

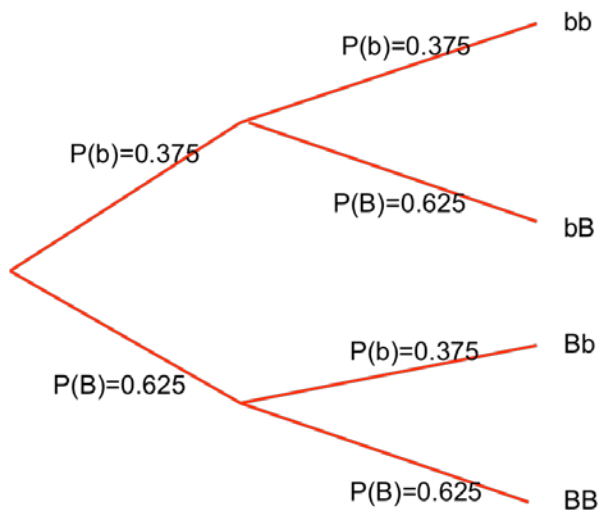
		Parent 1		
		<i>bb</i>		
Parent 2 <i>Bb</i>		<i>b</i>	<i>b</i>	
		<i>B</i>	<i>Bb</i> (brown hair)	<i>Bb</i> (brown hair)
		<i>b</i>	<i>bb</i> (blond hair)	<i>bb</i> (blond hair)

Now consider a population of many individuals of this species. Imagine that in the population the three different possible genotypes bb, Bb and BB are in the ratio 1:1:2, thus representing 25%, 25% and 50% of the population respectively.

Therefore by proportions there will be:

bb, bB, BB, BB

There are three b alleles and hence the probability of this allele being transmitted is $3/8 = 0.375$ as there are 8 in total. The probability of a B being transmitted is $5/8 = 0.625$. The probabilities for the allele pairs and hence the genotype for the next generation can be found using a probability tree diagram:



Multiplying the probabilities along the branches gives the respective probability of a pair. Hence:

$$P(bb) = 0.375 \times 0.375 = 0.140625 = 14.0625\%$$

$$P(bB) = 0.375 \times 0.625 = 0.234375 = 23.4375\%$$

$$P(Bb) = 0.625 \times 0.375 = 0.234375 = 23.4375\%$$

$$P(BB) = 0.625 \times 0.625 = 0.390625 = 39.0625\%$$

Hence the probability of having bB is $0.234375 + 0.234375 = 0.46875 = 46.875\%$. If we examine the next generation the proportion of b alleles in the population is 14.0625% plus the $\frac{1}{2}$ the proportion of the bB ($\frac{1}{2} (46.875) = 23.4375\%$). Hence the proportion of b alleles is $14.0625 + 23.4375 = 37.5\%$, which is exactly the same proportion in the original generation. This demonstrates the famous Hardy-Weinberg law.

Note that to perform these calculations we have treated the alleles as belonging to a 'gene pool' without considering the particular individuals carrying those genes. This is valid provided certain assumptions hold true, including the idea that there is no selective advantage to having one phenotype or another, and that breeding, as well as fertilisation, is a random event. If having blond hair makes those individuals more visible to predators, or if brown-haired individuals preferentially breed with other brown-haired individuals, these assumptions are broken and the Hardy-Weinberg law does not apply.

M1.5 Understand the principles of sampling as applied to scientific data

Learners may be tested on their ability to:

- analyse random data collected by an appropriate means, e.g. use Simpson's index of diversity to calculate the biodiversity of a habitat.

Mathematical concepts

AS/A Level biology learners will cover random and non-random (systematic) sampling.

In random sampling the positions at which samples are taken (when doing fieldwork) or more generally the sampling strategy used (when sampling in the lab or elsewhere) are intended to take an unbiased sub-set of the whole population (or the whole potential dataset).

In non-random sampling the positions at which samples are taken or the strategy for sampling are chosen by the investigator according to the question being investigated and some assumptions about the system being investigated.

The key thing to bear in mind when choosing how to perform sampling is the *purpose* of the investigation.

For example if the purpose of the work is to estimate the total population of a particular species in a defined geographical area a random sampling approach would be a natural choice, especially in an initial study where little was known about the distribution of that species. In this case positions for sampling should be chosen at random. A useful way to obtain this 'randomness' is to use as your coordinates for each sample, numbers from a random number generator function on a calculator or a list of random numbers.

On the other hand if the purpose of the work is to assess how distance from a freshwater source affects the distribution of the species a systematic sampling approach could be more appropriate with samples being taken at regular intervals moving further and further from the water's edge.

Contexts in Biology

Imagine a 1 m² area of the grass is to be sampled to count the number of worms. To sample without bias, the area needs to be split up into smaller squares of *equal* size (say for example 0.0625 m² squares labelled from 1 to 16):

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

If we want to take 4 samples then we require 4 random numbers to eliminate the bias caused if we were to choose the squares to be sampled. Using the RANDOM button on a calculator or using the '=randbetween(1,16)' function on a spreadsheet program these can be obtained. Below are four random numbers generated by Excel:

16, 7, 7, 6

Here 7 has been repeated so another random number has to be generated:

4

Therefore samples should be taken from squares 16, 7, 6 and 4.

M1.6 Understand the terms mean, median and mode

Learners may be tested on their ability to:

- calculate or compare the mean, median and mode of a set of data, e.g. height/mass/size of a group of organisms.

Mathematical concepts

The mean, median and mode are all measures of *central tendency* of a data set. They act as a representative value for the whole data set. These quantities are easy to calculate and should be well known to the learners but there are some subtleties.

The list below represents some data:

25,24,27,28,19,31,25,31

The mean has been mentioned in *M1.2* and is the sum of the data values divided by the number of data values:

$$\bar{x} = \frac{210}{8} = 26.25$$

To find the median the data have to be reordered:

19,24,25,25,27,28,31,31

The median is the *middle* value or in a formal manner the $\frac{n+1}{2}$ th piece of data where n is the number of pieces of data. In this example there are 8 pieces of data and hence the median lies on the $(8+1)/2$ th piece of data which is the 4.5th piece of data. This doesn't really make sense until you realise that the 4.5th data is halfway between the 4th and 5th items; 25 and 27 and hence the median (usually denoted by Q_2) is:

$$Q_2 = 26$$

The mode is the easiest to spot and is the most 'popular' item of data, the most frequent item. In the above example there is no single most frequent item of data with 25 and 31 both occurring twice. The data set is therefore *bi-modal* with modes 25 and 31.

As a general rule of thumb the mean is the most useful statistical measure and it uses *all* the items of data. If there are outliers however, then the median is more representative because it is less sensitive to outliers. If for instance a data value of 100 was added to the example above, the mean would change to 34.4 but the median would move to 27, a far more representative measure.

Contexts in Biology

Learners could be asked to calculate and/or compare these quantities for any biological context where measurements are taken, for example when measuring: the length of leaves, the mass of model cells, or the number of a group of organisms.

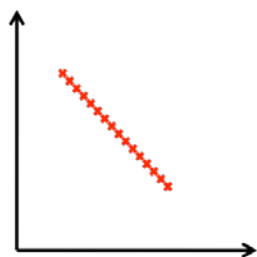
M1.7 Use a scatter diagram to identify a correlation between two variables

Learners may be tested on their ability to:

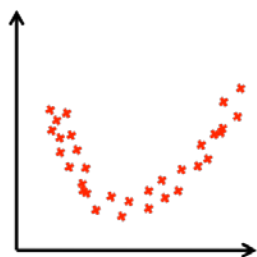
- interpret a scattergram, e.g. the effect of lifestyle factors on health.

Mathematical concepts

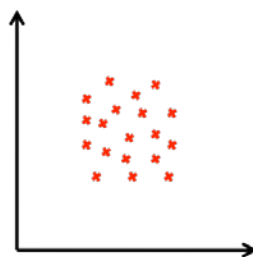
There are many types of correlation and in the absence of a mathematical procedure to calculate the correlation coefficient a lot of the interpretation comes down to judgement. The following graphs illustrate different types of correlation for two variables plotted against one another:



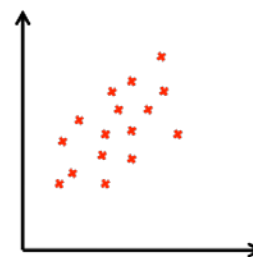
Perfect negative linear correlation



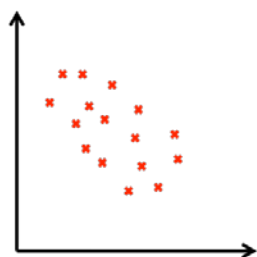
Quadratic correlation



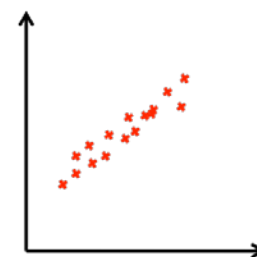
No correlation



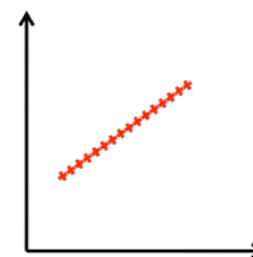
Weak positive linear correlation



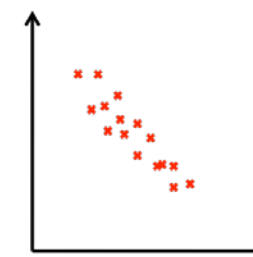
Weak negative linear correlation



Strong positive linear correlation



Perfect positive linear correlation



Strong negative linear correlation

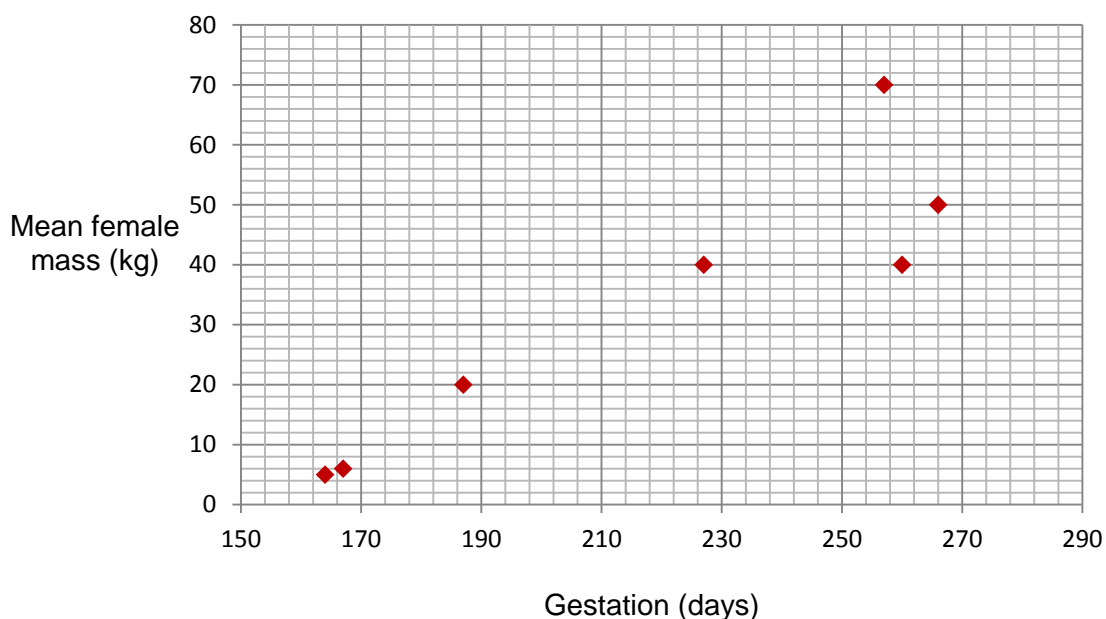
An incredibly important point however is that *correlation does NOT necessarily imply causation*. Even if two variables display a high level of linear correlation it does not mean that there is a link between them. For example if the average heights and average weights of different mammals were plotted you would expect there to be a positive correlation; the taller the mammal, the heavier it is. This could reasonably be expected to be causation; a tall animal will be heavier (in general) than a smaller one. However, if for a month we measured the temperature at midday and also the number of butterflies passing by a particular street there could feasibly be a strong correlation between the two measurements despite there being no causal link. There may be a link of course but this will require more thought, the point to make is that a strong correlation does not necessarily mean the variables are linked or dependent on one another.

Contexts in Biology

Learners could be asked to plot biological data and/or interpret a scattergram. For example, take the following table showing the mean gestation period and the mean mass of females of seven different species:

Species	Gestation (days)	Mean female mass (kg)
Human	266	50
Chimpanzee	227	40
Gorilla	257	70
Orangutan	260	40
Baboon	187	20
Monkey, <i>Rhesus</i>	164	5
Monkey, <i>Patas</i>	167	6

These can be plotted (using Excel or on paper) with mass on the y axis and gestation on the x axis as follows:



This *appears* to show a positive linear correlation (see *M1.9* for the correlation coefficient), therefore one could make a conjecture that the greater the mass of the species the longer the gestation period. However, this graph does not prove that there is a causal link between these two variables and so any probable causation would need to be justified.

M1.8 Make order of magnitude calculations

Learners may be tested on their ability to:

- use and manipulate the magnification formula,

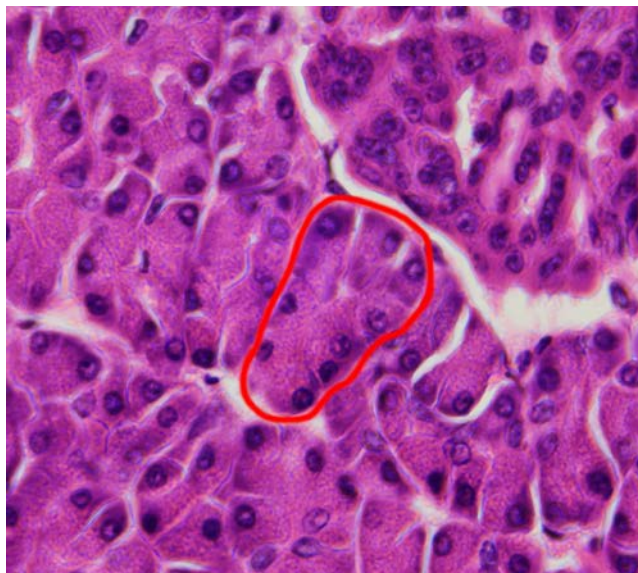
$$\text{Magnification} = \frac{\text{size of image}}{\text{size of real object}}$$

Mathematical concepts

Learners will need to be able to deal with different orders of magnitude. For example, in microscopy the size of a structure on a micrograph may be measured in mm but the actual size is six orders of magnitude smaller (nm). Order of magnitude calculations can be used to approximate values without extensive calculations and are also a useful check to ensure calculated values are reasonable (linking back to *MO.4*).

Contexts in Biology

Many biological calculations involve different orders of magnitude. For example, take the following micrograph showing a section of mouse pancreas at high power, one acinus is circled:



If the magnification is x400 and the length of the circled acinus is 32 mm, to calculate the actual size of the acinus,

$$\text{size of image} / \text{Magnification} = \text{size of real object}$$

$$32 / 400 = 0.080 \text{ mm} = 80 \mu\text{m}$$

However, if a learner had accidentally multiplied the two values together,

$$32 \times 400 = 12,800 \text{ mm} = 12,800,000 \mu\text{m}$$

Then an approximate value for the size of the acinus would be useful to show the learner that this calculated value is not reasonable.

M1.9 Select and use a statistical test

Learners may be tested on their ability to select and use:

- the chi squared (χ^2) to test the significance of the difference between observed and expected results
- the Student's *t*-test
- the Spearman's rank correlation coefficient.

Note: Please be aware that AS Level students need to be able to use all the statistical tests listed in this section: chi squared test, paired *t* test, unpaired *t* test and Spearman's rank correlation.

Mathematical concepts

In the context of OCR AS/A level Biology the statistical tests are relatively few and it should be straightforward to select the appropriate test. There is no need to memorise the formulae, as these will be provided in assessments, and it is not necessary to choose between one- and two-tailed tests: the appropriate critical value table(s) will be provided in the assessments. (Critical values tables are available in Appendix D - Statistical Tables.)

The chi squared (χ^2) test is used on data which has been observed given we know the expected values. In some cases our expected values are based on previous experimental or sample data. In the absence of such data, our expectation might simply be equal proportions in each category.

The Student's *t*-test is used when comparing two sets of the same type of data to see if there is a difference in the mean of each of them. This test can only be carried out if the data is assumed normal and the two distributions have the same standard deviation.

The correlation test is used to see if two different variables are correlated in a linear fashion in the context of a scatter-graph via either the Pearson's Product Moment Correlation coefficient or Spearman's rank correlation coefficient. The OCR GCE Biology specifications will only cover Spearman's rank correlation coefficient. Pearson's Product Moment Correlation coefficient does not need to be covered.

Contexts in Biology

Chi squared test

Suppose that the number of plant species in an area of woodland is *expected* in the quantities below and that the following data was collected showing the *observed* quantities:

Plant species	Expected Frequency	Observed Frequency
Dandelion	14	5
Buttercup	14	32
Daisy	14	17
Tulip	14	8
Crocus	14	8
Total	70	70

The question to answer is 'is there a significant difference between the expected and the observed frequencies?' To answer this we calculate the χ^2 statistic which is given by the formula:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Simply, we find the difference between the observed frequencies and the expected, square them, divide by the expected frequency and then sum them for each plant:

$$\chi^2 = \frac{(5-14)^2}{14} + \frac{(32-14)^2}{14} + \frac{(17-14)^2}{14} + \frac{(8-14)^2}{14} + \frac{(8-14)^2}{14} = 34.71$$

To perform the test we need a significance level (what percentage probability are we happy to accept for there to be a mistake in our conclusion) and the number of degrees of freedom.

There are 5 species (categories). The number of degrees of freedom is one less than the number of categories (see note on standard deviation M1.10) so in this case the degrees of freedom is 4. Choosing a significance level of 5% and looking up the critical value in the probability tables gives 9.488.

Our value of 34.71 is larger than the critical value of 9.488 so we can say there is evidence to suggest that there is a significant difference between the observed and expected frequencies. If the χ^2 statistic were smaller than the critical value then we would have concluded that there is no significant difference between the observed and the expected values.

Student's t-test

Take the following heights of 10 individuals of the same shrub species in an area where there is a lot of sunlight (data set **A**) and an area where there is little sunlight (data set **B**). An unpaired *t* test should be used in this case because the data collected are not paired measurements from the same individuals.

Individual	A height (cm)	B height (cm)
1	56	45
2	54	29
3	57	51
4	58	43
5	61	48
6	65	39
7	56	38
8	71	42
9	48	47
10	55	35
Mean	58.1	41.7

The null hypothesis is that the means are equal: $H_0 : \bar{x}_A = \bar{x}_B$

The alternative hypothesis is that the means are different: $H_1 : \bar{x}_A \neq \bar{x}_B$

This is a two-tailed test as the alternative hypothesis is 'not equal' and it will be tested at the 5% probability level.

The means and standard deviations are calculated as:

$$\bar{x}_A = 58.1 \quad s_A = 6.33$$

$$\bar{x}_B = 41.7 \quad s_B = 6.62$$

The t statistic is calculated using the formula:

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

hence:

$$t = 5.66$$

The number of *degrees of freedom* is given by $(n_A + n_B) - 2 = 18$ and using the table of critical values for a significance level of 5% for a two-tailed test:

Degrees of freedom	Critical value	Degrees of freedom	Critical value
4	2.78		
5	2.57	15	2.13
6	2.48	16	2.12
7	2.37	18	2.10
8	2.31	20	2.09
9	2.26	22	2.07
10	2.23	24	2.06
11	2.20	26	2.06
12	2.18	28	2.05
13	2.16	30	2.04
14	2.15	40	2.02

For degrees of freedom equal to 18 the critical value is 2.10. Our value of t is $5.66 > 2.10$ and hence the null hypothesis is rejected; there is sufficient evidence to reject H_0 .

Paired *t*-test

If two data sets are paired, so that two measurements are collected from the same individual, then a paired *t*-test should be used. So, for example, the effect of caffeine on the heart rate of daphnia was investigated using 15 individual daphnia ($n = 15$),

Daphnia	Heart rate before caffeine (bpm)	Heart rate after caffeine (bpm)	Difference (<i>d</i>)	How far is difference from mean difference? ($d - \bar{d}$)	($d - \bar{d}$) ²
1	180	220	40	0.3	0.09
2	178	221	43	3.3	10.89
3	185	219	34	-5.7	32.49
4	180	218	38	-1.7	2.89
5	179	222	43	3.3	10.89
6	181	220	39	-0.7	0.49
7	182	223	41	1.3	1.69
8	181	224	43	3.3	10.89
9	177	218	41	1.3	1.69
10	183	220	37	-2.7	7.29
11	182	223	41	1.3	1.69
12	184	224	40	0.3	0.09
13	180	219	39	-0.7	0.49
14	179	216	37	-2.7	7.29
15	183	222	39	-0.7	0.49
Total			595		89.35

Firstly, find the difference between the before and after values for each daphnia. The mean of the differences can then be calculated.

Mean difference $\bar{d} = 595 / 15 = 39.7$

Then for each daphnia calculate how far the difference in heart rates for that individual is from the mean difference. For example daphnia 1 has a 40 bpm difference in its heart rate which is just 0.3 bpm greater than the mean difference (39.7 bpm). Square these numbers (to give 0.09 in the case of daphnia 1, 10.89 for daphnia 2 and so on) and add them up to find the total (89.35 in this case).

The standard deviation can then be calculated:

$$\text{Standard deviation } s_d = \sqrt{\frac{\sum(d - \bar{d})^2}{n - 1}} = \sqrt{\frac{89.35}{14}} = 2.53$$

For the paired test the *t* statistic is calculated using the formula,

$$t = \frac{\bar{d}\sqrt{n}}{s_d} = \frac{39.7\sqrt{15}}{2.53} = 60.8$$

The number of degrees of freedom is $n - 1 = 14$. The *p* value at 5% for 14 degrees of freedom is 2.15. As our value of *t* is much greater than this we can say that the difference is not due to chance and that caffeine does have an effect on the heart rate of daphnia.

Correlation coefficient for a set of bivariate data

Take the previous data for gestation periods and weight (see M1.7):

Species	Gestation (days)	Mean female mass (kg)
Human	266	50
Chimpanzee	227	40
Gorilla	257	70
Orangutan	260	40
Baboon	187	20
Monkey, <i>Rhesus</i>	164	5
Monkey, <i>Patas</i>	167	6

Using Spearman's rank correlation coefficient, first the data have to be ranked:

Species	Gestation (days) x	Mean female mass (kg) y	Rank x	Rank y
Human	266	50	1	2
Chimpanzee	227	40	4	3.5
Gorilla	257	70	3	1
Orangutan	260	40	2	3.5
Baboon	187	20	5	5
Monkey, <i>Rhesus</i>	164	5	7	7
Monkey, <i>Patas</i>	167	6	6	6

They are ranked from 1 to 7 with 1 being the highest and 7 being the lowest. Notice that for the weights there are two 40 kg items of data occupying the 3rd and 4th highest ranks. Therefore they are given the average rank of $(3+4)/2=3.5$.

Next find the difference between the ranks d and then square d^2 :

Species	Gestation (days) x	Mean female mass (kg) y	Rank x	Rank y	d	d ²
Human	266	50	1	2	-1	1
Chimpanzee	227	40	4	3.5	0.5	0.25
Gorilla	257	70	3	1	2	4
Orangutan	260	40	2	3.5	-1.5	2.25
Baboon	187	20	5	5	0	0
Monkey, <i>Rhesus</i>	164	5	7	7	0	0
Monkey, <i>Patas</i>	167	5.5	6	6	0	0
					Total	7.5

The sum of $\sum d^2$ is 7.5. Finally use the formula:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 7.5}{7(7^2 - 1)} = 1 - \frac{45}{336} = 0.8661$$

To see if this indicates a significant correlation the test statistic r_s must be compared with the relevant critical value.

In this case n , the number of paired data items, is 7.

The critical value for the two-tailed test at the 5% level for $n=7$ is 0.7857 (see table p63)

As the test statistic $r_s = 0.8661 > 0.7857$ then we have evidence to reject H_0 and accept H_1 . Therefore we can say, with 95% confidence, that there is a positive correlation between gestation period and mass of the female.

M1.10 Understand measures of dispersion, including standard deviation and range

Learners may be tested on their ability to:

- calculate the standard deviation
- 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.

Mathematical concepts

Standard deviation is a quantitative measure of the spread of data about the mean. For a set of sample data, the formula for the standard deviation is:

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

Contexts in Biology

Learners could be asked to calculate standard deviation.

Taking the example of the gestation data (from M1.7 and M1.9):

Species	Gestation (days) (x)
Human	266
Chimpanzee	227
Gorilla	257
Orangutan	260
Baboon	187
Monkey, <i>Rhesus</i>	164
Monkey, <i>Patas</i>	167

The sum of the data values is:

Species	Gestation (days) (x)	$(x - \bar{x})^2$
Human	266	2275.29
Chimpanzee	227	75.69
Gorilla	257	1497.69
Orangutan	260	1738.89
Baboon	187	979.69
Monkey, <i>Rhesus</i>	164	2948.49
Monkey, <i>Patas</i>	167	2631.69
Total	1528	12147.43

Next the mean is calculated:

$$\bar{x} = \frac{1528}{7} = 218.3$$

This is used to find $(x - \bar{x})^2$ for each species and then the $\sum(x - \bar{x})^2$.

There are 7 species, so $n - 1 = 6$ and hence the standard deviation can be calculated:

$$s = \sqrt{\frac{12147.43}{6}} = 45.0$$

To interpret the standard deviation the following facts are useful for the expected distribution of the data:

- The middle 68% of the data lie within one standard deviation either side of the mean
- 95% of the data lie within 2 standard deviations either side of the mean.

M1.11 Identify uncertainties in measurements and use simple techniques to determine uncertainty when data are combined

Learners may be tested on their ability to:

- calculate percentage error where there are uncertainties in measurement.

Mathematical concepts

When a measurement is taken there will exist uncertainty due to the level of accuracy of the measuring apparatus. For example, when taking a temperature with a standard graduated thermometer you can only be correct to within 0.5 °C.

The *absolute* uncertainty is the extent a measurement could be 'out' by. In the case above this is 0.5 °C. The *relative* uncertainty is the ratio of the absolute uncertainty to the original measurement. For example if the temperature was measured as 67.5 °C ± 0.5 °C then the relative uncertainty is

$$\frac{0.5}{67.5} \times 100\% = 0.7\%$$

When measurements are added or subtracted, for example when measuring a change in temperature, the absolute uncertainties must be added to give the uncertainty in the combined measurement. So, if the above measurement of 67.5 °C is used to calculate a temperature change from a room temperature measurement of 21.0 °C, then the absolute uncertainty in the temperature change is

$$0.5\text{ °C} + 0.5\text{ °C} = 1.0\text{ °C}$$

The relative uncertainty in the measured temperature change is

$$\frac{1.0}{46.5} \times 100\% = 2.2\%$$

In general, the formula for calculating the relative uncertainty in a value calculated by difference is

$$\% \text{ uncertainty} = \frac{2 \times \text{absolute uncertainty}}{\text{quantity measured}} \times 100\%$$

In GCE Biology, learners do not need to be able to combine uncertainties in more complex operations, such as when multiplying or dividing.

Contexts in Biology

The above principles explained through the example of temperature measurements can be applied to any experimental results.

M2 – Algebra

M2.1 Understand and use the symbols: =, <, <<, >>, >, ∞, ~

Learners may be tested on their ability to:

- use these symbols appropriately and correctly in their given contexts
- understand these symbols in the contexts of formulae given.

Mathematical concepts

Learners should have had exposure to the symbols = (equals), < (less than), << (much less than), >> (much greater than) and > (greater than) from an early age, and should understand how and why they are used.

The symbol ∞ means 'is proportional to'. If two quantities *A* and *B* are *directly proportional* then the appropriate mathematical statement is,

$$A \propto B$$

If the two quantities are *inversely proportional* then the appropriate relationship is,

$$A \propto \frac{1}{B}$$

The symbol ~ means 'is roughly equal to' or 'of the same order'. This symbol may be used in the context of approximations made in calculations of quantities.

Learners will be required to understand and use these symbols as they arise in various contexts. The more important aspect here is that learners *understand* the symbols when they are used. When describing mathematical relationships, learners would often be able to use descriptions in place of the symbols, for example stating that one value is directly proportional to another rather than giving the formal mathematical statement. Conversely, learners might prefer to use symbols rather than descriptions for reasons of brevity. This is fine, but learners must be sure to use the correct symbol.

Contexts in biology

Learners will be expected to recognise these symbols in a range of biological contexts.

M2.2 Change the subject of an equation

Learners may be tested on their ability to:

- Use and manipulate equations, e.g. magnification.

Mathematical concepts

The most common equations to be rearranged can often be posed in the form of a formula triangle. This is where three quantities a , b and c are linked by the simple relationship:

$$ab = c$$

This is quite an easy equation to arrange for the other variables. For example dividing by b yields the formula for a :

$$a = \frac{c}{b}$$

whilst dividing by a will give the formula for b :

$$b = \frac{c}{a}$$

This mathematical principle should be familiar to the learners.

Contexts in Biology

Learners may need to recall and rearrange a biological formula within an assessment. For example in microscopy, to calculate magnification the equation is given by:

$$\text{magnification} = \frac{\text{size of image}}{\text{size of real object}}$$

This equation can be used to find the magnification given the size of the image and the size of the real object are known.

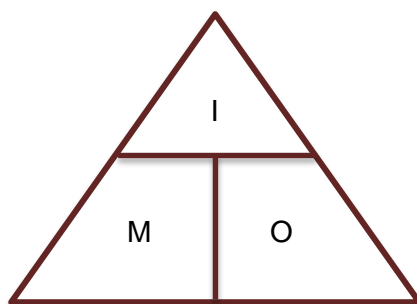
This can be rearranged to make the size of the real object the subject:

$$\text{size of real object} = \frac{\text{size of image}}{\text{magnification}}$$

or for the size of the image:

$$\text{size of image} = \text{magnification} \times \text{size of real object}$$

As mentioned above, learners often find the easiest way to visualise the relationship between three quantities is to use a triangle, e.g.



M2.3 Substitute numerical values into algebraic equations using appropriate units for physical quantities

Learners may be tested on their ability to:

- Use a given equation e.g. Simpson's index of diversity,

$$D = 1 - \left(\sum \left(\frac{n}{N} \right)^2 \right)$$

Mathematical concepts

Learners should be aware of the principles of algebraic equations from GCSE Maths but a few misconceptions may remain. The most common problem is when dealing with powers and negative quantities in formulae.

The expression x^2 for example, whilst innocuous enough can cause issues when a negative number is substituted. Substituting $x = -2$ for example should be calculated as $(-2)^2 = 4$ not $-2^2 = -4$.

There can be confusion when substituting numbers of different signs. The expression 'two negatives make a positive' is often over-used. It is *always* true that two negatives multiplied/divided equal a positive, so

$$-3 \times -5 = +15$$

whilst if only one of them is negative then the answer is negative

$$2 \times -3 = -6$$

The over-use of this rule arises when addition is involved. For example:

$$-3 + -5$$

Learners could think that because there are two negatives being added then they become positive and the answer is

$$3 + 5 = 8$$

Actually for addition the changes occur only when the signs are the same in the 'middle' of the sum. In the above sum one of the signs in the 'middle' is negative so it becomes negative. So the above example should actually be read as '-3 minus 5' which is -8. However, in the formula:

$$-3 - -5$$

Here there are two negatives in the 'middle' and there it becomes a plus. Hence the sum is '-3 plus 5' which is +2.

Additionally in the context of a formula such as $A = \pi r^2$ (for the area of a circle) learners should be aware that the radius squared is being *multiplied* by pi despite the absence of a multiplication sign.

In general the laws of BIDMAS should be adhered to where the operations should be completed in the order of Brackets, Indices, Division, Multiplication, Addition and Subtraction. Often the brackets are not present although implicitly are there as the example below illustrates.

Contexts in Biology

Simpson's Index of Diversity

Simpson's Index of Diversity states:

$$D = 1 - \sum \left(\frac{n}{N} \right)^2$$

where n is the number of a particular species and N is the total number of organisms. The following data is obtained from 2 different samples of forests:

Tree species	Number of individuals	
	Sample 1	Sample 2
Elm	250	15
Willow	385	61
Oak	365	924
Total	1000	1000

To calculate the index of diversity for Sample 1, we have to first find $\sum \left(\frac{n}{N} \right)^2$. For Sample 1 this is found by:

$$\left(\frac{250}{1000} \right)^2 + \left(\frac{385}{1000} \right)^2 + \left(\frac{365}{1000} \right)^2 = 0.34395$$

Substituting this into the formula gives:

$$D = 1 - 0.34395 = 0.65605$$

For Sample 2 this is:

$$D = 0.142278$$

The higher the value of D the higher the diversity and this is confirmed with these calculations.

Water Potential

When covering osmosis the equation for water potential (ψ) should be covered,

$$\psi = \psi_s + \psi_p$$

Where ψ_s is the solute potential and ψ_p is the pressure potential.

So for the following two adjacent cells,

Cell A	Cell B
$\psi_s = -250$ $\psi_p = 200$	$\psi_s = -300$ $\psi_p = 150$

$$\text{Cell A} = -250 + 200 = -50 \text{ kPa}$$

$$\text{Cell B} = -300 + 150 = -150 \text{ kPa}$$

So **Cell A** has the higher water potential (closer to zero) and water from **Cell A** will diffuse into **Cell B** via osmosis.

M2.4 Solve algebraic equations

Learners may be tested on their ability to:

- solve equations in a biological context, e.g.

$$\text{cardiac output} = \text{stroke volume} \times \text{heart rate}$$

Mathematical concepts

Solving an equation usually involves substituting values into a formula and realising that there is one unknown unaccounted for. Finding the value of this unknown is the same as solving the equation.

In order to calculate the unknown, it may be necessary to first rearrange the equation. Skills *M2.2* and *M2.3* are therefore often also needed in solving equations; indeed, the three skills are rarely encountered in isolation.

Take for example the formula:

$$E = U + pV$$

If we were to substitute some values in for E , U and p , the formula becomes an equation for V ; the only variable that remains unknown:

$$7 = 2 + 3V$$

To find V we have to 'unlock' what is happening to V . By this we mean we have to 'undo' the operations that link V to the other numbers. First we subtract the '2' from both sides to get the $3V$ by 'itself':

$$5 = 3V$$

To 'undo' the multiplication by 3 we divide by 3 and solve the equation:

$$V = \frac{5}{3} = 1.667$$

Contexts in Biology

Using the above biological context of cardiac output, if for example the cardiac output of an individual (volume of blood pumped by the heart per minute) is given as $10 \text{ dm}^3 \text{ min}^{-1}$ and the heart rate is measured as 185 bpm, we can solve the equation to find the stroke volume (volume of blood pumped per heart beat).

Substituting these values in:

$$10 \text{ dm}^3 \text{ min}^{-1} = \text{stroke volume} \times 185 \text{ bpm}$$

Dividing both sides by the heart rate:

$$\frac{10}{185} = \text{stroke volume} = 0.054 \text{ dm}^3$$

Then changing the units from dm^3 to cm^3 :

$$0.054 \text{ dm}^3 = 54 \text{ cm}^3$$

Hence the stroke volume is 54 cm^3 (54 ml).

M2.5 Use logarithms in relation to quantities that range over several orders of magnitude **A Level only**

Learners may be tested on their ability to:

- use a logarithmic scale in the context of microbiology, e.g. growth rate of a microorganism such as yeast.

Mathematical concepts

Logarithms are basically powers. If we take the following calculation:

$$10^2 = 100$$

this can be expressed as

the power of 10 that gives 100 is 2

or in formal notation

$$\log_{10} 100 = 2$$

and we usually drop the 10 as it is assumed to be base 10 unless stated otherwise:

$$\log 100 = 2$$

Logarithms provide a better scale when dealing with quantities that vary exponentially (get big/small very quickly). For example, imagine sketching a graph where the scale goes from 10, 100, 1000, 10 000, 100 000 and so on. This would be impossible to do on a standard graph. Taking the logarithms of these quantities gives 1, 2, 3, 4 and 5, which is far more manageable to handle and to spot trends.

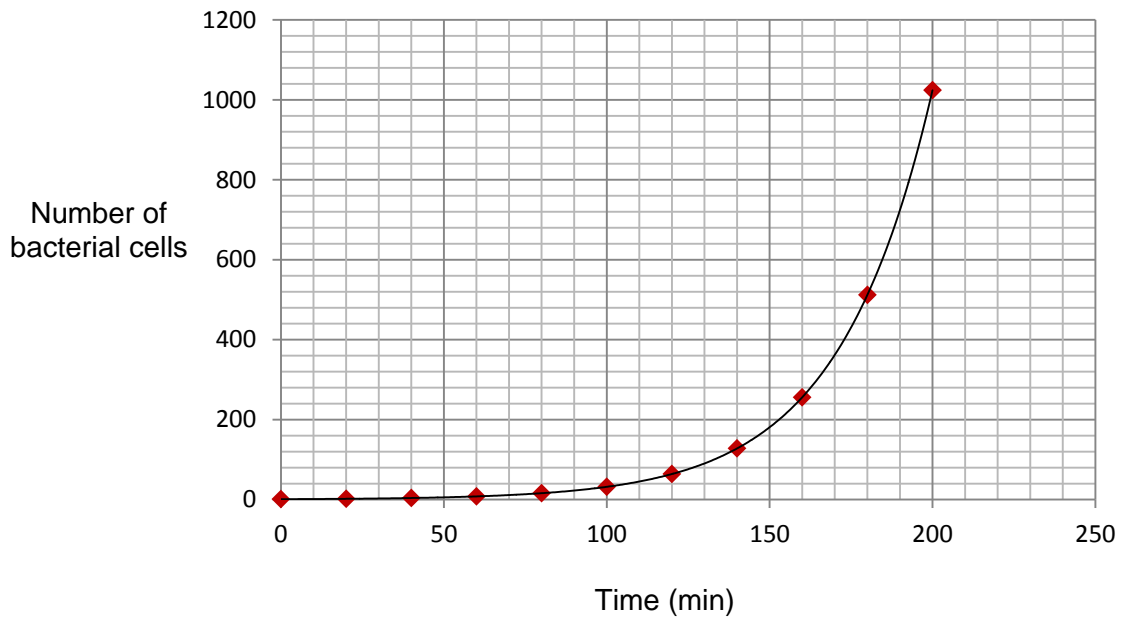
The natural logarithm is denoted by $\ln x$, which is shorthand for $\log_e x$. Here e is the mathematical constant approximately equal to 2.7182818. This number is of central importance in mathematics, and often occurs in situations where quantities change exponentially over time. Like π , which learners should be aware of, it is an irrational number, meaning it cannot be represented as a repeating decimal.

Contexts in Biology

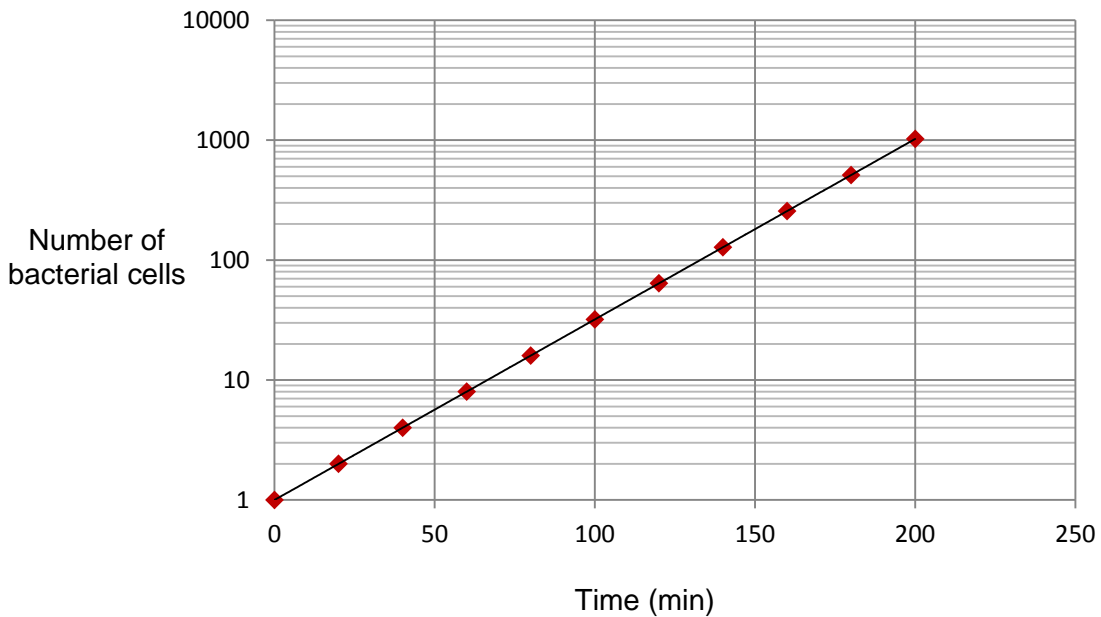
Learners following A Level Biology will cover the standard growth curve of a microorganism. As discussed in *M0.5*, bacterial cells multiply exponentially - one bacterial cell will usually divide about every 20 minutes under standard conditions. The cells produced will form a sequence of numbers:

1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024 ...

If you plot these data using standard paper you get the following exponential growth curve (the beginning of the classic sigmoidal curve),



However, if you plot the data using a logarithmic scale (or take the log of each number and plot these values) you get a straight line relationship:



This can be more manageable when visualising and interpreting the exponential relationship, or for extrapolating data from the graph.

M3 – Graphs

M3.1 Translate information between graphical, numerical and algebraic forms

Learners may be tested on their ability to:

- understand that data may be presented in a number of formats and be able to use these data, e.g. dissociation curves.

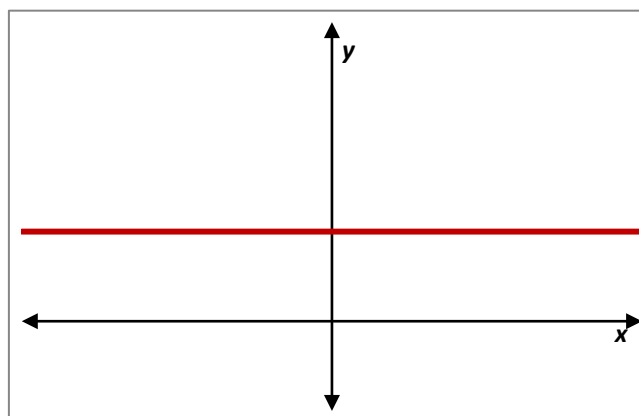
Mathematical concepts

There are several situations in GCE Biology where data may be presented graphically. Learners should be familiar with the types of graphical representations used, the conventions for variables used on the graph axes, and how to interpret the information provided in the graph.

Learners will need to be able to read co-ordinates for points on graphs (both x- and y-co-ordinates).

In biology, learners do not need an elaborate understanding of how to convert graphs into algebraic equations, but there are a few instances where it is useful if learners can judge the relationship between the plotted variables from the shape of the graph.

The following general graphs are useful for learners to know and recognise.



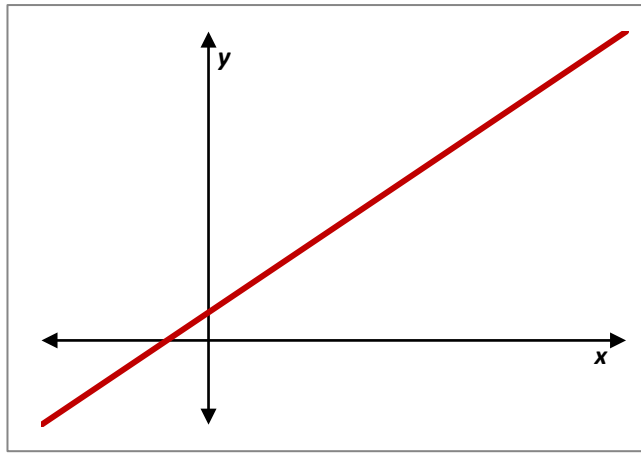
A graph showing a horizontal line parallel to the x-axis shows that the variable plotted on the y-axis is independent of the variable plotted on the x-axis. In mathematical terms, this relationship can be expressed as

$$y \propto x^0$$

or

$$y = \text{constant}$$

Where the constant in question is given by the y-axis value of any point on the line.



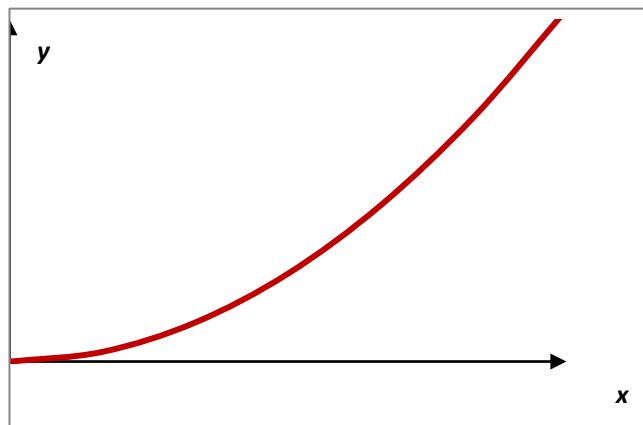
A non-horizontal straight line shows that the variable plotted on the y-axis is proportional to the variable plotted on the x-axis, or

$$y \propto x$$

In mathematical terms, this graph can be expressed as:

$$y = mx + c$$

m is the gradient of the graph, and c is the value of the intercept on the y-axis (link to M3.3).



A curved graph that passes through the origin indicates that the variable plotted on the y-axis is proportional to a power of the variable plotted on the x-axis that is greater than 1, or

$$y \propto x^n \text{ where } n > 1$$

Learners will not be required to determine the exact mathematical relationship of such graphs.

Contexts in Biology

Learners are expected to be able to interpret biological information presented in tables and graphs (for example bar charts, histograms, line graphs and pie charts) and to be familiar with experimental traces such as ECG and spirometer traces.

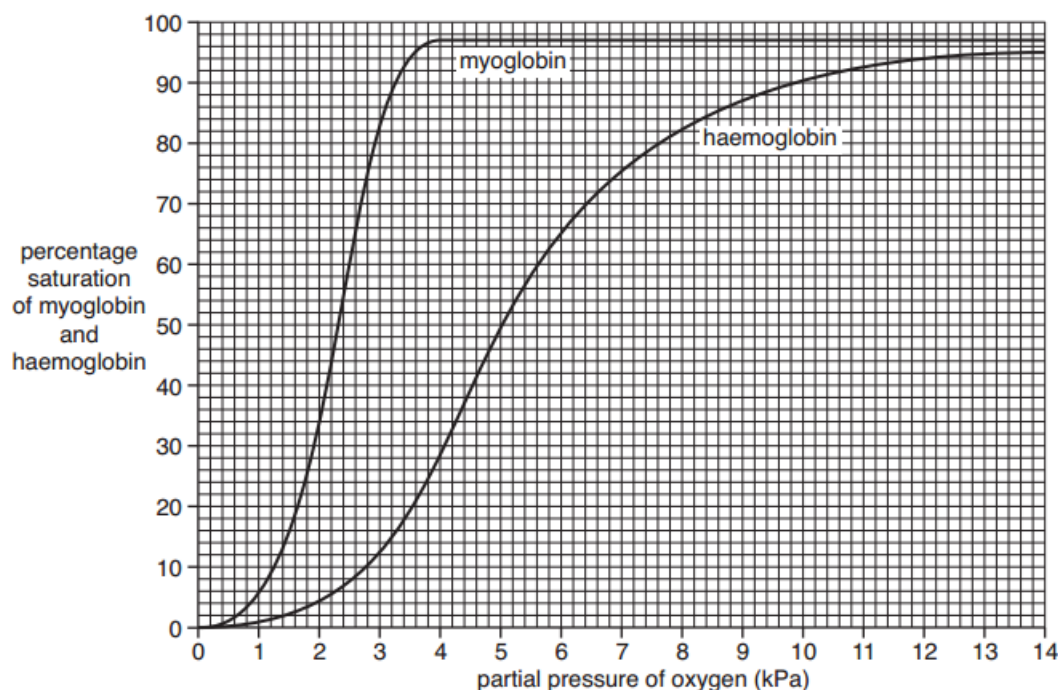
A key feature of any line graph displaying a relationship between 2 variables is that the gradient of the curve represents the rate at which the quantity on the y axis is increasing as the quantity on the x axis increases.

Learners find this concept easiest to understand and describe when the quantity on the x axis is time. The steeper the curve the more quickly the y axis quantity is changing and vice versa: the shallower the curve the more slowly the y axis quantity is changing.

When the x axis is not 'time' it can be harder for learners to accurately express their understanding. It is still acceptable to use the idea of quick or slow change but it should be made clear that this is with respect to the x axis quantity (rather than time). See below.

Dissociation curves

Both AS and A Level learners require an understanding of oxygen dissociation curves. For example, learners could be asked to find the percentage saturation of oxygen at a particular partial pressure, compare oxygen dissociation curves for different proteins or species, or to consider how the dissociation curve changes at different concentrations of carbon dioxide (the Bohr effect). An example figure is given below.



When describing features of the haemoglobin curve, and having observed a steep gradient in the middle part of the curve we could say:

“At pO_2 between 3 and 7 kPa small changes in pO_2 result in large changes of percentage saturation of haemoglobin.”

Or

“At pO_2 between 3 and 7 kPa the percentage saturation of haemoglobin increases rapidly as pO_2 increases.”

Both statements clearly refer to the x axis and y axis variables and accurately describe the relationship using data from the graph.

M3.2 Plot two variables from experimental or other data

Learners may be tested on their ability to:

- select an appropriate format for presenting data, e.g. bar charts, histograms, line graphs and scattergrams.

Mathematical concepts

There are a number of different graphs that can be plotted for different types of data. If there is a single variable being collected then a bar chart or histogram is used (e.g. number of leaves of a plant, height of plants). If the data that has been collected has two variables (e.g. time/concentration, oxygen/haemoglobin) then a line graph/scatter-plot is used. A line of best-fit can be constructed to identify trends (usually with the independent variable as the x axis).

Plotting a graph should be a straightforward concept but the following guidelines are useful:

- Points plotted must be within 1 square of the correct value
- Appropriate linear scale used on axes
- Graph should make good use of available space
- Scales should be 'sensible', i.e. using decimal or otherwise straightforward scale
- Scales must be chosen so that all points fall within the graph area – points must not be plotted outside the graph area
- Axes must be labelled, with units included
- There should be an informative title.

If there is a trend, a line (or curve) of best fit should be drawn (the points plotted should **not** be connected). The best smooth line must be drawn that achieves a balance of points above and below the line.

Learners must realise when lines of best fit need to be drawn through the origin; this is the case for certain relationships, such as in rate–concentration graphs.

Extrapolation is required in some instances, for example to determine the intercept with the y-axis. Learners will only be asked to extrapolate linear graphs; extrapolation should be achieved by extending the line of best fit to the appropriate point.

If no trend can be identified, or if the nature of the data means interpolation between data points is not appropriate, straight lines can be used to join points if this increases the clarity of presentation.

Please see the Practical Skills Handbook Appendix 6 for more on the topic of graphs including the use of error bars and range bars.

Contexts in Biology

Rate of reactions

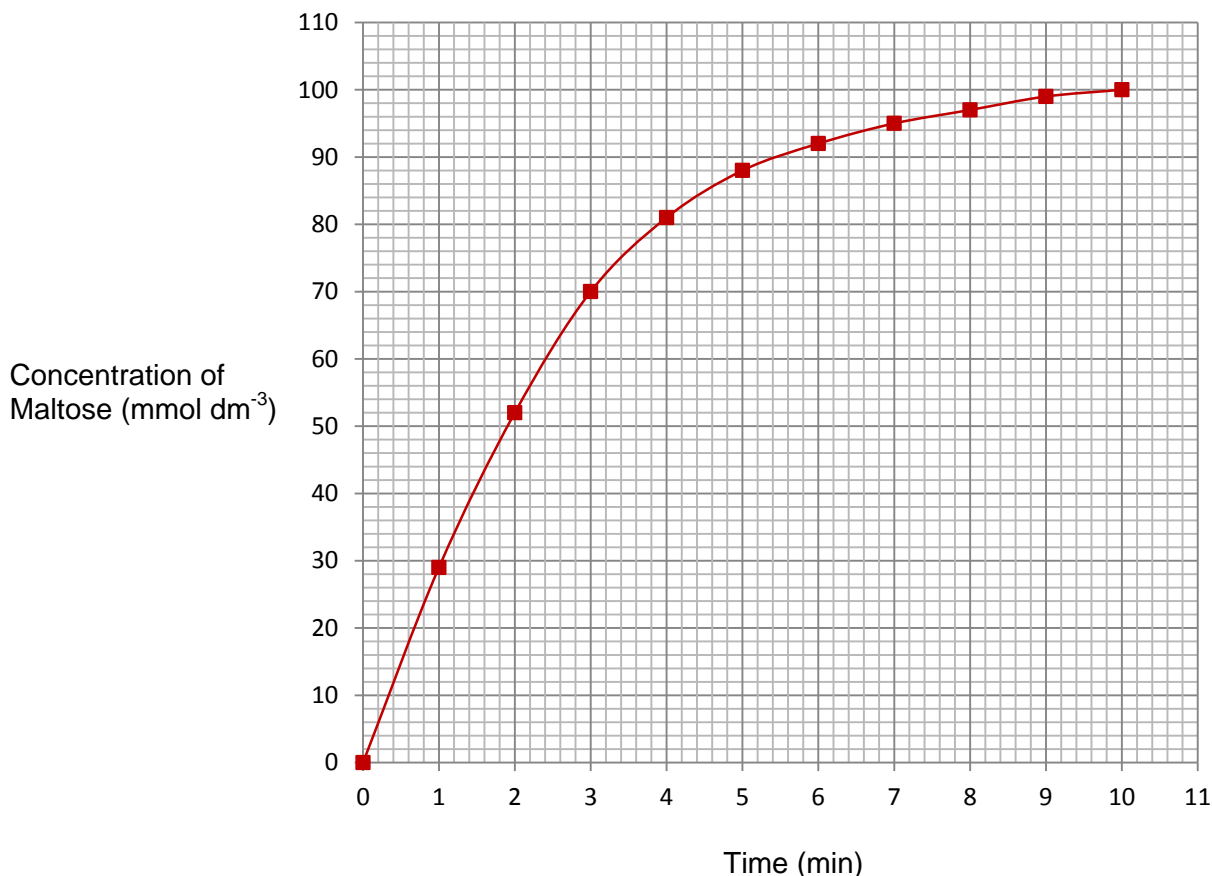
Data to be plotted would be recorded by the learners themselves, or provided in a table such as,

Time / min	0	1	2	3	4	5	6	7	8	9	10
Maltose / mmol dm ⁻³	0	29	52	70	81	88	92	95	97	99	100

Suitable axis scales will depend on the size of the graph paper available, but sensible scales would be:

- 0–11 min, with 1 min per major square, for the x-axis
- 0–110 mmol dm^{-3} , with 10 mmol dm^{-3} per major square, for the y-axis.

These scales should make good use of the space on an A4-sheet of graph paper with 2 mm squares:



Calibration curves

A calibration curve is a technique whereby a relationship is established between two variables, so that the resultant curve can be used to establish a property of an unknown sample.

For example, in colorimetry, a calibration curve is used to establish the relationship between the absorption of a particular wavelength of light by a solution, and the concentration of a coloured solute.

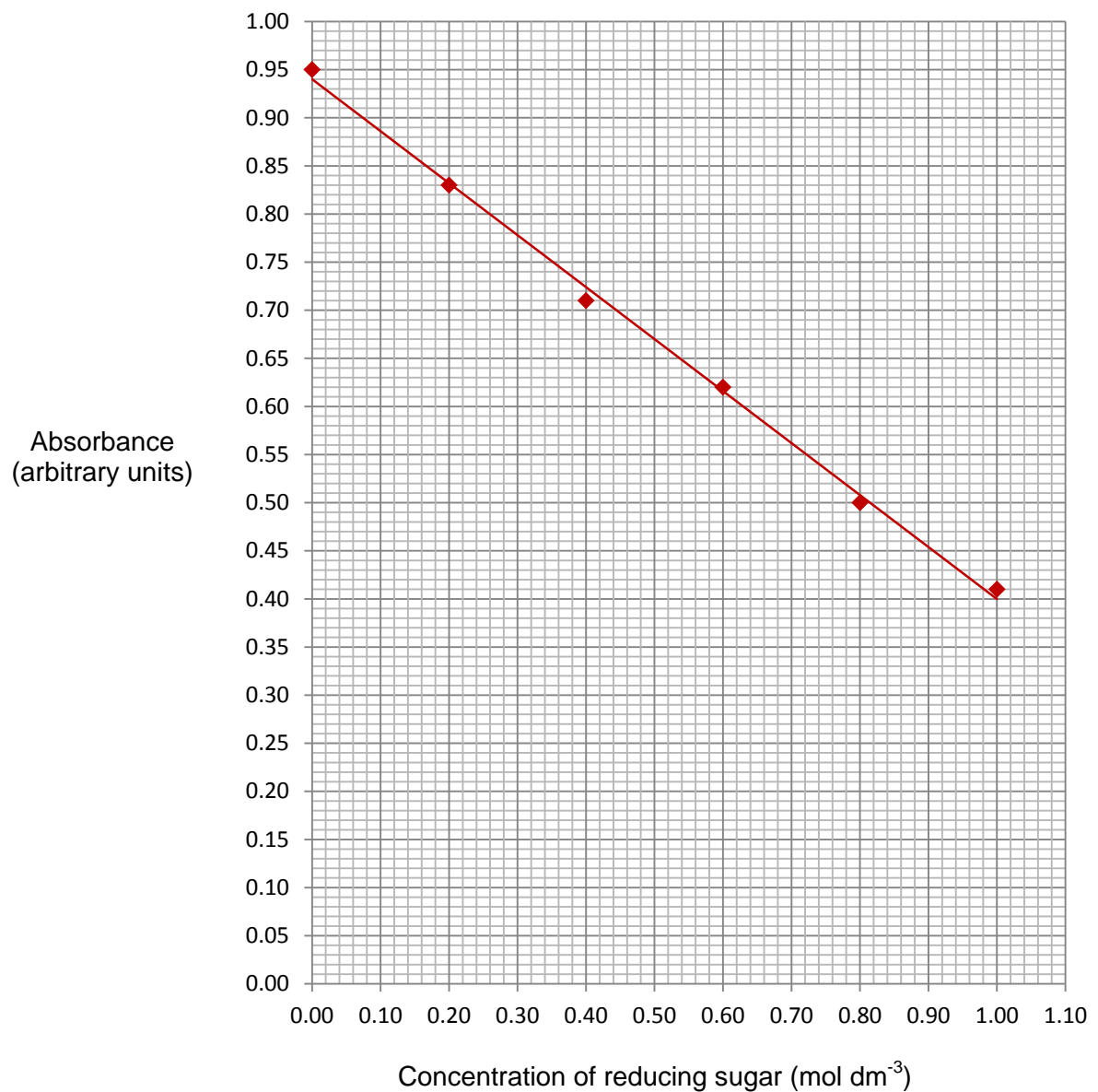
The following data might be produced for the absorption of a reducing sugar of known concentrations tested with Benedict's reagent:

Reducing sugar concentration (mol dm^{-3})	0.0	0.2	0.4	0.6	0.8	1.0
Absorbance (AU)	0.95	0.83	0.71	0.62	0.50	0.41

Suitable axis scales would be:

- 0–1.0 mol dm^{-3} , with 0.1 mol dm^{-3} per square, for the x-axis
- 0–1.0 absorbance (arbitrary units), with 0.05 arbitrary units per square, for the y-axis.

This would produce the following calibration curve:



The calibration curve is then used to determine the concentration of an unknown solution of reducing sugar. If such a solution has an absorbance of 0.56, the concentration according to the line of best fit is 0.70 mol dm⁻³.

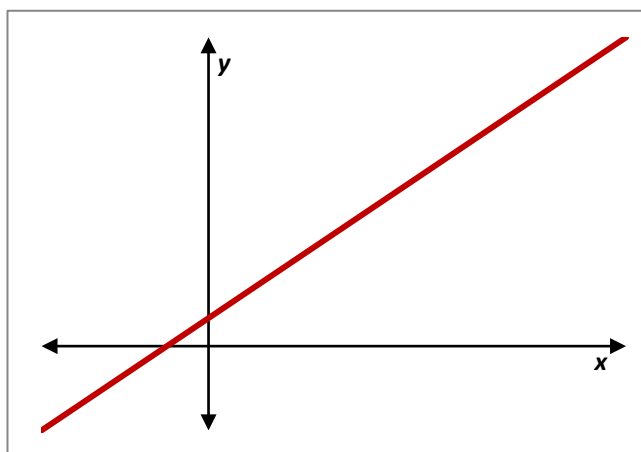
M3.3 Understand that $y = mx + c$ represents a linear relationship

Learners may be tested on their ability to:

- 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.

Mathematical concepts

This concept is a two-way process. As discussed in *M3.1*, learners should be able to work out a linear relationship given a graph and be able to sketch a graph given a linear relationship.



A non-horizontal straight line shows that the variable plotted on the y-axis is proportional to the variable plotted on the x-axis, or

$$y \propto x$$

In mathematical terms, this graph can be expressed as:

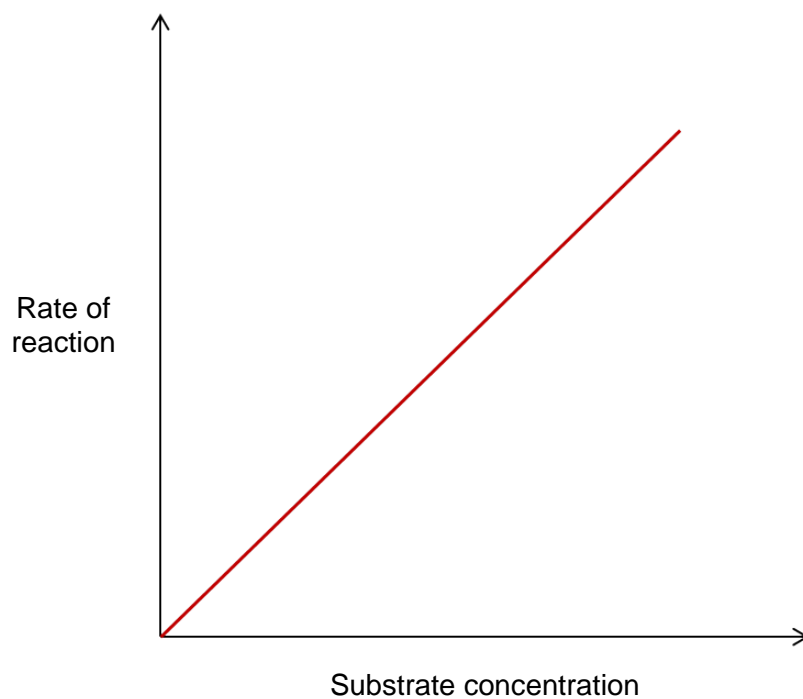
$$y = mx + c$$

m is the gradient of the graph, and c is the value of the intercept on the y-axis.

Learners should understand that a positive m represents a line going 'up' from left to right and a negative m represents a line going 'down' from left to right. When sketching the graph learners should always start from the y-intercept and then use the gradient to determine another point on the graph. Once this extra point has been determined the line can be drawn as it is only necessary for two points to be known to sketch a straight line.

Contexts in Biology

A graph of rate of reaction against substrate concentration for an enzyme is displayed below:



This is an example of a straight-line graph and hence will have the form $y = mx + c$. In fact because it goes through the origin, then the y intercept is 0 and will be of the form:

$$\text{Rate} = m \times \text{concentration}$$

where m is the gradient of the line (see M3.5).

M3.4 Determine the intercept of a graph **A Level only**

Learners may be tested on their ability to:

- Read off an intercept point from a graph, e.g. compensation point in plants.

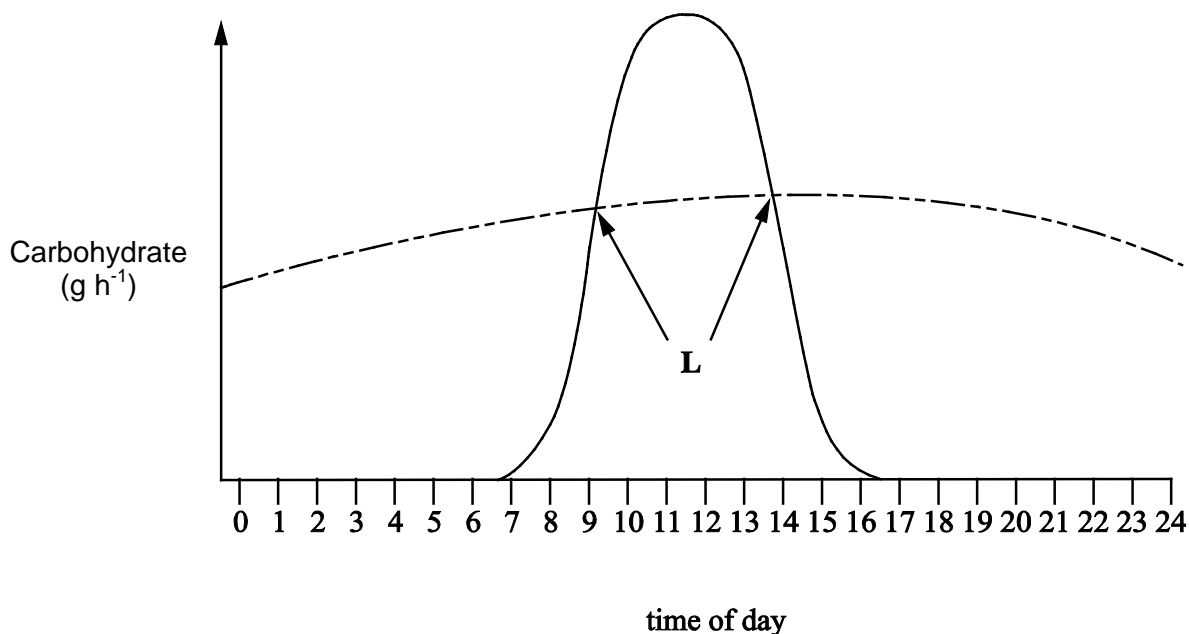
Mathematical concepts

Finding the y-intercept should be self-explanatory. Learners examine where the line crosses the y-axis and this value is the y-intercept. This is where $x=0$. For the x-intercept it is where the graph crosses the x-axis and is where $y=0$.

GCE Biology learners should also be able to find and read off the intersection of two or more data series from a graph.

Contexts in Biology

In biology learners may be asked for the y intercept in contexts such as osmosis. At A Level, learners could be asked for the intersection of two series of data, for example compensation points in plants:



Key

- rate of photosynthesis
----- rate of plant respiration

The compensation points (L) are where the rate of carbohydrate production from photosynthesis and the rate of carbohydrate use by respiration are equal.

M3.5 Calculate rate of change from a graph showing a linear relationship

Learners may be tested on their ability to:

- calculate a rate from a graph, e.g. rate of transpiration.

Mathematical concepts

To find the gradient the following formula is useful:

$$\text{Gradient} = \frac{\text{'change in y'}}{\text{'change in x'}}$$

The 'change in y' is the vertical step between two points and the 'change in x' is the horizontal step between the same two points. Both of these quantities could be negative and care has to be taken in these cases.

The principle is that two points are taken on the line of the graph. Measuring the horizontal distance between the points gives the change in x, and the vertical distance gives the change in y. The division according to the formula above gives the gradient.

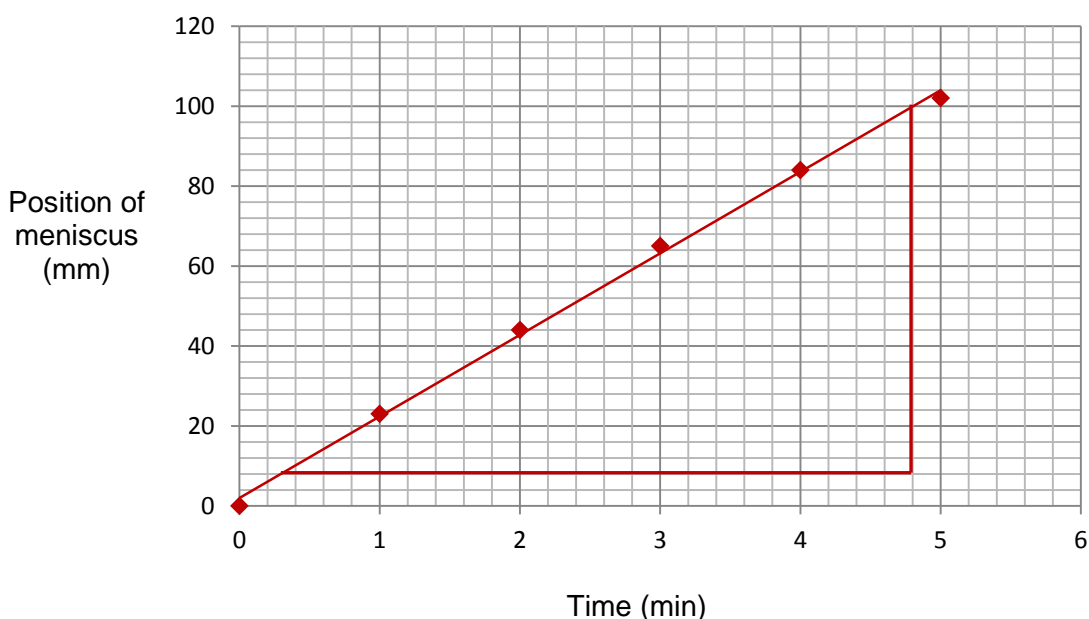
The gradient of a linear graph is always a measure of the rate of change of one variable with respect to another. The gradient of a linear graph with formula $y = mx + c$ measures the rate of change of y with respect to x. In words, the gradient expresses how quickly y changes as x changes.

A positive gradient means a quantity that increases as x increases whilst a negative gradient is a decreasing quantity as x increases.

Contexts in Biology

Biology learners may be asked to calculate the rate of change in various contexts, e.g. enzyme rate, diffusion rate or transpiration rate.

Transpiration was investigated using a potometer and the following graph was plotted:



To find the rate at which water is being lost the gradient has to be found. Choose two points on the line; (0.3, 8) and (4.8, 100).

The gradient can then be calculated by the gradient formula. The change in y is the vertical step to get from (0.3, 8) to (4.8, 100). This is the difference between the y coordinates; $100 - 8 = 92$. The change in x is the horizontal step to get from (0.3, 8) to (4.8, 100); the difference in the x coordinates, which is $4.8 - 0.3 = 4.5$. Therefore the gradient and therefore the rate can be calculated:

$$\text{Rate} = \text{Gradient} = \frac{92}{4.5} = 20.5 \text{ mm min}^{-1}$$

M3.6 Draw and use the slope of a tangent to a curve as a measure of rate of change

Learners may be tested on their ability to:

- use this method to measure the gradient of a point on a curve, e.g. quantity of product formed plotted against time when the concentration of enzyme is fixed.

Mathematical concepts

For linear graphs, the gradient is the same throughout and hence the rate of change is easy to accurately obtain (see previous sections). Non-linear graphs have an ever changing gradient and hence the rate of change will change from point to point. The rate of change at a particular point can be found mathematically using calculus, but this is not required in GCE Biology. Rather, a tangent is drawn by hand and eye to approximate the instantaneous rate of change at a particular point.

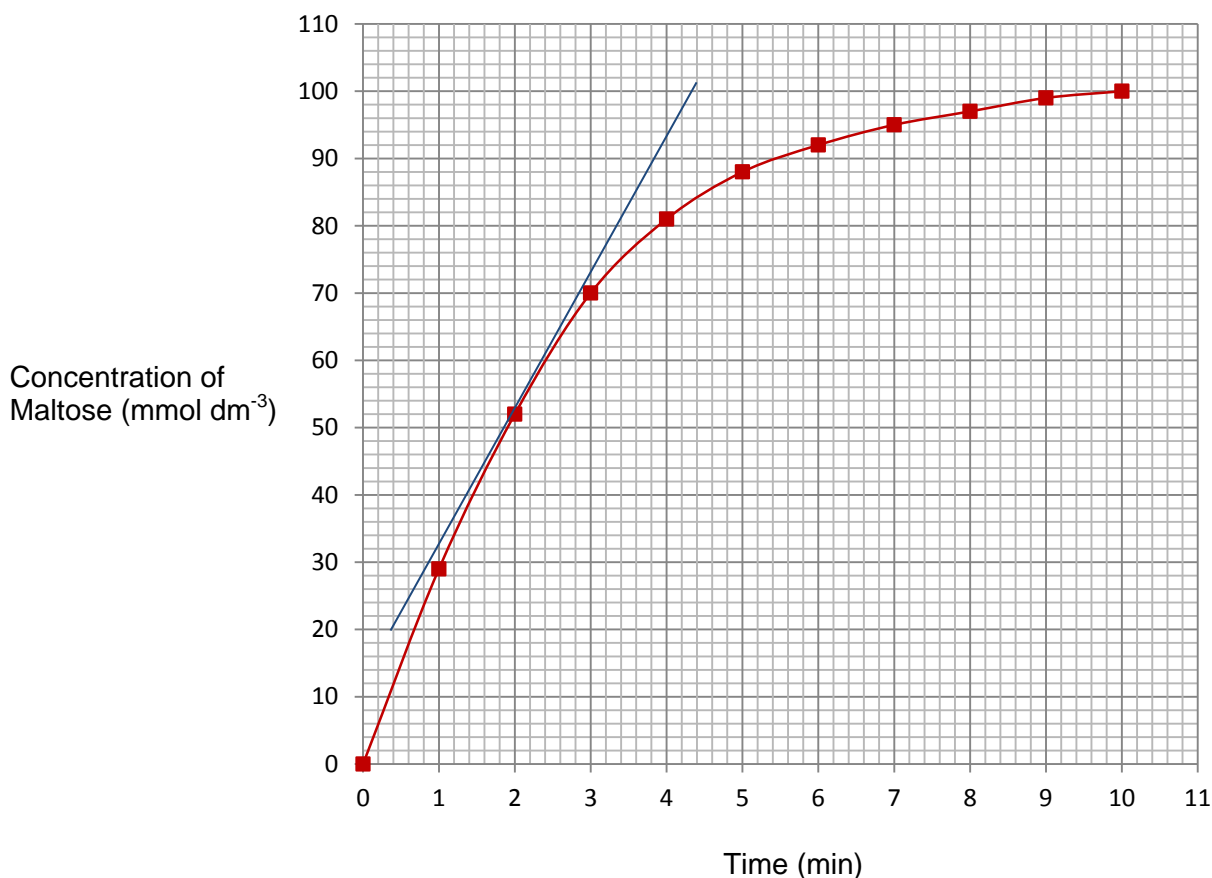
To draw a tangent accurately is tricky but there are a number of useful tips that can help:

- Use a ruler and pencil. This sounds obvious, but many learners will not do this naturally.
- Line the ruler up to the point where the tangent is supposed to be taken.
- While aligning the ruler, make sure that in the vicinity of the point **none** of the line of the curve is covered by the ruler. The aim is to have the entire curve visible as the line is drawn, otherwise the tangent will not be accurate.

Contexts in Biology

Rates of reaction

Below is a typical substrate concentration versus time curve for an enzyme controlled reaction.



To find the rate of reaction at 2 minutes, a tangent has to be drawn to approximate the gradient (which represents the rate) at that given time.

The gradient of the tangent is calculated as described in Section M3.5, by dividing the change in y by the change in x between two points.

For example, between the two points (1, 33) and (3.3, 80) on the tangent above which gives,

$$\text{gradient} = 47/2.3 = 20.4$$

So the rate of reaction at 2 minutes is $20.4 \text{ mmol dm}^{-3} \text{ min}^{-1}$.

M4 – Geometry and trigonometry

M4.1 Calculate the circumferences, surface areas and volumes of regular shapes

Learners may be tested on their ability to:

- Calculate the circumference and area of a circle
- Calculate the surface area and volume of rectangular prisms, of cylindrical prisms and of spheres
- e.g. calculate the surface area or volume of a cell.

Mathematical concepts

The calculations you need to know will have been covered in GCSE maths. You must note that the formulae for calculations involving circles and cuboids will not be provided in the exam, so you need to learn them!

So here are the geometry formulae you need to remember for the exam:

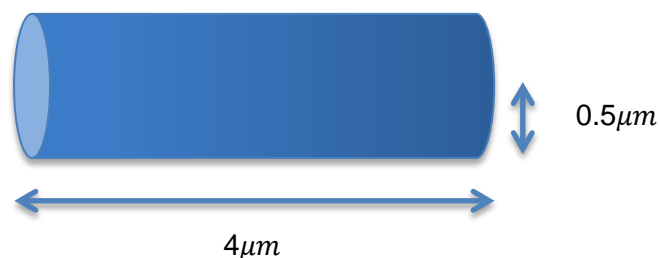
- 1) *Circumference* = $2\pi r$
- 2) *Area of circle* = πr^2
- 3) *Surface area of cuboid* = $2(bh + bl + hl)$
- 4) *Volume of Cuboid* = hbl

And here are the geometry formulae you will be given in the exam if you need them:

- 1) *Surface area of cylinder* = $2\pi r(r + l)$
- 2) *Volume of cylinder* = $\pi r^2 l$
- 3) *Surface area of sphere* = $4\pi r^2$
- 4) *Volume of sphere* = $\frac{4}{3}\pi r^3$

Contexts in Biology

Imagine a bacterial cell, for example the rod-shaped *Bacillus spp.*, is approximated as a cylinder with length $4\mu\text{m}$ and radius $0.5\mu\text{m}$:



The volume is calculated as:

$$V = \pi r^2 l = \pi \times 0.5^2 \times 4 = \approx 3.14 \mu\text{m}^3$$

The surface area is calculated as:

$$S.A = 2\pi r(r + l) = 2\pi \times 0.5(0.5 + 4) = \approx 14.1 \mu m^2$$

Therefore the surface area to volume ratio is 4.5:1.

Appendix A – Formulae learners need to be able to recall

GCSE (9–1) Mathematical formulae to recall

At AS and A Level Biology we assume knowledge of higher tier GCSE (9–1) Maths content. This includes (but is not limited to) the following list of formulae which learners will need to be able to recall.

Note that students should be familiar with the convention of using r for radius, h for height, b for base and l for length.

- Circumference of circle

$$\text{Circumference} = 2\pi r$$

- Area of circle

$$\text{Area of circle} = \pi r^2$$

- Surface area of cuboid

$$\text{Surface area of cuboid} = 2(bh + bl + hl)$$

- Volume of cuboid

$$\text{Volume of Cuboid} = hbl$$

- Mean

$$\bar{x} = \frac{\sum x}{n}$$

- Percentage (which can be used to calculate percentage change, percentage yield and percentage error)

$$\text{percentage change} = \frac{\text{new quantity} - \text{original quantity}}{\text{original quantity}} \times 100$$

$$\% \text{ yield} = \frac{\text{Actual Amount}}{\text{Theoretical Amount}} \times 100$$

$$\% \text{ error (uncertainty)} = \frac{2 \times \text{absolute uncertainty}}{\text{quantity measured}} \times 100\%$$

Biological formulae to recall

The following are the biological formulae learners will need to recall:

- Magnification (Bio A 2.1.1, Bio B 2.1.1)

$$\text{Magnification} = \frac{\text{size of image}}{\text{size of real object}}$$

- R_f (Bio A 2.1.2, Bio B 2.1.3)

$$R_f = \frac{\text{distance moved by the solute}}{\text{distance moved by the solvent}}$$

- Rates (e.g. enzymatic reactions, breathing (ventilation), transpiration, photosynthesis, respiration, reaction times, diffusion) (Bio A modules 2, 3 and 5, Bio B modules 2, 3 and 4)

$$\text{Rate} = \frac{\text{change in quantity}}{\text{time taken}}$$

- Surface Area to Volume ratio (Bio A 3.1.1, Bio B 2.2.1)

$$\text{Ratio} = \frac{\text{Surface Area}}{\text{Volume}}$$

- Genetic biodiversity (Bio A 4.2.1, Bio B 3.1.3)

$$\text{proportion of polymorphic gene loci} = \frac{\text{number of polymorphic gene loci}}{\text{total number of loci}}$$

- Cardiac output as a function of heart rate and stroke volume (Bio A 3.1.2, Bio B 2.2.1)

$$\text{cardiac output} = \text{heart rate} \times \text{stroke volume}$$

- Respiratory quotient (Bio A 5.2.2, Bio B 4.1.1)

$$RQ = \frac{\text{CO}_2 \text{ produced}}{\text{O}_2 \text{ consumed}}$$

- Microorganism population growth (Bio A only 6.2.1)

$$N = N_0 \times 2^n$$

- Efficiency of biomass transfers (Bio A 6.3.1, Bio B 4.3.1)

$$\text{efficiency} = \frac{\text{biomass transferred}}{\text{biomass intake}} \times 100$$

- Temperature coefficient (Q_{10}) (Bio A only 2.1.1)

$$Q_{10} = \frac{R_2}{R_1}$$

Appendix B – Formulae that will be provided in the assessments where needed

Learners should be aware of the following formulae. They should know how and when to use them.

Mathematical formulae that will need to be used but not recalled (provided in the assessment where needed).

- Surface area of a cylinder

$$\text{Surface area of cylinder} = 2\pi r(r + l)$$

- Volume of a cylinder

$$\text{Volume of cylinder} = \pi r^2 l$$

- Surface area of a sphere

$$\text{Surface area of sphere} = 4\pi r^2$$

- Volume of a sphere

$$\text{Volume of sphere} = \frac{4}{3}\pi r^3$$

- Chi squared

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

- Spearman's Rank Correlation Coefficient

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

- Standard Deviation

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

- Student's *t*-test – Unpaired

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

- Student's *t*-test – Paired

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

Note that critical values tables, or appropriate excerpts from these tables, will be provided in the assessment where needed.

Learners will need to be able to work out which 'degrees of freedom' or 'n' row, and which confidence column(s) is/are relevant to their analysis.

Biological formulae that will need to be used but not recalled (provided in the assessment where needed).

- The Hardy-Weinberg Equations

$$p^2 + 2pq + q^2 = 1$$

$$p + q = 1$$

- Simpson's Index

$$D = 1 - \left(\sum \left(\frac{n}{N} \right)^2 \right)$$

Appendix C – Key power laws

It is useful for learners to be aware of the following power laws to help in certain mathematical skills, as referenced in the text.

$$x^n \times x^m = x^{n+m} \quad \text{multiplicative rule}$$

$$\frac{x^n}{x^m} = x^{n-m} \quad \text{division rule}$$

$$(x^n)^m = x^{nm} \quad \text{power rule}$$

$$x^{-1} = \frac{1}{x^n} \quad \text{reciprocal rule}$$

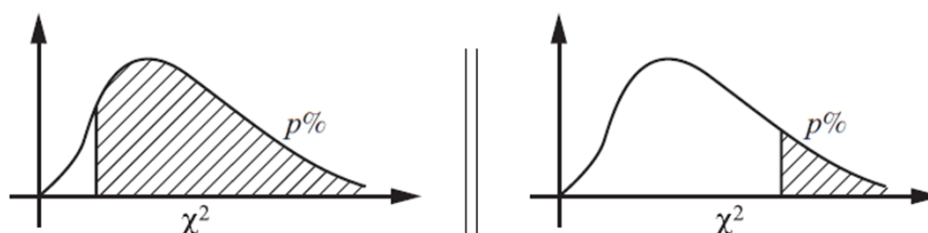
$$x^{n/m} = \sqrt[m]{x^n} \quad \text{root rule}$$

Appendix D – Statistical Tables

It is useful for learners to be aware of the following statistical tables and how to use them (M1.9).

Chi Squared

Percentage points of the χ^2 (chi-squared) distribution



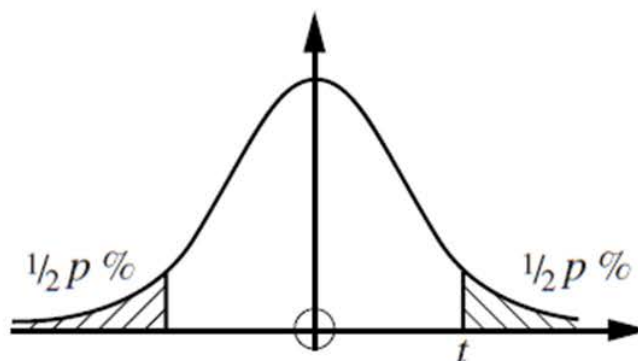
$p\%$	99	97.5	95	90	10	5.0	2.5	1.0	0.5
$v = 1$.0001	.0010	.0039	.0158	2.706	3.841	5.024	6.635	7.879
2	.0201	.0506	0.103	0.211	4.605	5.991	7.378	9.210	10.60
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34	12.84
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28	14.86
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09	16.75
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.646	2.180	2.733	3.490	13.36	15.51	17.53	20.09	21.95
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67	23.59
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21	25.19
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72	26.76
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22	28.30
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69	29.82
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14	31.32
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58	32.80
16	5.812	6.908	7.962	9.312	23.54	26.30	28.85	32.00	34.27
17	6.408	7.564	8.672	10.09	24.77	27.59	30.19	33.41	35.72
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.81	37.16
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
35	18.51	20.57	22.47	24.80	46.06	49.80	53.20	57.34	60.27
40	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
100	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2

Spearman's Rank Correlation Coefficient

Critical values for Spearman's rank correlation coefficient, r_s

n	5%	2½%	1%	½%	1-Tail Test	5%	2½%	1%	½%	
	10%	5%	2%	1%		10%	5%	2%	1%	
1	–	–	–	–	2-Tail Test	31	0.3012	0.3560	0.4185	0.4593
2	–	–	–	–		32	0.2962	0.3504	0.4117	0.4523
3	–	–	–	–		33	0.2914	0.3449	0.4054	0.4455
4	1.0000	–	–	–		34	0.2871	0.3396	0.3995	0.4390
5	0.9000	1.0000	1.0000	–		35	0.2829	0.3347	0.3936	0.4328
6	0.8286	0.8857	0.9429	1.0000		36	0.2788	0.3300	0.3882	0.4268
7	0.7143	0.7857	0.8929	0.9286		37	0.2748	0.3253	0.3829	0.4211
8	0.6429	0.7381	0.8333	0.8810		38	0.2710	0.3209	0.3778	0.4155
9	0.6000	0.7000	0.7833	0.8333		39	0.2674	0.3168	0.3729	0.4103
10	0.5636	0.6485	0.7455	0.7939		40	0.2640	0.3128	0.3681	0.4051
11	0.5364	0.6182	0.7091	0.7545		41	0.2606	0.3087	0.3636	0.4002
12	0.5035	0.5874	0.6783	0.7273		42	0.2574	0.3051	0.3594	0.3955
13	0.4835	0.5604	0.6484	0.7033		43	0.2543	0.3014	0.3550	0.3908
14	0.4637	0.5385	0.6264	0.6791		44	0.2513	0.2978	0.3511	0.3865
15	0.4464	0.5214	0.6036	0.6536		45	0.2484	0.2945	0.3470	0.3822
16	0.4294	0.5029	0.5824	0.6353		46	0.2456	0.2913	0.3433	0.3781
17	0.4142	0.4877	0.5662	0.6176		47	0.2429	0.2880	0.3396	0.3741
18	0.4014	0.4716	0.5501	0.5996		48	0.2403	0.2850	0.3361	0.3702
19	0.3912	0.4596	0.5351	0.5842		49	0.2378	0.2820	0.3326	0.3664
20	0.3805	0.4466	0.5218	0.5699		50	0.2353	0.2791	0.3293	0.3628
21	0.3701	0.4364	0.5091	0.5558		51	0.2329	0.2764	0.3260	0.3592
22	0.3608	0.4252	0.4975	0.5438		52	0.2307	0.2736	0.3228	0.3558
23	0.3528	0.4160	0.4862	0.5316		53	0.2284	0.2710	0.3198	0.3524
24	0.3443	0.4070	0.4757	0.5209		54	0.2262	0.2685	0.3168	0.3492
25	0.3369	0.3977	0.4662	0.5108		55	0.2242	0.2659	0.3139	0.3460
26	0.3306	0.3901	0.4571	0.5009		56	0.2221	0.2636	0.3111	0.3429
27	0.3242	0.3828	0.4487	0.4915		57	0.2201	0.2612	0.3083	0.3400
28	0.3180	0.3755	0.4401	0.4828		58	0.2181	0.2589	0.3057	0.3370
29	0.3118	0.3685	0.4325	0.4749		59	0.2162	0.2567	0.3030	0.3342
30	0.3063	0.3624	0.4251	0.4670		60	0.2144	0.2545	0.3005	0.3314

Percentage points of the *t*-distribution



<i>p</i> %	10	5	2	1	
<i>v</i> = 1	6.314	12.71	31.82	63.66	
2	2.920	4.303	6.965	9.925	
3	2.353	3.182	4.541	5.841	
4	2.132	2.776	3.747	4.604	
5	2.015	2.571	3.365	4.032	
6	1.943	2.447	3.143	3.707	
7	1.895	2.365	2.998	3.499	
8	1.860	2.306	2.896	3.355	
9	1.833	2.262	2.821	3.250	
10	1.812	2.228	2.764	3.169	
11	1.796	2.201	2.718	3.106	
12	1.782	2.179	2.681	3.055	
13	1.771	2.160	2.650	3.012	
14	1.761	2.145	2.624	2.977	
15	1.753	2.131	2.602	2.947	
20	1.725	2.086	2.528	2.845	
30	1.697	2.042	2.457	2.750	
50	1.676	2.009	2.403	2.678	
100	1.660	1.984	2.364	2.626	
∞	1.645	1.960	2.326	2.576	= Percentage points of the Normal distribution N(0, 1)



We'd like to know your view on the resources we produce. By clicking on the 'Like' or 'Dislike' button you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click 'Send'. Thank you.

Whether you already offer OCR qualifications, are new to OCR, or are considering switching from your current provider/awarding organisation, you can request more information by completing the Expression of Interest form which can be found here:

www.ocr.org.uk/expression-of-interest

OCR Resources: *the small print*

OCR's resources are provided to support the delivery of OCR qualifications, but in no way constitute an endorsed teaching method that is required by OCR. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources. We update our resources on a regular basis, so please check the OCR website to ensure you have the most up to date version.

This resource may be freely copied and distributed, as long as the OCR logo and this small print remain intact and OCR is acknowledged as the originator of this work.

OCR acknowledges the use of the following content:
Square down and Square up: alexwhite/Shutterstock.com

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications:
resources.feedback@ocr.org.uk

Looking for a resource?

There is now a quick and easy search tool to help find **free** resources for your qualification:

www.ocr.org.uk/i-want-to-find-resources/

www.ocr.org.uk/alevelreform

OCR Customer Contact Centre

General qualifications

Telephone 01223 553998

Facsimile 01223 552627

Email general.qualifications@ocr.org.uk

OCR is part of Cambridge Assessment, a department of the University of Cambridge. *For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored.*

© **OCR 2018** Oxford Cambridge and RSA Examinations is a Company Limited by Guarantee. Registered in England. Registered office 1 Hills Road, Cambridge CB1 2EU. Registered company number 3484466. OCR is an exempt charity.



Cambridge
Assessment

